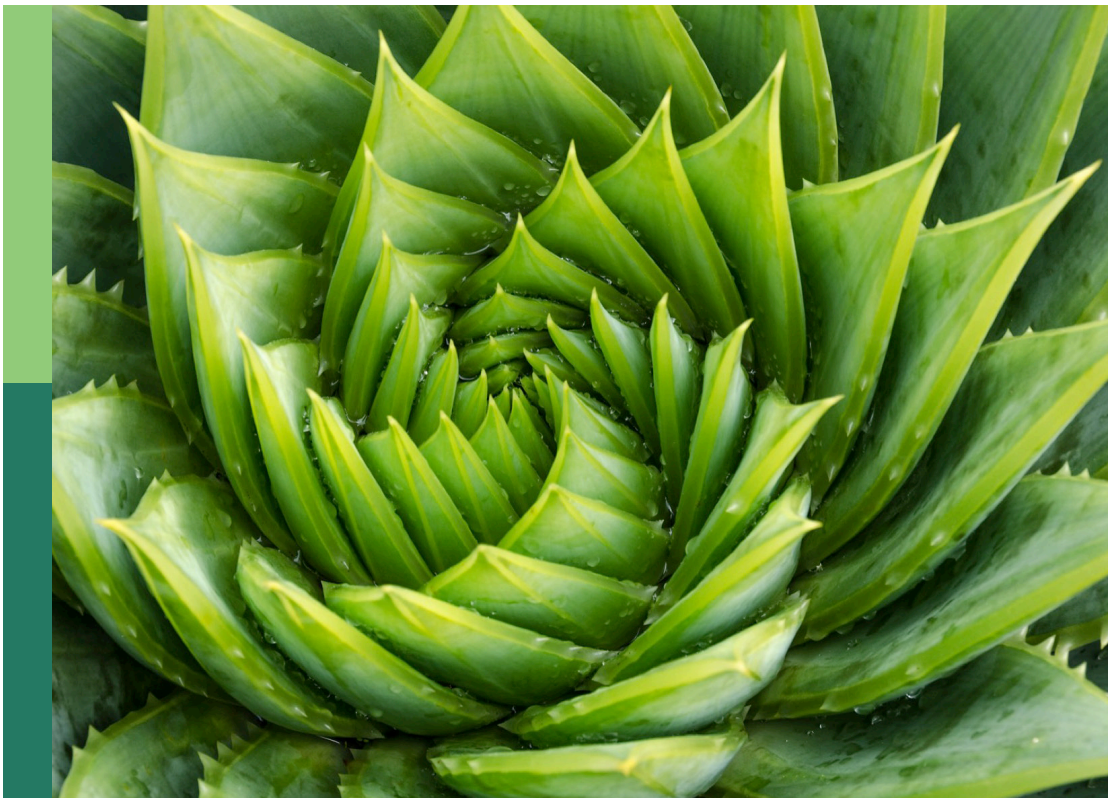
frontiers | Research Topics

Legumes for global food security, volume II

Edited by
Jose C. Jimenez-Lopez, Jaroslaw Czubinski, Karam B. Singh, Sergio J. Ochatt, Alfonso Clemente, Petr Smýkal and Eric Von Wettberg

Published in
Frontiers in Plant Science
Frontiers in Sustainable Food Systems



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-3604-9
DOI 10.3389/978-2-8325-3604-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Legumes for global food security, volume II

Topic editors

Jose C. Jimenez-Lopez — Tenured Scientist at CSIC Spanish National Research Council (CSIC). Estación Experimental del Zaidin; Department of Biochemistry, Cell & Molecular Biology of Plants, Spain

Jaroslaw Czubinski — Poznan University of Life Sciences, Poland

Karam B. Singh — Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Sergio J. Ochatt — INRA UMR1347 Agroécologie, France

Alfonso Clemente — Spanish National Research Council (CSIC), Spain

Petr Smýkal — Palacký University in Olomouc, Czechia

Eric Von Wettberg — University of Vermont, United States

Citation

Jimenez-Lopez, J. C., Czubinski, J., Singh, K. B., Ochatt, S. J., Clemente, A., Smýkal, P., Von Wettberg, E., eds. (2023). *Legumes for global food security, volume II*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-3604-9

Table of contents

06 Editorial: Legumes for global food security - volume II
Jose C. Jimenez-Lopez, Karam B. Singh, Alfonso Clemente, Jaroslaw Czubinski, Sergio Ochatt, Eric Von Wettberg and Petr Smýkal

11 Transcription Factor *GmWRKY46* Enhanced Phosphate Starvation Tolerance and Root Development in Transgenic Plants
Cheng Li, Kangning Li, Xinyi Liu, Hui Ruan, Mingming Zheng, Zhijie Yu, Junyi Gai and Shouping Yang

28 Morpho-Physiological Traits and Functional Markers Based Molecular Dissection of Heat-Tolerance in Urdbean
Debjyoti Sen Gupta, Partha S. Basu, J. Souframanien, Jitendra Kumar, P. Dhanasekar, Sanjeev Gupta, Muthaiyan Pandiyan, S. Geetha, P. Shanthi, Vaibhav Kumar and Narendra Pratap Singh

52 Lentil and Fava Bean With Contrasting Germination Kinetics: A Focus on Digestion of Proteins and Bioactivity of Resistant Peptides
Sara Bautista-Expósito, Albert Vandenberg, Elena Peñas, Juana Frias and Cristina Martínez-Villaluenga

71 Transient Nutrient Deficiencies in Pea: Consequences on Nutrient Uptake, Remobilization, and Seed Quality
Cécile Jacques, Marion Forest, Vincent Durey, Christophe Salon, Alain Ourry and Marion Prudent

85 Genetic Correlation Between Fe and Zn Biofortification and Yield Components in a Common Bean (*Phaseolus vulgaris* L.)
Santiago Diaz, Jose Polania, Daniel Ariza-Suarez, Cesar Cajiao, Miguel Grajales, Bodo Raatz and Stephen E. Beebe

98 Adaptive Mechanisms Make Lupin a Choice Crop for Acidic Soils Affected by Aluminum Toxicity
Miguel A. Quiñones, M. Mercedes Lucas and José J. Pueyo

106 A Bitter-Sweet Story: Unraveling the Genes Involved in Quinolizidine Alkaloid Synthesis in *Lupinus albus*
Claudia E. Osorio and Bradley J. Till

114 Effect of Drought and Low P on Yield and Nutritional Content in Common Bean
Millicent R. Smith, Eric Dinglasan, Erik Veneklaas, Jose Polania, Idupulapati M. Rao, Stephen E. Beebe and Andrew Merchant

128 Protein Biofortification in Lentils (*Lens culinaris* Medik.) Toward Human Health
Sonia Salaria, Jon Lucas Boatwright, Pushparajah Thavarajah, Shiv Kumar and Dil Thavarajah

142 Improving Association Studies and Genomic Predictions for Climbing Beans With Data From Bush Bean Populations
Beat Keller, Daniel Ariza-Suarez, Ana Elisabeth Portilla-Benavides, Hector Fabio Buendia, Johan Steven Aparicio, Winnyfred Amongi, Julius Mbiu, Susan Nchimbi Msolla, Phillip Miklas, Timothy G. Porch, James Burridge, Clare Mukankusi, Bruno Studer and Bodo Raatz

162 Genetic Mapping of the *Gmpgl3* Mutant Reveals the Function of GmTic110a in Soybean Chloroplast Development
Hui Yu, Qiushi Wang, Zhirui Zhang, Tao Wu, Xinjing Yang, Xiaobin Zhu, Yongheng Ye, Jiantian Leng, Suxin Yang and Xianzhong Feng

176 Breeding of Penta Null Soybean [*Glycine max* (L.) Merr.] for Five Antinutritional and Allergenic Components of Lipoxygenase, KTI, Lectin, 7S α′ Subunit, and Stachyose
Sang Woo Choi, Sarath Ly, Jeong Hwan Lee, Hyeon Su Oh, Se Yeong Kim, Na Hyeon Kim and Jong Il Chung

185 Prospects and Potential of *Bradyrhizobium diazoefficiens* Based Bio-Inoculants on Soybean Production in Different Agro-Ecologies of Mozambique
Canon Engoke Norris Savala, Alexander Nimo Wiredu, David Chikoye and Stephen Kyei-Boahen

201 Grain Nutrients Variability in Pigeonpea Genebank Collection and Its Potential for Promoting Nutritional Security in Dryland Ecologies
Dhanapal Susmitha, Thiyagarajan Kalaimagal, Ramachandran Senthil, Mani Vetriventhan, Swaminathan Manonmani, Prabhakaran Jeyakumar, Bellie Anita, Surender Reddymalla, Pushpajeet L. Choudhari, Chetna A. Nimje, Ovais H. Peerzada, Venkata Narayana Arveti, Vania C. R. Azevedo and Kuldeep Singh

221 Genome-wide association analysis of stress tolerance indices in an interspecific population of chickpea
Shweta Kalve, Krishna Kishore Gali and Bunyamin Tar'an

234 Estimates of genotypic and phenotypic variance, heritability, and genetic advance of horticultural traits in developed crosses of cowpea (*Vigna unguiculata* [L.] Walp)
Haitham E. M. Zaki and Khlode S. A. Radwan

251 A QTL approach in faba bean highlights the conservation of genetic control of frost tolerance among legume species
Estefanía Carrillo-Perdomo, Jean-Bernard Magnin-Robert, Blandine Raffiot, Chrystel Deulvot, Matthieu Floriot, Isabelle Lejeune-Hénaut, Pascal Marget, Judith Burstin, Nadim Tayeh and Grégoire Aubert

264 Adaptability and stability for soybean yield by AMMI and GGE models in Ethiopia
Mesfin Hailemariam Habtegebriel

283 Memory or acclimation of water stress in pea rely on root system's plasticity and plant's ionome modulation
Cécile Jacques, Sylvie Girodet, Fanny Leroy, Sylvain Pluchon, Christophe Salon and Marion Prudent

300 *SUPERMAN* strikes again in legumes
Ana L. Rodas, Edelín Roque, Rim Hamza, Concepción Gómez-Mena, José Pío Beltrán and Luis A. Cañas

307 Quantitative trait loci associated with amino acid concentration and *in vitro* protein digestibility in pea (*Pisum sativum* L.)
Junsheng Zhou, Zhongyang Wan, Krishna Kishore Gali, Ambuj Bhushan Jha, Michael T. Nickerson, James D. House, Bunyamin Tar'an and Thomas D. Warkentin

321 Iron accumulation and partitioning in hydroponically grown wild and cultivated chickpea (*Cicer arietinum* L)
Tamanna A. Jahan, Shweta Kalve, Zachery Belak, Christopher Eskiw and Bunyamin Tar'an

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 18 September 2023
DOI 10.3389/fpls.2023.1273600

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Paul Gepts,
University of California, Davis, United States

*CORRESPONDENCE
Jose C. Jimenez-Lopez
josecarlos.jimenez@eez.csic.es

RECEIVED 06 August 2023
ACCEPTED 07 September 2023
PUBLISHED 18 September 2023

CITATION
Jimenez-Lopez JC, Singh KB, Clemente A, Czubinski J, Ochatt S, Von Wettberg E and Smýkal P (2023) Editorial: Legumes for global food security - volume II. *Front. Plant Sci.* 14:1273600. doi: 10.3389/fpls.2023.1273600

COPYRIGHT
© 2023 Jimenez-Lopez, Singh, Clemente, Czubinski, Ochatt, Von Wettberg and Smýkal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Legumes for global food security - volume II

Jose C. Jimenez-Lopez[1]*, Karam B. Singh[2], Alfonso Clemente[3], Jaroslaw Czubinski[4], Sergio Ochatt[5], Eric Von Wettberg[6] and Petr Smýkal[7]

[1]Department of Stress, Development and Signaling in Plants, Estación Experimental del Zaidín, Spanish National Research Council (CSIC), Granada, Spain, [2]Agriculture and Food, Commonwealth Scientific and Industrial Research Organization (CSIRO), Perth, WA, Australia, [3]Department of Nutrition and Sustainable Animal Production, Estación Experimental del Zaidin, Spanish National Research Council (CSIC), Granada, Spain, [4]Department of Biochemistry and Food Analysis, Poznan University of Life Sciences, Poznan, Poland, [5]Agroécologie, Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement (INRAE), Institut Agro, University Bourgogne, University Bourgogne Franche-Comté, Dijon, France, [6]Department of Plant and Soil Science, University of Vermont, Burlington, VT, United States, [7]Department of Botany, Palacký University in Olomouc, Olomouc, Czechia

KEYWORDS

legumes, global food security, crop improvement, nutrient management, sustainable agriculture, food and feed security, climate resilience

Editorial on the Research Topic
Legumes for global food security - volume II

As the world faces an array of global challenges, including population growth, climate change, and the need for clean energy, the role of legumes in addressing these issues becomes increasingly significant. The Research Topic, *"Legumes for Global Food Security, Volume II"* aims to explore the diverse contributions of legumes in promoting sustainable agriculture and enhancing global food security.

Legumes play a crucial role in delivering vital services to societies worldwide. One of their primary contributions lies in their capacity to provide a wide range of food crops that serve as essential sources of plant-based proteins, addressing the challenge of food security for a rapidly growing population. Furthermore, grain legumes possess remarkable nutritional properties and act as cost-effective food choices, playing a pivotal role in achieving global food and feed security amid the growing world population. The significance of legumes extends beyond their role as a food source. Through rhizobial symbiosis, legumes have the unique ability to fix atmospheric nitrogen, enriching agro-ecosystems and boosting subsequent crop productivity by enhancing water and nutrient capture and recycling. Moreover, they play a key role in mitigating climate change, offering an alternative to synthetic nitrogenous fertilizers, which is energy-intensive to produce and release greenhouse gases upon breakdown. Furthermore, legumes contribute to the reduction of fossil fuel usage by providing biofuel feedstocks and industrial resources. Given the challenges posed by increasing climatic stresses, legumes' genetic diversity equips them to thrive in various environments, rendering them resilient and ideal for sustainable intensification on small-scale and resource-constrained farms. Moreover, they play a vital role as biocontrol agents, effectively combating pests and diseases that could otherwise cause significant agricultural losses.

This Research Topic aims to explore the multifaceted contributions of legumes in the development of robust and efficient agro-ecosystems, thereby enhancing global food security. Here we summarize some of the highlights derived from the 22 articles published in this Research Topic, dividing them in four main topics, in order to better understand how research on legumes and related crops is contributing to crop improvement, adaptation, and nutrient management, ultimately aiming to address food security challenges and support sustainable agriculture.

Regarding crop adaptation to abiotic stress and improvement of abiotic stress tolerance, Gupta et al. assessed heat tolerance in urdbean genotypes to identify heat-tolerant cultivars suitable for cultivation during the summer season. In this paper, 97 diverse genotypes of urdbean (*Vigna mungo* L. Hepper) were assessed for yield under heat stress and non-stress conditions to identify heat-tolerant genotypes. Eight genotypes were highly heat tolerant, while 35 were highly heat sensitive. Physiological and biochemical characterization of selected genotypes under heat stress revealed variations in leaf nitrogen balance index, chlorophyll content, flavonols, and anthocyanin contents. Heat-tolerant genotypes exhibited higher membrane stability index and superior photosynthetic ability. Molecular characterization distinguished genetic differences between heat-tolerant (UPU 85-86) and heat-sensitive (PKGU 1) genotypes.

Then, Diaz et al. investigated the genetic relationship between iron and zinc concentration and yield in common beans to develop biofortified varieties with higher micronutrient levels. In their study, they focused on common bean (*Phaseolus vulgaris* L.), an important legume that provides cost-effective proteins and micronutrients, especially iron and zinc. To combat malnutrition in developing countries, biofortification aims to develop varieties with higher Fe/Zn content. However, breeders faced challenges due to the negative correlation between Fe/Zn concentration and yield. Using QTL mapping and GWAS analysis on biofortified parent populations, they identified 79 QTLs and 23 hotspot regions that showed QTLs with opposing effects on yield components and Fe/Zn accumulation. This allowed them to select specific QTLs to enhance Fe/Zn levels without compromising yield in biofortified cultivars.

Alkaloids, a diverse group of cyclic nitrogen-containing secondary metabolites, are present in over 20% of plant species, including *Lupinus albus* (white lupin), which naturally contains quinolizidine alkaloids (QAs) in seeds. While QAs provide natural protection, lupin-breeding programs have unintentionally selected against them due to limited understanding of the QA biosynthetic pathway. The review by Osorio and Till discusses the current state of research, focusing on natural mutations like the pauper locus and the use of molecular markers and sequencing technology to identify candidate genes. Precision breeding of low-alkaloid, high-nutrition white lupin can be achieved by understanding QA biosynthesis, essential for sustainable agricultural productivity and high-quality protein for food and feed.

Then, research performed by Woo Choi et al. aimed to improve soybean food quality by breeding a new soybean line with penta null recessive alleles for five antinutritional and allergenic components: lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose. The breeding resulted in a soybean strain with the penta null (lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2) genotype, devoid of the mentioned proteins and with low stachyose content. The new strain exhibited desirable agronomic traits, including purple flowers, tawny pubescence, determinate growth habit, and yellow pods at maturity. The 100-seed weight was 31.1 g, and the yield was 2.80 t/ha, making it a promising soybean line with improved nutritional attributes.

Finally, the paper by Li et al. focused on the important area of plant responses phosphate deficiencies. Phosphorus is a key micronutrient for crops and large areas of arable land (30-40%) are limited by phosphorus availability. This study identified *GmWRKY46* as crucial for enhanced phosphate starvation tolerance in soybean. It was induced in response to phosphate starvation, particularly in roots. Overexpression of *GmWRKY46* in transgenic Arabidopsis and soybean composite plants showed its involvement in root development, leading to increased phosphate uptake and improved growth under phosphate starvation. RNAseq and ChIP-qPCR analysis identified differentially expressed genes, including *AtED1*, which improved tolerance to phosphate starvation when overexpressed in transgenic Arabidopsis. *GmWRKY46* directly binds to a W box in the *AtED1* promoter. Overall, *GmWRKY46* represents a promising target for conventional breeding and transgenic approaches to enhance soybean's phosphate starvation tolerance.

The second main theme of this Research Topic is genetic variability and yield improvement. In this regard, a study performed by Susmitha et al. assessed the genetic variability of grain nutrients in 600 pigeonpea germplasms from the RS Paroda Genebank, ICRISAT, India, aiming to identify potential sources for biofortification. Field trials in 2019 and 2020 revealed significant differences in agronomic traits and grain nutrients. Germplasms showed wide variation in days to 50% flowering, days to maturity, 100-seed weight, and grain yield per plant. Grain nutrients like protein, minerals (P, K, Ca, Mg, Fe, Zn, Mn, Cu) exhibited substantial variation. Germplasms from Asia had diverse grain nutrient levels, while those from Africa showed high nutrient density. Some germplasms exhibited favorable nutrient profiles and can serve as promising sources for developing biofortified lines with improved agronomic traits. The study's phenotypic data can aid genetic improvement efforts through GWAS and SNP/haplotype-based approaches.

Then, Habtegebriel performed a study focused on genotype by environment interaction (GEI) in soybean, which affects breeding progress by hindering the selection of superior cultivars. The objective was to identify adapted and stable genotypes and explore potential mega-environments for future testing. Experiments were conducted over two years, evaluating yield components and other traits. Stability analysis using GGE biplot and AMMI model identified five genotypes as top performers. Genotypes JM-CLK/CRFD-15-SD (G8) and 5002T (G1) displayed the highest seed yield, making them suitable as parents for hybridization and commercial production. These genotypes are

recommended for release as new soybean varieties for cultivation across various environments.

Rodas et al. investigated the *SUPERMAN* (*SUP*) gene, initially discovered in *Arabidopsis thaliana*, that maintains boundaries between reproductive organs, regulating stamen and carpel number in flowers. This study focuses on the *SUP* ortholog, *MtSUP*, in the legume *Medicago truncatula*. *M. truncatula* serves as a model to study unique developmental traits in legumes, like compound inflorescence and complex floral development. *MtSUP* plays a role in the genetic network controlling these processes, sharing conserved functions with *SUP* but also exhibiting context-specific novel functions in legumes. *MtSUP* controls flower number and determinacy of ephemeral meristems, providing insights into compound inflorescence and flower development in legumes. Understanding these genetic controls can benefit legume breeding for improved crop production and food security.

The interest in plant proteins for human and animal nutrition continues to grow with legume crops being at the forefront of efforts to breed crops with higher protein levels and better quality. The paper by Zhou et al. assessed two important protein quality traits in pea (*Pisum sativum* L.), a major grain legume grown widely worldwide. They used field trials spread over 3 years and in multiple locations to assess the amino acid profiles and protein digestibility in a pea recombinant inbred line (RIL) population consisting of 110 lines. While pea has reasonable levels of protein, it is low in some amino acids such as tryptophan, methionine and cysteine. The authors used near-infrared spectroscopy to measure the amino acid profiles in the RILs and showed that this method had advantages for such studies over other methods used previously, being a high-throughput, low cost and non-destructive method. The use of the method, coupled with the use of multiple sites and a three-year period, helped to identify QTLS associated with key amino acids as well as QTLS associated with protein digestibility. These QTLs and the improved method for measuring seed proteins will help with efforts to breed pea cultivars with improved nutritional qualities.

Given that their production in the tropics is located on marginal, rainfed lands and hence yield is low, Smith et al. assessed the adaptation of 12 bush bean (*Phaseolus vulgaris* L.) genotypes to individual and combined stress induced by drought and low P availability through their physiological and chemical responses. Thus, seed weight and aerial biomass were very significantly decreased by both stress factors, whether applied individually or in combination, and was coupled with an also decreased photosynthesis and stomatal conductance. Differences between genotypes were apparent though, and the common bean genotypes SEF60 and NCB226 were more resilient and better adapted to stress than the commercial control DOR390 providing a higher yield, but there were no statistically significant differences in the concentration of mineral nutrients and amino acids within the seed at harvest. Under water deficit, carbon assimilation and water use in leaves were reduced throughout development, while combined low P and water deficit resulted in significant changes in the concentration of key nutrients and amino acids in the soluble leaf fraction but did not impact the seed. Those results suggest that common bean genotypes have a certain degree of resilience in terms of yield as expressed through pod harvest index and conservation of seed nutritional content.

Keller et al. examined genetic variation for flowering time, yield, seed iron content, and growth habit in climbing common beans. In general, there has been paid less attention to the genetic basis of agronomic traits in climbing types than bush types of common bean, and the authors set out to correct this with an examination of 17 field trials and reanalysis of 16 previously published datasets from both the Meso-American and Andean genepools of common bean. In addition to finding a number of marker-trait associations, the authors show that genomic prediction is improved when bush bean information is included in models predicting climbing bean genotype-phenotype associations.

Finally, Zaki and Radwan studied the genotypic and phenotypic variance of five parental genotypes and six generated crosses in F1 and F2 generations. The results indicated a substantial degree of genotypic variability across the investigated variables, which was considerably greater than the phenotypic variance. The authors succeeded in obtaining crosses that possess promising potential for high yield.

The third topic includes five papers about improvement of nutritional quality of legumes. Salaria et al. reviewed the potential pathways for increase seed protein content, the role of anti-nutritional compounds, and the extent of genetic variation in cultivated lentil and its cross-compatible wild relatives. The authors discussed the need for careful phenotyping, and explored a range of breeding approaches considered as avenues for improvement including speed breeding, genomic selection, and genetic engineering.

To investigate the kinetics of iron (Fe) uptake and partitioning in chickpea, Jahan et al. used a hydroponic growth system and RNA-sequencing of six genotypes that vary in seed Fe content. A number of key transporters were found to be expressed in roots and leaves, with the genes *FRO2 and IRT1* being important in roots in the presence of Fe, and *GCN2* in low Fe conditions. Conversely, in leaves the genes *NRAMP3, V1T1, YSL1* along with storage gene *FER3* showed higher expression. The improved understanding of Fe dynamics provided by this work provides targets for efforts to increase chickpea seed Fe content under both high and low soil Fe conditions.

Carrillo-Pedromo et al. investigated cold tolerance in fava bean. They used two fava bean QTL mapping populations to investigate the genetic basis of cold tolerance. This investigation identified five genomic regions associated with improved overwintering tolerance. Investigation of synteny of these regions with the *Pisum* and *Medicago* genomes showed that these regions are also associated with cold tolerance across other closely related legumes, suggesting shared mechanisms of tolerance that could prove to be effective breeding targets.

Bautista-Expósito et al. analyzed the impact of germination on protein and phenolic compound profiles in lentil and fava bean seeds. In this paper, the authors decided to analyze how the profile of phenolic compounds in different seeds affects the duration of germination and digestibility of proteins. Hydrolysis of the main protein fractions (7S and 11S globulins) that occurred during germination resulted in a simultaneous increase in the content of

peptides and free amino acids. In addition, the products of protein hydrolysis were tested by the authors for potential health-promoting properties, i.e., antihypertensive and antioxidant activities. The results of this study clearly showed that regardless of the type of seed, the germination process contributes to increase of protease activity and reduction of the level of phytic acid, trypsin inhibitors, and tannins which are considered antinutrients. Furthermore, the key role of seed permeability on the speed of the germination process, which subsequently influences the degree of antinutrients and the change in seed microstructure and endogenous proteolysis, was clearly demonstrated.

Finally, the work performed by Kalve et al. evaluated an interspecific population of chickpeas in terms of genome-associated stress tolerance indices. The approach used by the authors assumed the application of wild relatives as a source of novel alleles for adaptation to suboptimal environments. In particular, in their work, an interspecific population derived from *Cicer reticulatum* accessions was used as a donor of introgression of heat and drought tolerance. From the initial 600 interspecific lines that were generated and tested in terms of resistance to ascochyta blight, a 195-line subset was selected and studied. Subsequent analyses helped establish a set of individual lines that perform better under suboptimal conditions. The additional outcome of this research can be associated with identifying specific SNP markers that can be used as marker-assisted selection that can help identify genes underlying tolerance to abiotic stress.

The last five papers of this Research Topic were focused on the use of bio-inoculants and soil fertility improvement. Norris Savala et al. assessed the efficacy of bio-inoculants in soybean production in Mozambique. In this context, to improve soybean yield, farmers use bio-inoculants from various sources and agroecological adaptability. However, these bio-inoculants are often unavailable during planting time and vary in yield based on their source and handling. Mozambique relies on imported bio-inoculants from neighboring countries and even South America. In this study, seven *Bradyrhizobium diazoefficiens* strain-based bio-inoculants with different carrier materials were evaluated for their performance, adaptability, and soybean productivity. Inoculation significantly improved plant growth, nodulation, and yield, suggesting the potential of bio-inoculants to enhance soybean production in Mozambique.

Jacques et al. conducted two studies related to pea plant responses. The first study focused on the impact of mineral deprivation on nutrient content and remobilization. They imposed transient deprivation of 13 mineral nutrients during vegetative growth and observed preferential allocation of dry weight and elements to shoots, particularly tendrils. Different remobilization strategies were identified, and the study suggests strategies to enhance seed quality through precise fertilization during periods of mineral nutrient deficiency. The second study by Jacques et al. investigated pea plant responses to various water stress types and their effects on nutrient uptake and remobilization. Pea plants, being nutritionally important, face vulnerability to water deficits induced by climate change. Common responses to all water stress types were observed in shoots, with manganese (Mn) playing a significant role. Under continuous stress, boron (B) impacted root architecture. An "ecophysiological imprint" in the root system was also observed, leading to increased nodule numbers during the recovery period. These findings provide insights into plant strategies to cope with water stress and contribute to global food security and nutrient deficiencies while reducing reliance on animal products.

Agricultural soils are affected worldwide by acidic pH and high levels of aluminum (Al) contamination due to several factors including agriculture practices and climate change. Quinones et al. have reported the ability of lupin to tolerate and accumulate Al in the rhizosphere and inside the root cells, suggesting its use in the restoration of marginal acid Al-rich soils in temperate zones where other legumes are unable to grow. The authors describe several physiological and molecular mechanisms of tolerance, uptake and accumulation of Al by lupin; among them, cluster roots which are able to exudate organic acids anions and polyphenols as well as Al-tolerant rhizobia strains with ability to produce abundant exopolysaccharides. These adaptive mechanisms make lupin a suitable crop for acidic soils affected by Al toxicity.

Finally, Yu et al. identified the *GmTic110a* gene's role in chloroplast development and its impact on soybean growth. More concretely, in this study, a Glycine max pale green leaf 3-1 (Gmpgl3-1) mutant with reduced chlorophyll content, chloroplast defects, decreased yields, and fewer pods per plant was isolated from soybean. Bulked segregant analysis and map-based cloning identified a mutation in the chloroplast development-related GmTic110a gene. Knockout plants exhibited similar phenotypes to the mutant. *GmTic110a* was highly expressed in leaves and localized to the inner chloroplast membrane. Interaction experiments revealed GmTic110a's interaction with other proteins involved in chloroplast development (GmTic20, GmTic40a, and GmTic40b). These findings suggest the crucial role of *GmTic110a* in chloroplast development, impacting photosynthesis and soybean growth.

In conclusion, the Research Topic *"Legumes for Global Food Security, Volume II"* delves into the immense potential of legumes, paving the way for innovative and transformative solutions to address pressing 21st-century challenges. These 22 papers highlight the crucial role of legumes in providing essential plant-based proteins, supporting food and feed security in a cost-effective manner. Moreover, the diverse topics covered offer valuable insights to enhance crop improvement, adaptation, and nutrient management, ultimately contributing to global food security and sustainable agriculture. As a collective, these articles make a substantial contribution to understanding and breeding resilient legumes capable of withstanding climate challenges. Our hope is that these findings will significantly enhance global food security in the foreseeable future.

Author contributions

JJ-L: Writing – original draft, Writing – review & editing. KS: Writing – original draft, Writing – review & editing. AC: Writing –

original draft, Writing – review & editing. JC: Writing – original draft, Writing – review & editing. SO: Writing – original draft, Writing – review & editing. EW: Writing – original draft, Writing – review & editing. PS: Writing – original draft, Writing – review & editing.

Funding

JJ-L thanks the support of the European Research Program MARIE CURIE (FP7-PEOPLE-2011-IOF),grant ref. PIOF-GA-2011-301550; the Spanish Ministry of Economy, Industry and Competitiveness (Ramon y Cajal Research Program), grant ref. RYC-2014-16536; the CSIC intramural research program, grant ref. 202240I002; and the Spanish Ministry of Science and Innovation, grant ref. number CPP2021-008989. AC was supported by PY20_00242 grant, financed by the Consejeria de Trasnformación económica, Industria, Conocimiento y Universidades de la Junta de Andalucía (Spain) and FEDER funds A way to make Europe.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 September 2021
doi: 10.3389/fpls.2021.700651

Transcription Factor *GmWRKY46* Enhanced Phosphate Starvation Tolerance and Root Development in Transgenic Plants

Cheng Li[1,2], *Kangning Li*[2,3], *Xinyi Liu*[1,2], *Hui Ruan*[1,2], *Mingming Zheng*[1,2], *Zhijie Yu*[1,2], *Junyi Gai*[1,2] *and Shouping Yang*[1,2]*

[1] Soybean Research Institute, National Center for Soybean Improvement, Key Laboratory of Biology and Genetic Improvement of Soybean (General, Ministry of Agriculture), Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, China, [2] State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China, [3] MOA Key Laboratory of Plant Nutrition and Fertilization in Lower-Middle Reaches of the Yangtze River, Nanjing Agricultural University, Nanjing, China

OPEN ACCESS

Edited by:
Karam B. Singh,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Reviewed by:
Shisong Ma,
University of Science and Technology of China, China
Yuanyuan Zhang,
Institute of Botany, Chinese Academy of Sciences, China

***Correspondence:**
Shouping Yang
spyung@126.com

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: 10 June 2021
Accepted: 20 August 2021
Published: 14 September 2021

Citation:
Li C, Li K, Liu X, Ruan H, Zheng M, Yu Z, Gai J and Yang S (2021) Transcription Factor GmWRKY46 Enhanced Phosphate Starvation Tolerance and Root Development in Transgenic Plants. Front. Plant Sci. 12:700651. doi: 10.3389/fpls.2021.700651

Phosphorus (P) is one of the essential macronutrients, whose deficiency limits the growth and development of plants. In this study, we investigated the possible role of *GmWRKY46* in the phosphate (Pi) starvation stress tolerance of soybean. GmWRKY46 belonged to the group III subfamily of the WRKY transcription factor family, which was localized in the nucleus and had transcriptional activator activity. *GmWRKY46* could be strongly induced by Pi starvation, especially in soybean roots. Overexpression of *GmWRKY46* significantly enhanced tolerance to Pi starvation and lateral root development in transgenic *Arabidopsis*. RNA-seq analysis showed that overexpression of *GmWRKY46* led to change in many genes related to energy metabolisms, stress responses, and plant hormone signal transduction in transgenic *Arabidopsis*. Among these differential expression genes, we found that overexpression of *AtAED1* alone could enhance the tolerance of transgenic *Arabidopsis* to Pi starvation. Y1H and ChIP-qPCR analyses showed that GmWRKY46 could directly bind to the W-box motif of the *AtAED1* promoter *in vitro* and *in vivo*. Furthermore, results from intact soybean composite plants with *GmWRKY46* overexpression showed that *GmWRKY46* was involved in hairy roots development and subsequently affected plant growth and Pi uptake. These results provide a basis for the molecular genetic breeding of soybean tolerant to Pi starvation.

Keywords: soybean (*Glycine max*), *GmWRKY46*, phosphate starvation, root development, RNA-Seq, *AED1*

INTRODUCTION

Phosphorus (P) is one of the most essential mineral nutrients required for the growth and development of plants and is a constituent of key molecules such as ATP, nucleic acids, and phospholipids (Chiou and Lin, 2011). It plays a crucial role in energy generation, photosynthesis, glycolysis, respiration, protein activation, and stability. Plants meet their P requirement only by taking up inorganic phosphate (Pi) from the soil. Although P is abundant in many soils, it is rarely present in the form of Pi that can be used by plants, so crop yield on 30–40% of the arable land of the world is limited by P availability (Vance et al., 2003). At present, intensive application of chemical

fertilizers containing Pi has become a standard agricultural practice to ensure crop productivity (Chiou and Lin, 2011). However, the excess Pi application is not a perfect solution. On the one hand, the excess dissolution of Pi pollutes water sources (MacDonald et al., 2011). On the other hand, most of the annually fertilized Pi is fixed in the soil in organic forms which are unavailable to plants in the absence of mineralization (Raghothama, 1999). To cope with Pi deficiency, plants themselves have evolved several ways to optimize Pi acquisition from soil (Cong et al., 2020). For example, stimulating lateral root and hairy root growth leads to profound changes in root structure, thereby increasing the Pi absorption surface of the root system (Williamson et al., 2001; Ticconi and Abel, 2004; Svistoonoff et al., 2007), releasing organic acids and phosphatases to release Pi by dissolving organic P (Jones, 1998; Hinsinger, 2001), and getting more Pi by symbiosis with mycorrhizal fungi (Javot et al., 2007). In-depth research on the mechanism of how plants resist Pi starvation can provide effective references to future plant production.

Phosphate starvation responses in plants are tightly regulated by an elaborate signaling network that comprises multiple components that are not fully understood. In this network, transcription factors (TFs) act as significant coordinators to transfer stress signals and to orchestrate the expression of their target genes (Huang et al., 2018). They play crucial roles in protecting against stress-associated damage by modulating the expression level of downstream target genes (Vigeolas et al., 2008). In recent years, researchers have used molecular biology methods to find that many families of TFs play important roles in plant signaling responses to low Pi including WRKY, MYB, AP2/ERF, and bHLH family members (Baek et al., 2013; Dai et al., 2015; Yang et al., 2016a; Chen et al., 2018). The WRKY TF family is one of the largest TF families in plants, which, based on the highly conserved WRKY domain, has 72 members in *Arabidopsis* (*Arabidopsis thaliana*) and 182 members in soybean (*Glycine max*) (Eulgem and Somssich, 2007; Bencke-Malato et al., 2014). The WRKY domain is a conserved DNA-binding region that includes highly conserved WRKYGQK peptide sequences and zinc finger motifs which can be either C_2H_2-type ($Cx_{4-5}Cx_{22-23}HxH$) or C_2HC-type ($Cx_7Cx_{23}HxC$) (Rushton et al., 2010). Members of the WRKY family have been found to contain at least one such domain. The WRKY domain generally binds to the promoter region of target genes containing the W-box(es) sequence (C/TTGACT/C), although alternative binding sites have been identified (Pandey and Somssich, 2009). WRKY TFs are involved in responses to biotic and abiotic stresses, and developmental processes (Eulgem and Somssich, 2007; Pandey and Somssich, 2009). It has been found that some WRKY TFs are particularly sensitive to Pi starvation in plants. For example, *AtWRKY45* overexpression in *Arabidopsis* increased the Pi content and uptake, while RNA interference suppression of *AtWRKY45* decreased the Pi content and uptake (Wang et al., 2014). Overexpression of *OsWRKY74* significantly enhanced tolerance to Pi starvation in rice (*Oryza sativa*), whereas transgenic lines with downregulation of *OsWRKY74* were sensitive to Pi starvation (Dai et al., 2015).

Aspartic proteinases (APs) are an important class of proteolytic enzymes, which occur in a wide variety of plants and participate in many important physiological processes (Mutlu and Gal, 1999). With the improvement of genome sequencing technology, more and more plant APs genes have been detected. About 69 genes are encoding APs in the *Arabidopsis* genome, 96 AP genes in the rice genome, and 50 AP genes in the grape (*Vitis vinifera*) genome (Takahashi et al., 2008; Chen et al., 2009; Guo et al., 2013). APs play important roles in the whole growth and development of plants, especially under stress response, sexual reproduction, aging, programmed cell death as well as processing and degradation of proteins. To date, it has been proved that some plant APs play an important role in abiotic stress. *ASPG1* (ASPARTIC PROTEASE IN GUARD CELL 1) in *Arabidopsis* is a typical example of plant APs participating in abiotic stress, which is involved in the ABA signaling pathway and mediates the response to plant drought stress by regulating the balance of ROS levels in cells (Yao et al., 2012). In addition, aspartic proteases responding to abiotic stress have also been identified in common bean (*Phaseolus vulgaris*), pineapple (*Ananas comosus*), buckwheat (*Fagopyrum esculentum Moench.*), and other species (Contouransel et al., 2010; Timotijevic et al., 2010; Raimbault et al., 2013). However, as far as we are aware, no studies have found evidence that APs are involved in Pi starvation.

Soybean (*Glycine max*) is one of the most important crops for oil and protein production, and its yield is severely affected by various environmental conditions. P deficiency is more likely to be a limiting factor for soybean yield because of the high demand for P in the nodule responsible for nitrogen fixation (Vance, 2001; Song et al., 2014). So far, botanists have identified hundreds of genes related to how plants adapt to Pi starvation, but most research has focused on model plants, with soybean relatively lagging (Zhang et al., 2014). Here, we investigated the function of *GmWRKY46* and characterized it as a regulator of Pi starvation responses. We found that *GmWRKY46* enhanced tolerance to Pi starvation through improving root development and direct interaction with the *AtAED1* promoter in transgenic *Arabidopsis*. Besides, we further found that overexpressing *GmWRKY46* in soybean transgenic hairy roots could enhance Pi starvation tolerance and root development. This study laid a foundation for the improvement of the low P tolerance of soybean in the future.

MATERIALS AND METHODS

Phylogenetic and Gene Structure Analysis

Plant Comparative Genomics portal (Department of Energy, Joint Genome Institute, https://phytozome.jgi.doe.gov/pz/portal.html) and National Biotechnology Information Center database (National Center of Biotechnology Information, www.ncbi.nlm.nih.gov/) were used for getting genetic information. BioXM 2.6 software was used to predict the molecular weight and isoelectric point of the gene, and GSDS (http://gsds.cbi.pku.edu.cn/) online program was used to predict the gene structure. Neighbor-joining phylogenetic trees were generated using the MEGA 5.1 program (Tamura et al., 2011). ClustalX 1.83 and GeneDoc were used for multiple alignments.

Subcellular Localization of *GmWRKY46*

The complete ORF sequences of *GmWRKY46* without a stop codon (1077 bp) were fused to the N-terminal of GFP reporter protein of *pJIT166* vector driven by CaMV35S promoter, generating a fusion construct *pJIT166-GmWRKY46-GFP*. The specific primers are listed in **Supplementary Table 1**. The fusion (*pJIT166-GmWRKY46-GFP*) and control (*pJIT166-GFP*) constructs were transformed into *Arabidopsis* protoplasts, respectively. *Arabidopsis* protoplast preparation and transformation were performed as described previously (Li et al., 2020). The GFP fluorescence in the transiently expressing protoplasts was imaged using a Zeiss LSM780 camera (Carl Zeiss, SAS, Jena, Germany).

Assays of Transcription-Activating Activity in Yeast

The full-length CDS of *GmWRKY46* was cloned into the pGBKT7 vector to create pGBKT7-*GmWRKY46* fusion vector, and the specific primers are listed in **Supplementary Table 1**. The fusion and pGADT7 vector were co-transformed into Y2HGold yeast. The transformants were screened on medium lacking Leu and Trp (SD/–Trp–Leu) or on medium lacking Leu, Trp, His, and adenine hemisulfate salt (SD/–Trp–Leu–His–Ade). Meanwhile, the empty vector was transformed as the control.

Plant Materials and Growth Conditions

The soybean genotype Williams 82 was used in these experiments. Seeds of soybean were germinated on vermiculite medium containing Hoagland's nutrient solution in a growth chamber (16 h light, 30°C/8 h dark, 20°C). Seven-day-old seedlings (removal of cotyledons) were transferred to Hoagland's nutrient solution and treated with two Pi levels (Pi-deficient, 2.5 μM KH_2PO_4; Pi-sufficient, 1 mM KH_2PO_4) for 11 days. At 10th day, the Pi-deficient group was resupplied with 1 mM Pi for 1 day, called R1d. The nutrient solution was refreshed every 3 days. Leaves and roots were harvested at the indicated times after initiation of Pi starvation treatment. Other soybean tissues were also collected at specific times. All samples were stored at −80°C prior to RNA extraction.

Reverse Transcription Quantitative Real-Time PCR

Reverse transcription quantitative real-time PCR (RT-qPCR) used three biological replicates, each containing three independent plants. Total RNA was extracted from the plant tissue using the plant RNA Extract Kit (TIANGEN Biotech, Beijing, China). The cDNA was synthesized with the HiScript II Q RT SuperMix (+gDNA wiper) for qPCR (Vazyme Biotech, Nanjing, China). The RT-qPCR was performed using a CFX96 Touch (Bio-Rad, Hercules, USA) with AceQ® qPCR SYBR Green Master Mix (Vazyme Biotech, Nanjing, China). The relative level of expression was calculated using the formula $2^{-\Delta Ct}$ or $2^{-\Delta\Delta Ct}$. The primers used for RT-qPCR analyses are listed in **Supplementary Table 2**.

Development of Transgenic *Arabidopsis*

The full-length target gene coding sequence (CDS) without the stop codon was inserted after the CaMV35S promoter, resulting in overexpression of the target gene constructs with GFP tags. The fusion constructs were confirmed by sequencing and then transformed into wild-type *Arabidopsis* (Col-0) plants. The gene-specific primers are listed in **Supplementary Table 1**. *Arabidopsis* plants were grown in a controlled environment at 23°C in a 16 h light/8 h dark photoperiod. The transformation and screening methods of *Arabidopsis* were performed as described previously (Li et al., 2020). Seeds used for phenotypic assays were harvested in the same environment. Homozygous T3 or T4 seeds were used.

RNA-Seq Library Construction, Sequencing, and Data Analysis

The seeds of WT and *GmWRKY46*-overexpressing transgenic *Arabidopsis* lines were grown on Hoagland's nutrient solution Phytagel plates containing 1 mM Pi, 0.2 mM Pi, and 62.5 μM Pi for 20 days. Whole seedlings of WT and transgenic *Arabidopsis* plants under three Pi conditions were collected, respectively, for total RNA isolation and RNA-seq analysis. Three independent biological replications were included for RNA-seq, and each biological replication contained five seedlings. Library construction was carried out by GENE DENOVO (Guang Zhou, China). The qualified cDNA libraries were ultimately sequenced by an Illumina HiSeq^{TM} 2500 instrument.

Gene expression was calculated using Fragments Per Kilobase of transcript per Million mapped reads (FPKM) (Griffith et al., 2015) and was compared between transgenic plants and WT control under three Pi conditions. Clustering software was used to perform cluster analysis of gene expression patterns. Assessment of RNA-seq quality, screening of differential expression genes (DEGs), expression pattern analysis, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs were carried out by GENE DENOVO (Guang Zhou, China). To identify differentially expressed genes across samples or groups, the edgeR package (http://www.r-project.org/) was used, and genes with a fold change ≥ 2 and a false discovery rate (FDR) <0.05 in comparison were listed as significant DEGs (Trapnell et al., 2012). For GO analysis, all DEGs were mapped to GO terms in the GO database (http://www.geneontology.org/), gene numbers were calculated for every term, and significantly enriched GO terms in DEGs compared with the genome background were defined by hypergeometric test. For KEGG analysis, pathway enrichment analysis identified significantly enriched metabolic pathways or signal transduction pathways in DEGs compared with the whole-genome background, and the calculating formula was the same as that in GO analysis. The raw transcriptome reads were submitted to the NCBI Sequence Read Archive (SRA) database under accession: PRJNA724748.

Yeast One-Hybrid Assay and Chromatin Immunoprecipitation

Promoter fragments were obtained and analyzed through NCBI (https://www.ncbi.nlm.nih.gov/), and specific primers were

designed for each W-box motif for yeast one-hybrid (Y1H) assay (**Supplementary Table 1**) and chromatin immunoprecipitation (ChIP-qPCR) (**Supplementary Table 3**). For Y1H, first, the *GmWRKY46* CDS without the stop codon was integrated into the sites of the pGADT7 to generate the effector vector pGADT7-*GmWRKY46*. The specific primers are listed in **Supplementary Table 1**. Then, the F1–F4 fragment from the AED1 (APOPLASTIC, EDS1-DEPENDENT 1) promoter region was integrated into the sites of the pAbAi to generate a reporter vector pAbAi-F1 to F4, respectively. Y1H assay was used to examine the interaction of *GmWRKY46* and the AED1 promoter fragment according to the manual provided by Matchmaker Gold Y1H Library Screening System (Clontech, Dalian, China). The yeast cells co-transformed with the prey, and either of the baits was cultured for 60 h on SD/-Ura/-Leu medium added with or without 300 ng/ml Aureobasidin A (AbA). For ChIP-qPCR, because the vector that we used to create the transgenic *Arabidopsis* has a GFP tag, the *GmWRKY46*-overexpressing transgenic *Arabidopsis* can be directly used for ChIP. ChIP assay was performed using the EpiQuikTM Plant ChIP Kit (Epigentek) in accordance with the instructions of the manufacturer. GFP Tag Monoclonal Antibody (Proteintech, Cat. No. 66002-1-Ig) was used to label the antibody. qPCR was performed with immunoprecipitated genomic DNA fragments, and enrichment was calculated as the ratio of immunoprecipitation to input and WT as control.

Induction of Transgenic Soybean Hairy Roots

The overexpression vector and empty vector were separately introduced into *Agrobacterium rhizogenes* K599, and the bacterium was used to infect 5 days old soybean seedling (Williams 82) hypocotyls by injection (Attila et al., 2007). About 14 days later, hairy roots were generated at the infected site. When transgenic hairy roots grew to approximately 8–10 cm long, a small part was harvested for RT-qPCR identification. The original main roots of the identified plants were removed and recovered in Hoagland's nutrient solution for 2 days, then the primary roots were cut off and transferred to Hoagland's vermiculite medium containing 1 mM Pi (Pi-sufficient) or 2.5 μM Pi (Pi-deficient) for 14 days. Each transgenic composite plant represented one independent transgenic line, and six independent transgenic lines were included for each Pi treatment. One independent soybean composite plant with transgenic hairy roots was considered a semi-biological replicate. A total of three replicates were included in this experiment.

Measurement of Total P and Soluble Pi Concentration

For the plant total P concentration assay, fresh soybean plant samples were heated at 75°C until completely dry and then ground into powder separately. Approximately 0.1 g of dry plant sample was weighed and digested by 2 ml HNO_3. After cooling, the digested samples were diluted to 25 ml with distilled water. Then, the concentration of P in the solution was determined by Optima 8000 ICP-OES (PerkinElmer, USA). For the measurement of soluble Pi concentration, the Tissue Inorganic Phosphorus Content Detection Kit (Solarbio, Beijing, China) was used. All experiments included three biological replicates.

Analysis of Root Development

The root development of soybean was analyzed using the scanner (Epson, Expression 21000XL, Japan) with a root analysis system (WinRHIZO 2020). Photography and root analysis of transgenic *Arabidopsis* were described previously (Li et al., 2020), seedlings were grown on square petri dishes under Pi-sufficient (1 mM Pi) or Pi-deficient (0.2 mM Pi and 62.5 μM Pi) condition for 15 days. All experiments included three biological replicates.

Statistical Analysis

The two-tailed Student's *t*-test ($P \leq 0.05$) was used to identify the statistical significance of any differences observed. The Microsoft Excel 2019 for Windows V10 was used for all statistical analyses.

RESULTS

Isolation and Bioinformatic Analysis of *GmWRKY46*

Studies have shown that *AtWRKY46* regulates lateral roots development in *Arabidopsis thaliana* under osmotic/salt stress (Ding et al., 2015), and we found in a report that *GmWRKY46* was highly expressed in soybean roots (Song et al., 2016). In osmotic/salt stress conditions, lateral root development is significantly reduced in loss-of-function *wrky46* mutants, while overexpression of *WRKY46* enhances lateral root development (Ding et al., 2015). In view of the close correlation between lateral roots development and Pi starvation response of plants (Péret et al., 2011), it can be speculated that *GmWRKY46* may be involved in the response of soybean to Pi starvation. *GmWRKY46* (GLYMA_08G021900) is located on chromosome 8 of soybean, position 1763171-1764910. Bioinformatics analysis showed that *GmWRKY46* contained a 1080 bp open reading frame (ORF) and encoded a predicted polypeptide of 359 amino acids with a calculated molecular mass of 89.7 kDa and an isoelectric point of 4.8. A phylogenetic tree was constructed based on the GmWRKY46, and using a total of 72 WRKYs from *Arabidopsis* showed that GmWRKY46, AtWRKY30, AtWRKY41, AtWRKY46, AtWRKY53, AtWRKY54, AtWRKY55, and AtWRKY70 had high homology (**Figure 1A**). Multiple sequence alignment analysis revealed that the N-terminal of GmWRKY46 protein contained a WRKYGQK domain and a C-X_7-C-X_{23}-H-X_1-H zinc-finger structure (**Figure 1B**), so GmWRKY46 belonged to the group III subfamily of the WRKY TF family (Chen et al., 2017a).

Analysis of Subcellular Localization and Transcription-Activating Activities of GmWRKY46

To investigate the subcellular localization of GmWRKY46, the recombinant constructs of the GmWRKY46-GFP fusion gene and GFP alone were introduced into *Arabidopsis* mesophyll protoplasts, respectively. Under a confocal laser microscope, the fluorescence of GMWRKY46-GFP was specifically detected

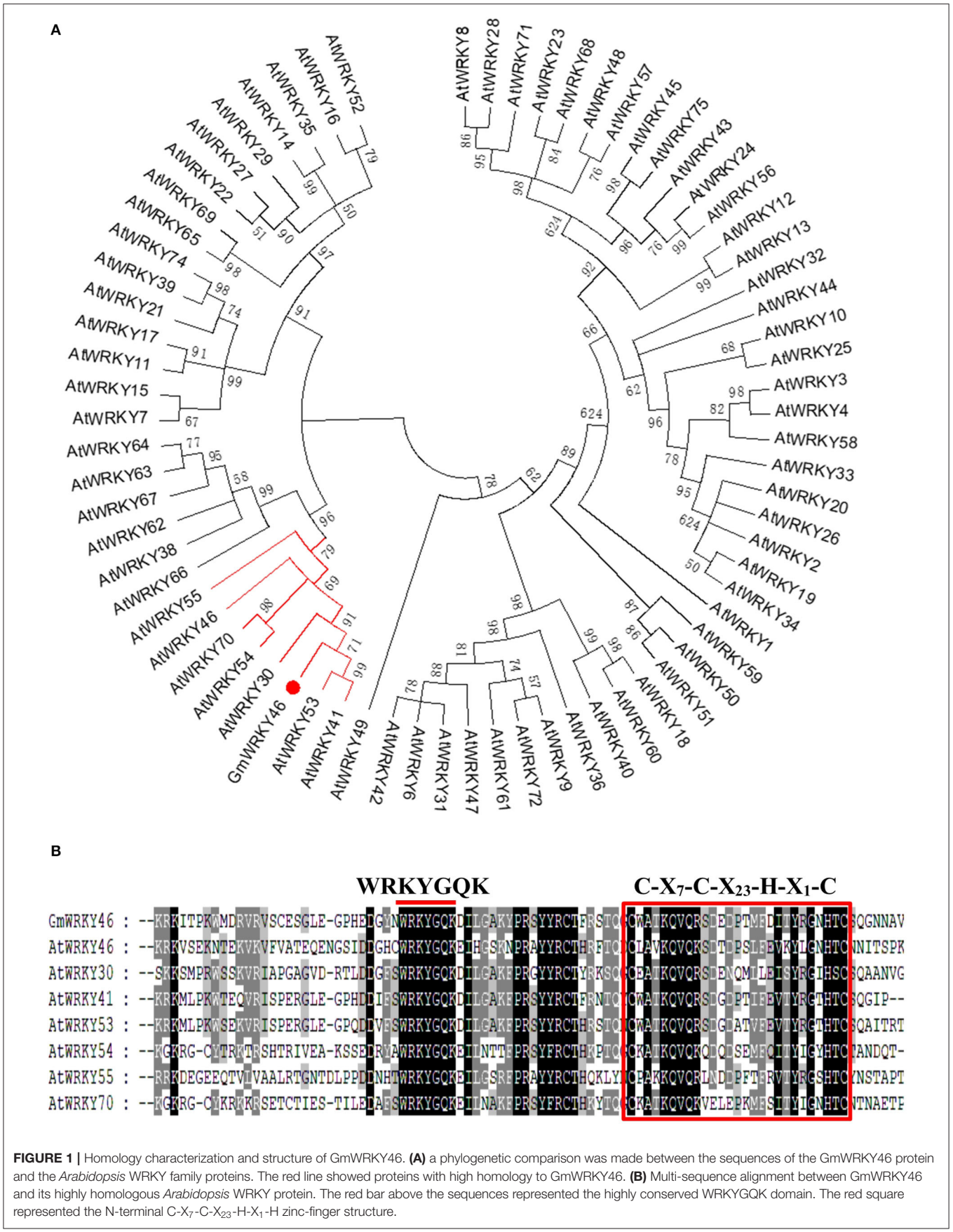


FIGURE 1 | Homology characterization and structure of GmWRKY46. **(A)** a phylogenetic comparison was made between the sequences of the GmWRKY46 protein and the *Arabidopsis* WRKY family proteins. The red line showed proteins with high homology to GmWRKY46. **(B)** Multi-sequence alignment between GmWRKY46 and its highly homologous *Arabidopsis* WRKY protein. The red bar above the sequences represented the highly conserved WRKYGQK domain. The red square represented the N-terminal C-X_7-C-X_{23}-H-X_1-H zinc-finger structure.

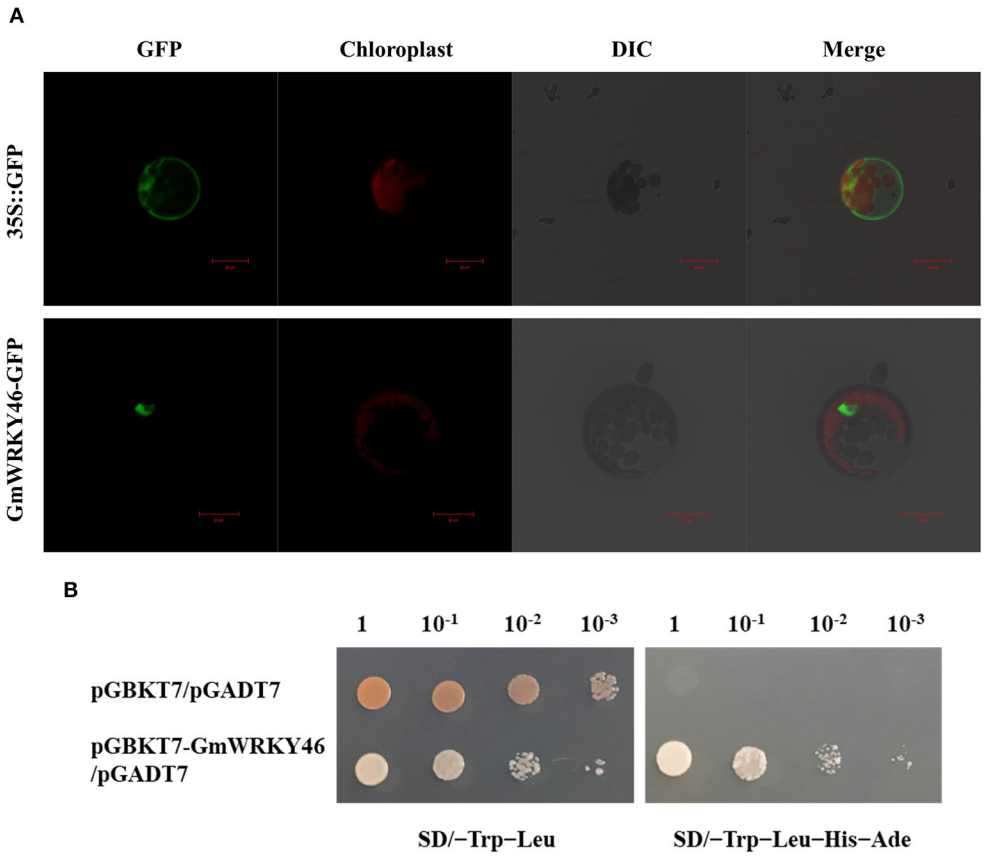


FIGURE 2 | Analysis of subcellular localization and transcription-activating activities of GmWRKY46. **(A)** Subcellular localization of GmWRKY46-GFP and 35S::GFP protein in *Arabidopsis* protoplasts. Scale bar = 20 μm. **(B)** Y2H assay for transcription-activating activities of GmWRKY46. Yeast cells co-transformed with pGBKT7-GmWRKY46/pGADT7 were grown on selective media SD/-Trip-Leu and SD/-Trip-Leu-His-Ade. Empty pGBKT7/pGADT7 was used as control.

in the nucleus, whereas the control construct was present throughout the whole cell (**Figure 2A**). These results indicate that GmWRKY46 was localized in the nucleus and was consistent with the predicted function as a transcription factor (TF).

Transcription factor is a protein molecule, and it usually contains a DNA-binding domain and a transcription activation domain. In order to verify whether GmWRKY46 has transcriptional activator activity, we used the yeast two-hybrid (Y2H) system. As shown in **Figure 2B**, the yeast cells transformed with the fusion construct pGBKT7-GmWRKY46 could grow normally in both SD/–Trp–Leu medium and SD/–Trp–Leu–His–Ade medium, while the growth of the control was inhibited in the SD/–Trp–Leu–His–Ade medium. These results indicated that the fusion proteins of GmWRKY46 could activate both HIS3 and AED2 reporter genes and could therefore act as transcriptional activators in yeast cells.

Expression Patterns of *GmWRKY46*

To investigate the transcript levels of *GmWRKY46* in specific tissues, the total RNA was extracted from root, stem, leaf, flower, pod, and the seed of soybean plants at first trifoliate (V1), full bloom (R2), and full seed (R6) stages. RT-qPCR analysis showed that the expression of *GmWRKY46* remained stable in the roots of the three stages. It should be noted that *GmWRKY46* expression in the leaves of the R2 stage was greatly increased compared with the other two stages. In addition, we noted that almost no expression of *GmWRKY46* was detected in all tissues at the R6 stage, except the roots (**Figure 3A**).

Then, the expression patterns of *GmWRKY46* were analyzed by RT-qPCR under Pi starvation in roots and leaves. As shown in **Figure 3B**, there was no significant change in *GmWRKY46* transcription level in leaves after Pi starvation except for the increase on the 5th day. In roots, *GmWRKY46* transcript levels quickly accumulated on the 1st day after Pi starvation, exhibited a dramatic decrease on the 5th day and then induced more than 9-fold of the normal level on the 10th day (**Figure 3B**). After resupplying Pi, the transcript level of *GmWRKY46* was close to the normal Pi level in both leaves and roots. These results indicated that *GmWRKY46* was involved in soybean response to Pi starvation.

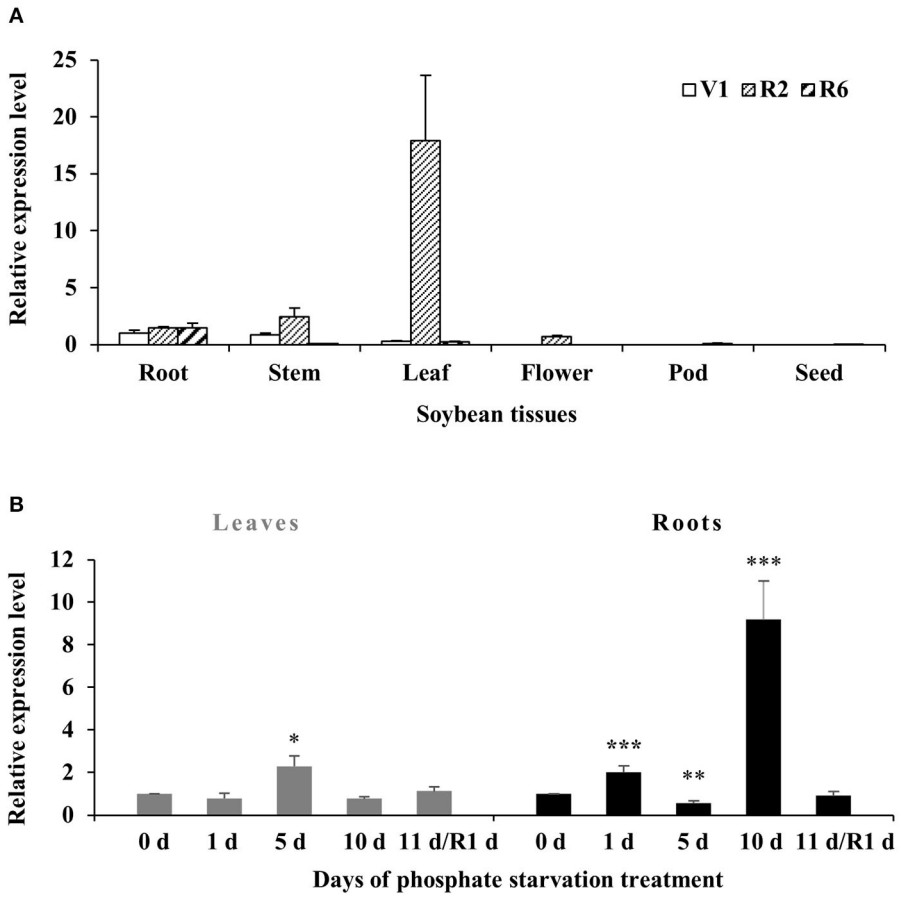


FIGURE 3 | Expression patterns of *GmWRKY46* in soybean. **(A)** The expression of GmWRKY46 in roots, stems, leaves, flowers, pods, and seeds at the first trifoliate stage (V1), full bloom stage (R2), and full seed stage (R6). **(B)** Time course of the expression level of *GmWRKY46* in leaves and roots. 0, 1, 5, 10, and 11 days, duration of Pi starvation (days); Recovery 11 days (R1d), 10 days of Pi starvation followed by 1 day on Pi-sufficient substrate. The expression level at 0 h was set as 1, and data represented means ± SD of three replicates. Data significantly different from the corresponding controls were indicated (Student's *t*-test, *****$P < 0.05$; ******$P < 0.01$; *******$P < 0.001$).

GmWRKY46 Enhancement of the Tolerance to Pi Starvation May Depend Partly on Improving Root Development in Transgenic *Arabidopsis*

To determine the functions of *GmWRKY46* in plant tolerance to Pi starvation, we attempted to use transgenic *Arabidopsis* due to the difficulty in obtaining transgenic soybean. Four *GmWRKY46*-overexpressing transgenic lines were identified by PCR and RT-qPCR (**Supplementary Figures 1A,B**), and three of them with higher expression levels of *GmWRKY46* were selected for further studies (**Supplementary Figure 1B**). Two-week-old seedlings of transgenic *Arabidopsis* lines and the wild type (WT) were grown in the greenhouse for 20 days under 1 mM Pi (Pi-sufficient) and 62.5 μM Pi (Pi-deficient) conditions. Under Pi-sufficient condition, the growth period of all plants was found to be the same, with all plants bolting (**Supplementary Figure 2**). However, the growth of WT plants was inhibited under Pi-deficient conditions, while the bolting of *GmWRKY46*-overexpressing transgenic *Arabidopsis* was not affected (**Supplementary Figure 2**). These results suggested that overexpression of *GmWRKY46* enhanced tolerance to Pi starvation in transgenic *Arabidopsis*.

Then, the three *GmWRKY46*-overexpressing transgenic *Arabidopsis* lines and WT were planted on vertical Phytagel plates containing three different Pi concentrations medium (Pi-sufficient, 1 mM Pi; Pi-deficient, 0.2 mM Pi, and 62.5 μM Pi, respectively) and cultured for 20 days to observe their phenotypes. We found that short-term Pi deficiency significantly inhibited the growth of *Arabidopsis*, especially the root growth, and the lower the environmental Pi concentration, the more obvious this inhibition (**Figure 4A**). However, this inhibitory effect was weakened in transgenic *Arabidopsis* (**Figure 4A**). We found that the primary and lateral root length of transgenic lines was significantly longer, and the lateral root number of transgenic lines was significantly more than those of WT under Pi-deficient conditions (**Figures 4B–D**). Furthermore, the lateral root length of transgenic lines was also significantly longer than that of

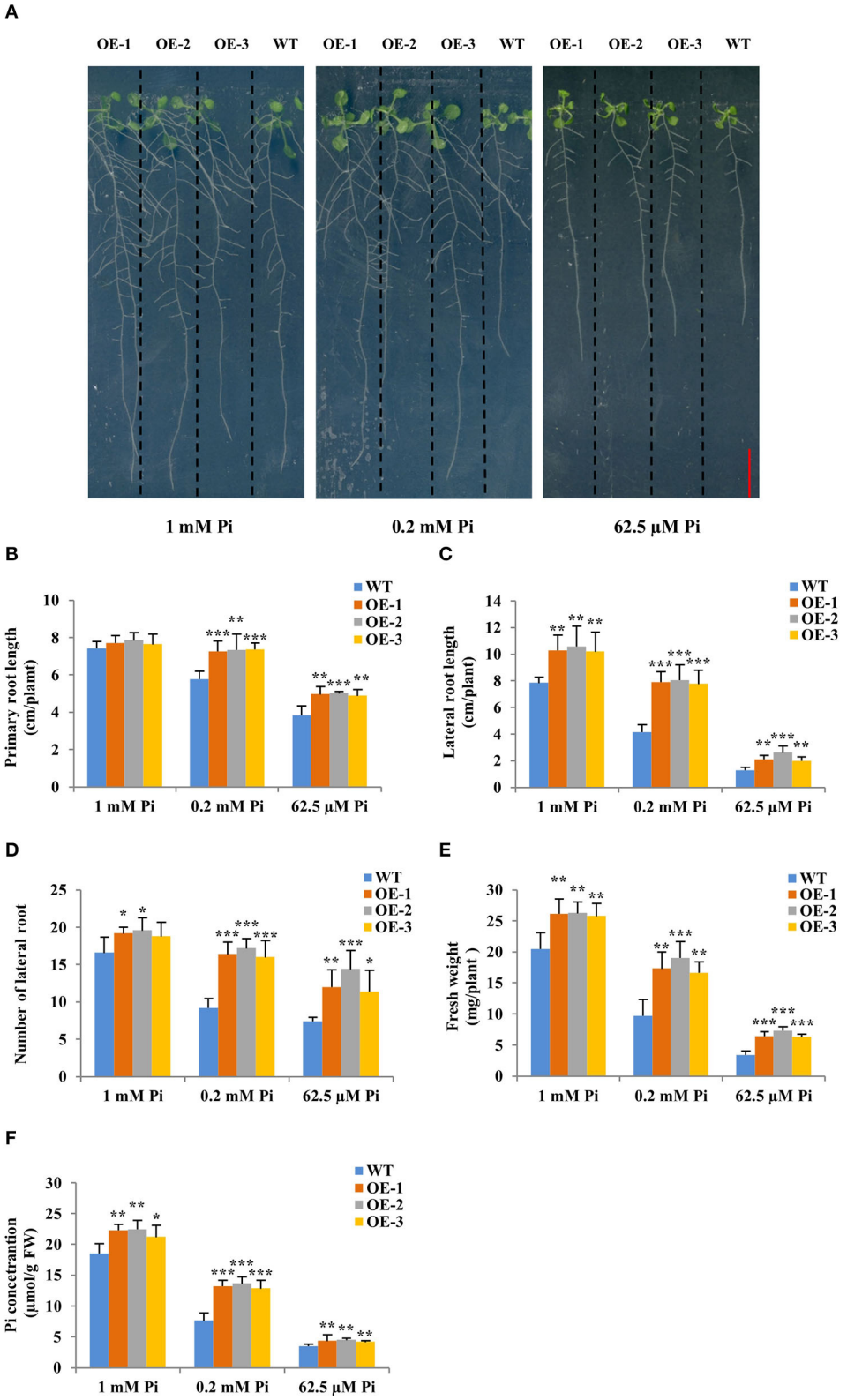


FIGURE 4 | Overexpression of *GmWRKY46* enhanced Pi starvation tolerance and lateral roots growth in transgenic *Arabidopsis*. **(A)** WT and GmWRKY46-overexpressing transgenic *Arabidopsis* were grown under three different Pi concentrations (Pi-sufficient, 1 mM Pi; Pi-deficient, 0.2 mM Pi and 62.5 μM Pi) for 15 days on vertically oriented petri plates, respectively. Primary root length **(B)**, lateral root length **(C)**, number of lateral root **(D)**, fresh weight **(E)**, and Pi concentration **(F)** were determined in WT and transgenic *Arabidopsis* plants on the 15th day. Values were mean ± SD ($n = 3$), and asterisks showed that the values were significantly different between the transgenic lines and the WT (Student's *t*-test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

WT under Pi-sufficient condition (**Figure 4B**). The lateral root number of all transgenic lines was significantly more than that of WT under Pi-sufficient condition, except OE-3 (**Figure 4D**), which we thought might be probably due to the relatively low expression of *GmWRKY46* in OE-3. There was no significant difference between the primary length of the transgenic lines and the WT under Pi-sufficient condition (**Figure 4B**). Additionally, the fresh weight and Pi concentration of the transgenic lines were significantly higher than those of the WT under three Pi level conditions (**Figures 4E,F**). Together, these results suggested that overexpression of *GmWRKY46* significantly improves the uptake of Pi in transgenic *Arabidopsis* by promoting root system development, and at the same time, its positive effect on lateral root development might be independent of the Pi level of the environment.

Differentially Regulated Genes Identified From RNA-seq Analysis

The Pi starvation tolerance phenotype of transgenic *Arabidopsis* also may be due to gene expression changes caused by the *GmWRKY46*. WRKY transcription factors were generally thought to regulate the expression of their target genes by binding to the W-box(es) in their target gene promoters (Rushton et al., 2010), and hence, it was reasonable to speculate that *GmWRKY46* might recognize heterologous promoters from a similar protein structure/DNA sequence interaction. To examine changes in gene expression patterns, RNA-seq analysis was conducted on 20-day-old plants under 1 mM Pi, 0.2 mM Pi, and 6.25 μM Pi conditions and compared the *GmWRKY46*-overexpressing *Arabidopsis* against the WT control. Under 1 mM Pi condition, the expression profile of *GmWRKY46*-overexpressing *Arabidopsis* compared with the WT control was hereafter referred to as comparison 1 (C1); likewise, under 0.2 mM Pi condition, OE-*GmWRKY46* compared with the WT was C2 and under 6.25 μM Pi condition was C3.

There were 27 differential expression genes (DEGs) in C1, 49 DEGs in C2, and 17 DEGs in C3 (**Supplementary Table 4**); 20, 34, and 8 genes were upregulated in C1, C2, and C3, respectively; and 7, 15, and 9 genes were downregulated in C1, C2, and C3, respectively (**Figure 5A**). DEGs were then subjected to enrichment analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (**Supplementary Table 4**). The results showed that the DEGs were commonly involved in energy metabolism (tricarboxylic acid and citrate metabolic process, organic substance metabolic process, and nitrogen and carbon metabolism), stress responses (response to endoplasmic reticulum stress, regulation of cell death, regulation of hydrogen peroxide and reactive oxygen species metabolic process, biosynthesis of amino acids, phenylalanine metabolism, and MAPK cascade), and plant hormone signal transduction (salicylic acid and jasmonic acid-mediated signaling pathway) (**Figures 5B,C**). Given that 11 DEGs were common to C1 and C2 or C1 and C3 or C2 and C3, we narrowed the focus to this smaller group. The 11 members were named DEG1 to DEG11 from top to bottom in the clustering analysis map (**Figure 5D**). Among these 11 DEGs, five are annotated to have molecular functions or transporter activity (**Supplementary Table 5**). In addition, to verify the RNA-seq analysis results, seven DEGs were chosen and evaluated in a qRT-PCR assay with the same tissues used for the RNA-seq analysis (**Figure 5E**). The expression patterns of the selected genes were consistent with the RNA-seq data.

Overexpression of *AtAED1* (DEG6) Enhanced Tolerance to Pi Starvation in Transgenic *Arabidopsis*

We obtained 11 DEGs through cluster analysis, among which 6 DEGs had upregulated performance in at least two cases of C1, C2, and C3 (**Figure 4D**), so we further focused on the analysis of these 6 DEGs (DEG1, DEG4, DEG5, DEG6, DEG8, and DEG9). Interestingly, promoter analysis showed that the promoter of the 6 DEGs contained at least one W-box motif, respectively (**Supplementary Figure 3**), suggesting that all of them might be regulated by *GmWRKY46*. So, we constructed transgenic *Arabidopsis* with overexpression of the 6 DEGs, respectively. At least two independent transgenic *Arabidopsis* lines were obtained for each gene and verified by PCR or RT-qPCR. Here, we only showed the data of *DEG6* (OE#1, OE#2, and OE#3) (**Supplementary Figure 4**). Two-week-old seedlings of all transgenic *Arabidopsis* lines and the WT were grown in the greenhouse for 20 days under Pi-sufficient 1 (mM Pi) and Pi-deficient (62.5 μM Pi) conditions. We found that bolting of transgenic *Arabidopsis* with overexpression of the *DEG6* was not inhibited by Pi starvation when compared with WT and was similar to the phenotype observed for *GmWRKY46* overexpression (**Figure 6A**, **Supplementary Figure 2**). Further data analysis showed that *DEG6*-overexpressing transgenic *Arabidopsis* had significantly higher fresh weight and shoot Pi concentrations than WT under low Pi condition (**Figures 6B,C**). However, *DEG6*-overexpressing transgenic *Arabidopsis* did not produce a root phenotype like that of *GmWRKY46*-overexpressing transgenic *Arabidopsis* in plate growth experiments (data not shown). Bioinformatics analysis showed that DEG6 was a eukaryotic aspartyl protease family protein with an Asp domain (**Figure 6D**), known as AED1 (Breitenbach et al., 2014). These results suggested that overexpression of *AtAED1* enhanced tolerance to Pi starvation in transgenic *Arabidopsis*.

GmWRKY46 Directly Interacts With the Promoter of *AtAED1 (DEG6)*

To determine whether *DEG6* was the target gene of *GmWRKY46*, four W-boxes (F1–F4) in the DEG6 upstream 2000-bp long promoter fragment were used for Y1H assay to examine the interaction between GmWRKY46 and the *DEG6* promoter (**Figure 7A**). Therefore, a 48 bp fragment containing three copies of F1–F4 with 5 bp promoter sequence on both sides, respectively, was used as bait and cloned into the reporter vector, while GmWRKY46 was used as prey. The yeast cells of control (pGADT7/pAbAi-F1 ∼ F4) and those co-transformed with prey (pGADT7-GmWRKY46)-bait (pAbAi-F1 ∼ F4) grew normally in the screening medium (**Figure 7B**).

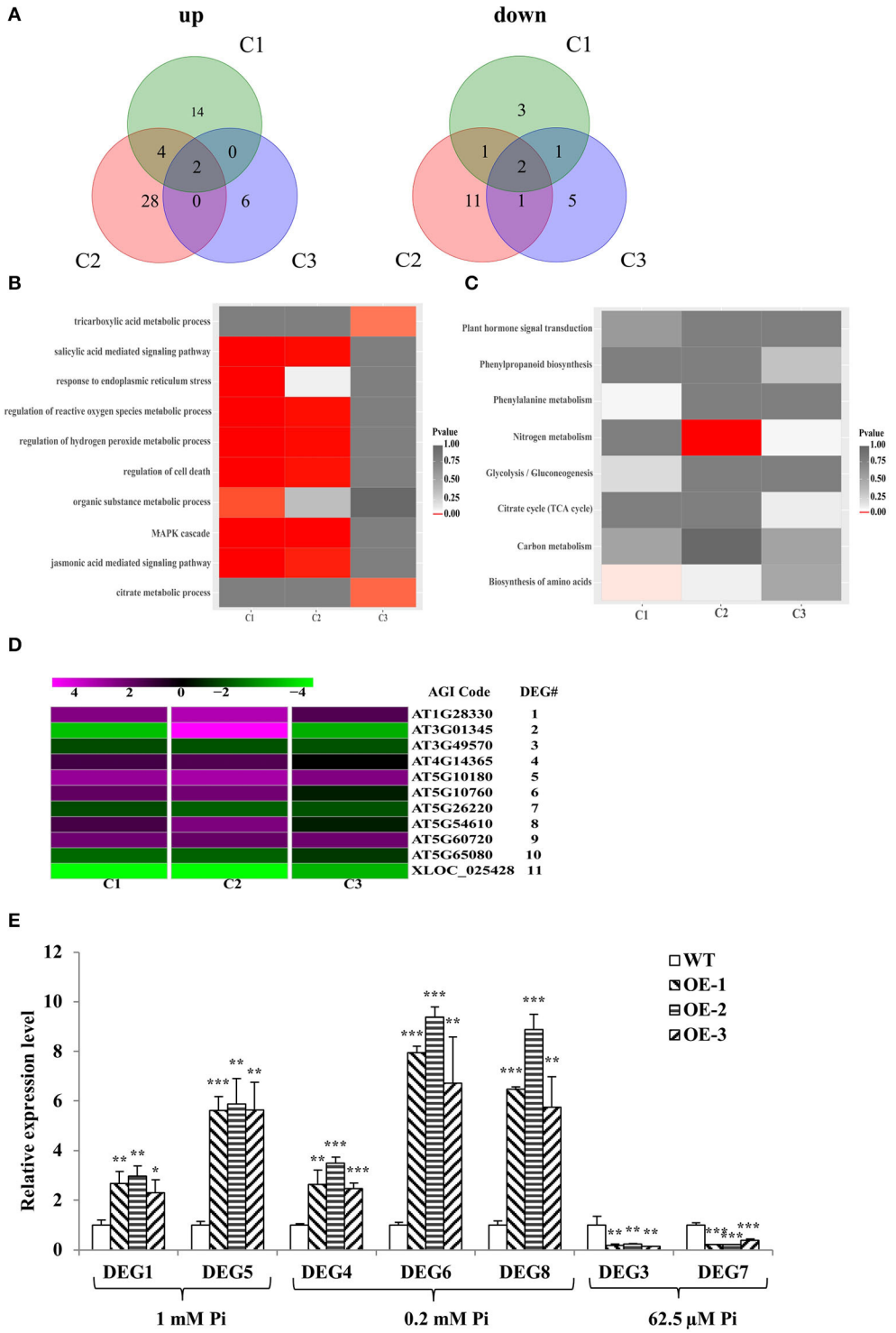


FIGURE 5 | RNA-seq analysis and RT-qPCR verification. **(A)** Venn diagrams of all genes exhibiting upregulated or downregulated expression between WT and *GmWRKY46*-overexpressing *Arabidopsis* plants. C1, C2, and C3 represent different levels of Pi. **(B,C)** Results of the GO enrichment analysis **(B)** and KEGG pathway enrichment analysis **(C)** of the DEGs between WT and *GmWRKY46*-overexpressing *Arabidopsis* plants. **(D)** Heat map of clustering analysis of the 11 DEGs in the intersection of C1, C2, or C3. Expression ratios are shown as $\log_2$ values. Magenta represented increased expression; green represented decreased expression; black represented no difference in expression compared with control. **(E)** Expression patterns of 7 DEGs candidates. Vertical axis showed fold enrichment of relative transcript levels between transgenic and WT plants. Expression was normalized to that of Actin2/8. Data were means ± SD ($n = 3$). Data significantly different from the WT were indicated (Student's *t*-test, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

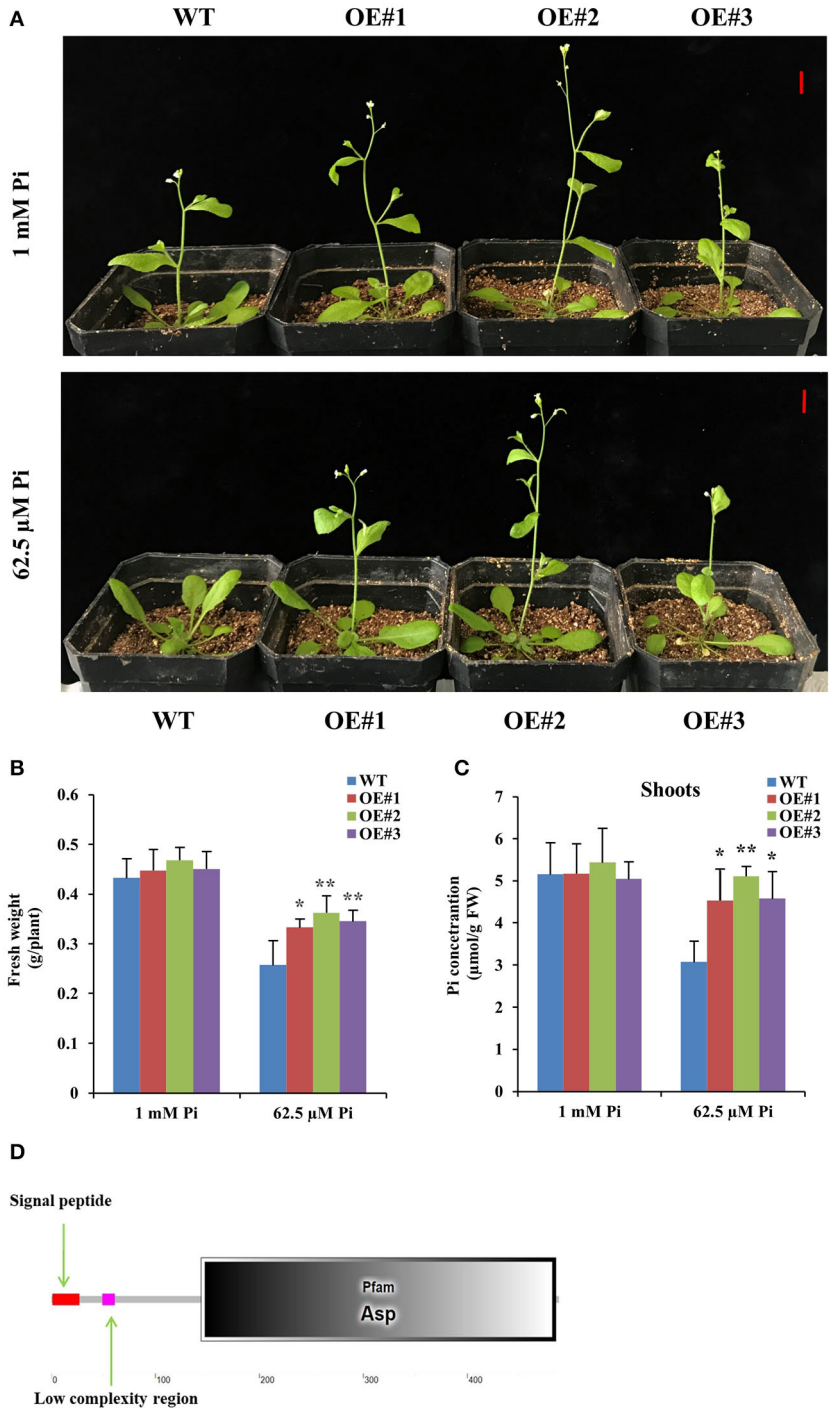


FIGURE 6 | Overexpression of *DEG6* enhanced tolerance to Pi starvation in transgenic *Arabidopsis*. **(A)** Two-week-old seedlings were grown in the greenhouse for 20 days under 1 mM Pi (Pi-sufficient) and 62.5 μM Pi (Pi-deficient) conditions. WT, wild type *Arabidopsis* lines; OE#1 to 3, independent *DEG6*-overexpresing *Arabidopsis* lines. Bars: 1 cm. **(B, C)** The fresh weight and shoots Pi concentration of WT and *DEG6*-overexpresing *Arabidopsis* on the 20th day. Data means ± SD ($n = 3$). Asterisks indicated significant differences between transgenic *Arabidopsis* and WT (Student's *t*-test, $^{*}P < 0.05$; $^{**}P < 0.01$). **(D)** Protein domain analysis of DEG6.

However, when 300 ng/ml AbA was added, growth of the control and pGADT7-GmWRKY46/pAbAi-F1 was completely inhibited, while pGADT7-GmWRKY46/pAbAi-F2, pGADT7-GmWRKY46/pAbAi-F3, and pGADT7-GmWRKY46/pAbAi-F4 transformants survived (**Figure 7B**). These results suggested that GmWRKY46 can specifically bind to three W-box cis-acting elements, F2, F3, and F4 on the *DEG6* promoter *in vitro*, but cannot bind F1.

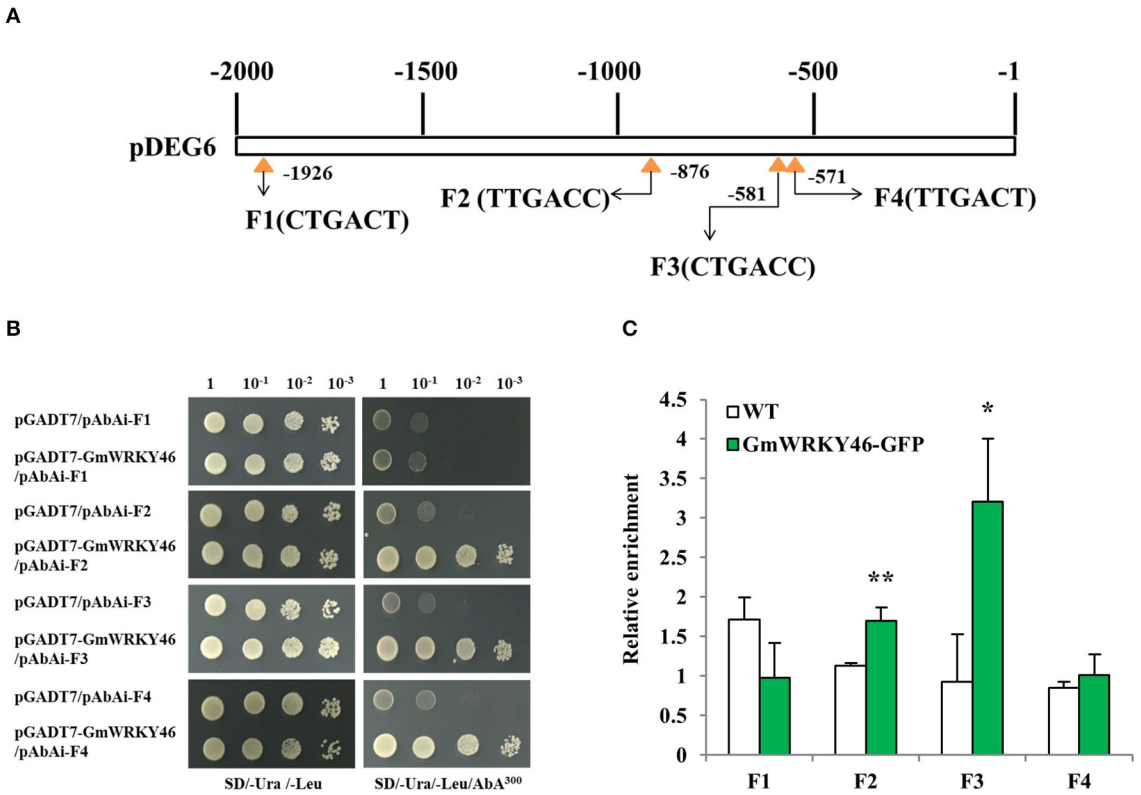


FIGURE 7 | GmWRKY46 binds to *DEG6* promoter region *in vivo* and *in vitro*. **(A)** Diagram of the 2000-bp promoter region of DEG6 showing the relative positions of the W-boxes. Four W-boxes F1, F2, F3, and F4 were marked by yellow triangle. Numbers indicated the position of starting nucleotide of each W-box relative to translation start. **(B)** GmWRKY46 binds to the *DEG6* promoter region in the Y1H assay. Yeast cells were transformed with a bait vector containing a promoter fragment F1, F2, F3, or F4 fused to pAbAi vector, and a prey vector containing GmWRKY46 fused to pGADT7 vector. Yeast cells were grown in liquid medium to an OD600 of 1.0 and diluted in a 10× dilution series (10^{-1}-10^{-3}). From each dilution, 5 μl was spotted onto SD/-Ura/-Leu medium to select for plasmids, and SD/-Ura/-Leu supplemented with 300 ng/ml Aureobasidin A (AbA) to select for interaction. Empty pGADT7 was used as control. **(C)** ChIP-qPCR analysis of GmWRKY46 binding to the DEG6 promoter region. *Arabidopsis* seeds of WT and Pro35S:GmWRKY46-GFP transgenic plants were germinated in medium and supplied with sufficient Pi. The whole plants were harvested for ChIP analysis. Enriched DNA fragments (F1 to F4) in the DEG6 promoter were quantified using RT-qPCR. Enrichment was calculated as the ratio of immunoprecipitation to input. Values represent means ± SD ($n = 3$). Data significantly different from the control are indicated (Student's *t*-test, $^{*}P < 0.05$; $^{**}P < 0.01$).

To further determine whether GmWRKY46 binds the *DEG6* promoter *in vivo*, specific primers were used against the *DEG6* promoter corresponding to fragments F1 ∼ F4 for ChIP-qPCR, respectively. The ChIP-qPCR showed that the relative DNA enrichment amount of GmWRKY46-GFP transgenic *Arabidopsis* plants at F2 and F3 was significantly higher than that of WT plants (**Figure 7C**), and the binding ability between GmWRKY46 and F3 was stronger than that between GmWRKY46 and F2. The results indicated that GmWRKY46 could specifically bind the *DEG6* (*AtAED1*) promoter in *Arabidopsis*. In addition, we observed that GmWRKY46 and F4 bind in the yeast system, but did not bind *in vivo*. The possible reason might be due to the closeness of F3 and F4 (interval 4 bp), thus influencing each other (Rushton et al., 2010). This interaction may also explain the stronger binding ability of GmWRKY46 and F3 than that of GmWRKY46 and F2.

Functional Analysis of *GmWRKY46* in Soybean Composite Plants

The "composite plant system" containing transgenic hairy root and their attached WT shoot can be obtained by genetic transformation of hairy roots, providing a rapid and effective method for functional analysis of genes expressed in soybean roots (Guo et al., 2011). Due to the strong response of *GmWRKY46* to the low Pi in the soybean root, the "composite plant system" was used for further studying the functions of *GmWRKY46* in the low Pi tolerance and root development in soybean. Increased expression of *GmWRKY46* in transgenic hairy roots was verified through RT-qPCR analysis (**Supplementary Figure 5**). The composite soybean plants with transgenic hairy roots were transferred to a nutrient solution containing 1 mM Pi (Pi-sufficient) or 2.5 μM Pi (Pi-deficient). After 14 days of Pi starvation, the control plants showed symptoms of P deficiency, such as small, narrow, and dark green

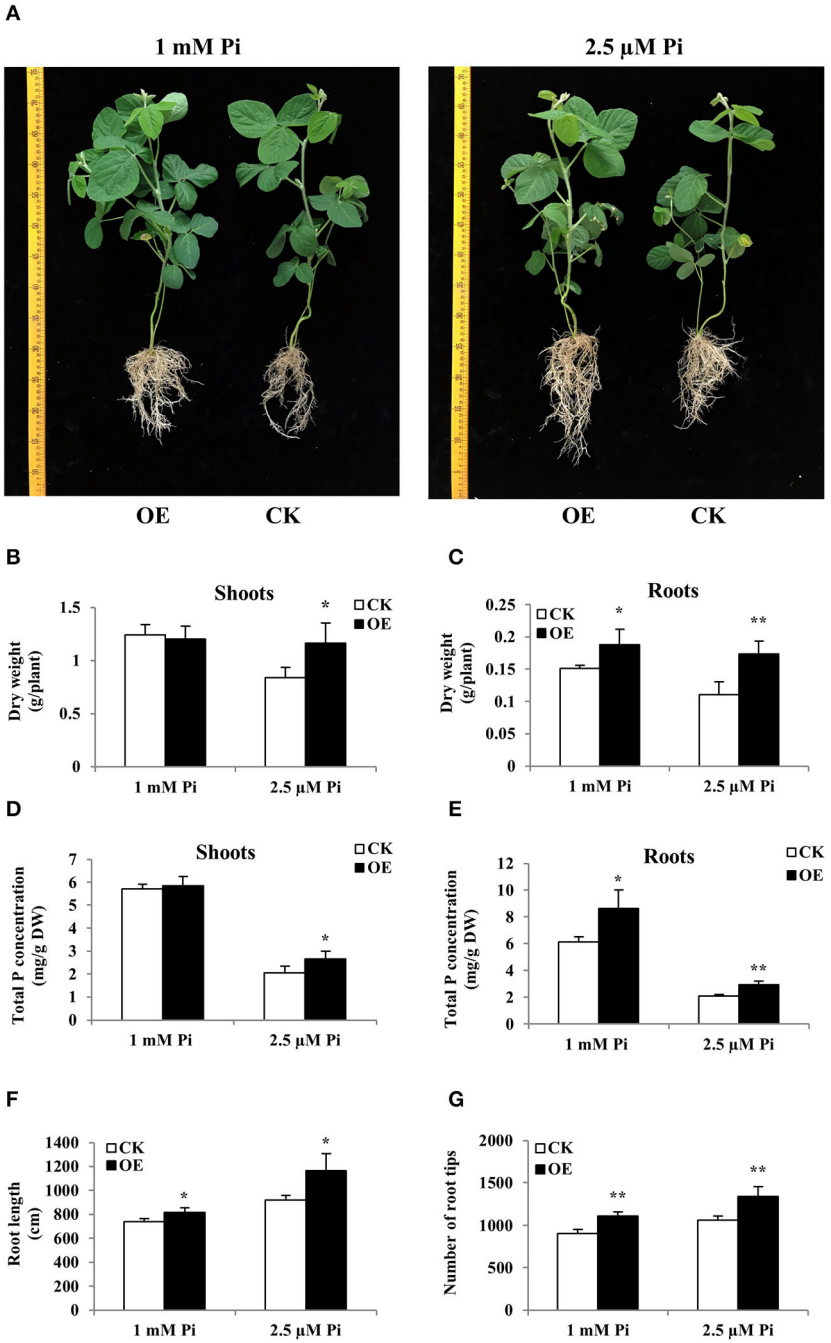


FIGURE 8 | Comparison of control (CK) and *GmWRKY46*-overexpressing composite soybean plants (OE) under two Pi level conditions. Composite soybean plants with transgenic hairy roots were transferred to a nutrient solution containing 1 mM Pi (Pi-sufficient) or 2.5 μM Pi (Pi-deficient). After a further 14 days, shoots and roots were separately harvested for analysis. **(A)** Phenotype of CK and OE. Scale unit: cm. **(B)** Dry weight of shoots. **(C)** Dry weight of roots. **(D)** Total P concentration of shoots. **(E)** Total P concentration of roots. **(F)** Root length of composite plants. **(G)** Number of root tips of composite plants. CK represents the composite plants transformed with the empty vector; OE indicates composite plants with *GmWRKY46*-overexpressing transgenic hairy roots. DW means dry weight. Data are means ± SD of three biological replicates. Asterisks indicate significant differences between OE and CK (Student's *t*-test, **P* < 0.05; ***P* < 0.01).

leaves, while the transgenic composite plants with *GmWRKY46*-overexpressing hairy roots were relatively normal (**Figure 8A**). At the same time, overexpression of *GmWRKY46* in hairy roots of soybean significantly increased the dry weight and total P concentration of transgenic composite plants compared with the control under Pi-deficient condition. More precisely, the dry weight rose by 38% in shoots and by 58% in roots, and the total P concentration rose by 29% in shoots and by 41% in

roots under Pi-deficient condition (**Figures 8B–E**). In addition, overexpression of *GmWRKY46* significantly elongated hairy roots length and increased the number of root tips compared with control plants regardless of Pi supply (**Figures 8F,G**). Under Pi-sufficient condition, we found that overexpression of *GmWRKY46* resulted in a significant increase in dry weight and total P concentration of roots compared with control plants (**Figures 8B–E**). The above results implied that the overexpression of *GmWRKY46* in soybean hairy roots could dramatically regulate root development, thus affecting the tolerance of soybean to Pi starvation.

DISCUSSION

As one of the largest TF families in plants, the WRKY TFs family plays an important role in the regulation of plant growth, development, and defense response mechanisms (Singh et al., 2002). Although *AtWRKY75* was reported as a member of the WRKY TFs family related to Pi acquisition and root development in *Arabidopsis* as early as 2007 (Devaiah et al., 2007), the role of WRKY TFs under Pi starvation in soybean and other higher crops remains unclear. Here, we functionally and mechanically characterize a WRKY TF, GmWRKY46, from soybean under low Pi stress. Overexpression of *GmWRKY46* promoted Pi starvation tolerance and root development in transgenic plants. In addition, we also further demonstrated that GmWRKY46 directly binds to the promoter of *AtAED1*, thereby enhancing the tolerance to Pi starvation in transgenic *Arabidopsis*. Our study provided a basis for the molecular genetic breeding of Pi starvation tolerance in soybean and novel insight into the regulatory pattern mediated by WRKYs in the stress response.

GmWRKY46 belonged to group III subfamily of the WRKY TF family, which was localized in the nucleus and had transcriptional activator activity. Interestingly, most of the *Arabidopsis* WRKY family of genes with high homology to *GmWRKY46* have been found to be associated with abiotic stress. For example, *AtWRKY30* enhanced abiotic stress tolerance during early growth stages in *Arabidopsis* (Scarpeci et al., 2013), *AtWRKY53* negatively regulated drought tolerance by mediating stomatal movement (Sun and Yu, 2015), and *AtWRKY46/54/70* involved in brassinosteroid-regulated plant growth and drought responses (Chen et al., 2017b). Our results showed that Pi starvation strongly induced *GmWRKY46* expression, particularly in roots, suggesting that it may play an important role in the root response to Pi starvation. It should be noted that the expression of *GmWRKY46* in soybean leaves of R2 stage was significantly increased compared with the other two stages, which might indicate that expression of the gene in soybean leaves was developmentally regulated. This expression pattern was like that of *GmWRKY58* and *GmWRKY76*, which were initially expressed at very low levels in leaves, then highly expressed in relatively young leaves, but then rapidly declined in old leaves (Yang et al., 2016b).

Arabidopsis has been commonly used in transgenic studies for stress-tolerant genes from crops that are not easy for gene transformation analysis, including soybean (Liao et al., 2008; Niu et al., 2012; Kong et al., 2018). Overexpression of *GmWRKY46* enhanced tolerance to Pi starvation in transgenic *Arabidopsis* plants as revealed from plant growth situation and changes in root development, fresh weight, and Pi concentration after low Pi treatments. Moreover, we further demonstrate the role of *GmWRKY46* in adaptation to Pi starvation in soybean by using a "composite plant system," which consisted of transgenic hairy roots attached to wild-type shoots for "whole-plant" functional analysis (Guo et al., 2011; Bian et al., 2020). Under Pi-deficient condition, the dry weight and total P concentration of *GmWRKY46*-overexpressing transgenic soybean composite plants were higher than those of the control plants. Root analysis showed that the root length and number of *GmWRKY46*-overexpressing transgenic *Arabidopsis* and transgenic soybean composite plants increased significantly under Pi-deficient conditions. Previous studies have shown that root systems play an important role in Pi uptake (Péret et al., 2011; Jia et al., 2018). Therefore, we speculate that *GmWRKY46* could affect plant P efficiency by regulating root adaptive changes in response to Pi starvation.

Interestingly, the lateral root development of transgenic *Arabidopsis* with *GmWRKY46* overexpression was still better than that of WT under Pi-sufficient condition. We speculated that *GmWRKY46* may regulate lateral root development independently of environmental Pi levels. Because plants get Pi mainly through lateral roots (Péret et al., 2011; Jia et al., 2018), better lateral root development should be one of the reasons *GmWRKY46*-overexpressing transgenic *Arabidopsis* has higher fresh weight and Pi concentration than the WT under Pi-sufficient and Pi-deficient conditions. The experiments with the "composite plant system" further confirmed our deduction. Regardless of Pi supply, overexpression of *GmWRKY46* in hairy roots increased root dry weight, root total P concentration, root length, and root tip number of transgenic soybean composite plants. In previous research reports, similar phenomena have been found in *Arabidopsis* and rice. Under both Pi-sufficient and Pi-deficient conditions, when *AtWRKY75* expression was suppressed, lateral root length and number were significantly increased (Devaiah et al., 2007). *OsMYB4P*-overexpressing transgenic rice had longer lateral roots than the WT, regardless of environmental Pi levels (Yang et al., 2014). However, the dry weight of the shoots and the total P concentration of transgenic soybean composite plants were not different from those of control plants under Pi-sufficient condition. We speculate that this may be a self-protective mechanism of soybean, because excessive accumulation of Pi in the shoots may be harmful to plants (Zhou et al., 2008; Wang et al., 2009). Although overexpression of *GmWRKY46* in the soybean hairy roots led to enhanced root development, which was beneficial to Pi acquisition, the upward transfer of Pi from transgenic roots may be limited by wild-type shoots in the "composite plant system."

Several studies have suggested that some of the basic functional modules of stress-responsive regulatory networks might be shared among higher plants (Chen and Zhu, 2004). Because WRKY TFs play a strict regulatory role in the specific recognition and binding of W-box or W-box similar sequences on downstream promoters, they have a promising application

prospect in crop improvement (Phukan et al., 2016). Although our goal was to find the target gene of *GmWRKY46* in soybean to analyze the associated signaling pathway, due to the difficulty of obtaining transgenic soybean, we regressed back to using transgenic *Arabidopsis* for a first analysis. As expected, the RNA-seq analysis of the overexpression of *GmWRKY46* revealed that there were a lot of differentially expressed genes (DEGs) in transgenic *Arabidopsis* compared with the WT at three Pi levels, with at least 60 of them upregulated. Many DEGs were involved in energy metabolism, stress response, and hormone synthesis and transport, and these biological processes have been found in relation to Pi starvation in previous studies (Woo et al., 2012; O'Rourke et al., 2013). These results suggest that *GmWRKY46* actively activates the response of transgenic *Arabidopsis* to Pi starvation.

In RNA-seq analysis, we set three Pi levels to capture the key target genes more accurately. However, only two genes were found to be co-upregulated at all three Pi levels, which might indicate that our experimental design is not perfect. The gene changes caused by Pi starvation may be transient and rapid, so many potentially relevant target genes may be lost by RNA-seq analysis when sampled after prolonged low Pi stress (Wu et al., 2003). In order to achieve our goal, we extended the scope to 6 DEGs upregulated at any two Pi levels and then overexpressed them in *Arabidopsis*. We found that overexpression of *DEG6* also improved the resistance of transgenic *Arabidopsis* to Pi starvation. Furthermore, Y1H and ChIP-qPCR assays showed that GmWRKY46 binds to the W-boxes on the DEG6 promoter, so we determined that DEG6 was a target of GmWRKY46 in response to Pi starvation. Bioinformatics analysis showed that DEG6 was a eukaryotic aspartic protease family protein called AED1, which was part of a homeostatic feedback mechanism regulating systemic immunity (Breitenbach et al., 2014). In this study, we provided direct evidence of its involvement in nutrient stress. We speculated that there might be a similar homolog that would exist in soybean and be regulated by *GmWRKY46*. However, some effort is needed to identify which of the many homologous genes in the database may play a similar role in responding to low Pi stress. Meanwhile, in order to better demonstrate *GmWRKY46* enhanced Pi starvation tolerance via regulating *AtAED1*, further research is needed to determine if overexpressing *GmWRKY46* in the mutant of *ataed1* will no longer lead to enhancement of Pi starvation tolerance.

In addition, *AtAED1* is induced by salicylic acid (SA) (Breitenbach et al., 2014), and our RNA-seq results also indicated that the genes involved in SA synthesis and transport are affected. Whether *AtAED1* affects the tolerance to low Pi by affecting SA is an interesting research direction. After all, some research has shown that SA played a role in the P uptake and root development of plants (Khorassani et al., 2011; Wang et al., 2020). Furthermore, we did not find the reason why overexpression of *GmWRKY46* in transgenic plants promoted root development in this study, which is an urgent problem to be solved in the future.

CONCLUSION

To conclude, we identified a WRKY TFs family gene from soybean, *GmWRKY46*, induced by Pi starvation. Overexpression of *GmWRKY46* enhanced tolerance to Pi starvation and affected root growth in transgenic plants. Also, *GmWRKY46* could enhance low Pi tolerance by activating *ATAED1*, a eukaryotic aspartyl protease family protein gene in transgenic *Arabidopsis*. Further studies should focus on the roles of *GmWRKY46* in soybean plants and find some similar *AtAED1* genes in soybean and their potential manipulation to improve Pi starvation tolerance or other agronomic traits.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession numbers can be found below: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA724748.

AUTHOR CONTRIBUTIONS

CL and SY planned the research. CL, KL, XL, HR, MZ, and ZY performed the experiments. CL analyzed the data and wrote the manuscript. SY revised the manuscript. JG reviewed the manuscript. All authors approved the final manuscript.

FUNDING

This work was supported by the National Transgene Science and Technology Major Program of China (2016ZX08004-005), the Fundamental Research Funds for the Central Universities (KYT201801), and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT_17R55).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.700651/full#supplementary-material

REFERENCES

Attila, K., Dongxue, L., Arief, I., Nguyen, C. D. T., Sureeporn, N., Mark, K., et al. (2007). Agrobacterium rhizogenes-mediated transformation of soybean to study root biology. *Nat. Protoc.* 2, 948–952. doi: 10.1038/nprot.2007.141

Baek, D., Kim, M. C., Chun, H. J., Kang, S., Park, H. C., Shin, G., et al. (2013). Regulation of miR399f transcription by AtMYB2 affects phosphate starvation responses in *Arabidopsis*. *Plant Physiol.* 161, 362–373. doi: 10.1104/pp.112.205922

Bencke-Malato, M., Cabreira, C., Wiebke-Strohm, B., Bücker-Neto, L., Mancini, E., Osorio, M. B., et al. (2014). Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. *BMC Plant Biol.* 14:236. doi: 10.1039/c5cp07272g

Bian, X. H., Li, W., Niu, C. F., Wei, W., Hu, Y., Han, J. Q., et al. (2020). A class B heat shock factor selected for during soybean domestication contributes to salt tolerance by promoting flavonoid biosynthesis. *New Phytol.* 225, 268–283. doi: 10.1111/nph.16104

Breitenbach, H. H., Wenig, M., Wittek, F., Jorda, L., Maldonado-Alconada, A. M., Sarioglu, H., et al. (2014). Contrasting roles of apoplastic aspartyl protease AED1 and legume lectin-like protein LLP1 in *Arabidopsis* systemic acquired resistance. *Plant Physiol.* 114.239665. doi: 10.1104/pp.114.239665

Chen, F., Hu, Y., Vannozzi, A., Wu, K., Cai, H., Qin, Y., et al. (2017a). The WRKY transcription factor family in model plants and crops. *CRC. Crit. Rev. Plant Sci.* 36, 311–335. doi: 10.1080/07352689.2018.1441103

Chen, J., Nolan, T. M., Ye, H., Zhang, M., Tong, H., Xin, P., et al. (2017b). *Arabidopsis* WRKY46, WRKY54, and WRKY70 transcription factors are involved in brassinosteroid-regulated plant growth and drought responses. *Plant Cell* 29, 1425–1439. doi: 10.1105/tpc.17.00364

Chen, J., Ouyang, Y., Wang, L., Xie, W., and Zhang, Q. (2009). Aspartic proteases gene family in rice: gene structure and expression, predicted protein features and phylogenetic relation. *Gene* 442, 108–118. doi: 10.1016/j.gene.2009.04.021

Chen, W. J., and Zhu, T. (2004). Networks of transcription factors with roles in environmental stress response. *Trends Plant Sci.* 9, 591–596. doi: 10.1016/j.tplants.2004.10.007

Chen, Y., Wu, P., Zhao, Q., Tang, Y., Chen, Y., Li, M., et al. (2018). Overexpression of a phosphate starvation response AP2/ERF gene from physic nut in *arabidopsis* alters root morphological traits and phosphate starvation-induced anthocyanin accumulation. *Front. Plant Sci.* 9:1186. doi: 10.3389/fpls.2018.01186

Chiou, T.-J., and Lin, S.-I. (2011). Signaling network in sensing phosphate availability in plants. *Annu. Rev. Plant Biol.* 62, 185–206. doi: 10.1146/annurev-arplant-042110-103849

Cong, W.-F., Suriyagoda, L. D., and Lambers, H. (2020). Tightening the phosphorus cycle through phosphorus-efficient crop genotypes. *Trends Plant Sci.* 25, 967–975 doi: 10.1016/j.tplants.2020.04.013

Contouransel, D., Torresfranklin, M. L., Zuilyfodil, Y., and De Carvalho, M. H. C. (2010). An aspartic acid protease from common bean is expressed 'on call' during water stress and early recovery. *J. Plant Physiol.* 167, 1606–1612. doi: 10.1016/j.jplph.2010.06.018

Dai, X., Wang, Y., and Zhang, W.-H. (2015). OsWRKY74, a WRKY transcription factor, modulates tolerance to phosphate starvation in rice. *J. Exp. Bot.* 67, 947–960. doi: 10.1093/jxb/erv515

Devaiah, B. N., Karthikeyan, A. S., and Raghothama, K. G. (2007). WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *arabidopsis*. *Plant Physiol.* 143, 1789–1801. doi: 10.1104/pp.106.093971

Ding, Z. J., Yan, J. Y., Li, C. X., Li, G. X., Wu, Y. R., and Zheng, S. J. (2015). Transcription factor WRKY46 modulates the development of *Arabidopsis* lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. *Plant J.* 84, 56–69. doi: 10.1111/tpj.12958

Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10, 366–371. doi: 10.1016/j.pbi.2007.04.020

Griffith, M., Walker, J. R., Spies, N. C., Ainscough, B. J., and Griffith, O. L. (2015). Informatics for RNA sequencing: a web resource for analysis on the cloud. *PLoS Comput. Biol.* 11:e1004393. doi: 10.1371/journal.pcbi.1004393

Guo, R., Xu, X., Carole, B., Li, X., Gao, M., Zheng, Y., et al. (2013). Genome-wide identification, evolutionary and expression analysis of the aspartic protease gene superfamily in grape. *BMC Genomics* 14:554. doi: 10.1186/1471-2164-14-554

Guo, W., Zhao, J., Li, X., Qin, L., Yan, X., and Liao, H. (2011). A soybean β-expansin gene GmEXPB2 intrinsically involved in root system architecture responses to abiotic stresses. *Plant J.* 66, 541–552. doi: 10.1111/j.1365-313X.2011.04511.x

Hinsinger, P. (2001). Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: a review. *Plant Soil* 237, 173–195. doi: 10.1023/A:1013351617532

Huang, K.-L., Ma, G.-J., Zhang, M.-L., Xiong, H., Wu, H., Zhao, C.-Z., et al. (2018). The ARF7 and ARF19 transcription factors positively regulate Phosphate Starvation Response1 in *Arabidopsis* roots. *Plant Physiol.* 178, 413–427. doi: 10.1104/pp.17.01713

Javot, H., Pumplin, N., and Harrison, M. J. (2007). Phosphate in the arbuscular mycorrhizal symbiosis: transport properties and regulatory roles. *Plant Cell Environ.* 30, 310–322. doi: 10.1111/j.1365-3040.2006.01617.x

Jia, X., Liu, P., and Lynch, J. P. (2018). Greater lateral root branching density in maize improves phosphorus acquisition from low phosphorus soil. *J. Exp. Bot.* 69, 4961–4970. doi: 10.1093/jxb/ery252

Jones, D. L. (1998). Organic acids in the rhizosphere—a critical review. *Plant Soil* 205, 25–44. doi: 10.1023/A:1004356007312

Khorassani, R., Hettwer, U., Ratzinger, A., Steingrobe, B., Karlovsky, P., and Claassen, N. (2011). Citramalic acid and salicylic acid in sugar beet root exudates solubilize soil phosphorus. *BMC Plant Biol.* 11:121. doi: 10.1186/1471-2229-11-121

Kong, Y., Li, X., Wang, B., Li, W., Du, H., and Zhang, C. (2018). The soybean purple acid phosphatase GmPAP14 predominantly enhances external phytate utilization in plants. *Front. Plant Sci.* 9:292. doi: 10.3389/fpls.2018.00292

Li, C., Liu, X., Ruan, H., Zhang, J., Xie, F., Gai, J., et al. (2020). GmWRKY45 Enhances tolerance to phosphate starvation and salt stress, and changes fertility in transgenic *Arabidopsis*. *Front. Plant Sci.* 10:1714. doi: 10.3389/fpls.2019.01714

Liao, Y., Zou, H.-F., Wang, H.-W., Zhang, W.-K., Ma, B., Zhang, J.-S., et al. (2008). Soybean GmMYB76, GmMYB92, and GmMYB177 genes confer stress tolerance in transgenic *Arabidopsis* plants. *Cell Res.* 18, 1047–1060. doi: 10.1038/cr.2008.280

MacDonald, G. K., Bennett, E. M., Potter, P. A., and Ramankutty, N. (2011). Agronomic phosphorus imbalances across the world's croplands. *Proc. Nat. Acad. Sci.* 108, 3086–3091. doi: 10.1073/pnas.1010808108

Mutlu, A., and Gal, S. (1999). Plant aspartic proteinases: enzymes on the way to a function. *Physiol. Plant.* 105, 569–576. doi: 10.1034/j.1399-3054.1999.105324.x,

Niu, C. F., Wei, W., Zhou, Q. Y., Tian, A. G., Hao, Y. J., Zhang, W. K., et al. (2012). Wheat WRKY genes TaWRKY2 and TaWRKY19 regulate abiotic stress tolerance in transgenic *Arabidopsis* plants. *Plant Cell Environ.* 35, 1156–1170. doi: 10.1111/j.1365-3040.2012.02480.x

O'Rourke, J. A., Yang, S. S., Miller, S. S., Bucciarelli, B., Liu, J., Rydeen, A., et al. (2013). An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants. *Plant Physiol.* 161, 705–724. doi: 10.1104/pp.112.209254

Pandey, S. P., and Somssich, I. E. (2009). The role of WRKY transcription factors in plant immunity. *Plant Physiol.* 150, 1648–1655. doi: 10.1104/pp.109.138990

Péret, B., Clément, M., Nussaume, L., and Desnos, T. (2011). Root developmental adaptation to phosphate starvation: better safe than sorry. *Trends Plant Sci.* 16, 442–450. doi: 10.1016/j.tplants.2011.05.006

Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. *Front. Plant Sci.* 7:760. doi: 10.3389/fpls.2016.00760

Raghothama, K. (1999). Phosphate acquisition. *Annu. Rev. Plant Biol.* 50, 665–693. doi: 10.1146/annurev.arplant.50.1.665

Raimbault, A., Zuilyfodil, Y., Soler, A., and De Carvalho, M. H. C. (2013). A novel aspartic acid protease gene from pineapple fruit (*Ananas comosus*): cloning, characterization and relation to postharvest chilling stress resistance. *J. Plant Physiol.* 170, 1536–1540. doi: 10.1016/j.jplph.2013.06.007

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006

Scarpeci, T. E., Zanor, M. I., Mueller-Roeber, B., and Valle, E. M. (2013). Overexpression of AtWRKY30 enhances abiotic stress tolerance during early growth stages in *Arabidopsis thaliana*. *Plant Mol. Biol.* 83, 265–277. doi: 10.1007/s11103-013-0090-8

Singh, K. B., Foley, R. C., and Oñate-Sánchez, L. (2002). Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* 5, 430–436. doi: 10.1016/S1369-5266(02)00289-3

Song, H., Wang, P., Hou, L., Zhao, S., Zhao, C., Xia, H., et al. (2016). Global analysis of WRKY genes and their response to dehydration and salt stress in soybean. *Front. Plant Sci.* 7:9. doi: 10.3389/fpls.2016.00009

Song, H., Yin, Z., Chao, M., Ning, L., Zhang, D., and Yu, D. (2014). Functional properties and expression quantitative trait loci for phosphate transporter GmPT1 in soybean. *Plant Cell Environ.* 37, 462–472. doi: 10.1111/pce.12170

Sun, Y., and Yu, D. (2015). Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Rep.* 34, 1295–1306. doi: 10.1007/s00299-015-1787-8

Svistoonoff, S., Creff, A., Reymond, M., Sigoillot-Claude, C., Ricaud, L., Blanchet, A., et al. (2007). Root tip contact with low-phosphate media reprograms plant root architecture. *Nat. Genet.* 39, 792–796. doi: 10.1038/ng2041

Takahashi, K., Niwa, H., Yokota, N., Kubota, K., and Inoue, H. (2008). Widespread tissue expression of nepenthesin-like aspartic protease genes in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 46, 724–729. doi: 10.1016/j.plaphy.2008.04.007

Tamura, K., Peterson, D. S., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Ticconi, C. A., and Abel, S. (2004). Short on phosphate: plant surveillance and countermeasures. *Trends Plant Sci.* 9, 548–555. doi: 10.1016/j.tplants.2004.09.003

Timotijevic, G., Milisavljevic, M., Radovic, S., Konstantinovic, M. M., and Maksimovic, V. R. (2010). Ubiquitous aspartic proteinase as an actor in the stress response in buckwheat. *J. Plant Physiol.* 167, 61–68. doi: 10.1016/j.jplph.2009.06.017

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7:562. doi: 10.1038/nprot.2012.016

Vance, C. P. (2001). Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiol.* 127, 390–397. doi: 10.2307/4280097

Vance, C. P., Uhde-Stone, C., and Allan, D. L. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New Phytol.* 157, 423–447. doi: 10.1046/j.1469-8137.2003.00695.x

Vigeolas, H., Chinoy, C., Zuther, E., Blessington, B., Geigenberger, P., and Domoney, C. (2008). Combined metabolomic and genetic approaches reveal a link between the polyamine pathway and albumin 2 in developing pea seeds. *Plant Physiol.* 146, 74–82. doi: 10.1104/pp.107.111369

Wang, C., Ying, S., Huang, H., Li, K., Wu, P., and Shou, H. (2009). Involvement of OsSPX1 in phosphate homeostasis in rice. *Plant J.* 57, 895–904. doi: 10.1111/j.1365-313X.2008.03734.x

Wang, H., Xu, Q., Kong, Y.-H., Chen, Y., Duan, J.-Y., Wu, W.-H., et al. (2014). *Arabidopsis* WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1; 1 expression in response to phosphate starvation. *Plant Physiol.* 164, 2020–2029. doi: 10.1104/pp.113.235077

Wang, Z., Rong, D., Chen, D., Xiao, Y., Liu, R., Wu, S., et al. (2020). Salicylic acid promotes quiescent center cell division through ROS accumulation and down-regulation of PLT1, PLT2, and WOX5. *J. Integr. Plant Biol.* 68, 583–596. doi: 10.1111/jipb.13020

Williamson, L. C., Ribrioux, S. P., Fitter, A. H., and Leyser, H. O. (2001). Phosphate availability regulates root system architecture in *Arabidopsis*. *Plant Physiol.* 126, 875–882. doi: 10.1104/pp.126.2.875

Woo, J., MacPherson, C. R., Liu, J., Wang, H., Kiba, T., Hannah, M. A., et al. (2012). The response and recovery of the *Arabidopsis thaliana* transcriptome to phosphate starvation. *BMC Plant Biol.* 12:62. doi: 10.1186/1471-22 29-12-62

Wu, P., Ma, L., Hou, X., Wang, M., Wu, Y., Liu, F., et al. (2003). Phosphate starvation triggers distinct alterations of genome expression in *Arabidopsis* roots and leaves. *Plant Physiol.* 132, 1260–1271. doi: 10.1104/pp.103.021022

Yang, T., Hao, L., Yao, S., Zhao, Y., Lu, W., and Xiao, K. (2016a). TabHLH1, a bHLH-type transcription factor gene in wheat, improves plant tolerance to Pi and N deprivation via regulation of nutrient transporter gene transcription and ROS homeostasis. *Plant Physiol. Biochem.* 104, 99–113. doi: 10.1016/j.plaphy.2016.03.023

Yang, W. T., Baek, D., Yun, D.-J., Hwang, W. H., Park, D. S., Nam, M. H., et al. (2014). Overexpression of OsMYB4P, an R2R3-type MYB transcriptional activator, increases phosphate acquisition in rice. *Plant Physiol. Biochem.* 80, 259–267. doi: 10.1016/j.plaphy.2014.02.024

Yang, Y., Chi, Y., Wang, Z., Zhou, Y., Fan, B., and Chen, Z. (2016b). Functional analysis of structurally related soybean GmWRKY58 and GmWRKY76 in plant growth and development. *J. Exp. Bot.* 67, 4727–4742. doi: 10.1093/jxb/erw252

Yao, X., Xiong, W., Ye, T., and Wu, Y. (2012). Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in *Arabidopsis*. *J. Exp. Bot.* 63, 2579–2593. doi: 10.1093/jxb/err433

Zhang, D., Song, H., Cheng, H., Hao, D., Wang, H., Kan, G., et al. (2014). The acid phosphatase-encoding gene GmACP1 contributes to soybean tolerance to low-phosphorus stress. *PLoS Genet.* 10:e1004061. doi: 10.1371/journal.pgen.1004061

Zhou, J., Jiao, F., Wu, Z., Li, Y., Wang, X., He, X., et al. (2008). OsPHR2 is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants. *Plant Physiol.* 146, 1673–1686. doi: 10.1104/pp.107.111443

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li, Li, Liu, Ruan, Zheng, Yu, Gai and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 September 2021
doi: 10.3389/fpls.2021.719381

Morpho-Physiological Traits and Functional Markers Based Molecular Dissection of Heat-Tolerance in Urdbean

Debjyoti Sen Gupta[1,2], Partha S. Basu[3], J. Souframanien[4], Jitendra Kumar[1]*, P. Dhanasekar[4], Sanjeev Gupta[2], Muthaiyan Pandiyan[5], S. Geetha[5], P. Shanthi[5], Vaibhav Kumar[3] and Narendra Pratap Singh[6]*

[1] Division of Crop Improvement, ICAR-Indian Institute of Pulses Research, Kanpur, India, [2] All India Coordinated Research Project on Mungbean, Urdbean, Lentil, Lathyrus, Rajmash, and Fieldpea, ICAR-Indian Institute of Pulses Research, Kanpur, India, [3] Division of Basic Sciences, ICAR-Indian Institute of Pulses Research, Kanpur, India, [4] Nuclear Agriculture & Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, India, [5] National Pulses Research Centre, Vamban, India, [6] Division of Biotechnology, ICAR-Indian Institute of Pulses Research, Kanpur, India

Urdbean (*Vigna mungo* L. Hepper) is one of the important pulse crops. Its cultivation is not so popular during summer seasons because this crop is unable to withstand excessive heat stress beside lack of humidity in the atmosphere. Therefore, a panel of 97 urdbean diverse genotypes was assessed for yield under stress and non-stress conditions with an aim to identify heat tolerant genotypes. This study identified 8 highly heat tolerant and 35 highly heat sensitive genotypes based on heat susceptibility index. Further, physiological and biochemical traits-based characterization of a group of six highly heat sensitive and seven highly heat tolerant urdbean genotypes showed genotypic variability for leaf nitrogen balance index (NBI), chlorophyll (SPAD), epidermal flavnols, and anthocyanin contents under 42/25°C max/min temperature. Our results showed higher membrane stability index among heat tolerant genotypes compared to sensitive genotypes. Significant differences among genotypes for ETR at different levels of PAR irradiances and PAR × genotypes interactions indicated high photosynthetic ability of a few genotypes under heat stress. Further, the most highly sensitive genotype PKGU-1 showed a decrease in different fluorescence parameters indicating distortion of PS II. Consequently, reduction in the quantum yield of PS II was observed in a sensitive one as compared to a tolerant genotype. Fluorescence kinetics showed the delayed and fast quenching of Fm in highly heat sensitive (PKGU 1) and tolerant (UPU 85-86) genotypes, respectively. Moreover, tolerant genotype (UPU 85-86) had high antioxidant activities explaining their role for scavenging superoxide radicals (ROS) protecting delicate membranes from oxidative damage. Molecular characterization further pinpointed genetic differences between heat tolerant (UPU 85-86) and heat sensitive genotypes (PKGU 1). These findings will contribute to the breeding toward the development of heat tolerant cultivars in urdbean.

Keywords: ***Vigna mungo*, heat tolerance, abiotic stress, membrane stability, electron transport rate, heat susceptibility index, chlorophyll fluorescence, molecular markers**

OPEN ACCESS

Edited by:
Jose C. Jimenez-Lopez, Estación Experimental de Zaidín, Consejo Superior de Investigaciones Científicas, Spain

Reviewed by:
Marek Zivcak, Slovak University of Agriculture, Slovakia
Krishnendu Chattopadhyay, National Rice Research Institute (ICAR), India
Jun-ichi Sakagami, Kagoshima University, Japan

****Correspondence:***
Debjyoti Sen Gupta
debgpb@gmail.com;
debjyoti.gupta@icar.gov.in
Jitendra Kumar
jitendra73@gmail.com

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *02 June 2021*
Accepted: *02 August 2021*
Published: *29 September 2021*

Citation:
Sen Gupta D, Basu PS, Souframanien J, Kumar J, Dhanasekar P, Gupta S, Pandiyan M, Geetha S, Shanthi P, Kumar V and Pratap Singh N (2021) Morpho-Physiological Traits and Functional Markers Based Molecular Dissection of Heat-Tolerance in Urdbean. Front. Plant Sci. 12:719381. doi: 10.3389/fpls.2021.719381

INTRODUCTION

Urdbean (*Vigna mungo* L. Hepper) is a popular food legume grown in many Asian countries including India, Pakistan, Myanmar, Bangladesh, Thailand, and China. India is the largest producer and consumer of urdbean. It produces about 3.36 million tons of urdbean (Project Coordinator's Report, 2019-2020) and imports another 0.5 million tons from other urdbean growing countries, particularly from Myanmar. During 2014-2015, Myanmar produced 1.51 million tons of urdbean that is locally known as black matpe bean. Nutritionally, urdbean is dense with protein (21-28%), dietary fiber (161-187 g/kg), iron (16-255 mg/kg), zinc (5-134 mg/kg), and other micronutrients like other pulses (Chitra et al., 1996; Sen Gupta et al., 2020). Therefore, its nutrient-dense profile has encouraged an introduction to many developed countries including the United States, Russia, and European nations as a potential pulse crop (Sen Gupta et al., 2020).

Urdbean is grown in different ecologies and seasons across the growing regions. In India, it is grown mainly in the rainy season (July-October) and in the southern part it is also cultivated as a winter season crop (November to February). However, its cultivation is not wide in the summer season due to excessive heat stress and a lack of humidity in the atmosphere. Thus, availability of heat tolerant cultivars can bring more areas under urdbean cultivation. Previously, genetic variability for heat tolerance was reported in many food legumes (Sita et al., 2017), but it is not yet explored in urdbean. It is a warm season food legume, which requires 25-35°C temperature along with high humidity for its normal growth and development. However, prevailing high temperature (>40°C) during flowering results in deformation of flower parts or flower drop leading to negative impact on yield. Similarly, in mungbean, higher temperatures of >38/25°C (day and night, respectively) markedly affected the yield under summer-season cultivation (Nayyar et al., 2017).

The effect of heat stress results in drastic yield losses due to pollen or ovule inactivity, flower abortion, and even post-fertilization impaired growth and development of embryo or seed in many pulses (Sita et al., 2017). Moreover, the current climate change scenario also leads to abrupt changes in mean temperature. Therefore, breeding of heat tolerant urdbean varieties becomes more relevant under such situations. Urdbean is a close relative of mungbean, which is extensively cultivated in identical ecologies. In this crop as well as in another *Vigna* pulse crop, cowpea, sources of heat tolerance have already been identified (Ehlers and Hall, 1998; Basu et al., 2019).

Knowledge of key traits imparting heat tolerance can help to improve the grain yield of urdbean (Scafaro et al., 2010). Therefore, physio-biochemical mechanisms underlying these key traits are essential to screen large numbers of germplasm at critical temperature under both field and controlled conditions (Gaur et al., 2019). In several other crops, various physiological traits such as photosynthetic activity, membrane stability, pollen viability, and phenolic compounds have been used to identify heat tolerant genotypes (Allakhverdiev and Murata, 2004; Asseng et al., 2015; Sita et al., 2017) and genetic variability has been reported for key physiological traits under heat stress conditions (Challinor et al., 2007).

Urdbean is a highly photothermo-sensitive crop. Therefore, its yield potential varies across locations due to variable daylength and thermal regimes. Thus, minimizing the genotype × environment interactions can help to achieve stable yield of urdbean. The high temperature stress above the threshold across the locations during the summer season could be the compounding effects of both heat and photosensitivity. One of the strategies for selecting photo-thermo insensitive lines is to evaluate different genotypes at multi-locations having varying daylength and thermal regimes. As a result, genotypes having stable yield across the locations could be identified as putative photo-thermo insensitive lines. This strategy should be made to screen thermo-tolerant lines from the panel of photo-thermo insensitive lines so that widely adapted stable heat tolerant lines could be identified having less influence of photo-thermoperiods. In the present investigation, this approach has been followed to identify contrasting genotypes having a high level of tolerance or sensitivity to high temperature.

Knowledge of genetics underlying key traits imparting heat tolerance helps the breeder to make genetic improvements more precisely. In recent years, molecular markers helped to decipher the genetics of complex key morpho-physiological traits imparting heat tolerance in several crops (Argyris et al., 2008; Roy et al., 2011; Paliwal et al., 2012). However, in urdbean, use of molecular markers for mapping and characterization of traits related to heat tolerance is poorly understood. Currently, simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) are available for molecular characterization in urdbean (Raizada and Souframanien, 2019; Souframanien et al., 2020; Pootakham et al., 2021). Hence, this experiment was designed with the following objectives: (i) to evaluate a set of urdbean genotypes under field conditions with natural heat stress conditions (flowering and podding stage coincides with high temperature), and to compare it with normal field conditions (comparatively less exposure to high temperature during flowering and podding), (ii) to precisely phenotype selected contrasting heat tolerant and sensitive genotypes for different physiological and biochemical traits, and (iii) to characterize heat tolerant and sensitive genotypes with heat-related genic markers.

MATERIALS AND METHODS

Plant Materials

Plant materials comprised of 97 urdbean genotypes, which were grown during the summer season of 2016 at the Main Research Farm of Indian Institute of Pulses Research (IIPR), Kanpur (26.28°N and 80.21°E), and National Pulses Research Centre of TNAU, Vamban (10.20°N, 78.50° E) in India. The tested urdbean genotypes comprised of germplasm, breeding lines, and cultivars of diverse origins (**Supplementary File 1**). Maturity duration for these genotypes ranged from 70 to 75 days. The field experiments were grown in augmented- randomized

complete block (RCB) design. Three checks (Uttara, Shekhar-2, IPU 02-43) were replicated with randomization in each one of the five blocks. Each plot consisted of double rows of 4 m length. Rows were spaced 30 cm apart and interplant distance was 10 cm. Two trials were conducted at each location and based on meteorological data and average yield of trial one was designated as "stress environment (SE)" and another was named as "non-stress environment (NSE)." Standard practices were followed to raise the rainfed crop excluding one pre-sown irrigation.

Heat Susceptibility Index

Heat susceptibility index (HSI) for each individual urdbean genotype was calculated using the equation by Fischer and Maurer (1978): HSI = (1-Yh/Y)/(1-Xh/X) where Yh and Y are the phenotypic means (Yield) for each genotype under heat stress and non-heat stress conditions, respectively, and Xh and X are the phenotypic means (Yield) for all lines under heat stressed and non-heat stress conditions, respectively.

Meteorological Data Collection

Weather data from Kanpur and Vamban locations were recorded throughout the growing period by the respective meteorological observatories present in both places.

Physiological Characterization of Heat Tolerant and Sensitive Genotypes

Plant Samples Under Controlled Environment

Seeds of selected contrasting urdbean genotypes were obtained from the urdbean breeding program of IIPR, Kanpur. Seeds were surface sterilized with 70% ethanol for 5 min, followed by treatment with 1% sodium hypochlorite (v/v) for 3 min. The sterilized seeds were rinsed 3 times with sterile Milli-Q (Merck Millipore, Germany) water under aseptic conditions and soaked overnight at room temperature.

The sterilized seeds were sown in cocopit–vermicompost–soil mixture (3:1:1 ratio) and irrigated with Hoagland solution. The plants were raised under a controlled environment chamber (Hi-point, Taiwan) and maximum minimum temperature 40/25°C with 14-h photoperiod was maintained. The light sources were RGB LED (Red-Green-Blue-White) having an irradiance level of 460 μmol photons $m^{-2}s^{-1}$ and relative humidity 80%. The required moisture and fertility of the soil compost was ensured by irrigating with water or Hoagland solution at regular intervals.

Nitrogen Balance Index, Chlorophyll, Anthocyanin, and Flavanols Contents

The physiological status of selected plant leaves was determined using a hand-held device DUALEX leaf clip device (Force-A, France), which enabled comparative values of leaf chlorophyll (leaf greenness) content (Chl), epidermal flavanols (Flv), nitrogen balance index (NBI), and anthocyanin content of leaves subjected to heat stress in a sensor-controlled chamber consistently maintained at 40/25°C (maximum/minimum).

Measurement of chlorophyll

DUALEX measures the chlorophyll content of a leaf based on the transmittance ratio at two different wavelengths. One in the far-red absorbed by chlorophyll and one in the near-infrared as a reference. The leaf chlorophyll content can rapidly and accurately be assessed from light transmittance. A first wavelength very close to the red quantifies the chlorophyll and a second in the near-infrared can take into account the effects of leaf structure.

Chlorophyll index = (Near-infrared transmittance – Red transmittance)/(Red transmittance)

Measurements of polyphenols (flavanols) and anthocyanin

DUALEX measures flavanols and anthocyanins content of the leave's epidermis based on differential ratio of chlorophyll fluorescence. Near-infrared chlorophyll fluorescence is measured under a first reference excitation light not absorbed by polyphenols. It is compared to a second sampling light specific to a particular type of polyphenols (e.g., green for anthocyanins or UV-A for flavanols). Only a fraction of this light reaches the chlorophyll in the mesophyll and can generate near-infrared fluorescence.

Flavanol index = Log (Near-infrared fluorescence excited red/ Near-infrared fluorescence excited UV-A)

Anthocyanin index = Log (Near-infrared fluorescence excited red/ Near-infrared fluorescence excited green)

Differential measurement of fluorescence emitted by chlorophyll

The difference in chlorophyll fluorescence measured in the near-infrared is thus directly proportional to the amount of polyphenols (flavanols) present in the epidermis of the leaf.

Measurement of nitrogen balance index

It is the ratio of chlorophyll to flavanol index. Polyphenols, specifically flavanols, are indicators of nitrogen status of plants. Indeed, when a plant is under optimal conditions, it favors its primary metabolism and synthesizes proteins (nitrogen-containing molecules) containing chlorophyll and a few flavanols (carbon-based secondary compounds). On the contrary, in case of nitrogen deficiency, the plant directs its metabolism toward an increased production of flavanols.

Membrane Stability

The membrane stability index (MSI) was determined using the electrolyte leakage (EL) method. For keeping uniformity among samples, the well-developed fully expanded fourth leaf from the top of test plants was collected, washed using distilled water, surface dried, and dipped in deionized water at 40°C for 1 h. The electrical conductivity (EC) of tissue leachates was measured using a conductivity meter (Model HI2300, Hanna, United States). The contents were incubated further by dipping the same leaf in deionized water at 100°C for 1 h and EC was measured. The MSI was calculated by the following formula:

MSI = C1/C2, where C1 = EC (EC μS) at 40°C for 1 h and C2 = EC (EC μS) at 100°C for 1 h (Blum and Ebercon, 1981)

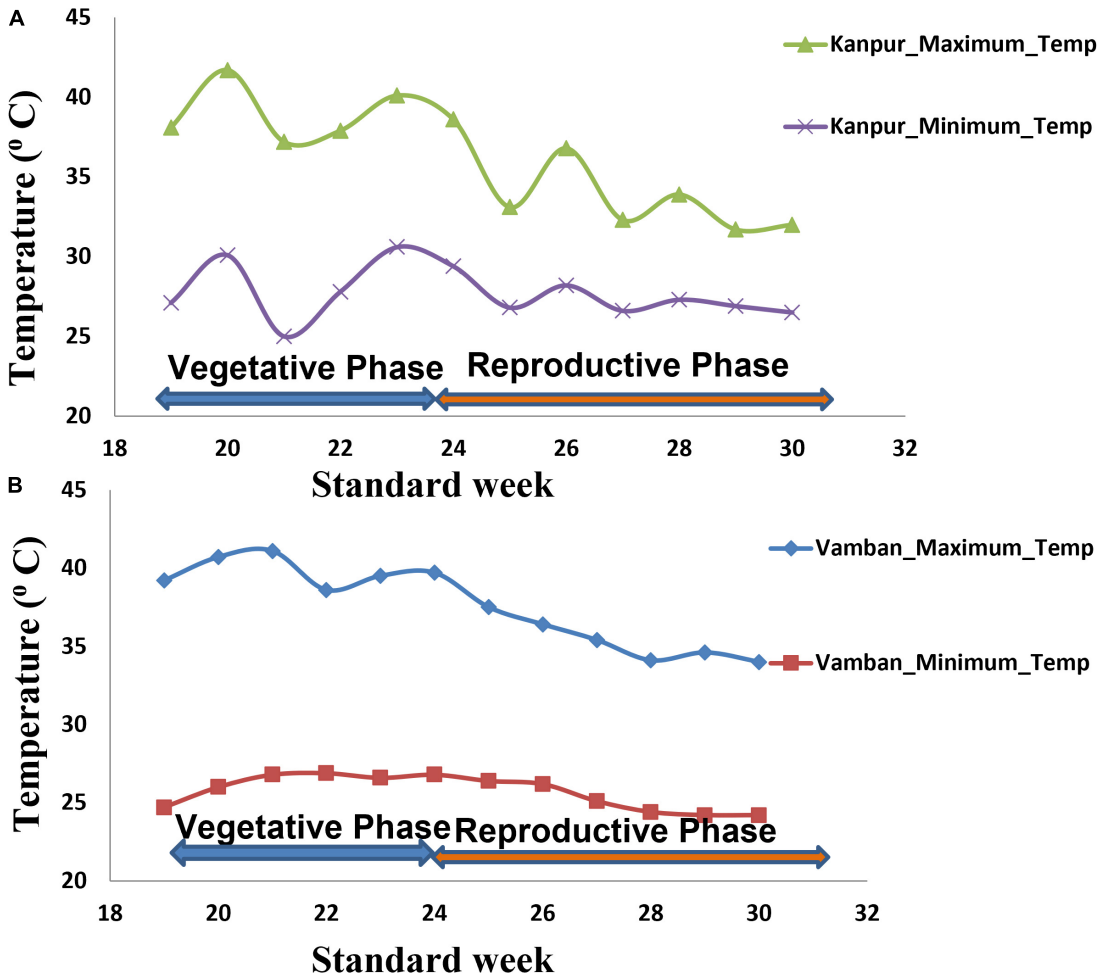


FIGURE 1 | Temperature regime at two field trial sites during crop growth period (Kanpur and Vamban). **(A)** Temperature regime at the Kanpur location. **(B)** Temperature regime at the Vamban location.

Fluorescence Image Analysis

Leaf samples of all high temperature (40/25°C; maximum/minimum) grown urdbean genotypes from both groups (heat tolerant and sensitive) were used for chlorophyll fluorescence studies as described by Schreiber and Bilger (1987). High temperature grown genotypes were given hot water heat shock at 43°C for 1 h and thereafter stressed leaves were dark-adapted for 10 min in a temperature-controlled chamber and image analysis was conducted. Photosynthetic response between the tolerant and sensitive lines was assessed using a fluorescence imaging system (Mess & Regeltechnik, Waltz, Germany). The dark-adapted leaves were subjected to 0.05 μmol weak 2 Hz modulated light for 100 μs followed by superimposing saturation light pulses of 4000 μmol $m^{-2}s^{-1}$ PAR for 400 ms to obtain quantum yield (Fv/Fm; variable to maximum fluorescence ratio) and fluorescence images were captured. Subsequently, leaves were exposed to actinic light of 200 μmol photons $m^{-2}s^{-1}$ for 2 min for light adaptation. Same saturated pulses were superimposed to obtain quantum yield in light-adapted leaves. Quantum yield (F_V/F_m), maximal fluorescence (F_m), minimum fluorescence (F_0), and quantum yield of non-regulated energy dissipation [Y(NO)] values were compared between heat tolerant and sensitive genotypes.

Photosynthetic Electron Transport Rate

All tested 13 genotypes were pretreated with thermal shock at 43°C for 1 h by inserting leaves in a circulating hot water bath. This temperature was considered detrimental for the photosynthetic membrane and induces disorganization of photosystems and membrane bound electron transport components. Light response of ETR representing the photosynthetic activity of leaves of all tested urdbean genotypes was studied using software ImagingWin (Walz-Imaging System, GmbH, Germany) employing an irradiance range of 200–700 $\mu molm^{-2}s^{-1}$. The light curve and initial fluorescence values (F_0 and F_m, respectively) of the dark-adapted leaves were used for calculation of ETR (ETR = Quantum yield × PAR × 0.5 × Absorptivity). Absorptivity describes the fraction of incident light, which is absorbed, and 0.5 indicates that only half of the

absorbed quanta is distributed to PS II (under steady state conditions). The light curve of an individual selection was obtained with increasing order of irradiance until ETR was light saturated.

Fluorescence Parameters During Light–Dark Transition

After measuring the F_0, F_m and F_v/F_m in dark-adapted leaves, the leaves were exposed to actinic light of irradiance 200 $\mu molm^{-2} s^{-1}$ and then saturated light pulse was triggered at every 50 s to obtain F_0, F_m and F_v/F_m in light-adapted leaves until 250 s of illumination. Thereafter, actinic light was switched off and F_0, F_m and F_v/F_m were measured at every 50 s in order to ascertain the restoration of normal F_0, F_m and F_v/F_m in heat-tolerant and sensitive lines during a dark cycle.

In another experiment, high temperature grown contrasting urdbean genotypes were allowed to adapt in the dark for 5 min and thereafter saturated light flash 4000 $\mu molm^{-2} s^{-1}$ was triggered for 100 ms to obtain F_0 and F_m. Then, leaves were exposed to actinic light 200 $\mu molm^{-2} s^{-1}$ for light adaptation. The light phase was continued until 350 s and then at every 15 s saturated pulse was applied to obtain F_0 and F_m. Thereafter, leaves were put into a dark phase for adaptation and in a similar manner a saturated pulse was applied at every 15 s to obtain F_0 and F_m. The only difference between these two events was fluorescence kinetics in light followed by in dark to see the recovery of F_0 and F_m in a dark phase.

Biochemical Parameters-*in vivo* Visualization of Superoxide Radical and Hydrogen Peroxide

In vivo visualization of superoxide radical

In vivo assay of superoxide radical in the leaf was carried out according to the method of Frahry and Schopfer (2001). Fresh leaf samples were collected and dipped in staining solution for 1 h. The staining solution was composed of 10 mM sodium azide, 100 mM potassium phosphate (pH 7.8), and 0.1 % Nitroblue tetrazolium. After 1 h, leaf samples were bleached by immersing them into boiling ethanol for 15 min. The bleaching solution decolorized the leaves except the dark blue insoluble formazan deposits formed by the reaction of NBT with a superoxide radical. The photographs of the stained samples were captured using a good quality camera for further use.

In vivo visualization of hydrogen peroxide

The visualization of hydrogen peroxide in the leaf samples was examined using the method of Christensen et al. (1997). The collected leaf samples were washed using double distilled water. The washed samples were dipped into a solution containing 0.1% 3,3′-diaminobenzidine (DAB) dissolved in HCl acidified water (pH 3.8). Then, it was incubated for 16 h to allow the uptake of DAB and its reaction with H_2O_2 and peroxidase. The leaf samples were bleached by immersing them in boiling ethanol for 15 min. The photographs of the stained samples were captured using a good quality camera for further use.

Molecular Characterization

Genic SSR markers were used to characterize eight heat-sensitive (IPU99-200, IC-21001, Shekhar-2, Uttara, PU-19, HPU-120, H-1, PKGU-1) and eight heat-tolerant (UPU-85-86, IPU94-2, IPU-98/36, No. 5/31, PGRU-95014, PGRU-95016, PLU-1, BGP-247) genotypes in the present study. Details of 21 genic-simple sequence repeat (SSR) markers were provided in **Table 13**.

DNA Extraction and PCR

DNA was extracted from 1-day-old seedlings by the Dellaporta et al. (1983) method. The SSR primer pairs for sequence-specific markers were designed from leguminous crops having relevance to abiotic stress tolerance (**Table 13**). PCR reactions were carried out in a 25-µl reaction volume in an Eppendorf Master Cycler (Eppendorf, Hamburg, Germany) with the following composition: 50 ng of genomic DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.08% Non-idet P40, 0.2 mM dNTPs, 1.5 pmoles each of forward and reverse primers, and 0.5 unit Taq DNA polymerase (Fermentas, Life Sciences). The amplification conditions were initial denaturation at 94°C for 3 min and 5 cycles at 94°C for 30 s, 56 to 46°C (-1°C each cycle), 72°C for 1 min followed by 35 cycles at 94°C for 30 s, 46°C for 1 min, 72°C for 1 min, and ends up with a final extension at 72°C for 7 min. PCR products were resolved on 3% agarose gel in TBE buffer at 80 V and the image was captured in a gel documentation system (Syngene, United Kingdom).

Statistical Analysis

Analysis of variance for yield at each growing environment (Kanpur-stress, Kanpur-nonstress, Vamban-stress, Vamban-non-stress) was performed using a statistical package augmented RCBD (Aravind et al., 2020) in RStudio application using R (Programming Language for Statistical Analysis) (R Core Team, 2019). The yield data across four growing environments were graphically analyzed for interpreting G × E interaction using the "GGEBiplotGUI" statistical package in RStudio software using R (Frutos et al., 2014). GGE biplot methodology, which is composed of two concepts, the biplot concept and the GGE concept, was used for yield analysis across locations (Gabriel, 1971; Yan, 2001). This methodology uses a biplot to show the factors (G and G × E) that are important for evaluation of genotypes and that are also the sources of variation in G × E interaction analysis of multi- location trial data (Yan et al., 2000; Yan, 2001). All the physiological and biochemical data points were subjected to statistical analysis using Microsoft Excel software. For molecular data, all gels were scored manually, and data were input into Excel (Microsoft) spreadsheets. The band data were scored as a 1/0 (presence/absence) matrix. Genetic similarity coefficients of pair-wise comparisons among the accessions analyzed were calculated based on Jaccard's similarity coefficient (Jaccard, 1908) within the Similarity for Qualitative Data (SIMQUAL) module of NTSYS 2.02i (Rohlf, 1998). The Unweighted Pair Group Method with

TABLE 1 | Yield of 97 urdbean genotypes grown in IIPR, Kanpur and TNAU, Vamban.

Sl. No.	Genotype	IIPR, Kanpur		TNAU, Vamban	
		SE#	NSE#	SE	NSE
1	IPU 91-7	862	598	858	1000
2	IPU 94-2	1016	998	1184	831
3	IPU 95-13	589	998	1011	1165
4	Pant U-30	516	1032	1302	725
5	LBG 20	529	1232	1030	876
6	UPU 97-10	429	732	929	471
7	NO 7668-4B	649	1132	1052	1020
8	PGRU 95018	722	865	1043	817
9	PGRU 95014	1036	1032	1002	595
10	PGRU 95016	1109	832	1787	1434
11	TU 99-293	756	1565	1086	1109
12	Pant U-19S	1056	1698	1445	1408
13	TU 99-2	1242	1832	859	1079
14	TU 91-22	902	1365	1506	1133
15	PLU-28	849	1498	1783	1401
16	PLU-1	1236	898	1541	1232
17	UH -177	1216	1498	1325	1683
18	BG-369	1400	1187	618	1702
19	BGP 21-28	1307	1420	388	954
20	U-9	780	1020	421	758
21	IC 106088	1700	1320	511	1023
22	UH 32-3	1427	1487	1254	959
23	STY 2868	1607	1187	178	773
24	UH 85-15	1220	920	537	847
25	IPU 90-32	1480	720	795	849
26	IPU 90-321	674	1720	947	1063
27	IPU 99-79	960	1520	421	1067
28	PLU-8	1140	1287	451	675
29	IPU 99-123	914	987	792	679
30	UH 80-26	1040	1387	661	1015
31	IPU 99-23	900	953	871	642
32	IC -21001	914	1253	447	1173
33	IPU 99-95	807	1187	245	518
34	IPU 99-40	1060	1387	325	526
35	PKGU-1	560	987	350	567
36	IPU 99-89	614	1353	217	587
37	NG-2119	860	1053	360	1119
38	NO- 5731	1347	765	838	537
39	Mash 1-1	880	965	722	396
40	UG 414	1454	1165	684	683
41	IPU 96-6	954	798	379	324
42	IPU 98/36	700	632	800	385
43	U 3108	1174	665	879	1130
44	DUS 34	540	932	472	781
45	STY 2289	1260	532	396	419
46	IC-65511	1047	165	171	677
47	UH 99-144	1607	432	790	1016
48	UH 86-5	520	632	1207	1014
49	STY-2834	680	498	1162	1383
50	PLU-429	860	1232	558	1377

(Continued)

TABLE 1 | (Continued)

Sl. No.	Genotype	IIPR, Kanpur		TNAU, Vamban	
		SE[#]	NSE[#]	SE	NSE
51	PLU-144	574	965	1167	1076
52	UPU 85-86	1100	532	1335	1020
53	STY 2115	1134	765	1026	1388
54	IPU 99-31	914	398	939	760
55	U-132	760	632	967	1109
56	N0 7368-15	587	765	709	1067
57	UH 80-38	890	1343	851	1248
58	PLU-65	477	1876	1372	1680
59	BGP-247	1130	876	1162	672
60	PLU 456	1190	1343	666	1208
61	UG -218	870	1576	1332	1285
62	PDU-3	2210	1976	746	627
63	PLU -328	917	1676	444	476
64	NHKD-31	950	1276	726	938
65	STY-2824	677	1876	928	998
66	IPU 96-1	744	1309	538	1015
67	IPU2K-21	850	1176	1912	426
68	IPU-722	784	1476	820	1311
69	IC-10703	850	1276	564	1012
70	IPU 96-12	844	1076	559	625
71	IPU 99-22	424	1143	1487	609
72	PLU-703	510	809	975	441
73	PLU-557	1190	1343	1064	513
74	UG-378	224	876	583	268
75	IPU 99-128	830	1343	672	894
76	H-1	624	909	429	824
77	IPU 99-40	760	1542	513	1189
78	UPU 83-3	333	1975	672	716
79	PLU-662	647	1142	645	1393
80	UH 87-7	387	1208	745	370
81	UH 84-4	713	1908	966	1399
82	IPU 99-43	787	1475	1146	1075
83	PDU-1	1093	1208	935	323
84	IPU 99-18	1380	1342	526	867
85	IPU 99-16	1293	1475	538	743
86	IPU 99-200	733	1942	702	1888
87	UH 85-3	633	2142	1169	786
88	HPU-120	1013	2342	444	663
89	JU 78-27	833	2208	947	1130
90	PU-19	493	2075	615	918
91	IPU 99-209	453	1542	1388	1224
92	IPU 99-232	867	1675	1166	1317
93	IPU 99-221	820	2008	990	665
94	IPU 99-179	820	1775	1222	533
95	Uttara	998	1348	1173	1246
96	Shekhar-2	871	1282	943	1030
97	IPU 02-43	944	964	1294	1186
	Mean	**910**	**1224**	**859**	**925**
	Standard error	**34**	**45**	**38**	**35**

[#]*SE denotes stressed environment, NSE denotes non-stressed environment.*

TABLE 2 | ANOVA of yield over stress and non-stress environments in IIPR, Kanpur and TNAU, Vamban.

Source	MS (Kanpur SE)	MS (Kanpur NSE)	MS (Vamban SE)	MS (Kanpur NSE)
Block (ignoring treatments)	342770**	685647**	143012**	207617**
Treatment (eliminating blocks)	100158**	111906*	121168**	115039**
Treatment: Check	20366	211098*	61912**	159219**
Treatment: Test and test vs. check	101754**	109922*	122365**	114146**
Residuals	11342	35865	1400	20663

** and ** denotes that mean square was significantly different at $p < 0.05$, $p < 0.01$, respectively.*

Arithmetic Mean (UPGMA) clustering method was used to construct the dendrogram.

RESULTS

Characterization of Heat Stress Conditions and Identification of Heat Tolerant Genotypes

In the present study, natural heat stress conditions were determined based on mean yield obtained over 97 genotypes at two different locations (Kanpur and Vamban). The Kanpur location is situated in the northern part of India (26.28°N and 80.21°E) where early sown (mid-May) genotypes experienced heat stress with a rise of temperature (>40°C) coinciding with the reproductive stage, whereas late sown genotypes received moderate temperature (<40°C) during onset of flower followed by pod setting to grain development (**Figure 1**). Mean yield of early sown trials at this location was low (910 kg/ha) compared to late sown trial (1224 kg/ha) ($p < 0.05$). Similarly, the Vamban location is the extreme southern part of India (10.20°N, 78.50°E) where early sown crops are usually subjected to stress conditions with a rise of temperature to the extent of about 40°C during the reproductive stage (**Figure 1**). Early sown trials of this location showed low average yield (895 kg/ha) compared to late sown trials that had low temperature (<40°C) during the reproductive stage and higher mean yield (925 kg/ha) for 97 genotypes (**Table 1**) ($p < 0.05$). Analysis of variation over 97 genotypes for yield showed significant genotypic differences at $p < 0.05$ and < 0.01 probabilities under stress (early) and non-stress (late) conditions at both locations (**Table 2**). The heat sensitive and tolerant genotypes were identified at a preliminary stage based on the heat susceptibility index (HSI) under a field trial conducted in two contrasting environments. Sensitive genotypes were characterized with an HSI ranging from 0.08 to 3.19 at the Kanpur location and from 0.37 to 13.75 at the Vambam location, while HSI varied from -0.01 to -20.48 at Kanpur and -0.03 to –62.29 at Vambam among tolerant genotypes (**Table 3**). GGE biplot analysis identified most stable genotypes over the locations for yield (**Figure 2**).

Physiological Characterization

Field trials identified 8 highly heat tolerant and 35 highly heat sensitive genotypes based on HSI (**Table 4**). Among these, six highly sensitive (IPU 99-200, IC 21001, Shekhar 2, PU 19, H-1, PKGU 1) and seven highly tolerant (UPU 85-86, IPU 94-2, IPU 98/36, NO- 5731, PGRU 95016, PLU 1, BGP 247) genotypes, showing stable HSI over both locations, were used for further physiological analyses (**Table 4**). These genotypes were grown under a controlled environment right from seedling stage until maturity in a high thermal regime (40/25°C: maximum/minimum) with high humidity and under optimum irrigation and soil fertility.

Physiological Status Under Stress Environment

Changes in the physiological status were observed between two contrasting groups having different degrees of heat sensitivity when they were grown under higher thermal regime (40/25°C). The genotypic variability in nitrogen balance index (NBI) and chlorophyll (Chl) was significant ($p < 0.05$). No significant differences were observed among tested genotypes for anthocyanin and flavanol content when subjected to heat stress (**Table 5**). However, group comparison (heat tolerant vs. heat sensitive) (*t*-test) revealed significant differences in leaf nitrogen status (NBI) and anthocyanin (Anth) at $p < 0.01$ (**Table 6**).

In the present study, a range of genetic variability was observed higher among seven heat tolerant genotypes (36.9-64.6 and 9.7-21.1) compared to six heat sensitive genotypes (23.0-48.7 and 6.2-19.4) for NBI and chlorophyll content, respectively (**Table 6**).

Membrane Stability

Analysis of variance showed significant differences among the studied urdbean genotypes for membrane stability index (**Table 5**). It ranged from 32.3% to 74.5% in sensitive genotypes while it ranged from 34.5 to 62.8% in tolerant genotypes (**Table 7**). Although membrane stability was observed significantly higher in the sensitive genotype IPU 99-200 (74.5%), membrane stability was on average higher among tolerant genotypes compared with sensitive genotypes (**Table 7**). Among tolerant genotypes, maximum membrane stability was observed in PLU-1 (62.8 %) followed by UPU 85-86 (60.7 %) (**Table 7**).

Correlation Analysis Among Nitrogen Balance Index, Chlorophyll, Flavanol, Anthocyanin Contents, and Membrane Stability

A highly significant ($p < 0.01$) positive correlation ($r^2 = 0.85$) was observed between NBI and chlorophyll

TABLE 3 | Heat susceptibility index (HSI) of 97 urdbean genotypes under IIPR, Kanpur and TNAU, Vamban condition.

Rank	IIPR, Kanpur (26.28°N and 80.21°E)		TNAU, Vamban (10.20°N, 78.50° E)	
	Genotype	HSI	Genotype	HSI
1	IC-65511	−20.48	IPU2K-21	−62.29
2	UH 99-144	−10.42	PDU-1	−33.83
3	STY 2289	−5.24	IPU 99-22	−25.74
4	IPU 99-31	−4.97	IPU 99-179	−23.08
5	UPU 85-86	−4.09	PLU-703	−21.62
6	IPU 90-32	−4.04	UG-378	−20.99
7	U 3108	−2.93	IPU 98/36	−19.25
8	NO- 5731	−2.91	PLU-557	−19.18
9	STY 2115	−1.85	UH 87-7	−18.1
10	IPU 91-7	−1.69	UPU 97-10	−17.36
11	PLU-1	−1.44	Mash 1-1	−14.7
12	STY-2834	−1.4	Pant U-30	−14.21
13	STY 2868	−1.36	BGP-247	−13.02
14	PGRU 95016	−1.28	PGRU 95014	−12.21
15	UH 85-15	−1.25	NO- 5731	−10.01
16	BGP-247	−1.11	IPU 99-221	−8.73
17	IC 106088	−1.1	UH 85-3	−8.7
18	UG 414	−0.95	IPU 94-2	−7.59
19	U-132	−0.78	IPU 99-23	−6.37
20	IPU 96-6	−0.75	TU 91-22	−5.88
21	BG-369	−0.69	UPU 85-86	−5.51
22	PDU-3	−0.45	UH 32-3	−5.49
23	IPU 98/36	−0.41	PGRU 95018	−4.94
24	IPU 99-18	−0.11	PLU-28	−4.87
25	IPU 94-2	−0.07	PLU-1	−4.48
26	PGRU 95014	−0.01	PGRU 95016	−4.4
27	IPU 02-43	0.08	IPU 99-31	−4.21
28	UH 32-3	0.15	UH 86-5	−3.4
29	IPU 99-23	0.21	PDU-3	−3.39
30	IPU 99-123	0.28	LBG 20	−3.14
31	BGP 21-28	0.3	IPU 96-6	−3.03
32	Mash 1-1	0.34	IPU 99-123	−2.97
33	PDU-1	0.36	IPU 99-209	−2.39
34	PLU 456	0.44	IPU 02-43	−1.63
35	PLU-557	0.44	PLU-144	−1.51
36	PLU-8	0.44	IPU 99-43	−1.18
37	IPU 99-16	0.47	UG -218	−0.65
38	PGRU 95018	0.63	NO 7668-4B	−0.56
39	UH 86-5	0.68	Pant U-19S	−0.47
40	NG-2119	0.7	UG 414	−0.03
41	UH -177	0.72	TU 99-293	0.37
42	IPU 96-12	0.83	STY 2289	0.98
43	N0 7368-15	0.89	Uttara	1.05
44	U-9	0.9	UPU 83-3	1.1
45	IPU 99-40	0.9	IPU 90-32	1.14
46	UH 80-26	0.96	PLU -328	1.2
47	NHKD-31	0.98	STY-2824	1.25
48	Uttara	0.99	Shekhar-2	1.51
49	IC -21001	1.04	IPU 96-12	1.89
50	IPU2K-21	1.06	IPU 90-321	1.95

(Continued)

TABLE 3 | (Continued)

Rank	IIPR, Kanpur (26.28°N and 80.21°E)		TNAU, Vamban (10.20°N, 78.50° E)	
	Genotype	HSI	Genotype	HSI
51	PLU-429	1.16	IPU 99-232	2.05
52	H-1	1.2	U-132	2.29
53	IPU 99-95	1.23	IPU 95-13	2.36
54	Shekhar-2	1.23	IPU 91-7	2.54
55	TU 99-2	1.23	STY-2834	2.85
56	IC-10703	1.28	JU 78-27	2.89
57	UH 80-38	1.29	PLU-65	3.27
58	TU 91-22	1.3	TU 99-2	3.64
59	IPU 99-79	1.41	UH -177	3.8
60	PLU-703	1.42	U 3108	3.97
61	Pant U-19S	1.45	UH 99-144	3.97
62	IPU 99-128	1.46	NHKD-31	4.04
63	PLU-144	1.55	IPU 99-128	4.43
64	IPU 95-13	1.57	STY 2115	4.66
65	UPU 97-10	1.59	IPU 99-16	4.93
66	DUS 34	1.61	UH 84-4	5.53
67	NO 7668-4B	1.63	UH 80-38	5.68
68	IPU 96-1	1.65	PU-19	5.89
69	PKGU-1	1.66	HPU-120	5.9
70	PLU-28	1.66	PLU-8	5.93
71	PLU-662	1.66	N0 7368-15	5.99
72	UG -218	1.72	UH 80-26	6.23
73	PLU -328	1.74	UH 85-15	6.54
74	IPU 99-43	1.79	IPU-722	6.69
75	IPU-722	1.8	IPU 99-40	6.82
76	IPU 99-232	1.85	PKGU-1	6.83
77	Pant U-30	1.92	IPU 99-18	7.02
78	IPU 99-40	1.94	DUS 34	7.07
79	TU 99-293	1.98	IC-10703	7.91
80	IPU 99-179	2.06	U-9	7.94
81	IPU 99-89	2.09	PLU 456	8.01
82	HPU-120	2.17	IPU 96-1	8.39
83	LBG 20	2.19	H-1	8.56
84	IPU 99-221	2.27	IC 106088	8.94
85	IPU 90-321	2.33	IPU 99-95	9.41
86	IPU 99-200	2.39	PLU-662	9.59
87	JU 78-27	2.39	IPU 99-40	10.15
88	UH 84-4	2.4	BGP 21-28	10.59
89	IPU 99-22	2.41	PLU-429	10.62
90	STY-2824	2.45	IPU 99-79	10.81
91	UH 87-7	2.6	IC -21001	11.05
92	UH 85-3	2.7	IPU 99-200	11.22
93	IPU 99-209	2.71	IPU 99-89	11.26
94	UG-378	2.85	BG-369	11.37
95	PLU-65	2.86	NG-2119	12.11
96	PU-19	2.92	IC-65511	13.35
97	UPU 83-3	3.19	STY 2868	13.75

content. Also, correlation of anthocyanin content with chlorophyll content ($r^2 = -0.72$) and NBI ($r^2 = -0.89$) was highly significant ($p < 0.01$) and negative in nature (**Table 8**).

Photosynthetic Electron Transport Rate

Photosynthetic electron transport rate (ETR) was analyzed among 13 heat sensitive and tolerant urdbean genotypes at increasing levels of PAR (photosynthetically active radiation).

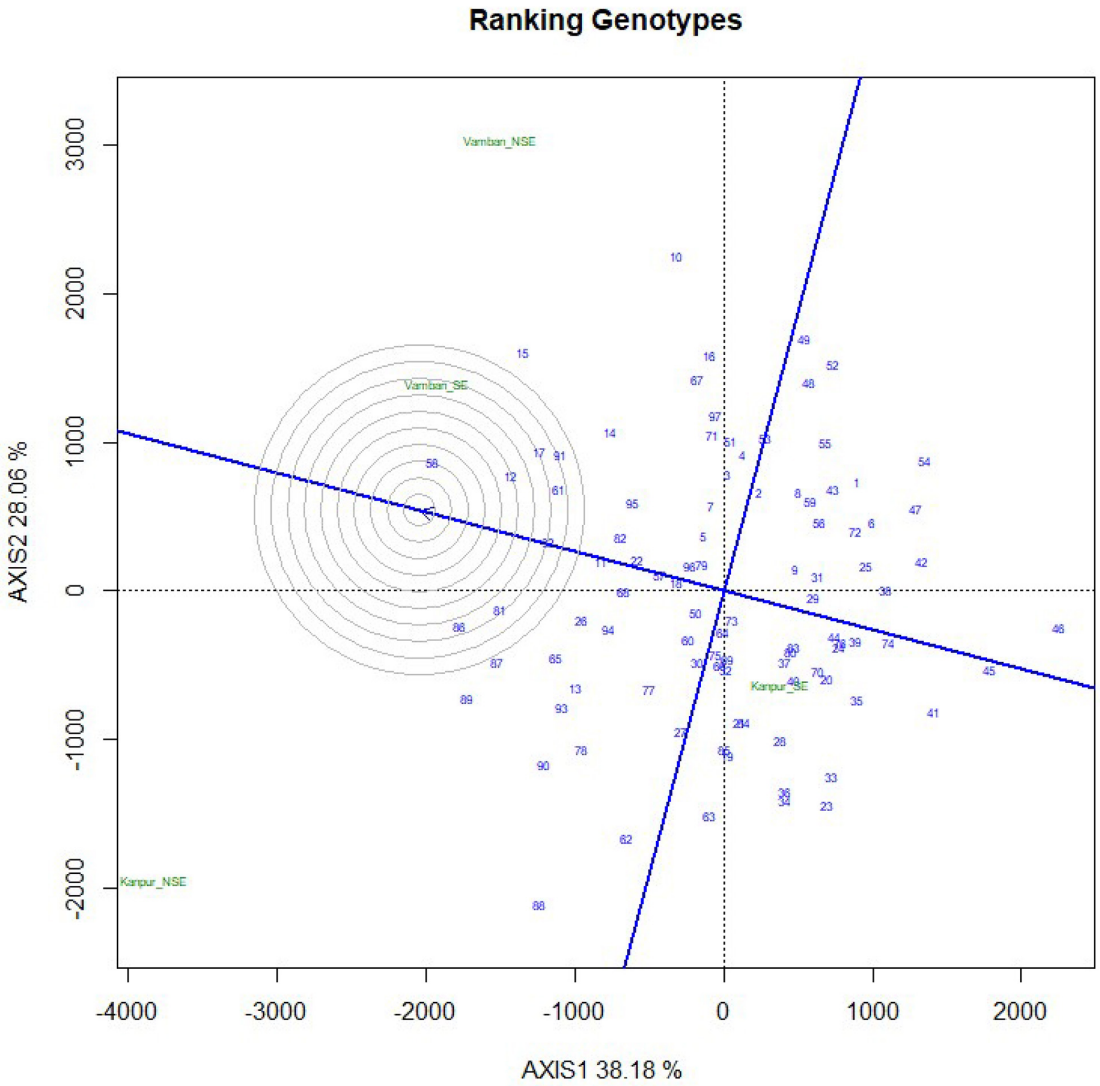


FIGURE 2 | Ranking of 97 urdbean geneotypes by GGEBiplot analysis on the basis of yield data across four growing environments.

The analysis of variance showed significant differences among genotypes for ETR at different levels of PAR irradiances (**Table 9**). These differences were more noticeable with progressive increase in the levels of PAR irradiances among test genotypes (**Figure 3**). The interaction of PAR irradiances with genotypes (PAR × genotypes) was also observed to be significant (**Table 9**). In the present study, higher levels of irradiances were found to be the main determinant of differentiating thermotolerance based on photosynthetic performance in all studied genotypes after heat shock (43°C for 1 h). The light-saturated ETR was obtained almost in all test genotypes within the PAR ranging from 400 to 500 μmol m^{-2} s^{-1} (**Figure 3**). The PAR irradiances exceeding the saturation range of 400-500 μmol m^{-2} s^{-1} pose damaging effects on photosynthetic systems due to excessive production of superoxide radicals. Perhaps tolerant genotypes exposed to combined stress of high PAR irradiances and heat shock that still maintain high ETR have alternate mechanisms scavenging harmful radicals. However, under light limiting conditions below 400-500 μmol m^{-2} s^{-1}, genotype performances were assessed primarily under single stress that was only heat shock. Therefore, the ability of heat tolerance can be detected but cannot be truly assessed under light limiting conditions. Realizing the facts under field conditions, actual heat stress is often combined or integrated with high solar radiation and the crop is forced to experience the combined effects of heat and high light stress, and assimilate production is virtually collapsed resulting in massive yield loss. In the present study, the light-saturated photosynthetic electron transport rate was observed higher than the mean of all test genotypes in most of the tolerant genotypes. Five out of seven heat tolerant genotypes (UPU 85-86, BG 247, PLU 1, PGRU 95016, and IPU 94-2) showed higher photosynthetic ETR than the rest of the tested urdbean genotypes.

TABLE 4 | Heat-tolerant and heat-sensitive urdbean genotypes over both the locations (IIPR, Kanpur and TNAU, Vamban) based on HSI.

Sl. No.	Genotypes[#]	HSI at IIPR, Kanpur	HSI at TNAU, Vamban
	Heat tolerant		
1	**UPU 85-86**	−4.09	−5.51
2	**IPU 94-2**	−0.07	−7.59
3	**IPU 98/36**	−0.41	−19.25
4	**NO- 5731**	−2.91	−10.01
5	**PGRU 95014**	−0.01	−12.21
6	**PGRU 95016**	−1.28	−4.40
7	**PLU-1**	−1.44	−4.48
8	**BGP-247**	−1.11	−13.02
	Heat sensitive		
1	DUS 34	1.61	7.07
2	**H-1**	1.20	8.56
3	**HPU-120**	2.17	5.90
4	**IC -21001**	1.04	11.05
5	IC-10703	1.28	7.91
6	IPU 90-321	2.33	1.95
7	IPU 95-13	1.57	2.36
8	IPU 96-1	1.65	8.39
9	IPU 96-12	0.83	1.89
10	IPU 99-128	1.46	4.43
11	**IPU 99-200**	2.39	11.22
12	IPU 99-232	1.85	2.05
13	IPU 99-40	1.94	6.82
14	IPU 99-79	1.41	10.81
15	IPU 99-89	2.09	11.26
16	IPU 99-95	1.23	9.41
17	IPU-722	1.80	6.69
18	JU 78-27	2.39	2.89
19	NO 7368-15	0.89	5.99
20	NHKD-31	0.98	4.04
21	**PKGU-1**	1.66	6.83
22	PLU -328	1.74	1.20
23	PLU-429	1.16	10.62
24	PLU-65	2.86	3.27
25	PLU-662	1.66	9.59
26	**Shekhar-2**	1.23	1.51
27	STY-2824	2.45	1.25
28	TU 99-2	1.23	3.64
29	U-9	0.90	7.94
30	UH -177	0.72	3.80
31	UH 80-26	0.96	6.23
32	UH 84-4	2.40	5.53
33	UPU 83-3	3.19	1.10
34	**Uttara**	0.99	1.05
35	**PU-19**	2.92	5.89

[#] *Genotypes in bold font were used for physiological, biochemical, and molecular characterization.*

Fluorescence Parameters in Highly Heat Tolerant and Sensitive Genotypes

Further studies remained confined to two extreme genotypes having a high degree of heat tolerance (UPU 85-86) and sensitivity (PKGU-1) based on the field trials and precision phenotyping. The different fluorescence parameters were analyzed to distinguish highly tolerant (UPU 85-86) and highly sensitive (PKGU-1) genotypes. Analysis of variance of fluorescence parameters between heat tolerant (UPU 85-86) and heat sensitive genotypes (PKGU-1) showed significant differences (**Table 10**). The mean value of these parameters is given in **Table 11**. The observed increase in average minimal

TABLE 5 | Analysis of variance of fluorescence parameters of 13 tested urdbean genotypes.

		Means of square				
Source	**Degrees of freedom**	**Leaf nitrogen balance index**	**Chlorophyll**	**Flavanol**	**Anthocyanin**	**Membrane stability**
Genotypes	12	295**	31**	0.003	0.002	452.9**
Error	14	75	5	0.002	0.001	58.10
Total	26	360	36	0.005	0.003	511

** *Significant at $p < 0.01$.*

TABLE 6 | Leaf nitrogen balance index (NBI), chlorophyll (SPAD), epidermal flavanols, and anthocyanins in different urdbean genotypes grown under 42/30°C max/min temperature.

Genotype	**NBI**	**Chlorophyll**	**Flavanol**	**Anthocyanin**
Heat tolerant genotypes				
UPU 85-86	64.6 ± 1.25	21.1 ± 0.05	0.3 ± 0.01	0.00 ± 0.00
IPU 94-2	49.7 ± 6.70	17.2 ± 2.75	0.3 ± 0.01	0.00 ± 0.00
IPU 98/36	36.9 ± 5.60	9.7 ± 1.30	0.3 ± 0.08	0.03 ± 0.02
NO- 5731	49.4 ± 7.05	12.7 ± 0.45	0.2 ± 0.02	0.00 ± 0.00
PGRU 95016	56.8 ± 0.25	16.8 ± 1.15	0.3 ± 0.04	0.00 ± 0.00
PLU 1	56.1 ± 1.90	15.3 ± 0.30	0.3 ± 0.00	0.00 ± 0.00
BGP 247	61.9 ± 1.90	18.0 ± 0.00	0.3 ± 0.01	0.00 ± 0.00
Range	**36.9-64.6**	**9.7-21.1**	**0.2-0.3**	**0.00-0.00**
Heat sensitive genotypes				
IPU 99-200	48.7 ± 2.75	19.4 ± 0.85	0.4 ± 0.01	0.00 ± 0.01
IC 21001	51.7 ± 11.70	15.6 ± 3.35	0.3 ± 0.00	0.02 ± 0.02
Shekhar 2	37.3 ± 15.65	13.4 ± 2.65	0.3 ± 0.03	0.06 ± 0.06
PU 19	30.3 ± 5.30	13.1 ± 2.65	0.3 ± 0.06	0.08 ± 0.01
H-1	43.7 ± 3.60	13.9 ± 1.50	0.3 ± 0.01	0.01 ± 0.01
PKGU 1	23.0 ± 0.70	6.2 ± 0.25	0.3 ± 0.00	0.09 ± 0.00
Range	**23.0-48.7**	**6.2-19.4**	**0.3-0.4**	**0.00-0.09**
C.D. (5%)	18.52	4.82	N/A	N/A
SE(m)	6.13	1.59	0.03	0.02
SE(d)	8.66	2.25	0.04	0.03
C.V.	19.07	15.76	12.72	133.77
t-Value	**2.57**	**–**	**–**	**-3.045**
p-Value ($p < 0.01$)	**0.0129**	**–**	**–**	**0.0062**

fluorescence (F_0) and corresponding decline in the maximal fluorescence (F_m) and variable fluorescence (Fv) in preheated leaves of sensitive genotype PKGU-1 was the strong indicator of distortion of PS II. Consequently, reduction in the quantum yield of PS II was evident in the sensitive one as compared to the tolerant genotype. The decrease in the quantum yield with concomitant rise in the quantum yield of non-regulated energy dissipation [Y(NO) = 0.439] in the sensitive genotype compared to the tolerant genotype [Y(NO) = 0.253] suggested dissipation of absorbed light energy into wasteful thermal or fluorescence quenching, leading to reduction in the photosynthetic efficiency especially targeting the light reaction. While average maximal fluorescence, variable fluorescence, and quantum yield were higher in the tolerant genotype (Fm = 0.277, Fv = 0.215, and F_v/F_m = 0.749, respectively) than the sensitive one (Fm = 0.257, Fv = 0.155, and F_v/F_m = 0.544, respectively). Despite the differences in average values of these two genotypes, analysis of variance showed significant differences only for minimal fluorescence (F_0), quantum yield (F_v/F_m), and quantum yield of non-regulated energy dissipation [Y(NO)] at $p =< 0.01$ (**Table 10**). The significant differences for quantum yield suggest that these two test genotypes responded differently under heat stress conditions as depicted in fluorescence images for heat tolerant (UPU 85-86) and heat sensitive genotypes (PKGU-1) (**Figure 4** and **Table 10**).

Phenotyping Heat Tolerant and Sensitive Genotypes Using Chlorophyll Fluorescence Image-Based Diagnostics

The quantitative values of fluorescence parameters as shown in **Table 11** were transformed into color fluorescence images and

TABLE 7 | Membrane stability of 13 urdbean genotypes.

Genotype	Membrane stability (%)
Heat sensitive genotypes	
IPU 99-200	74.5 ± 4.3
IC 21001	44.3 ± 3.7
H-1	55.3 ± 3.9
PKGU-1	32.3 ± 3.5
Shekhar 2	38.4 ± 3.4
PU 19	42.3 ± 4.5
Heat tolerant genotypes	
UPU 85-86	60.70 ± 5.4
IPU 94-2	34.50 ± 4.6
IPU 98/36	49.60 ± 7.4
NO- 5731	42.90 ± 4.2
PGRU 95016	56.40 ± 3.9
PLU-1	62.80 ± 3.7
BGP-247	55.80 ± 4.8

TABLE 8 | Correlation analysis of nitrogen balance index (NBI), chlorophyll, flavanol, anthocyanin contents, and membrane stability.

	NBI	Chlorophyll	Flavanol	Anthocyanin
NBI				
Chlorophyll	0.85**			
Flavanol	−0.01	0.34		
Anthocyanin	−0.89**	−0.72**	−8.051E-17	
Membrane stability	0.02	0.04	3.655E-02	0.18

***Significant at $p < 0.01$.*

TABLE 9 | Analysis of variance for electron transport rate.

Source	Degrees of freedom	Means of square
Genotypes	12	0.766**
PAR	12	1.746**
Genotypes × PAR	144	0.013**
Error	676	0.003
Total	844	

***Significant at $p < 0.01$.*

the differences in the image pattern between heat tolerant and sensitive genotypes could be easily distinguishable by different shades of color as indicated in the color code bar appended with **Figure 4** having low or high values. For example, a deep blue color of quantum yield as shown in the heat tolerant genotype UPU 85-86 is attributed to high quantum yield of PS II, while similar

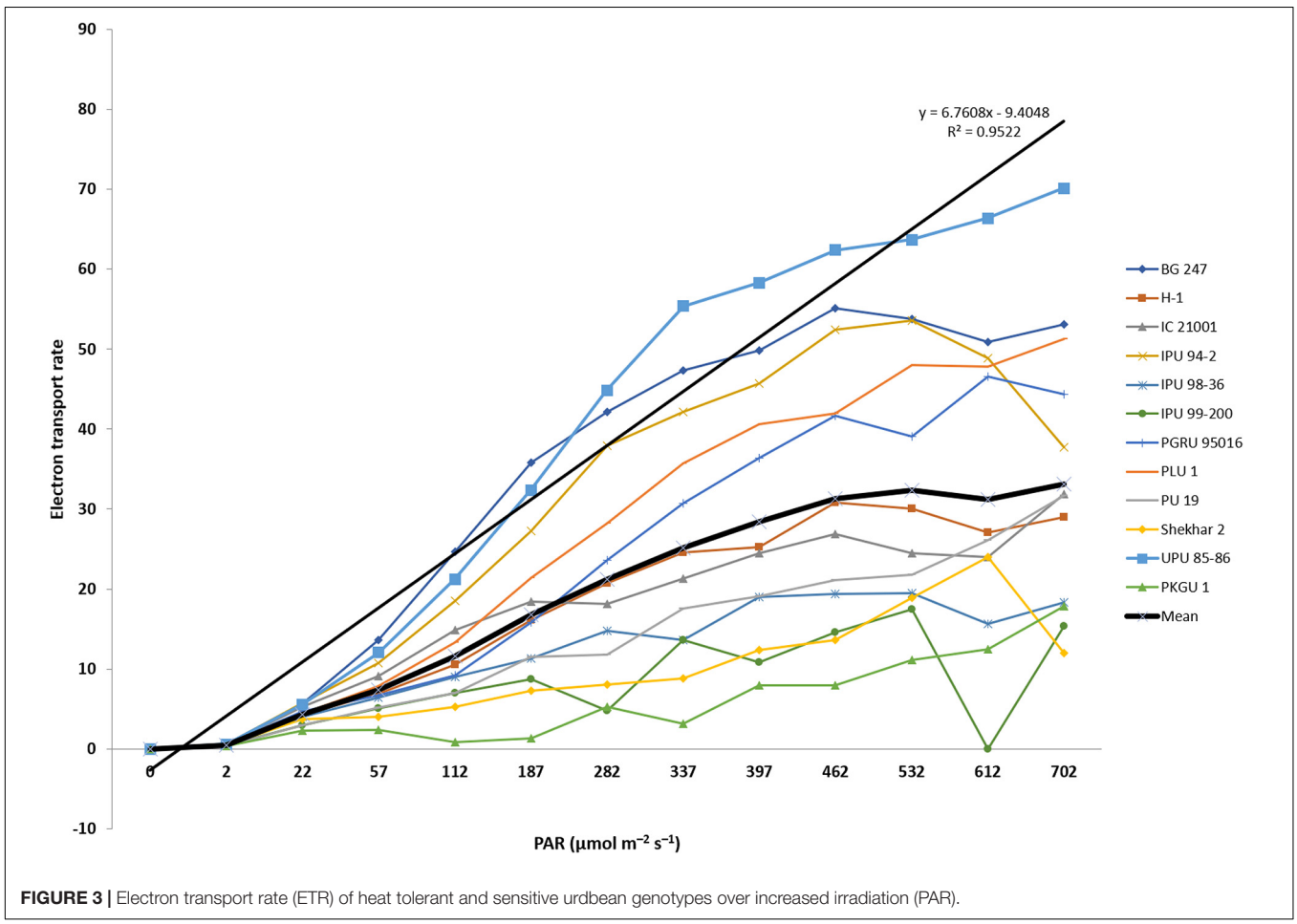


FIGURE 3 | Electron transport rate (ETR) of heat tolerant and sensitive urdbean genotypes over increased irradiation (PAR).

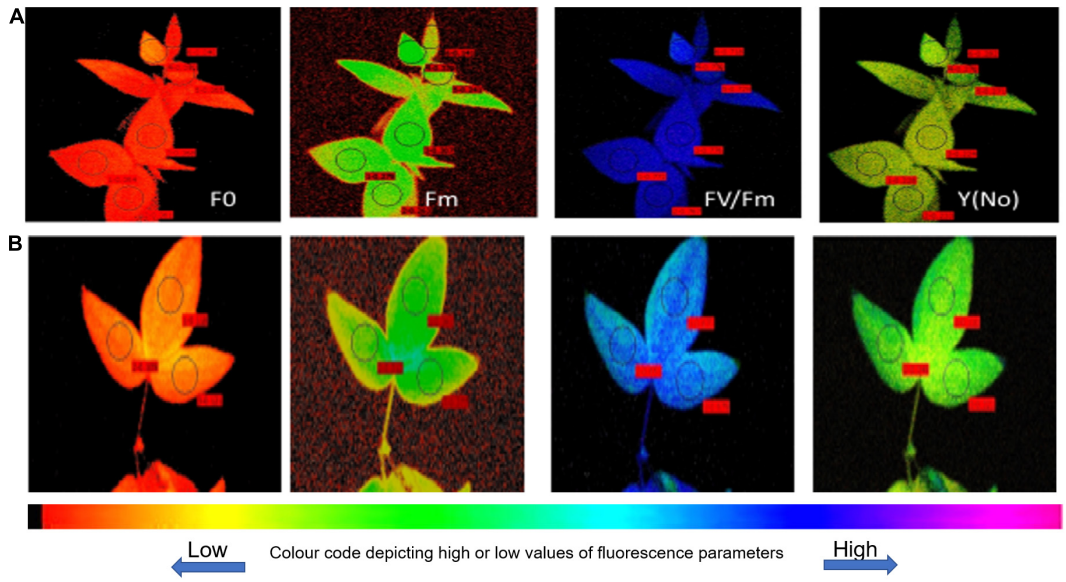


FIGURE 4 | Fluorescence images for heat-tolerant **(A)** (UPU85-86) and heat-sensitive genotypes **(B)** (PKGU-1).

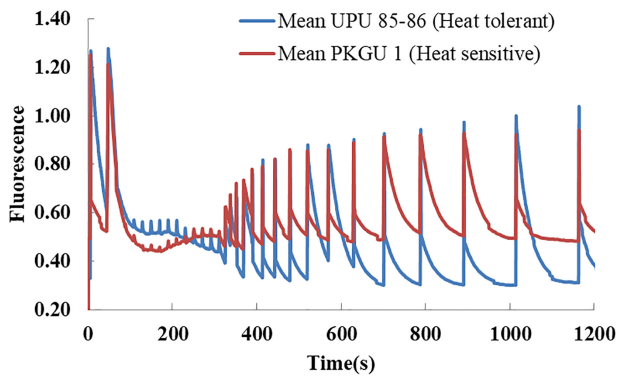


FIGURE 5 | Relative auto-recovery of altered fluorescence during light to dark transition in two contrasting urdbean genotypes (heat-tolerant UPU85-86 and heat-sensitive PKGU-1).

treatment resulted in fading of the blue color to convert to sky blue, representing reduction in the quantum yield. In a similar manner, the heat sensitive genotype PKGU-1 had higher values of minimal fluorescence (F_0) and quantum yield of non-regulated energy dissipation [Y(NO)], which has been well depicted by changes in the color of the corresponding fluorescence images largely differing from the heat tolerant genotype (UPU 85-86). The maximum fluorescence F_m also decreased compared to the heat tolerant ones, which could be easily defined by changes in the color of fluorescence images between these two categories.

Auto-Recovery of Fluorescence Parameters During Light-Dark Transition

The repeated flashes of saturated pulses were triggered at regular intervals upon leaves of heat tolerant (UPU 85-86) and sensitive genotype (PKGU-1) adapted to actinic light (200 μmol m^{-2} s^{-1}) continuously for 300 s to obtain F_v/F_m, F_m, and F_0 after each saturation pulse. Thereafter, actinic light switched off to allow leaves for light to dark transitions to assess the recovery of F_v/F_m, F_m, and F_0 (**Figure 5**). The results showed that quantum yield (F_v/F_m) decreased drastically in the heat sensitive genotype (PKGU-1) when leaves were continuously exposed to actinic light and dark transition. The quantum yield (F_v/F_m) could not recover to the pre-illumination value of 0.50 (**Figure 5**). Notably minimal fluorescence F_0 remained higher and unaltered during the entire dark period suggesting damage or distortion of PS II in sensitive genotypes. Whereas, quantum yield F_v/F_m remained higher in the tolerant one (UPU 85-86) during light phase and completely and reversibly recovered to a pre-illumination value of 0.7 in the dark phase (**Figure 5**).

Fluorescence kinetics of high temperature grown heat sensitive (PKGU 1) and heat tolerant (UPU 85-86) genotypes were studied during light to dark transition (**Figure 6**). Maximum fluorescence (F_m) peak was observed in both contrasting lines immediately after dark adaptation and thereafter the time course trend revealed faster quenching or declining of F_m in both heat sensitive and tolerant genotypes along with shorter peaks of F_m (**Figure 6**) throughout the period until leaves were exposed to light. At the beginning of the dark phase starting after 350 s, the F_m values started rising and the time taken to decrease in the F_m in these two contrasting genotypes could differentiate them on the basis of their differential sensitivity toward heat stress (**Table 12**).

Biochemical Analysis of Heat Sensitive and Tolerant Genotypes

Antioxidative enzymes such as superoxide dismutase (SOD) and peroxidase (POX) play important roles in protecting cellular

TABLE 10 | Analysis of variance of fluorescence parameters between heat-tolerant (UPU85-86) and heat-sensitive genotypes (PKGU-1).

Source	Degrees of freedom	Means of square				
		F_0 (Minimal fluorescence)	F_m (Maximal fluorescence)	F_v (Variable fluorescence)	F_v/F_m (Quantum yield)	Y(NO) (Quantum yield of non-regulated energy dissipation)
Genotypes	1	0.005**	0.001	0.011	0.126**	0.104**
Error	10	0	0.009	0.007	0.006	0.007
Total	11	0.005	0.1	0.018	0.132	0.111

***Significant at $p < 0.01$.*

TABLE 11 | Fluorescence parameters to differentiate two contrasting heat tolerant and sensitive urdbean genotypes grown at 42/25°C max/min temperature.

Treatment	Minimal fluorescence, F_0	Maximal fluorescence, Fm	Variable fluorescence, Fv	Quantum yield, Fv/Fm	Quantum yield of non-regulated energy dissipation, Y(NO)
	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE
UPU 85-86	0.062 ± 0.007	0.277 ± 0.036	0.215 ± 0.031	0.749 ± 0.010	0.253 ± 0.011
PKGU 1	0.102 ± 0.008	0.257 ± 0.043	0.155 ± 0.036	0.544 ± 0.045	0.439 ± 0.049
C.D.	0.024	N/A	N/A	0.105	0.113
SE(m)	0.008	0.039	0.034	0.033	0.035
SE(d)	0.011	0.056	0.048	0.047	0.050
C.V.	22.542	36.094	44.872	12.464	25.015

systems like membranes, proteins, and enzymes by scavenging superoxide radicals and hydrogen peroxides produced during detrimental temperature beyond the threshold level of tolerance which is shown by *in vivo* visualization of superoxide radicals and hydrogen peroxides in **Figure 7**. High antioxidant activity confers tolerance to heat stress, which was represented by less blue color staining zones (formazan deposits) in the leaf (superoxide radicals) or lack of dark brown staining (hydrogen peroxides) as observed in the heat tolerant genotype UPU 85-86 (**Figure 7**). On the contrary, more intense blue crystal patches over leaf surfaces (superoxide radicals) and intense brown coloration (hydrogen peroxides) were the indicators of low antioxidative enzyme activities in heat sensitive genotypes (PKGU-1).

Molecular Characterization

Twenty heat related polymorphic SSR markers were able to group the 16 urdbean genotypes into three major clusters as shown in **Figure 8**. The representative amplification profiles of the 16 urdbean genotypes using SSR markers are illustrated in **Figure 9**. Polymorphic information content ranged from 0.23 to 0.88 with an average value of 0.55 and one to three alleles were amplified by markers (**Table 13**). Cluster I is comprised of a single genotype UPU 85-86. Cluster II consisted of a mixture of six heat tolerant (IPU94-2, NO.5/31, PLU1, IPU98-36, PGRU-95014, PGRU-95016) and six heat sensitive genotypes (HPU120, H1, IC21001, PU19, UTTARA, IPU99-200). Cluster III housed two sensitive (SHEKHAR-2 and PKGU-1) and one heat tolerant genotype (BGP-247). The heat tolerant genotype UPU 85-86 and the heat sensitive genotype PKGU-1 were genetically distinct and were resolved at the extremes of the dendrogram. Thus, UPU 85-86 and PKGU-1 are genetically distinct as well as contrasting for heat tolerance.

DISCUSSION

In the present study, a panel of 97 urdbean genotypes was assessed under heat stress and non-heat stress conditions at two field locations. Stress conditions of a location have been decided based on average yield of trials and high temperature during early sown trials compared to lower temperature during late sown trials in the present study. The significant genotypic differences among tested urdbean genotypes for yield indicated the availability of heat tolerant genotypes. In other *Vigna* species, genetic variability for yield and yield contributing traits have also been observed under heat stress conditions (Basu et al., 2019). Heat susceptibility index (HSI) based on yield potential of a particular genotype under heat stress and non-stress conditions helped to distinguish heat sensitive and tolerant genotypes. This led to the identification of 8 highly heat tolerant and 35 highly heat sensitive genotypes. In the present study, tolerant genotypes had negative HSI due to higher yield under stress conditions compared to non-stress conditions, while highly sensitive genotypes had positive high HSI at both locations. HSI is a widely used method for identification of heat tolerant genotypes and has been used to identify heat tolerant genotypes in other crops (Pandey et al., 2015; Bhandari et al., 2017; Sita et al., 2017). Further,

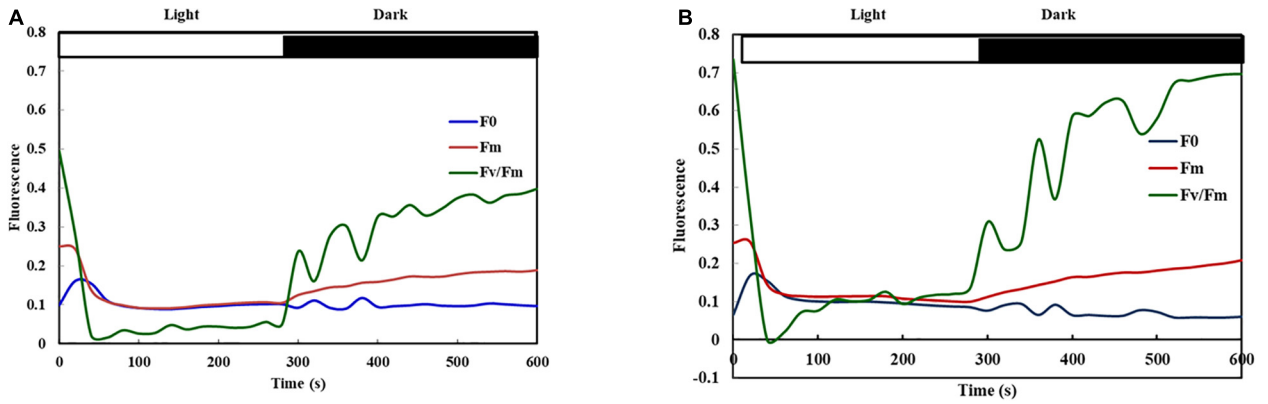


FIGURE 6 | Fluorescence kinetics of heat sensitive (PKGU-1; **A**) and heat tolerant (UPU 85-86; **B**) urdbean genotypes during light to dark transition. The dark phase started from 350 s.

TABLE 12 | Half quenching time of Fm for heat tolerant (UPU85-86) and sensitive (PKGU-1) urdbean genotypes.

Dark phase starting point (s)	Half quenching time of Fm of heat tolerant genotype UPU 85-86 (s)	Half quenching time of Fm of heat sensitive genotype PKGU -1 (s)
400	0.0	0
420	0.0	0
440	0.0	0
520	0.0	0
580	0.0	0
700	0.0	50.0
800	0.0	50.0
900	0.0	50.0
1000	25.0	0.0
1200	0.0	0.0

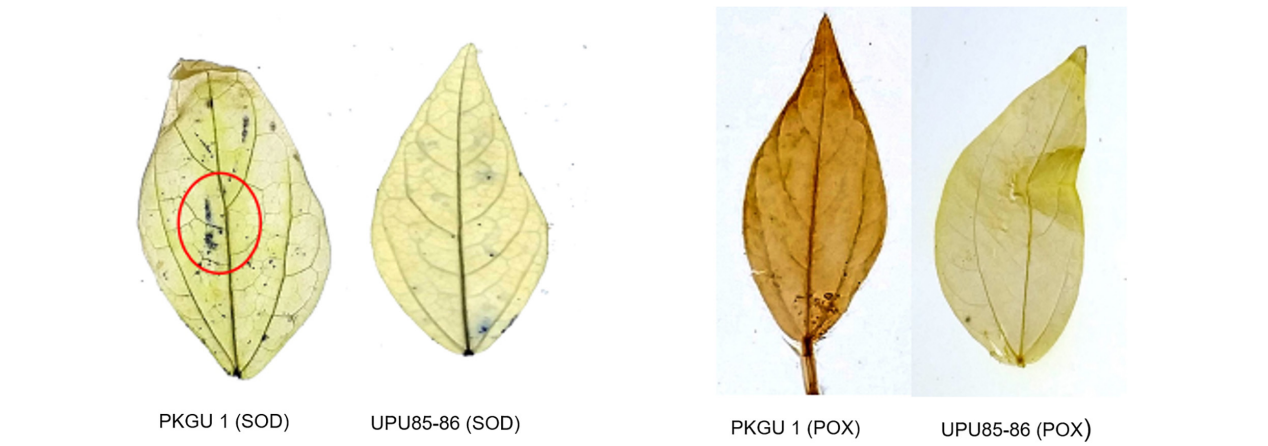


FIGURE 7 | *In vivo* visualization of antioxidant activity for superoxide radicals and hydrogen peroxides in heat sensitive and tolerant genotypes.

GGE biplot analysis identified most stable genotypes over the locations for yield.

The leaf NBI and chlorophyll content based on SPAD value showed significant differences among genotypes and both these parameters were higher in seven heat tolerant genotypes (36.9-64.6 and 9.7-21.1) compared to six heat sensitive genotypes (23.0-48.7 and 6.2-19.4). These results indicated enhanced chlorophyll synthesis and thereby maintaining higher leaf nitrogen balance in the heat tolerant genotype when grown at high temperature. The decrease in chlorophyll content in the heat sensitive genotypes reduced photosynthetic capacity and induced faster senescence due to high temperature as reported earlier in

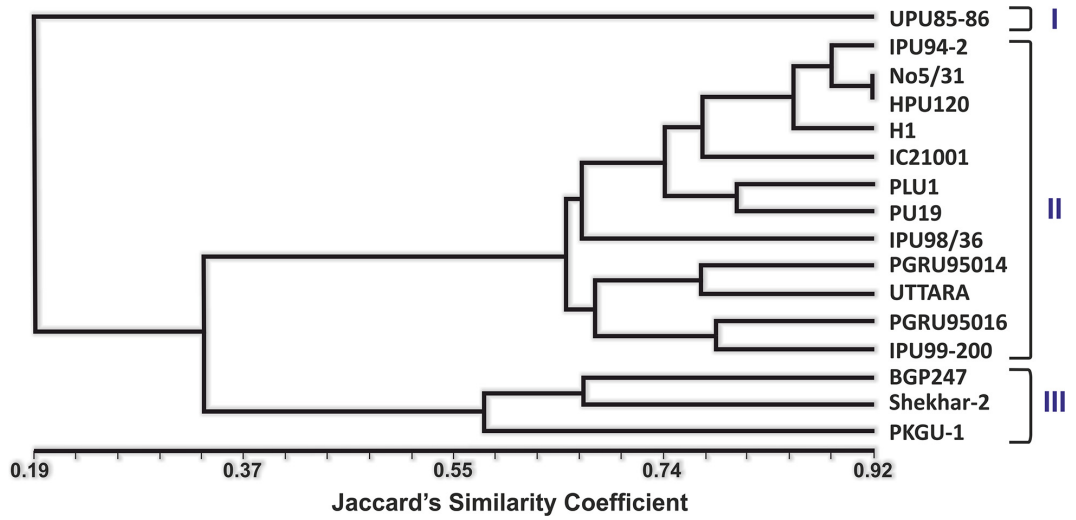


FIGURE 8 | Unweighted Pair Group Method with Arithmetic Mean-based dendrogram showing the clustering of different urdbean genotypes.

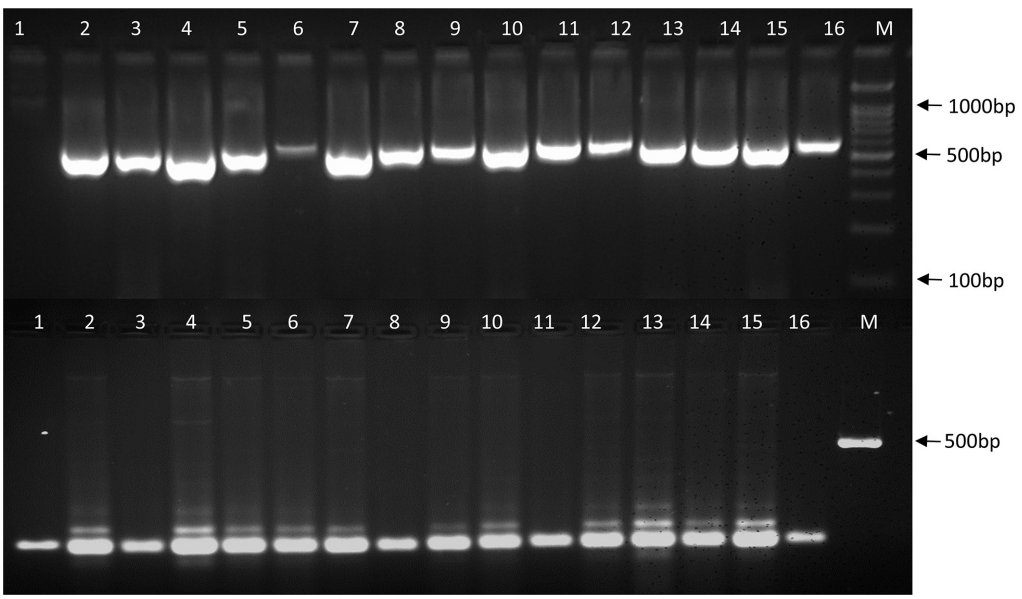


FIGURE 9 | Amplification profiles of 16 urdbean genotypes using SSR primers (DR04 and YMVSSR74). The numbers represent the genotypes 1: UPU-85-86; 2: IPU94-2; 3: IPU-98/36; 4: NO. 5/31; 5: PGRU-95014; 6: PGRU-95016; 7: PLU-1, 8: BGP-247; 9: IPU99-200; 10: IC-21001; 11: SEKHAR-2; 12: UTTARA; 13: PU-19; 14: HPU-120; 15: H-1; 16: PKGU-1; M: DNA marker.

wheat and cucumber (Tewari and Tripathy, 1998). In the present study, no significant differences were observed among the heat sensitive and tolerant genotypes for leaf anthocyanin pigment. However, significantly higher leaf anthocyanin content ($p < 0.01$) was found in heat sensitive genotypes indicating the role of anthocyanin pigment for protecting the survival of sensitive genotypes from high temperature stress (**Table 6**). In other crops, the role of anthocyanin pigment accumulation has also been shown in response to various abiotic stresses (Castellarin et al., 2007) due to its antioxidant properties and photoprotection ability (Abdel-Aal et al., 2008).

Membrane stability index (MSI) under heat stress is one of the important physiological traits for identification of heat tolerant genotypes (Sikder et al., 2001; Dhanda and Munjal, 2006; Ashraf and Foolad, 2007) because high temperature affects several physiological processes such as photosynthesis and respiration through conformational changes in cell membrane bound proteins (Blum et al., 2001). In cowpea and *Brassica,* this trait has been used to identify potential heat tolerant genotypes (Ismail and Hall, 1999; Ram et al., 2012). In the present study, significant genotypic differences have been observed among genotypes for MSI. However, the membrane stability index

TABLE 13 | Details of genic-SSRs used for genotyping 16 urdbean genotypes.

Sl. no.	Primer	Forward sequence Reverse sequence	Tm values	Product size (bp)	No. of alleles amplified	PIC	Function annotation	References
1	TWSSR14	CCGGAAAAGGGAAAACTACATT GCAGAACAGCAGAAACCTCTTT	56.5 58.4	300	1	0	*Vigna angularis var. angularis* DNA, chromosome 1	Raizada, 2020
2	TWSSR15	TCCTGTTCATCCTGATCTTCTTC TAACAAACCCCAAACACACAAC	58.9 56.5	100	1	0.44	Vacuolar sorting receptor	Raizada, 2020
3	TWSSR1	AGAGGGATGGGAGAGGGAT GAAGAAATTGGTGAGACCCAAA	58.8 56.5	180	1	0.61	Protein ABCI7	Raizada, 2020
4	TWSSR34	CGTGCTCGCAACTTCTCTC TCACCACTCTTCTTGTTGTGCT	58.8 58.4	600	1	0.44	60S ribosomal protein L29–1	Raizada, 2020
5	TWSSR4	AACCTTGTCGTGTTCAATCCTT CAAAGATCAGTGTTTCCCACAA	56.5 56.5	220	1	0.75	Transcription factor bHLH143-like	Raizada, 2020
6	TWSSR20	TCGTTAAGAAGGTCAAATGGGT GGCTCGATTGATGAAGAAGGT	56.5 57.9	170	1	0.23	Transcription factor 25	Raizada, 2020
7	TWSSR24	AGTGTTTTGGATTATGGATGGG TCACCAGTTTTATGCACCAGAG	56.5 58.4	200	1	0.61	Glycerol-3-phosphate transporter 4	Raizada, 2020
8	TWSSR72	GGAAAGAGCAGACCTTGACATC CCCAACAAAGCACAGAAACAA	60.3 55.9	280	1	0.23	*Vigna angularis* uncharacterized LOC108326918	Raizada, 2020
9	TWSSR12	GAACTGTATGTAGCAGGGGCTC AGAGGAGACAAAACGCAGAGAT	62.1 58.4	280-300	2	0.61	*Vigna radiata* var. *radiata* uncharacterized LOC106765753	Raizada, 2020
10	YMVSSR74	GAGAGTTTGAGGAGCGGTTG TTGACCTCGTGCAAGCATAG	59.3 57.3	200-220	2	0.53	*Glycine max* heat shock protein (SB100)	Raizada, 2020
11	D102666	TACGAGGCATTTGGTTTGACAGTG AGCCGGTTCCTCCATTTCTT	61 57.3	500-600	2	0.75	*Vigna radiata* sucrose synthase	Venkatesha et al., 2007
12	AF077224	AGCTGAAGCCGCCACCATA AGCAGCAGCCTTAAACTCATCAA	58.8 58.9	700-800	3	0.88	*Glycine max* Fe-super oxide dismutase	Venkatesha et al., 2007
13	AB056453	CCCTCGGCTATAGCATTGAAGAC ACGCATAAACAAAGAGGCTGGACT	62.4 61	600-700	2	0.76	*Vigna unguiculata* S-adenosylmethionine decarboxylase	Venkatesha et al., 2007
14	CA906101	AACACGCGGTACTACGAAATCCTC CTCCGCGTCTCTGTCTCCTACCTC	62.7 67.8	400-600	3	0.86	*Vigna radiata* var. *radiata* dnaJ Protein	Venkatesha et al., 2007
15	CLM446	TCCTCTGTCCTTTCTTTCTCTTT TGGAAGTTAAGACCCACCAG	57.1 57.3	200	1	0.61	*Vigna unguiculata* alpha/beta hydrolase domain-containing protein	Xu et al., 2011
16	X91836-5C	CCGGAAACATGGCATTATTATTAG CCATTGCCTCGTTCCCATCTT	57.6 59.8	500	1	0.44	*Vigna unguiculata* extension 2 like	Gowda, 2008
17	CLM438	TAAAGCCTCCACCCTTCTTT TTCCATGAGTCACCCACTTT	55.3 55.3	200	1	0.34	myb-related transcription factor [*Arabidopsis thaliana*]	Xu et al., 2011
18	CLM443	GGATGCGTCTAAGCCTGTTA CACATGACGAAAGAGATGGA	57.3 55.3	300-400	3	0.79	Putative serine acetyltransferase [*Oryza sativa* (japonica cultivar-group)]	Xu et al., 2011
19	CLM447	GGAAACATGACCTTGACGTT GACAGATGCGTGTGTCCATA	55.3 57.3	250	1	0.33	Putative nuclear ribonuclease Z [*Oryza sativa* (japonica cultivar-group)]	Xu et al., 2011
20	CLM451	ACAATGGACACAACCAACCT CTTGAAGACAGGTTCCTGAAA	55.3 55.9	300	1	0.44	Leaf senescence-associated receptor-like protein kinase (*Phaseolus vulgaris*)	Xu et al., 2011
21	CLM1000	GAGTCTATCGCTTTCTCAGTC CAGTAGGAACCCTCTTGATTT	57.9 55.9	300	1	0.44	Putative uncharacterized protein	Xu et al., 2011

did not correlate strongly with heat tolerance as few sensitive genotypes (i.e., IPU 99-200) also showed higher membrane stability index. In general, a higher membrane stability index has been observed among tolerant genotypes compared to sensitive genotypes under heat stress in the present investigation. In wheat, genetic variability has been observed for this trait, which could be exploited in the development of a heat tolerant wheat variety (Kumar et al., 2013). Since different physiological stages affect this trait, a particular physiological stage that has maximum correlation of MSI with the heat tolerance is needed to be identified for screening the diverse genotypes under heat stress (Hemantaranjan et al., 2014). Therefore, combinations of different physiological traits can be useful for harnessing higher yield under heat stress conditions as observed in an earlier study (Kumar et al., 2018).

Photosynthetic ETR is a potential physiological trait for screening the heat tolerant genotypes. It determines photosynthetic functionality of plants under high temperature and excessive irradiances. Both are often detrimental for plants due to excess generation of toxic superoxide radicals responsible for damaging the functionality of the photosynthetic system (Allakhverdiev and Murata, 2004). In the present study, photosynthetic ETR was significantly different among studied genotypes. Yamada et al. (1996) reported enhanced physiological efficiency of a genotype under high temperature if it had the ability to maintain higher photosynthetic ETR with increasing PAR. In our study, all heat tolerant genotypes generally showed a curvilinear relationship of photosynthetic ETR with increasing PAR but responses of ETR beyond light saturation (ETR_{max}) remained significantly higher in highly tolerant genotype-UPU 85-86 (**Figure 3**) indicating its greater radiation use efficiency even under higher temperature exposure. Whereas the light-saturation point of ETR in highly sensitive one-PKGU-1 was very low and as a result it could not sustain photosynthesis at combined stresses such as high irradiance and high temperature. Thus, sensitive genotypes are more prone to heat stress than tolerant genotypes primarily due to substantial reduction of electron transport and damage of photosystems as reported in earlier studies (Song et al., 2014; Brestic et al., 2016; Chovancek et al., 2019).

The ratio of variable to maximum fluorescence (Fv/Fm), which is known as quantum yield, and the minimal fluorescence (F_0) show their correlation with heat tolerance (Yamada et al., 1996). These parameters are associated with photosystem II (PSII) and carbon fixation. Heat stress affects the photosynthesis process due to inhibition of the activity of PSII (Camejo et al., 2005; Allakhverdiev et al., 2008; Yamamoto, 2016). In winter wheat significant differences among genotypes have been observed for thermostability of PSII and its acclimation effects on PSII photochemical efficiency (Brestic et al., 2012). However, in another study, high temperature stress also affected PSI due to a non-stomatal limitation of photosynthesis by decreasing the activity of rubisco and other parameters of photochemistry (Salvucci and Crafts-Brandner, 2004; Chovancek et al., 2019). In the present study, decreasing the variable fluorescence (F_V = Fm – F_0) leads to decrease in the quantum yield (Fv/Fm) due to inhibition of PSII under heat stress (Kumar et al., 2018). However, quantum yield varied among the studied genotypes and most of the heat tolerant urdbean genotypes had higher quantum yield when grown under high temperature (**Figure 4**). This indicates that heat tolerant genotypes have superior photosynthetic activity under stress than heat sensitive ones. This could be due to increased activity of antioxidative enzymes SOD and peroxidase, the inherent ability to have higher membrane stability, higher chlorophyll retention capacity, or development of certain compounds in heat tolerant genotypes that protect PSII as reported in earlier studies (Murata et al., 2012).

Different fluorescence parameters were recorded to study the effect of temperature on the photosynthetic activities in two contrasting genotypes having different sensitivity to heat stress. Initial fluorescence intensity (F_0) measured in the dark-adapted state, when all PSII reaction centers are open, has been used as a thermo-injury index. The increase in the F_0 in sensitive genotype PKGU-1 as shown in the image (**Figure 4**) was evident from the color code bar toward the higher side as well as its corresponding numerical value (**Table 11**). This sudden change in F_0 is associated with photosynthetic membranes that had suffered irreversible injury. These findings have been supported further from earlier reports by Georgieva and Yordanov (1993). In contrast, heat shocked leaf of tolerant genotype had much lower values of initial fluorescence (F_0). The maximum F_m and variable fluorescence showed no significant difference between sensitive and tolerant genotypes. However, quantum yield (F_v/F_m) image and its numerical values were distinctly different, suggesting that altered quantum yield was largely affected by initial fluorescence (F_0). The higher thermal injury or rise of F_0 was observed in the sensitive genotype as compared to the tolerant ones.

The significant decrease in F_v/F_m at high temperature in sensitive genotype (PKGU-1) indicated that plants were under severe stress and that the photochemical efficiency of PSII was severely impaired. This revealed that high temperature significantly affected the photochemistry of PSII leading to photoinhibition (Baker and Rosenqvist, 2004). Furthermore, the sharp decrease in the F_v/F_m at high temperature was due to the increase in F_0 under the stress condition. Our results are consistent with earlier reports indicating the decline in F_v/F_m that involves an increase in F_0 (Yamada et al., 1996).

The fluorescence parameters such as maximum fluorescence (F_m), quantum yield (F_v/F_m), and minimal fluorescence (F_0) during light to dark transition phases demonstrate the potential ability of a photosynthetic system to recover to normal values that were observed before the illumination by actinic light. In the present study, the decrease in the effective quantum yield F_v/F_m was more pronounced in heat-shocked leaves of the sensitive genotype exposed to light condition compared to the tolerant ones which could likely to be due to photoinhibition of PSII associated with increase in F_0. The photoinhibition is often reversible during light to dark transition but it depends on sensitivity of genotype to heat

stress and hence recovery might be delayed. The heat tolerant genotype showed complete recovery in F_v/F_m in the dark after 500 s, suggesting that reversible changes of photosystems occurred during continuous illumination up to 250 s (**Figure 5**). However, in the case of sensitive ones, it appeared to undergo an irreversible change for a longer period and could not recover in the dark phase even after 500 s (**Figure 5**). The delayed recovery of F_v/F_m could likely be associated with major conformational changes in photosystems to operate in a normal manner.

Maximum decrease in quantum yield indicates damage to the photosynthetic apparatus of the plants (Van der Westhuizen et al., 2020). Many studies have reported variation in the tolerance to high-temperature stress among genotypes of wheat, chickpea, lentil, and mungbean on the basis of pollen sterility, seed abortion, maintenance of photosynthesis, chlorophyll content, and an extended grain-filling duration at elevated temperatures (Viswanathan and Khanna-Chopra, 2001; Tahir and Nakata, 2005; Hays et al., 2007; Krishnamurthy et al., 2011; Kumar et al., 2018; Basu et al., 2019). Heat stress sensitivity of photosynthesis (Singh and Thakur, 2018) due to the inactivation of photosystem II (PSII) (Rustioni et al., 2015) leads to the decrease in variable chlorophyll fluorescence (F_v). This is the most thermolabile component of the photosynthetic electron transport chain (Camejo et al., 2005). Therefore, the detection and quantification of temperature-induced changes in the photosynthetic apparatus is an important tool to distinguish genotypes for their heat stress tolerance (Baker and Rosenqvist, 2004). In the present investigation, significant increase in quantum yield of non-regulated energy dissipation was also observed in the highly heat sensitive genotype PKGU-1 (**Table 11**). Moreover, more time was required toward quenching of maximum fluorescence F_m (i.e., delay in quenching of F_m) and high values of quenching of F_0 indicated severe photo-inactivation of PS II in the sensitive genotype (PKGU-1). This was in complete agreement to the fact that greater thermo-tolerance is associated with faster recovery of photo-damage to PSII. Therefore, rapid overnight recovery of photo-inhibition was observed in tolerant genotype UPU 85-86 (**Figure 6** and **Table 12**).

The qualitative analysis was done to demonstrate *in vivo* visualization of oxidants such as superoxide radicals and hydrogen peroxides, which clearly elucidated the differences in enzymatic activities of superoxide dismutase (SOD) and peroxidase (POX) in heat shocked leaves of extreme heat tolerant and sensitive genotypes (**Figure 7**). The presence of blue crystalline formazan deposits and dark brown precipitates indicated low activities of SOD and POX enzymes in the sensitive genotype. As a result, the sensitive genotype failed to scavenge the harmful radicals, which caused damage to membranes due to heat stress.

Further, in the case of SSR marker data based dendrogram, a highly heat tolerant genotype (UPU 85-86) was distinctly clustered from the highly heat sensitive genotype (PKGU-1). Heat tolerance being a trait governed by several genes, it becomes very difficult to categorize them solely based on SSR markers unless the markers are highly linked to the heat tolerance trait. Since the primers were designed based on their relevance to abiotic stress tolerance like drought, salinity, and so on, in addition to heat tolerance, the clustering based on their amplification profiles holds importance. The dendrogram obtained by Sun et al. (2015) in tall fescue also showed the heat tolerant genotypes to be strewn across the dendrogram. In addition, a low correlation was found between morpho-physiological heat tolerance traits and SSR markers by the Mantel test (data not shown). The identified genetically diverse and high temperature tolerant lines would be useful in designing breeding programs for developing heat stress tolerance in urdbean.

CONCLUSION

Based on field evaluation of 97 urdbean genotypes over two locations under two different growing conditions, a panel of heat tolerant and sensitive genotypes was identified which were stable in yield. Genotypic differences existed for physiological traits like leaf NBI, chlorophyll (SPAD), epidermal flavanols and anthocyanin contents among the tested heat tolerant and sensitive genotypes. The genotypic variation in the membrane stability was evident, which defined the variation in the heat tolerance but to a lesser extent. The high antioxidant activities were shown by heat tolerant genotype (UPU 85-86) explaining their role for scavenging superoxide radicals (ROS) protecting delicate membranes from oxidative damage. Perhaps the higher photosynthetic activities including ETR, quantum yield, and lesser photoinhibition as observed in the heat tolerant genotype UPU 85-86 are associated with inherent stable membranes and higher expression of antioxidative enzymes during exposure to high temperature enabling the plant to maintain optimum functionality. Molecular characterization further pinpointed genetic differences between heat tolerant (UPU 85-86) and heat sensitive genotypes (PKGU-1).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

DS conceived the idea, conducted the experiments, analyzed data, and wrote the draft manuscript. PB contributed to the writing, reviewing, editing, physiological data, and analysis of experimental data. JS and PD carried out the molecular work. JK conceived the idea and wrote the drafted manuscript. SaG,

MP, SG, and NP reviewed and edited the manuscript. VK contributed to the antioxidant profile of two urdbean genotypes. PS conducted the field trials. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported partially by the research grant from the ICAR, New Delhi in the form of a project "National Innovations in Climate Resilient Agriculture (NICRA)." ICAR-Indian Institute of Pulses Research also funded the experiments. No separate funding was received for publication in peer-reviewed journals.

ACKNOWLEDGMENTS

We thank Director, ICAR-Indian Institute of Pulses Research, Head, Basic Science Division and Head, Crop Improvement Division of this institute for extending the laboratory and field facilities to conduct research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.719381/full#supplementary-material

REFERENCES

Abdel-Aal, E. S. M., Abou-Arab, A. A., Gamel, T. H., Hucl, P., Young, J. C., and Rabalski, I. (2008). Fractionation of blue wheat anthocyanin compounds and their contribution to antioxidant properties. *J. Agric. Food Chem.* 56, 11171–11177. doi: 10.1021/jf802168c

Allakhverdiev, S. I., and Murata, N. (2004). Environmental stress inhibits the synthesis de novo of proteins involved in the photodamage–repair cycle of Photosystem II in Synechocystis sp. PCC 6803. *Biochim. Biophys. Acta BBA.Bioenerget.* 1657, 23–32. doi: 10.1016/j.bbabio.2004.03.003

Allakhverdiev, S. I., Kreslavski, V. D., Klimov, V. V., Los, D. A., Carpentier, R., and Mohanty, P. (2008). Heat stress: an overview of molecular responses in photosynthesis. *Photosyn. Res.* 98, 541–550. doi: 10.1007/s11120-008-9331-0

Aravind, J., Mukesh Sankar, S., Wankhede, D. P., and Kaur, V. (2020). *augmentedRCBD: Analysis of augmented randomised complete block designs*. R package version 0.1.*2*. ICAR-NBPGR, New Delhi

Argyris, J., Dahal, P., Hayashi, E., Still, D. W., and Bradford, K. J. (2008). Genetic variation for lettuce seed thermoinhibition is associated with temperature-sensitive expression of abscisic acid, gibberellin, and ethylene biosynthesis, metabolism, and response genes. *Plant Physiol.* 148, 926–947. doi: 10.1104/pp.108.125807

Ashraf, M. F. M. R., and Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ. Exp. Bot.* 59, 206–216. doi: 10.1016/j.envexpbot.2005.12.006

Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., et al. (2015). Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* 5, 143–147. doi: 10.1038/nclimate2470

Baker, N. R., and Rosenqvist, E. (2004). Applications of chlorophyll fluorescence can improve crop production strategies: an examination of future possibilities. *J. Exp. Bot.* 55, 1607–1621. doi: 10.1093/jxb/erh196

Basu, P. S., Pratap, A., Gupta, S., Sharma, K., Tomar, R., and Singh, N. P. (2019). Physiological traits for shortening crop duration and improving productivity of greengram (*Vigna radiata* L. Wilczek) under high temperature. *Front. Plant Sci.* 10:1508. doi: 10.3389/fpls.2019.01508

Bhandari, K., Sharma, K. D., Rao, B. H., Siddique, K. H., Gaur, P., Agrawal, S. K., et al. (2017). Temperature sensitivity of food legumes: a physiological insight. *Acta Physiol. Plant.* 39:68. doi: 10.1007/s11738-017-2361-5

Blum, A., and Ebercon, A. (1981). Cell membrane stability as a measure of drought and heat tolerance in wheat. *Crop Sci.* 21, 43–47. doi: 10.2135/cropsci1981.0011183X002100010013x

Blum, A., Klueva, N., and Nguyen, H. T. (2001). Wheat cellular thermotolerance is related to yield under heat stress. *Euphytica* 117, 117–123. doi: 10.1023/A:1004083305905

Brestic, M., Zivcak, M., Kalaji, H. M., Carpentier, R., and Allakhverdiev, S. I. (2012). Photosystem II thermostability in situ: environmentally induced acclimation and genotype-specific reactions in *Triticum aestivum* L. *Plant Physiol. Biochem.* 57, 93–105. doi: 10.1016/j.plaphy.2012.05.012

Brestic, M., Zivcak, M., Kunderlikova, K., and Allakhverdiev, S. I. (2016). High temperature specifically affects the photoprotective responses of chlorophyll b-deficient wheat mutant lines. *Photosyn. Res.* 130, 251–266. doi: 10.1007/s11120-016-0249-7

Camejo, D., Rodríguez, P., Morales, M. A., Dell'Amico, J. M., Torrecillas, A., and Alarcón, J. J. (2005). High temperature effects on photosynthetic activity of two tomato cultivars with different heat susceptibility. *J. Plant Physiol.* 162, 281–289. doi: 10.1016/j.jplph.2004.07.014

Castellarin, S. D., Pfeiffer, A., Sivilotti, P., Degan, M., Peterlunger, E., and Di Gaspero, G. (2007). Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant Cell Environ.* 30, 1381–1399. doi: 10.1111/j.1365-3040.2007.01716.x

Challinor, A., Wheeler, T., Garforth, C., Craufurd, P., and Kassam, A. (2007). Assessing the vulnerability of food crop systems in Africa to climate change. *Clim. Change* 83, 381–399. doi: 10.1007/s10584-007-9249-0

Chitra, U., Singh, U., and Rao, P. V. (1996). Phytic acid, in vitro protein digestibility, dietary fiber, and minerals of pulses as influenced by processing methods. *Plant Foods Hum. Nutr.* 49, 307–316. doi: 10.1007/BF01091980

Chovancek, E., Zivcak, M., Botyanszka, L., Hauptvogel, P., Yang, X., Misheva, S., et al. (2019). Transient heat waves may affect the photosynthetic capacity of susceptible wheat genotypes due to insufficient photosystem I photoprotection. *Plants* 8:282. doi: 10.3390/plants8080282

Christensen, T. H., Zhang, Z., Wei, Y., and Collinge, D. B. (1997). Subcellular localization of H2O2 in plants. H2O2 accumulation in papillae and hypersensitive response during the barley—powdery mildew interaction. *Plant J.* 11, 1187–1194.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* 1, 19–21. doi: 10.1007/BF02712670

Dhanda, S. S., and Munjal, R. (2006). Inheritance of cellular thermotolerance in bread wheat. *Plant Breed.* 125, 557–564. doi: 10.1111/j.1439-0523.2006.01275.x

Ehlers, J. D., and Hall, A. E. (1998). Heat tolerance of contrasting cowpea lines in short and long days. *Field Crops Res.* 55, 11–21. doi: 10.1016/S0378-4290(97)00055-5

Fischer, R. A., and Maurer, R. (1978). Drought resistance in spring wheat cultivars. 1. Grain-yield responses. *Aust. J. Agric. Res.* 29, 897–912. doi: 10.1071/AR9780897

Frahry, G., and Schopfer, P. (2001). NADH-stimulated, cyanide-resistant superoxide production in maize coleoptiles analyzed with a tetrazolium-based assay. *Planta* 212, 175–183. doi: 10.1007/s004250000376

Frutos, E., Galindo, M. P., and Leiva, V. (2014). An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stoch. Environ. Res. Risk Assess.* 28, 1629–1641.

Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58, 453–467. doi: 10.1093/biomet/58.3.453

Gaur, P. M., Samineni, S., Thudi, M., Tripathi, S., Sajja, S. B., Jayalakshmi, V., et al. (2019). Integrated breeding approaches for improving drought and heat adaptation in chickpea (*Cicer arietinum* L.). *Plant Breed.* 138, 389–400. doi: 10.1111/pbr.12641

Georgieva, K., and Yordanov, I. (1993). Temperature dependence of chlorophyll fluorescence parameters of pea seedlings. *J. Plant Physiol.* 142, 151–155. doi: 10.1016/S0176-1617(11)80955-7

Gowda, M. B. (2008). *Genetic enhancement of Dolichos bean through integration of conventional breeding and molecular approaches*. Final Project Report. Witney: The Kirkhouse Trust.

Hays, D. B., Do, J. H., Mason, R. E., Morgan, G., and Finlayson, S. A. (2007). Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased maturation in a susceptible cultivar. *Plant Sci.* 172, 1113–1123. doi: 10.1016/j.plantsci.2007.03.004

Hemantaranjan, A., Bhanu, A. N., Singh, M. N., Yadav, D. K., Patel, P. K., Singh, R., et al. (2014). Heat stress responses and thermotolerance. *Adv. Plants Agric. Res.* 1, 1–10. doi: 10.15406/apar.2014.01.00012

Ismail, A. M., and Hall, A. E. (1999). Reproductive-stage heat tolerance, leaf membrane thermostability and plant morphology in cowpea. *Crop Sci.* 39, 1762–1768. doi: 10.2135/cropsci1999.3961762x

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bull.Soc. Vaudoise des Sci. Nat.* 44, 223–270.

Krishnamurthy, L., Gaur, P. M., Basu, P. S., Chaturvedi, S. K., Tripathi, S., Vadez, V., et al. (2011). Large genetic variation for heat tolerance in the reference collection of chickpea (*Cicer arietinum* L.) germplasm. *Plant Genet. Resour.* 9, 59–69. doi: 10.1017/S1479262110000407

Kumar, J., Basu, P. S., Gupta, S., Dubey, S., Gupta, D. S., and Singh, N. P. (2018). Physiological and molecular characterisation for high temperature stress in *Lens culinaris*. *Funct. Plant Biol.* 45, 474–487. doi: 10.1071/FP17211

Kumar, S., Kumari, P., Kumar, U., Grover, M., Singh, A. K., Singh, R., et al. (2013). Molecular approaches for designing heat tolerant wheat. *J. Plant Biochem. Biotechnol.* 22, 359–371. doi: 10.1007/s13562-013-0229-3

Murata, N., Allakhverdiev, S. I., and Nishiyama, Y. (2012). The mechanism of photoinhibition in vivo: re-evaluation of the roles of catalase, α-tocopherol, non-photochemical quenching, and electron transport. *Biochim. Biophys. Acta BBA Bioenerget.* 1817, 1127–1133. doi: 10.1016/j.bbabio.2012.02.020

Nayyar, H., Gaur, P. M., Kumar, S., Singh, S., Bindumadhava, H., Nair, R. M., et al. (2017). "How rising temperatures would be detrimental for cool and warm-season food legumes," in *Proceedings of the InterDrought-V, February 21-25, 2017*, (Hyderabad: HICC).

Paliwal, R., Röder, M. S., Kumar, U., Srivastava, J. P., and Joshi, A. K. (2012). QTL mapping of terminal heat tolerance in hexaploid wheat (*T. aestivum* L.). *Theor. Appl. Genet.* 125, 561–575. doi: 10.1007/s00122-012-1853-3

Pandey, G. C., Mamrutha, H. M., Tiwari, R., Sareen, S., Bhatia, S., Siwach, P., et al. (2015). Physiological traits associated with heat tolerance in bread wheat (*Triticum aestivum* L.). *Physiol. Mol. Biol. Plants* 21, 93–99. doi: 10.1007/s12298-014-0267-x

Pootakham, W., Nawae, W., Naktang, C., Sonthirod, C., Yoocha, T., Kongkachana, W., et al. (2021). A chromosome-scale assembly of the black gram (*Vigna mungo*) genome. *Mol. Ecol. Resour.* 21, 238–250. doi: 10.1111/1755-0998.13243

Project Coordinator's Report (2019). *All India Coordinated Research Project on Mungbean, Urdbean, Lentil, Lathyrus, Field Pea*. Kanpur: Indian Institute of Pulses Research.

R Core Team (2019). *R: a Language and Environment for Statistical computing*. Vienna: R Foundation for Statistical Computing.

Raizada, A. (2020). *Development of genic-markers for yellow mosaic virus and bruchid resistance traits in blackgram (Vigna mungo (L.) Hepper)*. Ph.D. Thesis. Mumbai: Homi Bhabha National Institute, 214–265.

Raizada, A., and Souframanien, J. (2019). Transcriptome sequencing, de novo assembly, characterisation of wild accession of blackgram (*Vigna mungo* var. silvestris) as a rich resource for development of molecular markers and validation of SNPs by high resolution melting (HRM) analysis. *BMC Plant Biol.* 19:358. doi: 10.1186/s12870-019-1954-0

Ram, B., Singh, B. K., Singh, M., Singh, V. V., and Chauhan, J. S. (2012). *Physiological And Molecular Characterization Of Indian Mustard (b. Juncea l.) Genotypes for High Temperature Tolerance*. London: Crop Improvement (ICSA), 5–6.

Rohlf, F. J. (1998). *NTSYS-Pc: Numerical taxonomy and Multivariate Analysis System, Version 2.0*. Setauket: Exeter publications.

Roy, S. J., Tucker, E. J., and Tester, M. (2011). Genetic analysis of abiotic stress tolerance in crops. *Curr. Opin. Plant Biol.* 14, 232–239. doi: 10.1016/j.pbi.2011.03.002

Rustioni, L., Milani, C., Parisi, S., and Failla, O. (2015). Chlorophyll role in berry sunburn symptoms studied in different grape (*Vitis vinifera* L.) cultivars. *Sci. Hortic.* 185, 145–150. doi: 10.1016/j.scienta.2015.01.029

Salvucci, M. E., and Crafts-Brandner, S. J. (2004). Relationship between the heat tolerance of photosynthesis and the thermal stability of *Rubisco activase* in plants from contrasting thermal environments. *Plant Physiol.* 134, 1460–1470. doi: 10.1104/pp.103.038323

Scafaro, A. P., Haynes, P. A., and Atwell, B. J. (2010). Physiological and molecular changes in *Oryza meridionalis* Ng., a heat-tolerant species of wild rice. *J. Exp. Bot.* 61, 191–202. doi: 10.1093/jxb/erp294

Schreiber, U., and Bilger, W. (1987). "Rapid assessment of stress effects on plant leaves by chlorophyll fluorescence measurements," in *Plant Response To Stress*, eds J. D. Tenhunen, F. M. Catarino, O. L. Lange, and W. C. Oechel (Berlin: Springer), 27–53. doi: 10.1007/978-3-642-70868-8_2

Sen Gupta, D., Singh, U., Kumar, J., Shivay, Y. S., Dutta, A., Sharanagat, V. S., et al. (2020). Estimation and multi-variate analysis of iron and zinc concentration in a diverse panel of urdbean (Vigna mungo L. Hepper) genotypes grown under differing soil conditions. *J. Food Comp. Anal.* 93:103605. doi: 10.1016/j.jfca.2020.103605

Sikder, S., Ahmed, J. U., and Hossain, T. (2001). Heat tolerance and relative yield performance of wheat varieties under late seeded conditions. *Indian J. Agric. Res.* 35, 141–148.

Singh, J., and Thakur, J. K. (2018). "Photosynthesis and abiotic stress in plants," in *Biotic And Abiotic Stress Tolerance in Plants*, ed. S. Vats (Singapore: Springer), 27–46. doi: 10.1007/978-981-10-9029-5_2

Sita, K., Sehgal, A., HanumanthaRao, B., Nair, R. M., Vara Prasad, P. V., Kumar, S., et al. (2017). Food legumes and rising temperatures: effects, adaptive functional mechanisms specific to reproductive growth stage and strategies to improve heat tolerance. *Front. Plant Sci.* 8:1658. doi: 10.3389/fpls.2017.01658

Song, Y., Chen, Q., Ci, D., Shao, X., and Zhang, D. (2014). Effects of high temperature on photosynthesis and related gene expression in poplar. *BMC Plant Biol.* 14:111. doi: 10.1186/1471-2229-14-111

Souframanien, J., Raizada, A., Dhanasekar, P., and Suprasanna, P. (2020). Draft genome sequence of the pulse crop blackgram [*Vigna mungo* (L.) Hepper] reveals potential R-genes. *Sci. Rep.* 11:11247. doi: 10.1038/s41598-021-90683-9

Sun, X., Xie, Y., Bi, Y., Liu, J., Amombo, E., Hu, T., et al. (2015). Comparative study of diversity based on heat tolerant-related morpho-physiological traits and molecular markers in tall fescue accessions. *Sci. Rep.* 5, 1–14. doi: 10.1038/srep18213

Tahir, I. S. A., and Nakata, N. (2005). Remobilization of nitrogen and carbohydrate from stems of bread wheat in response to heat stress during grain filling. *J. Agron. Crop Sci.* 191, 106–115. doi: 10.1111/j.1439-037X.2004.00127.x

Tewari, A. K., and Tripathy, B. C. (1998). Temperature-stress-induced impairment of chlorophyll biosynthetic reactions in cucumber and wheat. *Plant Physiol.* 117, 851–858. doi: 10.1104/pp.117.3.851

Van der Westhuizen, M. M., Oosterhuis, D. M., Berner, J. M., and Boogaers, N. (2020). Chlorophyll a fluorescence as an indicator of heat stress in cotton (*Gossypium hirsutum* L.). *South Afr. J. Plant Soil* 37, 116–119. doi: 10.1080/02571862.2019.1665721

Venkatesha, S. C., Gowda, M. B., Mahadevu, P., Rao, A. M., Kim, D. J., Ellis, T. H. N., et al. (2007). Genetic diversity within Lablab purpureus and the application of gene-specific markers from a range of legume species. *Plant Genet. Resour.* 5, 154–171. doi: 10.1017/S1479262107835659

Viswanathan, C., and Khanna-Chopra, R. (2001). Effect of heat stress on grain growth, starch synthesis and protein synthesis in grains of wheat (*Triticum aestivum* L.) varieties differing in grain weight stability. *J. Agron. Crop Sci.* 186, 1–7. doi: 10.1046/j.1439-037x.2001.00432.x

Xu, P., Wu, X., Wang, B., Liu, Y., Ehlers, J. D., Close, T. J., et al. (2011). A SNP and SSR based genetic map of asparagus bean (*Vigna. unguiculata* ssp. sesquipedialis) and comparison with the broader species. *PLoS One* 6:e15952. doi: 10.1371/journal.pone.0015952

Yamada, M., Hidaka, T., and Fukamachi, H. (1996). Heat tolerance in leaves of tropical fruit crops as measured by chlorophyll fluorescence. *Sci. Hortic.* 67, 39–48. doi: 10.1016/S0304-4238(96)00931-4

Yamamoto, Y. (2016). Quality control of photosystem II: the mechanisms for avoidance and tolerance of light and heat stresses are closely linked to membrane fluidity of the thylakoids. *Front. Plant Sci.* 7:1136. doi: 10.3389/fpls.2016.01136

Yan, W. (2001). GGEbiplot—A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agron. J.* 93, 1111–1118. doi: 10.2134/agronj2001.9351111x

Yan, W., Hunt, L. A., Sheng, Q., and Szlavnics, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605. doi: 10.2135/cropsci2000.403597x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Sen Gupta, Basu, Souframanien, Kumar, Dhanasekar, Gupta, Pandiyan, Geetha, Shanthi, Kumar and Pratap Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 October 2021
doi: 10.3389/fpls.2021.754287

Lentil and Fava Bean With Contrasting Germination Kinetics: A Focus on Digestion of Proteins and Bioactivity of Resistant Peptides

*Sara Bautista-Expósito[1], Albert Vandenberg[2], Elena Peñas[1], Juana Frias[1] and Cristina Martínez-Villaluenga[1]**

[1] Department of Food Characterization, Quality and Safety, Institute of Food Science, Technology and Nutrition (ICTAN-CSIC), Madrid, Spain, [2] Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Saskatoon, SK, Canada

OPEN ACCESS

Edited by:
Jaroslaw Czubinski, Poznan University of Life Sciences, Poland

Reviewed by:
Petr Smýkal, Palacký University, Czechia
Ryszard Amarowicz, Institute of Animal Reproduction and Food Research (PAS), Poland

***Correspondence:**
Cristina Martínez-Villaluenga
c.m.villaluenga@csic.es

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: 06 August 2021
Accepted: 20 September 2021
Published: 25 October 2021

Citation:
Bautista-Expósito S, Vandenberg A, Peñas E, Frias J and Martínez-Villaluenga C (2021) Lentil and Fava Bean With Contrasting Germination Kinetics: A Focus on Digestion of Proteins and Bioactivity of Resistant Peptides. Front. Plant Sci. 12:754287. doi: 10.3389/fpls.2021.754287

Germination offers advantages to improve legume protein digestibility as it disintegrates seed structure and hydrolyzes proteins and anti-nutrients. Seed permeability (related to polyphenol content of seed coats) is an important factor affecting the duration of seed germination and its impact on protein digestibility and bioactivity. The objective was to compare the effect of seed germination on protease activity, structure, and proteolysis of four selected legumes with contrasting seed coat polyphenol profiles (gray zero-tannin lentil [GZL], beluga lentil [BL], and dehulled red lentil [DL]; and zero tannin/low vicine–convicine fava bean [ZF]). Protein hydrolysis was characterized during germination and digestion with respect to proteins, peptides, and free amino acids (FAAs). *In vitro* antihypertensive and antioxidant activities of digests were investigated, and the peptidomic characterization [high-performance liquid chromatography coupled to tandem mass spectrometry (HPLC-MS/MS)] and identification of bioactive fragments in intestinal digests were performed. Regardless of the seed type, germination increased protease activity and reduced the levels of phytic acid, trypsin inhibitors, and tannins (only in BL). A significant proteolysis of the 7S and 11S globulins and a concomitant increase of peptides and FAAs were observed in all sprouted legumes. Digestion kinetics in sprouts revealed a faster generation of FAAs and peptides than in dry seeds, with changes being more evident for DL, associated with a faster imbibition, germination, and sprout growth. In contrast, BL sprouts showed the lowest protein digestibility, likely due to a lower protease activity, seed structure disintegration, and higher anti-nutrient levels in comparison to GZL, DL, and ZF. Moreover, the digestion of sprouts resulted in a higher number of resistant peptides in DL and ZF that matched with previously reported bioactive sequences, suggesting a promising health potential of legume sprouts that was confirmed *in vitro*. The results suggested that the germination process improved protein digestibility and the health-promoting potential of lentil and fava bean proteins although these changes were more evident in DL due to its rapid imbibition, faster germination, and sprout development. This study will provide important information for either plant

breeders to develop legume varieties with permeable seed coats or food producers that could use dehulled seeds for efficient production of sprouts as sustainable food sources of plant proteins with improved nutritional and healthy properties.

Keywords: seed imbibition, germination kinetics, peptidome analysis, bioactivity, lentil, fava bean, digestion, proteins

INTRODUCTION

Protein is one of the major food components for balanced human diets. There is an increasing interest in exploring plant-based protein sources to encourage their consumption as an alternative to animal proteins motivated by the global awareness on food security, environmental sustainability, and health benefits (Loveday, 2019). Among plant protein sources, legume crops are the main dietary sources for >1 billion people, providing high protein yield (18–32%) and generous economic and environmental benefits (Saricaoglu, 2020; Rahate et al., 2021). Legume crops provide symbiotically fixed nitrogen, thus reducing the contribution to atmospheric greenhouse gases and improving soil fertility for successive crops.

Some pulse crops (pea, fava bean, and lentil) are excellent sources of high-quality proteins for human nutrition. In general, legume proteins contain all the essential amino acids (EAA) and are rich in lysine (K), leucine (L), aspartic acid (D), glutamic acid (E), and arginine (R); whereas cysteine (C), methionine (M), tryptophan (W), and threonine (T) are considered to be limited (Boye et al., 2010; Joshi et al., 2017). Relatively little information is available in the literature regarding the behavior of legume proteins during human digestion and the characterization of the resulting products of protein hydrolysis. Knowledge on how plant proteins behave in the human gut and their associated nutritional value will eventually contribute to the development of high-quality and healthier plant-based food products.

It is known that legume protein digestibility and the bioavailability of amino acids are lower in comparison with animal proteins (Ohanenye et al., 2020). The seed structure, protein primary structure, and the presence and quantity of anti-nutrients (trypsin and chymotrypsin inhibitors, tannins, and phytic acid) are the major determinants of seed-based protein quality. Food processing techniques involving the changes of seed structure, removal of anti-nutrients as well as solubilization, fractionation, and concentration of plant proteins may improve legume protein quality. Germination is an ancient sustainable bioprocess that provides a great opportunity for improving legume protein digestibility as it involves enzymatic hydrolysis that modifies the seed structure, resulting in the breakdown of proteins stored within plant cell vacuoles into free amino acids (FAAs) (Sjöö and Nilsson, 2017; Chen et al., 2019) and reducing anti-nutrient content (Nosworthy et al., 2018). Nevertheless, the behavior of proteins during human gastrointestinal digestion of sprouted legumes in terms of the distribution of nitrogen fraction and identification of hydrolysis products remains unknown. Improved digestibility of proteins in sprouted legumes may concomitantly facilitate the release of bioactive peptides, providing additional health benefits (Ohanenye et al., 2020).

The seed coat not only has protective and structural functions but also provides a crucial role in timing of the seed germination of legume crops by regulating water uptake (Smýkal et al., 2014). Despite not being formally demonstrated, the main testa pigments, proanthocyanidins, are hypothesized to play a role in seed coat permeability (Nakayama and Komatsu, 2008; Smýkal et al., 2014). The hydrophobic nature of pigmented seed coats, likely due to the presence of proanthocyanidins (condensed tannins) in the epidermis, was presumed to be responsible for the delayed permeability of seeds and germination. The modification of seed permeability to water through the reduction of tannin content of seed coats may provide additional benefits in legume protein digestibility based on a reduced anti-nutrient content and speeding up the germination-induced mobilization of seed storage proteins. Evaluating specific legume crops with known genetically based contrasting seed coat biochemical profiles could identify genetic variability in the effect of their germination on the quality and the health-promoting potential of the protein fraction. The objective was to study a variation in the kinetics of imbibition, radicle emergence, and sprout elongation of the selected legume types representing contrasting seed permeability (dehulled, zero, low, and high tannin seeds) to establish associations with differences in germination-induced changes on the anti-nutrient and protein content, amino acid profile, proteolysis, cotyledon cellular microstructure, and their influence on *in vitro* protein digestibility and bioactivity. Kinetics and yields of the protein breakdown products and, for the first time, the peptidome generated by legume flours at the end of intestinal phase of digestion and the identification of bioactive peptides are presented. This study will provide important information for either plant breeders to develop legume varieties with permeable seed coats or food producers that could use seeds with improved permeability for efficient production of sprouts as sustainable

Abbreviations: A, alanine; ACE, angiotensin converting enzyme; ABTS, 2,2′-azino-bis-3-ethylbenzothiazoline-6-sulfonic acid; ANOVA, One-way ANOVA; BL, beluga lentil; BBB, blood brain barrier; C, cysteine; CW, cell wall; D, aspartic acid; DH, degree of hydrolysis; DL, dehulled red lentil; E, glutamic acid; EAA, essential amino acids; F, phenylalanine; FAA, free amino acids; G, glycine; GZL, gray zero-tannin lentil; H, histidine; I, isoleucine; HIA, human intestinal absorption; K, lysine; L, leucine; M, methionine; N, asparagine; NEAA, non-essential amino acids; MW, molecular weight; ORAC, oxygen radical absorbance capacity; OPA, *o*-phthaldialdehyde; P, proline; PBS, phosphate buffer saline; PCA, principal component analysis; PSV, protein storage vacuoles; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Q, glutamine; R, arginine; SIF, simulated intestinal fluid; SG, starch granules; SGF, simulated gastric fluid; SSF, simulated salivary fluid; T, threonine; TE, Trolox equivalents; TIA, trypsin inhibitory activity; TIU, trypsin inhibitory units; V, valine; UPLC-MS/MS, ultra-high performance liquid chromatography-tandem mass spectrometry; Y, tyrosine; ZF, low tannin/zero vicine–convicine fava bean.

food sources of plant proteins with improved nutritional and healthy properties.

MATERIALS AND METHODS

Materials and Seed Germination

Legume seed types representing low tannin [gray zero-tannin lentil (GZL) breeding line from the University of Saskatchewan (UoS), genetically related to the cultivar CDC Gold (GZL) and low tannin/low vicine–convicine fava bean (ZF) breeding line 951-1-11] and high tannin/anthocyanin [Indianhead lentil, a cultivar developed at UoS marketed as beluga lentil (BL)] were selected to sample seed coat biochemical diversity. An additional seed type representing a decorticated legume sample [dehulled commercial sample of CDC Maxim red lentil from UoS (DL)] was also selected for additional comparison. All seed types were stored at 4°C in vacuum-sealed plastic bags. For germination trials, 150 g of each seed type were disinfected for 10 min with 0.5% sodium hypochlorite before soaking in sterile water (1 : 6, w : v) at 20°C and then rinsed three to four times with tap water. Soaking time was established for each seed type by plotting the timeline of seed weight during the imbibition phase. Soaked seeds were covered by a moist filter paper and placed in a thermostatically controlled climatic chamber (model EC00-065, Snijders Scientific, Tilburg, The Netherlands), with a water circulating system to maintain air humidity ≥90%. Sprout growth was monitored over time measuring the length of rootless and shoots. Germination trials were performed at 20°C at two endpoints (96 and 144 h) in the dark. Under these experimental conditions, the germination rate reached values over 90% in all the cases (**Supplementary Table 1**). Samples were immediately frozen at −80°C and freeze-dried using a lyophilizer (Virtis Company, Inc., Gardiner, NY, USA). Flours from dry seeds and sprouts were obtained using an Ika M20 laboratory mill (IKA-Werke, Staufen, Germany). After milling, flours were sieved (mesh size 300 μm) and stored at −20°C in vacuum-sealed plastic bags.

In vitro Simulated Gastrointestinal Digestion

Flours were digested using INFOGEST 2.0 protocol of static *in vitro* gastrointestinal digestion (Brodkorb et al., 2019). First, 1.25x stocks of simulated salivary fluid (SSF), simulated gastric fluid (SGF), and simulated intestinal fluid (SIF) were prepared as given in detail in the INFOGEST 2.0 method. Raw and germinated flours (2 g) were dispersed in 2 ml of SSF containing 75 U/ml of human salivary amylase (E.C.3.2.1.1, Sigma-Aldrich, St. Louis, MO, USA), and then incubated for 2 min at 37°C and pH 7. The oral digest was immediately incubated with SGF (50:50, v:v) containing 2,000 U/ml of pepsin porcine from gastric mucosa (E.C. 3.4.23.1, Sigma-Aldrich, St. Louis, MO, USA) at pH 3 and 37°C for 2 h in an orbital shaker. Gastric digestion was stopped for adjusting pH at 7 with NaOH 1 M. The intestinal phase was initiated by the addition of SIF (50:50, v:v) containing 100 U/ml of pancreatin from porcine pancreas (Sigma-Aldrich, St. Louis, MO, USA) and porcine bile extract (2.5 mM final volume) for 2 h at 37°C at pH 7. The intestinal phase was stopped by thermal treatment at 80°C for 10 min. Gastric (g) and intestinal (i) digests were freeze-dried and stored at −20°C in vacuum-sealed plastic bags.

Determination of Anti-nutrients

Trypsin inhibitory activity (TIA) was determined as described previously by Sueiro et al. (2015). Flours (100 mg) were dispersed in 5 ml of 0.01 M NaOH (pH 8.4–10.0) and incubated for 3 h at 20°C in a Thermomixer C (Eppendorf, Thermo Fisher Scientific, Waltham, MA, USA). The final volume of the resulting solution was adjusted to 10 ml with distilled water, shaken, and left standing for 15 min, after which aliquots of 1 ml were withdrawn and diluted to cause 40–60% inhibition of the trypsin activity. TIA was expressed as trypsin inhibitory units (TIU)/mg of flour.

Phytic acid content was determined by the Phytic Acid (Phytate)/Total Phosphorus Assay kit (Megazyme, Wicklow, Ireland). Phytic acid content was expressed as mg/g flour.

An analysis of condensed tannins was carried out as described previously by Zurita et al. (2012). Briefly, 200 mg of flour were hydrolyzed with 10 ml of hydrochloric acid (HCl)/n-butanol (5:95, v:v) containing 0.7 g/L of iron (III) chloride at 100°C for 1 h. Samples were centrifuged (14,000 × g for 10 min) and supernatants were washed two times with 10 ml of BuOH:HCl:$FeCl_3$. After adjusting the final volume to 25 ml, absorbance was measured at 550 nm in a Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA). Delphinidin chloride (processed in the same way as samples) was used as standard (Extrasynthese, Genay, France). External calibration curve in the linear range from 0 to 40 μg/ml was plotted to calculate condensed tannin content. The values were expressed in mg/g of flour.

Determination of Protein, Peptides, and Amino Acids

The total protein content in dry seeds and germinated legume flours (0.5 g) was determined by the Dumas method using a Trumac nitrogen analyzer (Leco Corporation, St Joseph, MI, USA). Nitrogen content was transformed using a conversion factor of 6.25 and data were expressed as g protein/100 g flour.

Molecular weight (MW) distribution of peptides (<3, 3–10, and 10–30 kDa) in the total protein fraction was measured by the Quantitative Colorimetric Peptide Assay kit (Pierce, Thermo Fisher Scientific, Waltham, MA, USA) in permeates obtained by ultrafiltration through Amicon units with the cellulose membranes of 3, 10, and 30 kDa MW cutoff (Merck, Darmstadt, Germany). Absorbance was read at 480 nm using a Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA). The results were expressed as g/100 g of flour or g/100 g of protein for gastric and intestinal digests.

Total amino acids (TAAs) were determined in flours after acid hydrolysis using 6 N HCl at 110°C under vacuum for 22 h containing 100 μM of internal standard (Norleucine, Sigma-Aldrich, St. Louis, MO, USA). FAAs were extracted by dispersing 100 mg of flours or freeze-dried digests in 1 ml of 0.1 N HCl for 1 h at 4°C. The purification, derivatization, and analysis of TAAs and FAAs were performed using the EZ:faast™ GC-FID Physiological (Free) Amino Acids (Phenomenex, Torrance, CA,

USA) in an Agilent gas chromatograph 7820 A (Agilent, Santa Clara, CA, USA) equipped with a flame ionization detector (FID), an automatic injector, and a Zebron ZBAAA capillary column (10 m x 250 μm x 0.25 μm; Phenomenex, Torrance, CA, USA). Initially, the temperature of the oven was set at 110°C, and the temperature increased using a ramp of 32°C/min up to 320°C. Derivatized samples were injected (2 μl) in a split mode (15:1) at 250°C. The FID detector was set to 320°C, and the carrier helium gas flow rate was maintained at 1.5 ml/min. External calibration was carried out using a multistandard solution of 26 amino acids, and the results were corrected according to the internal standard recovery. TAA profile was expressed as % of TAAs. FAA content was expressed as g/100 g of flour or g/100 g of protein for gastric and intestinal digests.

Protease Activity and Degree of Hydrolysis

Protease activity was determined in dry seeds and germinated legume flours. Protein extracts were obtained by an incubation of 200 mg of flour in 1 ml of double distilled water at 25°C and 1,500 rpm of agitation speed for 60 min using a Thermomixer C orbital shaker (Eppendorf, Thermo Fisher Scientific, Waltham, MA, USA). After incubation, the samples were centrifuged at 14,000 *g*, 25°C for 10 min. The Fluorescent Protease Assay kit (Pierce, Thermo Fisher Scientific, Waltham, MA, USA) was used for evaluating total protease activity in protein extracts. Fluorescence was read in a microplate reader (Biotek Instruments, Winooski, VT, USA) at excitation and emission wavelengths of 485 and 538 nm, respectively. The increase of relative fluorescence units within 10 min of reaction was plotted as a function of trypsin (standard) concentration. Different concentrations of trypsin ranging from 0 to 1,500 ng/ml (in distilled water) were used to plot an external calibration curve. The results were expressed as ng of trypsin activity equivalents/mg of flour.

The degree of hydrolysis (DH) was monitored using the reported *o*-phthaldialdehyde (OPA) spectrophotometric assay in 96-well plates (Nielsen et al., 2001). Briefly, 50 mg of flour were dispersed in 50 ml of bidistilled water and centrifuged at 10,000 *g* for 5 min at 23°C. Supernatant (30 μl) was placed into each well and mixed in 225 μl of OPA reagent (Sigma-Aldrich, St. Louis, MO, USA). After incubation at 23°C for 2 min, the absorbance was measured at 340 nm using a Synergy HT microplate spectrophotometer (BioTek Instruments, Winooski, VT, USA). Serine was used as a standard at the concentration of 0.1 mg/ml of deionized water (0.9516 meqv/L). The DH was calculated based on the total number of peptide bonds per protein equivalents and the number of hydrolyzed bonds according to theoretical general values (Nielsen et al., 2001).

$$\mathrm{DH}\ (\%) = h/h_{\mathrm{tot}} * 100$$

where h_{tot} was 8.0 and h was expressed as

$$h = \text{Ser-NH}_2 - \beta/\alpha \text{meqv/g protein}$$

Here, β was 0.40, α was 1.00, and Ser-NH_2 was determined as

$$\text{Ser-NH}_2 = \mathrm{Abs}_{\mathrm{sample}} - \mathrm{Abs}_{\mathrm{blank}}/\mathrm{Abs}_{\mathrm{standard}} - \mathrm{Abs}_{\mathrm{blank}} * 0.9516\ \mathrm{meqv/L} * 0.1 * 100/\ \mathrm{X}$$

where X was a 0.05-g sample and 0.05 was the sample volume in liter.

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

Soluble protein extracts obtained from flours were diluted (1:4, v:v) in NuPAGE® dodecyl lithium sulfate sample buffer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). After heating at 70°C for 10 min, 20 μg protein/well were loaded on NuPAGE® Novex 4–12% Bis-Tris gels (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), and run in a Mini Gel Tank (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) with NuPAGE® 2-(N-morpholino) ethanesulfonic acid-SDS (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) at 200 V for 35 min. Novex® Sharp Prestained Protein Standard (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) containing polypeptides from 3.5 to 260 kDa was used as an MW marker.

In digestion experiments, extracts were diluted in Tricine SDS sample buffer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) and loaded (13 μg protein/well) onto Novex 16% Tricine gels (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Electrophoretic separation was performed for 65 min at 125 V. PageRuler^{TM} Unstained Low Range Protein Ladder (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) containing polypeptides from 3.4 to 100 kDa was used as the MW standard. Gels were stained with SimplyBlue SafeStain (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) for 1 h and distained in deionized water for 2 h. Images were obtained using the ChemiDoc MP Imaging System (Bio-Rad, Hercules, CA, USA). Band optical density and MW were measured using the Quantity One software (Bio-Rad, Hercules, CA, USA).

Transmission Electron Microscopy

Seeds that were imbibed and germinated for 96 h were selected for observation using the JEOL 1230 transmission electron microscope (Jeol Ltd., Tokyo, Japan). Seeds were dissected into several small pieces with a scalpel and fixed in 2.5% glutaraldehyde in phosphate buffer saline (PBS, pH 7.4) containing 1.5% *p*-formaldehyde for 2 h at 4°C under vacuum. Subsequently, the samples were imbibed in 5% glutaraldehyde in PBS containing 4% *p*-formaldehyde at room temperature for 2 h. After several washes with PBS, the seed tissues were dehydrated gradually using ethanol solutions from 30 to 100% followed by Spurr's resin infiltration at 4°C for 3 days. For each experimental group, a minimum of three seed replicates were examined under electron microscopy. Ultra-thin sections of embedded tissues were prepared using a diamond knife of a rotary microtome and collected on 200 mesh copper grids. The sections were double stained with 5% uranyl acetate and lead citrate for 20 and 2 min, respectively, at room temperature. The stained sections were examined in the transmission electron microscope at 80 kV, and the images were captured with a TVIPS 4x4K camera (Jeol Ltd., Tokyo, Japan).

Bioactivity of Peptide Fractions

Angiotensin converting enzyme (ACE) inhibitory activity of peptide fractions <3 kDa was determined as previously reported in Garcia-Mora et al. (2014). Briefly, 50 μl of 3mU/ml of ACE from rabbit lung (Sigma-Aldrich, Madrid, Spain) was mixed with 200 μl of Abz-Gly-Phe (NO_2)-Pro (Cymit Quimica, Barcelona, Spain) in the absence (control) or presence of diluted peptide fractions (50 μl). Reactions were carried out in 96-well plates at 37°C for 30 min. Fluorescence was read every minute for 30 min at emission and excitation wavelengths of 335 and 405 nm, respectively, in a Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA). The sample concentration (in mg/ml) that inhibited 50% of enzyme activity (IC_{50}) was calculated by plotting the nonlinear regression sigmoidal dose-response curves in GraphPad Prism 4.00 (GraphPad Software, Inc., San Diego, CA, USA).

The antioxidant activity was determined by the two different methods: the oxygen radical absorbance capacity (ORAC) and 2,2′-azino-bis-3-ethylbenzothiazoline-6-sulfonic acid (ABTS) radical scavenging assay (Martín-Diana et al., 2021). In the ORAC method, 180 μl of 70 nM fluorescein was mixed with 90 μl of 12 mM AAPH and 30 μl of diluted sample. Reaction mixtures were placed in a black 96-well plate (Fisher Scientific, Waltham, MA, USA) and the fluorescence was measured in a Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA) every minute at excitation and emission wavelengths of 485 and 520 nm, respectively. An external calibration curve using Trolox as standard in a linear concentration range from 0 to 160 μM was prepared from a freshly made 1 mM stock solution. The results were expressed as mM of Trolox equivalents (TE)/g of intestinal digest. In the ABTS assay, a stock ABTS solution was prepared by mixing a 7 mM aqueous ABTS solution with 2.45 mM $K_2O_8S_2$ in a 1:1 (v/v) ratio. Then, the stock ABTS solution was diluted with phosphate buffer (75 mM, pH 7.4) to obtain a working solution with an absorbance value of 0.70 ± 0.02 at 734 nm. A volume of 20 μl of diluted samples was mixed with 200 μl of ABTS working solution in a 96-well microplate. The absorbance was measured over 30 min at 734 nm in a Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA), and a Trolox calibration curve was used in the concentration range from 0 to 800 μM. The results were expressed as mmol TE/g intestinal sample.

Peptide Identification

For peptide identification, intestinal digests of seeds were analyzed by ultra-high performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) using a nano Easy-nLC 1000 (Proxeon, Odense, Denmark), connected to an LTQ–Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a nanoelectrospray ionization source. Lyophilized samples were suspended in 0.1% acetic acid until reaching the concentration of 0.5 mg/ml. Intestinal digests were purified using a C18 ZipTip (Millipore, Bellerica, MA, USA) reversed phase columns and eluted using 2% acetonitrile in water containing 0.1% formic acid. Sample aliquots of 4 μl were injected into the nLC-MS/MS equipment and desalted in Acclaim™ PepMap™ 100 precolumn and loaded onto a 0.075 mm x 500 mm, 2 μm bead Acclaim™ PepMap™ 100 C18 column (Thermo Fisher Scientific, Waltham, MA, USA) for analytical separation. Peptides were eluted using a gradient of 0.1% formic acid in water (solvent A) and 99.9% acetonitrile containing 0.1% formic acid (solvent B) at a flow rate of 250 nl/min. The mass spectra corresponding to the complete scan (m/z 400–2,000) were obtained with a resolution of 70,000, and the 15 most intense ions of each scan were selected for fragmentation by means of collision-induced dissociation in the ionic trap, with the collision energy normalized to 35%. Ions with single charge or no charge assignment were discarded. A dynamic exclusion was carried out with the duration of 20 s. Lentil and fava bean databases were prepared by combining the NCBI (https://www.ncbi.nlm.nih.gov) and Uniprot (https://www.uniprot.org) "*Lens*" and "*Vicia faba*" protein sequences. The peptide mass spectra were compared with protein database using Proteome Discoverer (version 1.4.1.14, Thermo Scientific, Waltham, MA, USA). Carbamidomethylation of cysteine and oxidation of methionine as fixed and dynamic modifications, and proteases trypsin and chymotrypsin (allowing 2 "miscleavages") were selected in the search parameters. The selection of precursors and product ions was set to be tolerated at 10 ppm and 0.05 Da, respectively. The identification of the peptides was validated by the Percolator algorithm using the value of $q \leq 0.01$ (Käll et al., 2007). The proteomics analysis was carried out in duplicate at the Proteomics and Genomics Facility of the Center for Biological Research (CIB-CSIC), a member of ProteoRed-ISCIII network. Each detected peptide was associated with its annotated sequence, intensity, parent protein accessions, and positions in the parent protein.

Computational Analysis

Peptide patterns after intestinal digestion were represented by using the online software tool Peptigram (Manguy et al., 2017). The potential of peptides to be bioactive was predicted using the PeptideRanker tool (http://distilldeep.ucd.ie/PeptideRanker/) in the Discovery@Bioware server, and their theoretical bioactivity was expressed as the score values calculated (from 0 to 1, with 1 being the most likely to be bioactive). Moreover, the prediction of the bioactive profile was performed using the BIOPEP-UWM database (Minkiewicz et al., 2019). Absorption, Distribution, Metabolism, Excretion, and Transport (ADMET) for the selected peptides was predicted using the AdmetSAR server. Brain blood barrier (BBB), human intestinal absorption (HIA), *Salmonella typhimurium* reverse mutation assay (AMES) toxicity, carcinogenic potential, acute oral toxicity, and hepatotoxicity were predicted with the help of the AdmetSAR tool (http://lmmd.ecust.edu.cn/admetsar2/).

Statistical Analysis

Germination and digestion time course experiments were carried out three times. Experimental data represent the mean and SD of the three replicates ($n = 3$). One-way ANOVA using Duncan's test was employed to compare the mean values and identify statistical differences ($p \leq 0.05$) using Statgraphics Centurion XVIII (Statgraphics Technologies, The Plains, VA, USA). Correlations between traits were analyzed by a Pearson

correlation coefficient (r). Principal component analysis (PCA) was performed to evaluate associations among germination time, legume type, anti-nutrient content, and proteolysis using Statistica 8.0 (Statsoft, Inc, Tulsa, OK, USA).

RESULTS AND DISCUSSION

Imbibition, Germination, and Sprout Development Are Faster in Dehulled Lentil

There were remarkable variations in the duration of imbibition among legume seeds. Clearly, DL showed the shortest imbibition time (1 h, **Supplementary Figure 1C**). The reason for this observation might be an increased water permeability due to the absence of a seed coat that caused a rapid swelling of the seed. The duration of the imbibition phase in GZL, BL, and ZF was 3-, 2-, and 4-fold longer than that found in DL probably due to the presence of a seed coat (**Supplementary Figures 1A–D**). Additionally, the duration of soaking time in the group of legumes with a seed coat was positively correlated to seed weight ($p \leq 0.05$). Further germination trials were performed using 3, 2, 1, and 4 h as the optimum soaking times for GZL, BL, DL, and ZF.

Radicle protrusion indicates the end of germination in "*sensu stricto*" that involves the activation of the endogenous metabolism necessary for mobilizing reserve material and radicle growth (Nonogaki et al., 2010). From a visual inspection, it was observed that radicle protrusion occurred after 24 h for GZL, BL, and ZF (**Supplementary Figures 2A–D**), in contrast the germination of DL occurred faster (15 h, **Supplementary Figure 2C**). These results indicated that post-germination phase, in which the mobilization of seed storage reserves is mainly mobilized, initiates earlier in DL compared to the other seed types. In consistency with the faster imbibition and germination, the development of DL sprouts occurred earlier (between 24 and 89 h from the end of imbibition) (**Supplementary Figure 3C**) compared to GZL, BL, and ZF in which a rapid sprout development occurred between 72 and 144 h (**Supplementary Figures 3A–D**). These results indicated that seed dehulling speeds up the germination and post-germination events. Opposite to DL, ZF showed the slowest sprout development probably associated to its lower rate of water uptake determined by a bigger size of the seed and a reduced surface area.

There is limited mobilization of reserves during germination compared to reserve utilization during post-germination (Nonogaki et al., 2010). Based on this information as an indicative of the period of major mobilization of the reserve material, we selected 96 and 144 h as sampling points to study the nutritional changes on seed proteins in terms of the digestibility, amino acid composition, and bioaccessibility of amino acids and peptides with potential health implications.

Germination Progressively Reduced Anti-nutrient Content in Lentil and Fava Bean Seed Types

The comparison among lentil types before germination showed that GZL and BL were characterized by higher amounts of TIA values (≈11 TIU/mg) but by a lower phytic acid content (0.8 and 0.9 mg/g, respectively) (**Table 1**). In contrast, DL showed a significantly ($p \leq 0.05$) lower amount of TIA (10.2 TIU/mg) but higher phytic acid concentrations (1.1 mg/g) than GZL and BL. Condensed tannins were absent in GZL (zero-tannin variety) and DL in which the removal of a seed coat was the main determinant to this feature (Mirali et al., 2016), whereas BL showed values (4.3 mg/g) comparable to the literature data for tannin content in lentil seeds (Irakli et al., 2021). In comparison with lentil seeds, ZF was characterized by higher TIA (16.1 TIU/mg) and phytic acid (2.6 mg/g) values and 4-fold less concentration of condensed tannins than BL.

There was a general decreasing trend in the amount of anti-nutrients after germination in all legumes studied in agreement with existing evidence (Ohanenye et al., 2020). This effect is related to the activation of dormant enzymes, a physiological response that is influenced by germination conditions, sprouting time, and the type of legume (Donkor et al., 2012). Overall, TIA was significantly ($p \leq 0.05$) decreased after 96 h of germination by 14.0–26.2% in most legumes' studies from which ZF showed the highest reduction. An exception to this observation was DL in which TIA levels were slightly increased by 11%. The extension of germination time from 96 to 144 h did not cause additional reductions of this activity in most of the seeds, except for ZF that showed a further decrease in TIA up to 35%. The partial degradation of trypsin inhibitors during germination as a result of endogenous protease activation has been recently reviewed (Ohanenye et al., 2020). For instance, cowpea, fava bean, Indian bean, lentil, mung bean, and soybean germinated for 2–5 days decreased TIA by 21.9–23.5, 64.0–65.0, 51.0–83.4, 7.0–45.0, 53.6, and 24.0–29.3%, respectively.

Phytic acid content was reduced time-dependently during germination in all the four legumes studied although the extent of degradation was dependent on the legume type and sprouting time. Specifically, phytic acid content was reduced by 10.2–57.8 and by 23.9–63.8% after 96 and 144 h of germination, respectively. The highest decrease in phytic acid was noticed in DL in which sprouts reached the lowest amount of this anti-nutrient (0.4 mg/g). The increase in endogenous phytase activity has been recognized as the primary cause of phytate degradation during the germination and improvement of mineral, protein, and starch bioaccessibility in germinated cereals, pseudocereals, and legumes (Montemurro et al., 2019). Relatively long sprouting times (3–5 days) are needed to lower the phytate concentration by more than 30% (Ohanenye et al., 2020).

Condensed tannins were significantly ($p \leq 0.05$) reduced in BL (7.6%) only when germination was extended up to 144 h. An opposite trend was observed for ZF in which the concentration of condensed tannins was 1.4-fold higher regardless of germination time. Previous studies found a reduction in tannin content of common bean, fava bean, Indian bean, peanut, and soybean subjected to germination (Ohanenye et al., 2020) due to the fact that tannins may leach into liquid medium during the soaking phase or can be enzymatically degraded (Chinma et al., 2021).

TABLE 1 | Anti-nutrient and total protein contents in dry seeds (0 h) and sprouts of lentil and fava bean at different times of germination (96 and 144 h).

Legume type	GT (h)	TIA (TIU/mg)	Phytic acid (mg/g)	Condensed tannins (mg/g)	Protein (g/100 g)
GZL	0	$11.41 \pm 0.34^{a,B}$	$0.75 \pm 0.03^{a,D}$	0	$27.16 \pm 0.15^{c,B}$
	96	$9.09 \pm 0.35^{c,C}$	$0.60 \pm 0.01^{b,C}$	0	$29.27 \pm 0.59^{b,A}$
	144	$9.71 \pm 0.11^{b,B}$	$0.47 \pm 0.01^{c,C}$	0	$30.13 \pm 0.19^{a,A}$
BL	0	$10.96 \pm 0.78^{a,B}$	$0.93 \pm 0.01^{a,C}$	$4.27 \pm 0.12^{a,A}$	$27.57 \pm 0.09^{b,A}$
	96	$9.42 \pm 0.39^{b,C}$	$0.84 \pm 0.02^{b,B}$	$4.43 \pm 0.17^{a,A}$	$29.94 \pm 0.24^{a,A}$
	144	$9.42 \pm 0.55^{b,B}$	$0.71 \pm 0.02^{c,B}$	$3.95 \pm 0.13^{b,A}$	$30.11 \pm 0.45^{a,A}$
DL	0	$10.17 \pm 0.79^{b,C}$	$1.12 \pm 0.02^{a,B}$	0	$24.51 \pm 0.08^{c,D}$
	96	$11.28 \pm 0.37^{a,B}$	$0.47 \pm 0.02^{b,D}$	0	$27.50 \pm 0.17^{b,B}$
	144	$11.20 \pm 0.69^{a,A}$	$0.40 \pm 0.01^{c,D}$	0	$28.07 \pm 0.21^{a,B}$
ZF	0	$16.14 \pm 0.47^{a,A}$	$2.59 \pm 0.04^{a,A}$	$0.98 \pm 0.04^{b,B}$	$24.74 \pm 0.04^{c,C}$
	96	$11.91 \pm 0.36^{b,A}$	$1.17 \pm 0.03^{b,A}$	$1.37 \pm 0.03^{a,B}$	$25.76 \pm 0.49^{b,C}$
	144	$10.03 \pm 0.24^{c,B}$	$1.10 \pm 0.04^{c,A}$	$1.40 \pm 0.06^{a,B}$	$26.53 \pm 0.40^{a,C}$

Data are the mean ± SD of the three replicates.
Different lowercase letters within a column indicate statistical differences among different germination endpoints ($p \leq 0.05$ post-hoc Duncan's test).
Different uppercase letters denote statistical differences among legumes at the same germination time ($p \leq 0.05$ post-hoc Duncan's test).
BL, beluga lentil; DL, dehulled red lentil; GT, germination time; GZL, gray zero-tannin lentil; TIA, trypsin inhibitory activity; TIU, trypsin inhibitory units; ZF, low tannin/zero vicine–convicine fava bean.

Germination Increased Progressively Protein Content in Lentil and Fava Bean Seed Types

Table 1 shows the protein content of dry seeds and sprouted legumes at different times of germination (96 and 144 h). The protein content in lentil types ranged from 24.5 to 27.6% in line with the literature data (Subedi et al., 2021) with the highest values observed for GZL and BL and the lowest for DL ($p \leq 0.05$). ZF showed a protein content (24.7%) comparable to lentil seeds, which was in agreement with previous studies (Warsame et al., 2018). Germination brought about a gradual and significant time-dependent increase in the protein content in lentil and fava bean types (except BL in which there was no significant ($p > 0.05$) difference between 96 and 144 h). This effect was more pronounced in DL with a 12.2% and 14.6% higher protein content after 96 and 144 h of germination, respectively. These results are in accordance with a previous study reporting that germination slightly increased the protein content of chickpea, lentil, and yellow pea by 3.4, 2.9, and 2.8%, respectively (Xu et al., 2019). The observed effect could be explained by the *de novo* synthesis of proteins during germination or a seed compositional change following the degradation of other constituents such as carbohydrates through respiration (Ohanenye et al., 2020).

Germination Slightly Reduced EAA to Non-essential Amino Acids Ratio in Lentil and Fava Bean Seed Types

The distribution of EAA and non-essential amino acids (NEAA) in dry seeds and sprouted legumes at different times of germination (96 and 144 h) is shown in **Table 2**. E, D, and R were found to be the major NEAA in lentil and fava bean seeds, whereas sulfur amino acids C and M were limiting compared with the Joint FAO/WHO Ad Hoc Expert Committee on Energy and Protein Requirements, World Health Organization, and Food and Agriculture Organization of the United Nations (1973) reference pattern. These results are in agreement with the values reported by several authors (Khazaei et al., 2019; Warsame et al., 2020). EAA/NEAA ratio was similar between GZL and DL (45%:55% = 0.81) and higher than the ratio calculated for BL and ZF (39%:61% = 0.64 and 38%:61% = 0.62, respectively). In general, this parameter was minimally reduced in GZL, DL, and ZF after 96 h of germination (EAA/NEAA ratio: 0.76, 0.77, and 0.58, respectively), except for BL in which the distribution of amino acids was not significantly ($p > 0.05$) modified. This decreasing trend observed for the EAA/NEAA ratio reached the lowest values after 144 h in GZL, DL, and ZF (0.67, 0.69, and 0.56, respectively).

Germination Increased Protease Activity and Activated Proteolysis of Storage Proteins in Lentil and Fava Bean Seed Types

Protease activity in dry seeds varied from 0.9 to 1.9 ng trypsin equivalents/mg flour, with the highest values observed for GZL (**Figure 1A**, blue bars). A gradual increase in the protease activity over germination time was observed in most of the cases studied, except for ZF that showed a drop of this parameter by 63% from 96 to 144 h. The highest increase in protease activity during germination was observed for DL (3.3- and 5.6-fold after 96 and 144 h, respectively). Our results are in agreement with previous studies showing a 2- to 3-fold increase in protease activity in soybean when germinated for 2–6 days at 25°C (Zou and Hou, 2017). Maximum values were reached at longer germination times (5 and 6 days). Germination is associated with the activation of endogenous papain- and legumain-type cysteine proteases that are essential for seedling development as they degrade storage proteins to release FAAs for *de novo* synthesis of proteins (Ohanenye et al., 2020).

TABLE 2 | The distribution of essential amino acids (EAA) and non-essential amino acids (NEAA) in raw and sprouted legume flours at different times of germination (96 and 144 h).

	AA	GT (h)	% of total AA			
			GZL	BL	DL	ZF
EAA	Arg	0	6.32	5.25	5.91	5.85
		96	5.35	4.81	4.92	5.13
		144	4.91	4.57	4.47	4.97
	Thr	0	4.20	3.21	4.14	2.56
		96	4.15	3.60	3.99	2.59
		144	3.87	3.61	3.80	2.55
	His	0	2.92	2.76	2.87	2.74
		96	2.92	2.87	2.82	3.17
		144	2.95	2.97	2.74	3.27
	Lys	0	8.86	8.62	8.24	8.10
		96	9.07	8.56	7.89	8.23
		144	7.85	8.68	7.55	7.95
	Phe	0	5.17	5.99	5.31	6.86
		96	5.15	5.71	5.54	6.87
		144	5.13	5.42	5.26	6.55
	Leu	0	7.60	6.37	7.98	5.83
		96	7.13	6.53	7.98	5.77
		144	6.57	6.31	7.69	5.67
	Met	0	0.62	0.31	0.63	0.37
		96	0.58	0.39	0.66	0.34
		144	0.63	0.33	0.55	0.33
	Ile	0	4.07	2.62	4.28	2.62
		96	3.71	2.85	4.41	2.67
		144	3.35	2.77	3.94	2.52
	Val	0	5.12	3.90	5.26	3.37
		96	5.05	4.18	5.41	4.01
		144	4.82	4.21	5.00	4.20
NEAA	Cys	0	0.55	0.78	0.65	0.93
		96	0.65	0.59	0.60	0.93
		144	0.71	0.41	0.50	0.92
	Tyr	0	2.37	1.78	2.51	3.33
		96	2.19	1.52	2.51	3.26
		144	2.39	1.32	2.14	3.07
	Pro	0	4.43	4.51	4.10	3.93
		96	4.51	4.81	4.32	4.54
		144	5.06	4.99	4.63	4.88
	Ala	0	4.99	5.85	5.09	6.03
		96	4.98	5.35	4.82	6.81
		144	5.00	5.13	5.03	6.58
	Gly	0	4.93	6.21	5.03	7.05
		96	4.56	4.96	4.48	6.88
		144	4.45	4.82	4.61	6.65
	Ser	0	6.30	7.15	6.16	6.12
		96	6.58	7.15	6.10	6.97
		144	6.53	7.05	5.90	7.11
	Glu+Gln	0	18.17	20.40	18.82	20.87
		96	16.48	18.14	16.47	21.48
		144	15.89	17.02	17.00	20.83
	Asp + Asn	0	13.40	14.31	13.03	13.43
		96	16.95	17.99	17.08	15.52
		144	19.90	20.40	19.18	17.94

(Continued)

TABLE 2 | Continued

Data are the mean of the three replicates.
AA, amino acids; Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; BL, beluga lentil; Cys, cysteine; DL, dehulled red lentil; Gln, glutamine; Glu, glutamic acid; Gly, glycine; GT, germination time; GZL, gray zero-tannin lentil; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Tyr, tyrosine; Val, valine; ZF, low tannin/zero vicine–convicine fava bean.

The proteolysis of the four studied legume types occurred after germination for 96 and 144 h was monitored by the analysis of protein degradation by the DH (**Figure 1A**, yellow lines), by SDS-PAGE (**Figure 1B**), and the release of peptides of different MW and FAAs (**Figure 1C**). The results showed that there was a time-dependent release of free amino groups in the four legumes studied, in consistency with an increase in protease activity during germination. A clear variation was observed in the kinetic behavior of protein hydrolysis among species. Lentil samples (GZL, BL, and DL) showed a higher release of free amino groups compared to ZF that showed a lower DH at each germination time point. These results could be associated with contrasting imbibition rates and speed of sprout development (**Supplementary Figures 1**, **3**) between lentil and fava bean.

As shown in **Figure 1B**, the three lentil types had a typical electrophoretic profile including polypeptides corresponding to 2S albumins (4–15 kDa), vicilin precursor (20 kDa), legumins (31–42 kDa), vicilins (49 and 53 kDa), convicilins (60–70 kDa), and lipoxygenase isoforms (90–110 kDa) according to the literature (Scippa et al., 2010; Ialicicco et al., 2012). Fava bean seeds showed 14 major protein bands in the MW range from 6 to 111 kDa. The most abundant proteins in ZF were unknown peptide (6 kDa), albumin-1 polypeptides (11 and 14 kDa), polypeptides belonging to legumin β (16, 20, and 43 kDa), vicilin (30 and 48 kDa), and convicilin (54, 61, and 79 kDa), heat shock protein (70 kDa), and linoleate 9S-lipoxygenase-3 isoforms (111 and 94 kDa) (Warsame et al., 2020).

Germination decreased the optical density of major storage protein bands (vicilins, convicilins, and legumins) concurrently with the increased intensity of low MW bands in all lentil types studied, being these effects more pronounced at 144 h of germination (**Figure 1A**). Densitometric analysis of protein bands indicated that after germination vicilin, convicilin, legumin, and lipoxygenase polypeptides are reduced between 11–25, 1–21, 11–41, and 2–21% in GZL, BL, DL, and ZF, respectively (**Supplementary Figure 4**). Moreover, 7S globulins were earlier and greater degraded than 11S globulins during germination. At the same time, intensities of bands with MW between 4–30 kDa increased by 7–22, 7–31, 8–33, and 2–19% in GZL, BL, DL, and ZF, respectively. In consistency with protease activity, proteolysis was more evident in DL, whereas

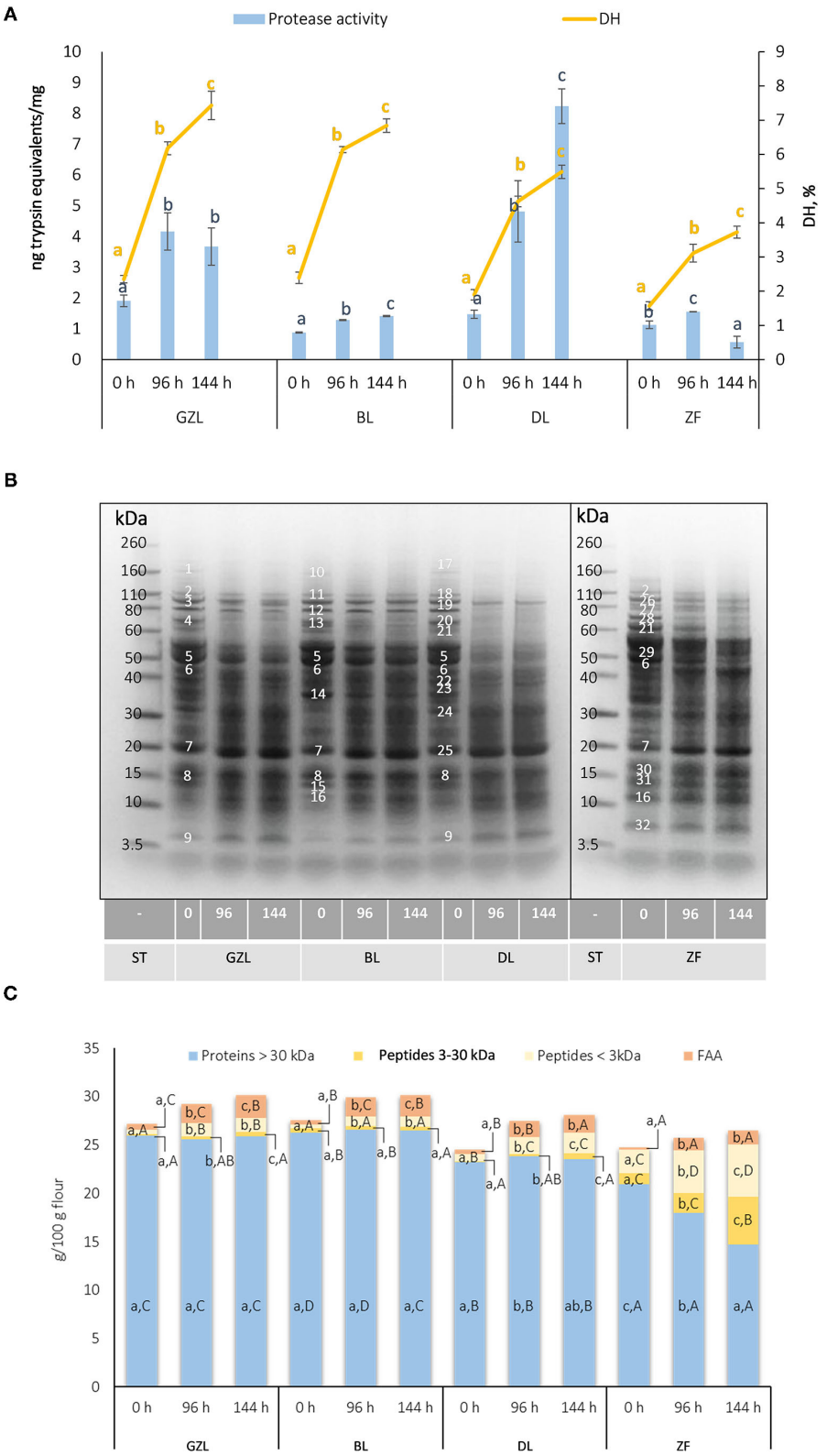


FIGURE 1 | (A) Protease activity and the degree of protein hydrolysis (DH) in ungerminated (0 h) and sprouted lentil and fava bean seeds at different germination times (96 and 144 h). Data are the mean ± SD of the three replicates. Different uppercase letters denote statistical differences among legumes at the same germination time *(Continued)*

FIGURE 1 | ($p \leq 0.05$ *post-hoc* Duncan's test). **(B)** A pattern of proteolysis in ungerminated seeds (0 h) and sprouts of lentil and fava bean at different germination times (96 and 144 h). Representative Coomasie-blue stained gels. 1: Unknown protein, 164 kDa; 2: Lipoxygenase isoform 1/Linoleate-9s-lipoxygenase-3 isoform 1, 111 kDa; 3: Lipoxygenase isoform 2, 93 kDa; 4: Convicilin, 68 kDa; 5: Vicilin, 53 kDa; 6: Vicilin, 48 kDa; 7: Vicilin precursor/Legumin β, 20kDa; 8: 2S albumin, 15 kDa; 9: 2S albumin, 4 kDa; 10: Unknown protein, 115 kDa; 11: Lipoxygenase isoform 1, 107 kDa; 12: Lipoxygenase isoform 2, 90 kDa; 13: Convicilin, 66 kDa; 14: Legumin, 35 kDa; 15: 2S albumin, 13 kDa; 16: 2S albumin/Albumin 1, 11kDa; 17: Unknown protein, 170 kDa; 18: Lipoxygenase isoform 1, 109 kDa; 19: Lipoxygenase isoform 2, 91 kDa; 20: Convicilin, 67 kDa; 21: Convicilin, 61 kDa; 22: Legumin, 42 kDa; 23: Legumin, 39 kDa; 24: Legumin, 31 kDa; 25: Vicilin precursor, 19 kDa; 26: Linoleate-9s-lipoxygenase-3 isoform 2, 94 kDa; 27: Convicilin, 79 kDa; 28: Heat shock protein, 70 kDa; 29: Convicilin, 54 kDa; 30: Legumin β, 16 kDa; 31: Albumin 1, 14 kDa; 32: Unknown peptide, 6 kDa; ST, Molecular weight marker. **(C)** The distribution of protein content between proteins >30 kDa, peptides 3–30 kDa, peptides <3 kDa, and FAA in dry seeds and sprouted lentil and fava bean at different times of germination (96 and 144 h). Data are the mean ± SD of the three replicates. Different lowercase letters indicate statistical differences among different germination endpoints ($p \leq 0.05$ *post-hoc* Duncan's test). Different uppercase letters denote statistical differences among legumes at the same germination time ($p \leq 0.05$ *post-hoc* Duncan's test). FAA, free amino acids; GZL, gray zero-tannin lentil; BL, beluga lentil; DL, dehulled red lentil; ST, molecular weight standard; ZF, low tannin/zero vicine–convicine fava bean.

it was less pronounced in ZF. Therefore, the extent of protein degradation and the generation of new protein fragments with lower MW seem to be linked to seed permeability, the speed of imbibition, germination, and post-germination (Zahir et al., 2020). In consistency with our results, several studies have shown that the degradation of legumin starts later during germination, after vicilin; as such, legumin represents the bulk reserve source for amino acids in the legume seeds. For instance, electrophoretic profiles of yellow pea proteins showed the disappearance of 7S globulins (convicilin and vicilin) earlier in the germination than 11S globulins (legumin) (Di Stefano et al., 2019). Similarly, β-conglycinin was degraded earlier than the acidic glycinin subunit while the basic chain was more resistant (Kim et al., 2011).

The distribution of the nitrogen fraction in dry seeds and sprouted legumes at different times of germination is shown in **Figure 1C**. FAAs, peptides <3 kDa, and polypeptides from 3 to 30 kDa represent a small portion of the total nitrogen fraction (1.1–2.3, 1.6–10.7, and 0.7–4.9%, respectively) in the four dry seed types. ZF showed a lower content of FAA but the highest amounts of peptides with MW <3 kDa and from 3 to 30 kDa compared to lentil seeds, differences that could be associated to species-specific profiles of the nitrogen fraction. After germination, an increase of low molecular protein fragments and FAAs was observed in all the samples studied. There was a time-dependent increase of protein fragments with MW from 3 to 30 kDa and below 3 kDa in DL and ZF, whereas FAAs increased gradually up to 144 h of germination only in BL. The highest increase of FAAs caused by germination was observed for ZF (5- to 6-fold change after 96 and 144 h vs. dry seeds). On the other hand, GZL and DL sprouts showed the highest increases (almost a 3-fold change regardless of time) in the concentration of peptides <3 kDa. In terms of absolute concentration, GZL sprouts had the highest FAA content (2.0 and 2.3 g/100 g flour after 96 and 144 h of germination, respectively), whereas ZF sprouts were characterized for the highest levels of peptides with MW <3 kDa (4.4 and 5.3 g/100 g flour after 96 and 144 h of germination, respectively) and from 3 to 30 kDa (2.1 and 5.0 g/100 g flour after 96 and 144 h of germination, respectively).

Germination-Induced Changes on Protein Fraction and Anti-nutrient Content Varied With Time and Among Legume Type

Principal component analysis was applied to explore a variation in the effect of germination on the protein content, anti-nutrient content, and markers of proteolysis (DH, FAAs, peptides with MW <3 kDa and between 3 and 30 kDa) as influenced by post-germination time and legume type factors. The first two components accounted for 63.71% of the total variance (**Supplementary Figure 5**). The first component (factor 1) explained 44.09% of the variance and had positive loadings for TIA and phytic acid, whereas it was negatively influenced by the total protein content, FAAs, DH, and germination time (**Supplementary Figure 5A**). The second component (factor 2, 19.62% of the variability) had positive loadings for peptide content (<3 kDa and between 3 and 30 kDa). Furthermore, the protein content and markers of proteolysis such as FAAs and DH were positively correlated with germination time ($r = 0.61, 0.91, 0.82$, $p \leq 0.05$). Furthermore, DH was positively correlated with FAAs and negatively correlated with TIA and phytic acid content ($r = 0.97, -0.63, -0.64$, $p \leq 0.05$).

The score plot of the two principal components (**Supplementary Figure 5B**) showed a high variation in the extent of proteolysis and anti-nutrient content between legumes and germination time. The proximity among cases reflected similarities in composition. When analyzing this plot for the discrete variable "germination time", it was clear that the great differences between dry and germinated seeds although extending the germination time from 96 to 144 h did not result in greater changes in the composition of sprouts with the exception of ZF. Dry seeds clustered on the right side (positive) of a bi-plot, indicating their higher content in TIA and phytic acid and lower protein content and markers of proteolysis than sprouts regardless of sprouting time; however, germination changes differed significantly among seed types. ZF sprouts decreased in principal components 1 and 2, indicative of a slower rate of proteolysis based on a progressive increase in the protein content, a degradation of anti-nutrients, and an increase of protein breakdown products with higher MWs (peptides < 3 kDa and between 3 and 30 kDa) up to 144 h of germination. These changes were also observed in the case of lentil types although they were located on the left side of the score plot, due to a greater protein content, a higher inactivation of anti-nutrients, and the accumulation of the proteolysis end products (FAAs) compared to ZF sprouts. GZL and BL experienced similar nutritional changes after germination in terms of anti-nutrients and protein mobilization, as confirmed by their proximity in the score plot regardless of sprouting time; however, DL was clustered with different coordinates due to its lower anti-nutrient content and more extensive proteolysis.

Due to the similarities found in legumes germinated for 96 and 144 h in most of the seed types included in the

current study, the shorter germination time was selected as an representative to observe a variation in the effect of germination on seed microstructure, protein digestibility, and bioactivity across seed types.

Germination Degrades Cellular Structure in a Variable Extent Depending on the Legume Seed Type

As mentioned earlier, germination triggers partial digestion of seed proteins, which can be translated into higher digestibility and nutritional value when consumed by humans. Nonetheless, an intact cellular structure has been identified as a critical factor affecting the rate and extent of seed protein digestion (Becker and Yu, 2013). A thick and compact seed coat and cell walls (CWs) of cotyledon represent the physical barriers that limit the activity of digestive enzymes although germination may improve amino acid bioavailability through modifications on the integrity of plant CWs and the dissociation of nutrient–matrix complexes (Zahir et al., 2020). Therefore, changes in the seed microstructure of lentil and fava bean types as affected by germination for 96 h were investigated.

Changes observed in the cellular structure of parenchyma cells in the mid-region of cotyledons of the soaked and germinated legume seeds are shown in **Figure 2**. The two main nutrient reserves of legume seeds, protein and starch, were stored in the specialized organelles of seed cotyledon parenchyma cells such as protein storage vacuoles (PSV) and starch granules (SG), respectively. An examination of the parenchyma cells after soaking of GZL, BL, DL, and ZF seeds (**Figures 2A–G**, respectively) revealed a compact cellular structure covered by an intact CW that was filled with numerous PSV of various sizes and large-size SG. Other organelles, including rough endoplasmic reticulum and Golgi apparatus, were not observed in these cells. The observation of parenchyma cells of sprouted legumes at 20°C for 96 h (**Figures 2B–H**) presented a different anatomy compared with cotyledon images of non-germinated seeds. It could be clearly seen that cotyledons of sprouted seeds showed a lower density of PSV with heterogeneous sizes. In a particular case of DL, the depletion of PSV and emptying of these organelles were more evident (**Figure 2F**), suggesting a more extensive mobilization of storage proteins as compared to GZL, BL, and ZF. This observation was consistent with shorter imbibition, germination, and post-germination phases observed for DL (**Supplementary Figures 1**, **3**) that resulted in higher protease activity and protein degradation (**Figures 1A,B**) in comparison with other seed types. By comparing the micrographs of the other seed types having a seed coat, minor changes were observed for BL and ZF after 96 h from the initiation of germination. These results were, indicative of delayed post-germination changes in seed microstructure compared to GZL that could be attributed to low seed permeability for BL and a bigger seed size in the case of ZF. Similarly, a study on five legumes (fava bean, yellow field pea, green lentil, chickpea, and kidney bean) demonstrated that microstructure changes induced by germination did not follow the same degradation pattern in all legume types (Di Stefano et al., 2019). The germination of yellow field pea led to parenchyma cell disappearance, an increase in CW thickness and a depletion in SG protein bodies, whereas in green lentil both the seed coat and CW appeared thin after germination.

Based on existing evidence, changes in the cellular microstructure observed in our study suggest that germination may have initiated the permeabilization of CWs and the predigestion of storage proteins attached to cell structure (**Table 1**). This hypothesis is supported by the study of Zahir et al. (2020) who demonstrated that partial degradation/solubilization of pectin and other polymers of the CW in germinated soybean seeds led to extensive cellular permeability to dextrans of high molecular sizes (up to 150 kDa). It is important to highlight that, in agreement with our results, the extent of these changes varies among legume types. For instance, a higher microstructure degradation and consequently the degree of proteolysis were reported for green lentil, yellow pea, chickpea, and fava bean after 5 days of germination compared to a kidney bean, which showed the lowest protein bioaccessibility (Di Stefano et al., 2019).

Germination Favored Protein Digestion in Lentil and Fava Bean Seed Types

Digestion kinetics of proteins in dry and sprouted seeds was followed by SDS-PAGE. A representative gel is illustrated in **Figure 3A** where a gradual degradation of the main storage proteins (vicilin, convicilins, and legumins) and the generation of smaller breakdown products in dry and sprouted seeds can be seen. Regardless of sample treatment, vicilin- and legumin-type globulins were partially resistant to pepsin digestion in lentil and fava bean types. Clear differences in the protein profile of lentil and fava bean were observed between dry and sprouted seeds at the end of the gastric phase (g lanes). A lower intensity of bands corresponding to convicilins, vicilins, and legumins and a higher intensity of peptide bands of MW <10 kDa were noticed in sprouted samples as compared to dry seeds (u lanes). These results suggested a more rapid digestion of proteins by pepsin in GZL, BL, DL, and ZF sprouted seeds. As it is well known that the hydrolysis of proteins by pepsin is very limited (Santos-Hernández et al., 2020), the higher protein hydrolysis in the gastric phase demonstrates the role of the germination process in "pre-digesting" proteins into smaller peptides and, eventually, leading to more digestible proteins in the gastric phase.

At 120 min of intestinal digestion (i lanes), most of the protein bands had fully disappeared and only those with MW under 15 kDa became more intense, which probably correspond to incomplete protein degradation products. Furthermore, new bands having MWs ∼24, 36, 38, and 52 kDa, also visible in the blank of digestion (blank of digestion (BLK) lane), might belong to digestive enzymes pepsin (36 kDa) and pancreatin (23–27, 38, and 51–54 kDa) including trypsin, amylase, lipase, ribonuclease, and protease (Santos-Hernández et al., 2020). Similar to the results found in the gastric phase, visible differences were observed between intestinal digesta of dry and sprouted seeds in which higher intensities in bands with MWs <15 kDa could be seen. Similar observations were described by Santos-Hernández et al. (2020) and Sousa et al. (2020) who reported

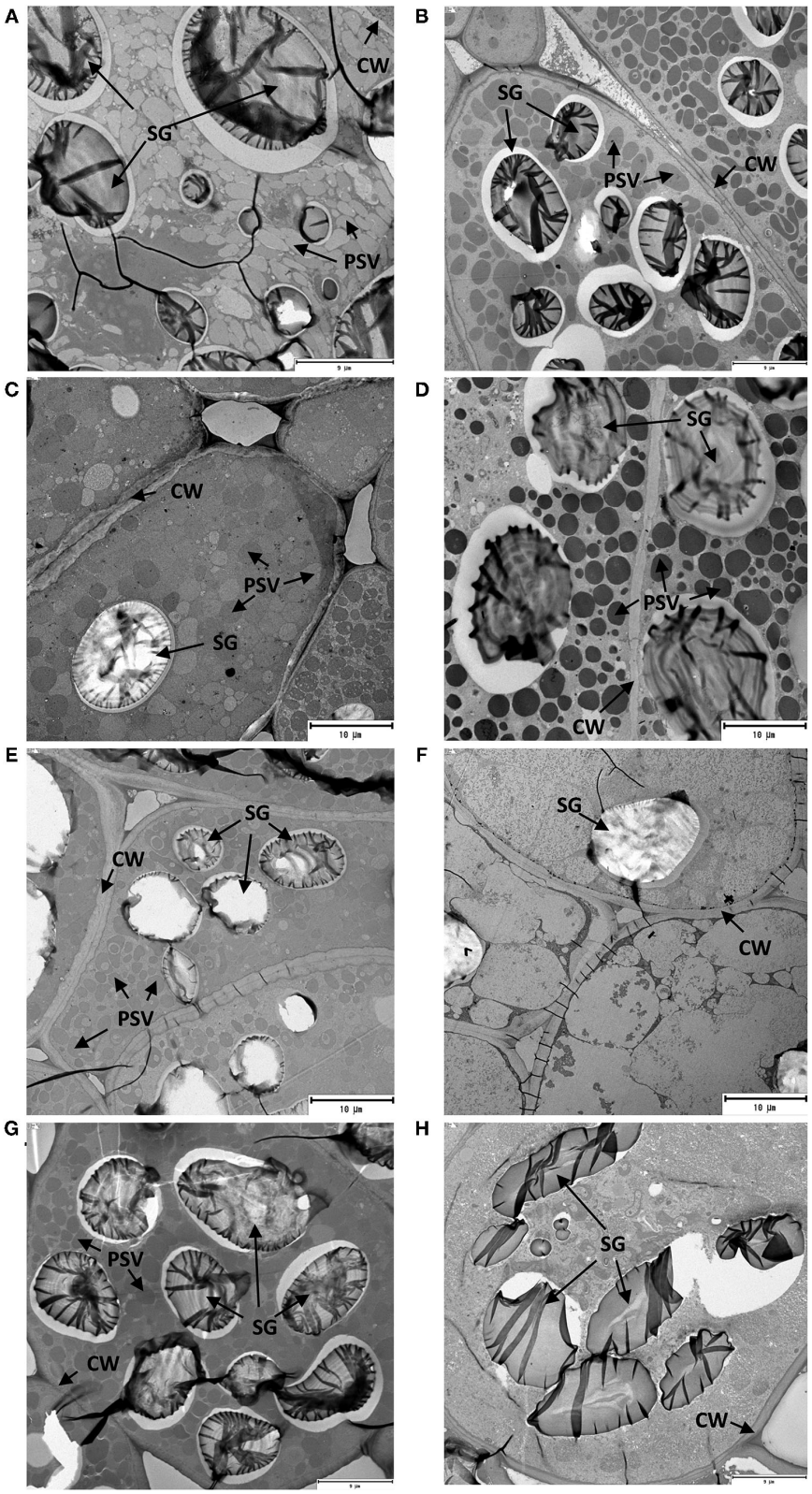


FIGURE 2 | Transmission electron micrographs of cotyledons from the soaked (**A,C,E,G** corresponding to GZL, BL, DL, and ZF, respectively) and sprouted lentil and fava bean seeds at 20○C for 96 h (**B,D,F,H** corresponding to GZL, BL, DL, and ZF, respectively).

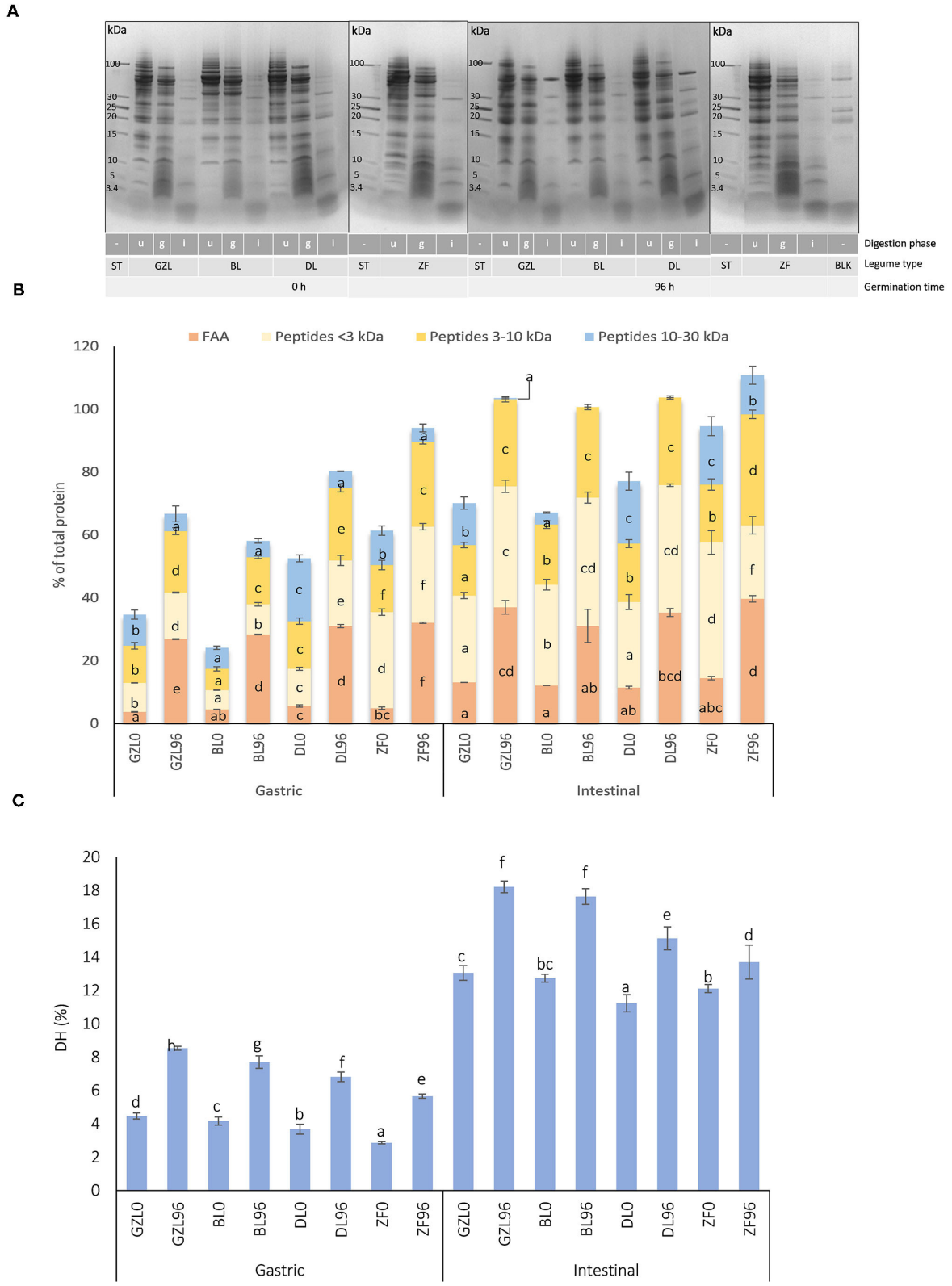


FIGURE 3 | (A) SDS-PAGE protein profiles of dry and sprouted lentil and fava bean seeds (20°C, 96 h) at different times of *in vitro* gastrointestinal digestion. **(B)** The distribution of nitrogen fraction after the simulated gastric and intestinal digestion of lentil and fava bean seed types before (GZL0, BL0, DL0, and ZF0) and after seed

(Continued)

FIGURE 3 | sprouting at 20°C for 96 h (GZL96, BL96, DL96, and ZF96). **(C)** Degree of hydrolysis (DH) after the simulated gastric and intestinal digestion of dry and sprouted lentil and fava bean seeds at 20°C for 96 h. Data and error bars represent the mean and SD, respectively, of the three replicates. In each analysis performed, different letters indicate statistical differences among legumes at the same digestion phase ($p \leq 0.05$ *post-hoc* Duncan test). BL, beluga lentil; BLK, blank of digestion (contains all reagents and enzymes used in digestion; DL, dehulled red lentil; ZF, low tannin/zero vicine–convicine fava bean; GZL, gray zero-tannin lentil; g, end of gastric digestion (120 min); i, end of digestion in the small intestine (240 min); ST: molecular weight marker; u, undigested sample.

that the partial resistance of some legume proteins such as vicilin, provicilin, convicilin, lipoxygenase, legumin A, and legumin A2 during the gastric phase and an almost complete degradation at the end of the intestinal phase joined to the appearance of new electrophoretic bands below 15 kDa.

The evolution of MW distribution of the nitrogen fraction up to 30 kDa in dry and sprouted seeds during digestion is shown in **Figure 3B**. Compared with the intestinal phase, nitrogen fraction of gastric digesta was composed of a lower percentage of FAAs and small peptides independent of the legume type and germination. For instance, after gastric digestion of ungerminated GZL, 3.7, 21.0, and 9.9% of the total nitrogen fraction were comprised of FAAs, peptides (up to 10 kDa), and proteins between 10 and 30 kDa, respectively; and this percentage increased up to 13.1, 43.7, and 13.3%, respectively, at the end of intestinal digestion. A similar response was obtained for BL, DL, and ZF. Higher peptide solubilization and FAA release in the intestinal phase are typically observed with other legume protein isolates from garden pea, grass pea, soybean, and lentil submitted to the INFOGEST method (Santos-Hernández et al., 2020).

In line with peptides and FAA release, DH (**Figure 3C**) reached the highest levels in the intestinal phase regardless of seed type and germination. By comparing dry and sprouted seeds for 96 h, a significantly higher DH was observed after germination for all the legume types either in gastric and intestinal phases of digestion. Our results are consistent with those reported by Ghavidel and Prakash (2007) who showed that the *in vitro* protein digestibility increased by a range of 14–18% after the germination of green gram, cowpea, lentil, and chickpea.

Because the seed coat and CW are the primary barriers to protein digestibility, a relatively compact intracellular environment may serve as an additional barrier for enzyme permeability and diffusion in intact seed cells (Zahir et al., 2020). Thus, loosely packed parenchyma cells that are observed after sprouting may have favored the permeability and diffusion of digestive enzymes, explaining a greater proteolysis rate observed for lentil and fava bean types in this study together with anti-nutrient reduction, protein solubilization, and predigestion, compared to dry seeds. In addition, a lower content of β-sheet and higher contents of the random coil, α-helix, and β-turn have been reported in germinated legumes as compared to dry seeds (Zahir et al., 2021). The helical structure may expose more peptide bonds of proteins for enzymatic hydrolysis compared to the sheet structure (Martinez-Velasco et al., 2018). Moreover, less folded proteins are more digestible than highly folded proteins (Joye, 2019). Therefore, in this study, protein structural changes that occurred after seed germination may provide additional support to the increased *in vitro* protein digestibility observed for lentil and fava bean sprouts.

By comparing the four sprouted legumes, at the end of the gastric phase, DL and ZF showed higher percentages of FAA (31 and 32%, respectively) and peptides (44 and 57%) than GZL and BL in which the total fraction was made up of 26 and 28% of FAAs and 34 and 24% of peptides up to 10 kDa, respectively. At the end of intestinal digestion, a similar distribution of the nitrogen fraction was found for the three sprouted lentils in which peptides <3 kDa were the most abundant protein breakdown products (38–41% of total nitrogen) followed by FAA (31–36% of total nitrogen) and peptides between 3 and 10 kDa (28–29% of total nitrogen). A different profile was observed for sprouted ZF at the end of intestinal digestion that showed the highest percentage of FAAs (40% of total nitrogen) and peptides in the range 3–10 kDa (35% of total nitrogen), the lowest percentage of peptides <3 kDa (23% of total nitrogen) and a considerable percentage of proteins between 10 and 30 kDa (12% of total nitrogen). In line with these results, sprouted ZF showed the lowest DH (13.7%) as compared to GZL, BL, and DL (18.2, 17.6, and 15.2%, respectively) (**Figure 3C**), which could be associated with a slow germination process as compared to lentil seed types. By comparing the distribution of nitrogen fraction between dry and sprouted lentil and fava bean types, a positive effect of germination in protein digestibility was evident. Percentages of protein breakdown products in sprouted lentil and fava bean at the end of intestinal phase were higher compared to recently reported values for intestinal digests of garden pea, grass pea, soybean, and lentil (66–76% and 20–30% of soluble nitrogen are peptides and FAAs) (Santos-Hernández et al., 2020).

Germination Improved Bioactive Traits of Lentil and Fava Bean Types Before and During Digestion

To get additional insights into the antihypertensive and antioxidant properties of peptides generated by germination and subsequent gastrointestinal digestion, their ability to inhibit ACE activity and to scavenge peroxyl (ORAC method) and ABTS radicals (ABTS method) was screened (**Table 3**). There was a clear variation in the ACE inhibitory activity of dry seeds, of which GZL and ZF showed the highest antihypertensive potential (IC_{50} = 0.50 and 0.63 mg/ml). *In vitro* digestion affected differently this parameter depending on the legume type, digestion phase, and germination. In general, ACE inhibitory activity significantly increased in the gastric phase of digestion (except for the dry seeds of BL, BL0) and was maintained up to the end of intestinal digestion with the exception of GZL and ZF sprouts in which IC_{50} values significantly increased. These results indicate that simulated gastrointestinal digestion of different legume types generates peptides with different sizes, sequences,

and structures that potentially explain the differences observed in terms of bioactivity (Dugardin et al., 2020). By comparing dry and sprouted seeds, clear differences were observed at each phase of digestion. A similar antihypertensive potential was noticed between dry and germinated seeds at the end of gastric digestion. However, a higher ACE inhibitory activity was shown for GZL and BL sprouts as compared to dry seeds at the end of intestinal digestion, whereas this parameter remained without any changes for DL or even decreased for ZF.

Although the generation of ACE inhibitory peptides has been reported after legume sprouting in some studies, legume seed germination may not always lead to the production of ACE inhibitors (Maleki and Razavi, 2020). Time and temperature of germination as well as legume type are the important factors that influence the production of antihypertensive peptides in pulses. Bamdad et al. (2009) demonstrated that by increasing germination time, more peptides with a lower MW and ACE inhibitory properties were produced in lentils. Mamilla and Mishra (2017) demonstrated that red lentil germination at 40°C could show a higher ACE inhibitory activity than sprouts obtained at 30°C. IC_{50} values for ACE activity obtained in this study are comparable to IC_{50} values of various intestinal digests of sprouted legumes listed in a recent review ranging from 0.2 to 0.8 mg/ml (Maleki and Razavi, 2020). As an example, in the literature, IC_{50} values for digested chickpea, cowpea, black pea, and white pea sprouts (3 days) of 0.6, 0.6, 0.7, and 0.3 mg/ml, respectively, are reported (Maleki and Razavi, 2020).

With respect to antioxidant activity, BL and ZF dry seeds showed the highest values of radical scavenging activity measured by ORAC (62.7 and 60.1 mmol TE/g, respectively) and ABTS methods (41.1 and 59.2 mmol TE/g, respectively) (**Table 3**). A higher antioxidant activity of these seed types as compared to GZL and DL could be associated with the presence of tannins (proanthocyanidins), main contributors to the antioxidant capacity although other polyphenols may also play a supporting role (Elessawy et al., 2021). Overall, antioxidant activity in the four legume samples increased gradually during digestion with the independence of germination treatment. Interestingly, ZF sprouts showed the highest antioxidant activity values for ORAC and ABTS methods (186.1 and 152.1 mmol TE/g, respectively).

In general, sprouting increases the antioxidant activity of legumes. Various studies report a significant increase in antioxidant activity in lentils (Aguilera et al., 2015), fava bean (Mekky et al., 2020), and pigeon pea (Sharma et al., 2019) when seeds were germinated from 1 to 10 days at 20–35°C. The higher antioxidant activity of sprouted seeds in comparison to dry seeds is attributed not only to the production of peptides but also to the accumulation of other antioxidant compounds like vitamins (E and C), polyphenols, and FAA (Lemmens et al., 2019). Plant-derived peptides and FAA such as Y, M, H, K, P, and W are accepted as antioxidants that act as metal-chelating and hydrogen-/electron-donating agents through interactions with free radicals, thereby terminating the radical chain reaction or preventing its formation (Matemu et al., 2021). Peptides may also exert antioxidant activity *via* certain amino acids although this property is determined by the amino acid sequence and composition (Nwachukwu and Aluko, 2019). When considering the differences observed in the protein, peptide, and FAA profiles, and seed microstructures after germination and during the gastrointestinal digestion of lentil and fava bean types, it is not surprising that the bioactivity changed from one pulse to another and from the gastric to the intestinal phase of digestion. Higher amounts of peptides and FAAs in intestinal digests could explain the powerful antioxidant effect of ZF sprouts at the end of the intestinal phase.

Differences in the Peptidome of Intestinal Digests of Dry Seeds and Sprouts of Lentil and Fava Bean Types

Peptidome analyses were undertaken to explore the impact of germination and legume type on peptide profiles generated during digestion. The total number of peptides identified in the four legume samples before (GZL0, BL0, DL0, and ZF0) and after the germination for 96 h (GZL96, BL96, DL96, and ZF96) is shown in **Supplementary Data 1**. The method used in this study resulted in a window of identification of peptides between 814.46 and 4,172.18 Da to obtain an efficient fragmentation in the collision cell and a minimum peptide length of six amino acids that can be unambiguously matched to a single protein sequence. Some trends emerged when we compared the number of the identified peptides common to each sample and the median MW of the peptide population (**Supplementary Figure 6**). Focusing on a comparison between intestinal digests before germination, DL and ZF showed a lower number of peptides (39 and 68, respectively) and median MWs (1,323.6 and 1,421.1 Da, respectively). After germination, GZL and DL stood out with a comparable number of peptides (80 and 76, respectively) but a clear decrease in the median MW (1,285.1 and 1,190.5 Da, respectively) as compared to the intestinal digests of ungerminated seeds. These results led us to hypothesize that the proteins of GZL and DL sprouts are more rapidly and extensively digested than BL and ZF sprouts in agreement with conclusions drawn by other authors when comparing the intestinal peptidomes of milk proteins and enriched legume pasta (Berrazaga et al., 2020).

Most of the peptides in lentil intestinal digests belong to convicilin (≈30% of the total number of peptides) followed by allergen Len c 1.0101, allergen Len c 1.0102, and vicilin type C; whereas the main parent proteins in ZF were vicilin (40% of the total number of peptides) followed by legumin-type B and favin. The identified peptides were represented using the Peptigram web application (**Supplementary Figure 7**) that visualizes the digestion-resistant regions within a protein. The plots generated by Peptigram show the amino acids identified as part of a peptide sequence (vertical bars), the count of peptides overlapping this position (the height of bars), and a green color intensity that is proportional to peptide abundance. The blank areas in the plots correspond to protein regions completely degraded by the action of digestive enzymes. The first interesting observation was that allergen Len c 1.0101 and allergen Len c 1.0102 were more resistant to digestion than convicilin and vicilin in the three lentil types studied either before or after the germination (**Supplementary Figures 7A–C**). Similarly, vicilin in ZF showed

TABLE 3 | ACE inhibitory activity, ORAC, and ABTS scavenging activity of raw and sprouted legumes at different phases of gastrointestinal digestion.

Legume	GT (h)	Digestion phase	ACE inhibitory activity (IC_{50}, mg/mL)	ORAC (mmol TE/ g)	ABTS (mmol TE/ g)
GZL	0	u	$0.50 \pm 0.02^{cd,F}$	$35.57 \pm 2.93^{f,E}$	$30.05 \pm 2.91^{e,D}$
		g	$0.50 \pm 0.02^{cd,BC}$	$69.72 \pm 6.61^{d,F}$	$67.28 \pm 3.77^{d,D}$
		i	$0.63 \pm 0.04^{b,A}$	$118.38 \pm 10.86^{b,C}$	$97.15 \pm 2.55^{b,CD}$
	96	u	$1.01 \pm 0.01^{a,E}$	$45.32 \pm 3.44^{e,D}$	$31.47 \pm 3.06^{e,D}$
		g	$0.44 \pm 0.01^{d,CD}$	$89.81 \pm 6.77^{c,BC}$	$78.57 \pm 6.80^{c,C}$
		i	$0.56 \pm 0.01^{c,ABC}$	$150.69 \pm 14.62^{a,B}$	$107.11 \pm 2.89^{a,C}$
BL	0	u	$0.63 \pm 0.02^{b,F}$	$62.67 \pm 5.08^{c,B}$	$41.14 \pm 4.36^{d,C}$
		g	$0.66 \pm 0.04^{b,A}$	$82.04 \pm 3.78^{b,D}$	$61.83 \pm 5.72^{c,E}$
		i	$0.61 \pm 0.00^{b,AB}$	$158.75 \pm 11.75^{a,B}$	$88.49 \pm 4.51^{b,D}$
	96	u	$4.24 \pm 0.07^{a,A}$	$49.79 \pm 4.28^{d,C}$	$26.58 \pm 2.61^{e,E}$
		g	$0.57 \pm 0.00^{bc,ABC}$	$83.14 \pm 6.73^{b,D}$	$58.52 \pm 5.11^{c,EF}$
		i	$0.47 \pm 0.02^{c,C}$	$159.12 \pm 12.88^{a,B}$	$101.62 \pm 8.80^{a,C}$
DL	0	u	$2.48 \pm 0.02^{b,C}$	$34.68 \pm 3.37^{f,E}$	$26.82 \pm 2.64^{c,E}$
		g	$0.57 \pm 0.10^{c,ABC}$	$75.31 \pm 6.21^{d,E}$	$54.55 \pm 4.74^{b,F}$
		i	$0.51 \pm 0.02^{c,AB}$	$127.87 \pm 12.16^{b,C}$	$98.60 \pm 5.00^{a,CD}$
	96	u	$4.05 \pm 0.05^{a,B}$	$45.08 \pm 4.49^{e,D}$	$23.19 \pm 2.29^{c,F}$
		g	$0.58 \pm 0.02^{c,AB}$	$87.49 \pm 7.10^{c,CD}$	$58.42 \pm 5.16^{b,EF}$
		i	$0.48 \pm 0.03^{c,C}$	$154.05 \pm 14.98^{a,B}$	$99.00 \pm 9.66^{a,CD}$
ZF	0	u	$0.63 \pm 0.02^{b,F}$	$60.13 \pm 5.30^{f,B}$	$59.15 \pm 3.78^{f,B}$
		g	$0.32 \pm 0.00^{c,D}$	$94.58 \pm 7.49^{d,B}$	$89.21 \pm 8.78^{d,B}$
		i	$0.37 \pm 0.05^{c,D}$	$151.53 \pm 13.30^{b,B}$	$131.99 \pm 11.98^{b,B}$
	96	u	$1.29 \pm 0.07^{a,D}$	$87.80 \pm 7.88^{e,A}$	$78.13 \pm 4.48^{e,A}$
		g	$0.44 \pm 0.01^{c,CD}$	$128.97 \pm 11.92^{c,A}$	$117.01 \pm 6.67^{c,A}$
		i	$0.61 \pm 0.03^{b,AB}$	$186.13 \pm 10.33^{a,A}$	$152.10 \pm 14.73^{a,A}$

Data are the mean and SD of the three replicates.
Different lowercase letters within a column indicate statistical differences among different germination endpoints ($p \leq 0.05$ post-hoc Duncan's test).
Different uppercase letters denote statistical differences among legumes at the same germination time ($p \leq 0.05$ post-hoc Duncan's test).
ABTS, 2,2′-azino-bis (3-ethylbenzthiazoline-6-sulfonic acid); ACE, angiotensin I converting enzyme; BL, beluga lentil; DL, dehulled red lentil; GZL, gray zero-tannin lentil; g, end of gastric digestion (120 min); GT, germination time; i, end of digestion in the small intestine (240 min); IC_{50}, half maximal inhibitory concentration; ORAC, oxygen radical absorbance capacity; TE, Trolox equivalents; u, undigested sample; ZF, low tannin/zero vicine–convicine fava bean.

more resistant regions to digestion than legumin-type BL and favin (**Supplementary Figure 7D**). By comparing lentil types, the main difference was observed for DL in which fewer resistant regions in the main storage proteins were found at the end of intestinal digestion. The absence of the seed coat in DL could have contributed to a higher digestibility of lentil storage proteins due to a rapid germination of this seed type that resulted in a more extensive seed endogenous proteolysis and the degradation of cell microstructure, favoring the diffusion of digestive enzymes (Zahir et al., 2021).

Variations in the Bioactive Potential of Digestion-Resistant Peptides in Dry Seeds and Sprouts of Lentil and Fava Bean Seed Types

To identify bioactive peptide sequences in intestinal digests obtained from dry and sprouted lentils and fava bean, we used a combination of computational tools. At first, all identified peptide sequences in the intestinal digests of dry and sprouted legumes were processed with the Peptide Ranker that gave a range of score values from 0.0140 to 0.8654 (**Supplementary Data 2**). Peptides predicted with scores higher than 0.5 (threshold level set in the Peptide Ranker to label a peptide as bioactive) were selected and included in **Supplementary Data 2**. In total, 97 peptides released from convicilin, vicilin, allergen Len c 1.0101, allergen Len c 1.0102, legumin-type B, and favin were predicted to be bioactive and found to display two or more activities. The bioactivity that appeared with more frequency in all samples was ACE inhibitory activity (in consistency with the experimental results of **Table 3**) and dipeptidyl peptidase IV inhibitory activity. Moreover, the peptide sequence DLAIPVNRPGQL identified in DL sprouts has reported ACE inhibitory activity (García-Mora et al., 2017). In line with antioxidant activity shown by the intestinal digests of the studied legume types, a total of nine identified sequences (six in lentil and three in fava bean) were predicted to display antioxidant activity. In particular, the fava bean peptide LSPGDVLVIPAGYPVAIK shares the amino acid sequence VIPAGYP with an antioxidant peptide identified in fava bean hydrolysates (Samaei et al., 2020).

Differences in the number of bioactive peptides were observed among legume types and between dry and sprouted seeds (**Supplementary Data 3**). GZL and BL showed a higher number of bioactive peptides (21 and 20, respectively) that decreased after germination (7 and 12, respectively). A lower number of bioactive peptide sequences were found for DL (3) and ZF (7) although the germination of these legumes increased the number of bioactive sequences for DL (seven in total), whereas it remained unaltered for ZF. In the sprouts of GZL, BL, DL, and ZF, a total of 1, 4, 6, and 1 unique bioactive peptides were identified.

Absorption, distribution, metabolism, excretion, and transport (ADMET) properties are the crucial features of bioactive peptides. Therefore, BBB and HIA, as the two major factors of well-absorbed peptides, were predicted in this study using the AdmetSAR server. Herein, all predicted bioactive peptides had a high probability of permeability for BBB with values between 0.77 and 0.96 (**Supplementary Data 2**). In the prediction of HIA, 15 of 21 peptides in GZL, 15 of 24 in BL, 3 of 9 in DL, and 2 of 7 in ZF had the ability to be absorbed. These results indicate that a large amount of the identified peptides can be absorbed, showing a high probability to be bioavailable. Probability scores for AMES mutagenesis and carcinogenesis indicated that lentil and fava bean bioactive peptides are not mutagenic or carcinogenic. This observation is supported by a recent study, showing a high correlation between the ADMET predicted values and experimental results on the bioactivity and intestinal absorption of ACE inhibitory peptides (Zhao et al., 2019). Regarding hepatotoxicity, the peptides SDLFENL, LSPGDVFVVPAGHPVA, LSPGDVFVVPAGHPV, FLPQFTDADF in lentil and FFEITPK and ATPADVLANAFGL in fava bean showed low probability values (<0.6) that were smaller than those previously calculated for cocoa peptides (Coronado-Cáceres et al., 2020). The peptide SDLFENL identified in all lentil types and the peptide QQQPDSHQKIRR in ZF showed the lowest values of acute oral toxicity (2.1 and 2.3 kg/mol, respectively), whereas LSPGDVFVVPAGHPV found in GZL and BL lentils and LSPGDVLVIPAGYPVAIK in ZF exhibited the highest acute oral toxicity (3.4 and 3.2 kg/mol, respectively).

CONCLUSIONS

This study provides significant insights into the role of the germination process in inducing distinct changes in the seed cellular structure, anti-nutrient profiles, as well as nutritional and healthy traits of lentil and fava bean proteins that are translated into an increment in their digestibility, the bio accessibility of peptides/amino acids, and the bioactivity of protein digestion end products. This study showed that seed permeability had an influence on the speed of the germination process and consequently on the degree of anti-nutrient and seed microstructure degradation and endogenous seed proteolysis, and ultimately in protein digestibility and the release of bioactive peptides. DL was characterized by a rapid imbibition rate, radicle emergence, and sprout development, which was associated with a greater cellular structure degradation, a lower anti-nutrient content, and a higher protein digestibility, and amino acid bioaccessibility as compared to the rest of the studied seed types. Although low tannin/zero vicine and convicine fava bean sprouts had a lower protein digestibility than lentil sprouts due to a slow germination, a higher anti-nutrient content, and a more intact cellular structure, this legume type showed a greater bioactive potential before and during the digestion. The intestinal phase of digestion was an important stage of peptide differentiation among seed types and between dry seeds and sprouts. Peptidome analysis revealed differences in the number of peptides, MW distribution, and bioactive profile of the nitrogen fraction and allowed the identification of protein regions resistant to gastrointestinal digestion. Several peptides resistant to gastrointestinal digestion were found to match with previously reported bioactive peptides, suggesting their potential health effects. Legume types with a higher protein digestibility showed a smaller number of total and bioactive peptides with a lower mean MW. In addition, lentil and fava bean vicilins were more resistant to digestion than legumins. This study provides new knowledge about the role of germination on protein digestibility and the bioactivity of pulses, suggesting the selection of specific types that could be employed to improve the nutritional value and health properties of foods.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: http://www.ebi.ac.uk/pride, PXD027847.

AUTHOR CONTRIBUTIONS

JF, AV, EP, and CM-V: conceptualization. SB-E and EP: formal analysis and data curation. CM-V: project administration. SB-E and CM-V: writing—original draft. SB-E, AV, EP, JF, and CM-V: writing—reviewing and editing. AV: funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This research was carried out with the financial support of the UoS grant Lentils for Sustainable and Healthy Proteins (Grant No. 20196324) through the SPG-NSERC Industrial Research Chair Program in Lentil Genetic Improvement. SB-E is a Postdoctoral Researcher funded by a grant Lentils for Sustainable and Healthy Proteins (Grant No. 20196324).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.754287/full#supplementary-material

REFERENCES

Aguilera, Y., Herrera, T., Liébana, R., Rebollo-Hernanz, M., Sanchez-Puelles, C., and Martín-Cabrejas, M. A. (2015). Impact of melatonin enrichment during germination of legumes on bioactive compounds and antioxidant activity. *J. Agric. Food Chem.* 63, 7967–7974. doi: 10.1021/acs.jafc.5b03128

Bamdad, F., Dokhani, Sh., Keramat, J., and Zareie, R. (2009). The impact of germination and in vitro digestion on the formation of Angiotensin Converting Enzyme (ACE) inhibitory peptides from lentil proteins compared to whey proteins. *World Acad. Sci. Eng. Technol.* 3, 109–119.

Becker, P. M., and Yu, P. (2013). What makes protein indigestible from tissue-related, cellular, and molecular aspects? *Mol. Nutr. Food Res.* 57, 1695–1707. doi: 10.1002/mnfr.201200592

Berrazaga, I., Bourlieu-Lacanal, C., Laleg, K., Jardin, J., Briard-Bion, V., Dupont, D., et al. (2020). Effect of protein aggregation in wheat-legume mixed pasta diets on their *in vitro* digestion kinetics in comparison to "rapid" and "slow" animal proteins. *PLoS ONE.* 15, e0232425. doi: 10.1371/journal.pone.0232425

Boye, J., Zare, F., and Pletch, A. (2010). Pulse proteins: Processing, characterization, functional properties and applications in food and feed. *Food Research International.* 43, 414–431. doi: 10.1016/j.foodres.2009.09.003

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Ballance, S., et al. (2019). INFOGEST static *in vitro* simulation of gastrointestinal food digestion. *Nat. Protocols.* 14, 991–1014. doi: 10.1038/s41596-018-0119-1

Chen, L., Wu, J. E., Li, Z., Liu, Q., Zhao, X., and Yang, H. (2019). Metabolomic analysis of energy regulated germination and sprouting of organic mung bean (*Vigna radiata*) using NMR spectroscopy. *Food Chemist.* 286, 87–97. doi: 10.1016/j.foodchem.2019.01.183

Chinma, C. E., Abu, J. O., Asikwe, B. N., Sunday, T., and Adebo, O. A. (2021). Effect of germination on the physicochemical, nutritional, functional, thermal properties and *in vitro* digestibility of Bambara groundnut flours. *LWT- Food Sci. Technol.* 140, 110749. doi: 10.1016/j.lwt.2020.110749

Coronado-Cáceres, L. J., Rabadán-Chávez, G., Mojica, L., Hernández-Ledesma, B., Quevedo-Corona, L., and Lugo Cervantes, E. (2020). Cocoa (*Theobroma cacao* L.) seed proteins' anti-obesity potential through lipase inhibition using *in silico*, *in vitro* and *in vivo* models. *Foods.* 9, 1359. doi: 10.3390/foods9101359

Di Stefano, E., Tsopmo, A., Oliviero, T., Fogliano, V., and Udenigwe, C. C. (2019). Bioprocessing of common pulses changed seed microstructures, and improved dipeptidyl peptidase-IV and α-glucosidase inhibitory activities. *Scientific Rep.* 9, 15308. doi: 10.1038/s41598-019-51547-5

Donkor, O. N., Stojanovska, L., Ginn, P., Ashton, J., and Vasiljevic, T. (2012). Germinated grains – Sources of bioactive compounds. *Food Chemistry.* 135, 950–959. doi: 10.1016/j.foodchem.2012.05.058

Dugardin, C., Cudennec, B., Tourret, M., Caron, J., Guérin-Deremaux, L., Behra-Miellet, J., et al. (2020). Explorative screening of bioactivities generated by plant-based proteins after *in vitro* static gastrointestinal digestion. *Nutrients.* 12, 3746. doi: 10.3390/nu12123746

Elessawy, F. M., Vandenberg, A., El-Aneed, A., and Purves, R. W. (2021). An untargeted metabolomics approach for correlating pulse crop seed coat polyphenol profiles with antioxidant capacity and iron chelation ability. *Molecules.* 26, 3833. doi: 10.3390/molecules26133833

García-Mora, P., Martín-Martínez, M., Angeles Bonache, M., González-Múniz, R., Peñas, E., Frias, J., et al. (2017). Identification, functional gastrointestinal stability and molecular docking studies of lentil peptides with dual antioxidant and angiotensin I converting enzyme inhibitory activities. *Food Chemist.* 221, 464–472. doi: 10.1016/j.foodchem.2016.10.087

Garcia-Mora, P., Peñas, E., Frias, J., and Martínez-Villaluenga, C. (2014). Savinase, the most suitable enzyme for releasing peptides from lentil (*Lens culinaris* var. Castellana) protein concentrates with multifunctional properties. *J. Agri. Food Chemistry.* 62, 4166–4174. doi: 10.1021/jf500849u

Ghavidel, R. A., and Prakash, J. (2007). The impact of germination and dehulling on nutrients, antinutrients, in vitro iron and calcium bioavailability and *in vitro* starch and protein digestibility of some legume seeds. *LWT - Food Sci. Technol.* 40, 1292–1299. doi: 10.1016/j.lwt.2006.08.002

Ialicicco, M., Viscosi, V., Arena, S., Scaloni, A., Trupiano, D., Rocco, M., et al. (2012). Lens culinaris Medik. seed proteome: Analysis to identify landrace markers. *Plant Sci.* 197, 1–9. doi: 10.1016/j.plantsci.2012.08.010

Irakli, M., Kargiotidou, A., Tigka, E., Beslemes, D., Fournomiti, M., Pankou, C., et al. (2021). Genotypic and environmental effect on the concentration of phytochemical contents of lentil (*Lens culinaris* L.). *Agronomy.* 11, 1154. doi: 10.3390/agronomy11061154

Joint FAO/WHO Ad Hoc Expert Committee on Energy and Protein Requirements, World Health Organization, and Food and Agriculture Organization of the United Nations (1973). *Energy and Protein Requirements: Report of a Joint FAO/WHO ad hoc Expert Committee [Meeting Held in Rome from 22 March to 2 April 1971]*. World Health Organization. Available online at: https://apps.who.int/iris/handle/10665/41042

Joshi, M., Timilsena, Y., and Adhikari, B. (2017). Global production, processing and utilization of lentil: A review. *J. Integrative Agri.* 16, 2898–2913. doi: 10.1016/S2095-3119(17)61793-3

Joye, I. (2019). Protein Digestibility of Cereal Products. *Foods (Basel, Switzerland).* 8, 199. doi: 10.3390/foods8060199

Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and Maccoss, M. J. (2007). Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods.* 4, 923–925. doi: 10.1038/nmeth1113

Khazaei, H., Subedi, M., Nickerson, M., Martínez-Villaluenga, C., Frias, J., and Vandenberg, A. (2019). Seed Protein of Lentils: Current Status, Progress, and Food Applications. *Foods.* 8, 391. doi: 10.3390/foods8090391

Kim, H. T., Choi, U. K., Ryu, H. S., Lee, S. J., and Kwon, O. S. (2011). Mobilization of storage proteins in soybean seed (Glycine max L.) during germination and seedling growth. *Biochim. Biophys. Acta.* 1814, 1178–1187. doi: 10.1016/j.bbapap.2011.05.004

Lemmens, E., Moroni, A. V., Pagand, J., Heirbaut, P., Ritala, A., Karlen, Y., et al. (2019). Impact of cereal seed sprouting on its nutritional and technological properties: a critical review. *Comprehensive Rev. in Food Sci. Food Safety.* 18, 305–328. doi: 10.1111/1541-4337.12414

Loveday, S. M. (2019). Food Proteins: technological, nutritional, and sustainability attributes of traditional and emerging proteins. *Ann. Rev. Food Sci. and Technol.* 10, 311–339. doi: 10.1146/annurev-food-032818-121128

Maleki, S., and Razavi, S. H. (2020). Pulses' germination and fermentation: Two bioprocessing against hypertension by releasing ACE inhibitory peptides. *Critical Rev. Food Sci. Nutr.* 1–18. doi: 10.1080/10408398.2020.1789551

Mamilla, R., and Mishra, V. (2017). Effect of germination on antioxidant and ACE inhibitory activities of legumes. *LWT - Food Sci. Technol.* 75. doi: 10.1016/j.lwt.2016.08.036

Manguy, J., Jehl, P., Dillon, E. T., Davey, N. E., Shields, D. C., and Holton, T. A. (2017). Peptigram: a web-based application for peptidomics data visualization. *J. Proteome. Res.* 16, 712–719. doi: 10.1021/acs.jproteome.6b00751

Martín-Diana, A. B., Tomé-Sánchez, I., García-Casas, M. J., Martínez-Villaluenga, C., Frías, J., and Rico, D. (2021). A novel strategy to produce a soluble and bioactive wheat bran ingredient rich in ferulic acid. *Antioxidants.* 10, 969. doi: 10.3390/antiox10060969

Martinez-Velasco, A., Alvarez-Ramirez, J., Rodriguez-Huezo, E., Meraz-Rodriguez, M., Vernon-Carter, E. J., and Lobato-Calleros, C. (2018). Effect of the preparation method and storage time on the *in vitro* protein digestibility of maize tortillas. *J. Cereal Science.* 84, 7–12. doi: 10.1016/j.jcs.2018.09.016

Matemu, A., Nakamura, S., and Katayama, S. (2021). Health benefits of antioxidative peptides derived from legume proteins with a high amino acid score. *Antioxidants.* 10, 316. doi: 10.3390/antiox10020316

Mekky, R. H., Thabet, M. M., Rodríguez-Pérez, C., Elnaggar, D. M. Y., Mahrous, E. A., Segura-Carretero, A., et al. (2020). Comparative metabolite profiling and antioxidant potentials of seeds and sprouts of three Egyptian cultivars of *Vicia faba* L. *Food Res Int.* 136, 109537. doi: 10.1016/j.foodres.2020.109537

Minkiewicz, P., Iwaniak, A., and Darewicz, M. (2019). BIOPEP-UWM Database of bioactive peptides: current opportunities. *Int. J. Molecular Sci.* 20, 5978. doi: 10.3390/ijms20235978

Mirali, M., Purves, R. W., Stonehouse, R., Song, R., Bett, K., and Vandenberg, A. (2016). Genetics and biochemistry of zero-tannin lentils. *PLoS ONE.* 11, e0164624. doi: 10.1371/journal.pone.0164624

Montemurro, M., Pontonio, E., Gobbetti, M., and Rizzello, C. G. (2019). Investigation of the nutritional, functional and technological effects of the sourdough fermentation of sprouted flours. *Int. J. Food Microbiol.* 302, 47–58. doi: 10.1016/j.ijfoodmicro.2018.08.005

Nakayama, N., and Komatsu, S. (2008). Water uptake by seeds in yellow-seeded soybean (*Glycine max* (L.) merrill) cultivars with contrasting imbibition behaviors. *Plant Prod. Sci.* 11, 415–422. doi: 10.1626/pps.11.415

Nielsen, P. M., Petersen, D., and Dambmann, C. (2001). Improved method for determining food protein degree of hydrolysis. *J. Food Sci.* 66, 642–646. doi: 10.1111/j.1365-2621.2001.tb04614.x

Nonogaki, H., Bassel, G. W., and Bewley, J. D. (2010). Germination—Still a mystery. *Plant Sci.* 179, 574–581. doi: 10.1016/j.plantsci.2010.02.010

Nosworthy, M. G., Medina, G., Franczyk, A. J., Neufeld, J., Appah, P., Utioh, A., et al. (2018). Effect of processing on the in vitro and in vivo protein quality of red and green lentils (*Lens culinaris*). *Food Chemistry.* 240, 588–593. doi: 10.1016/j.foodchem.2017.07.129

Nwachukwu, I. D., and Aluko, R. E. (2019). Structural and functional properties of food protein-derived antioxidant peptides. *J. Food Biochem.* 43, e12761. doi: 10.1111/jfbc.12761

Ohanenye, I., Tsopmo, A., Ejike, C., and Udenigwe, C. (2020). Germination as a bioprocess for enhancing the quality and nutritional prospects of legume proteins. *Trends Food Sci. Technol.* 101, 213–222. doi: 10.1016/j.tifs.2020.05.003

Rahate, K. A., Madhumita, M., and Prabhakar, P. K. (2021). Nutritional composition, anti-nutritional factors, pretreatments-cum-processing impact and food formulation potential of faba bean (*Vicia faba* L.): A comprehensive review. *LWT-Food Sci Technol.* 138, 110796. doi: 10.1016/j.lwt.2020.110796

Samaei, S. P., Ghorbani, M., Tagliazucchi, D., Martini, S., Gotti, R., Themelis, T., et al. (2020). Functional, nutritional, antioxidant, sensory properties and comparative peptidomic profile of faba bean (*Vicia faba*, L.) seed protein hydrolysates and fortified apple juice. *Food Chemist.* 330, 127120. doi: 10.1016/j.foodchem.2020.127120

Santos-Hernández, M., Alfieri, F., Gallo, V., Miralles, B., Masi, P., Romano, A., et al. (2020). Compared digestibility of plant protein isolates by using the INFOGEST digestion protocol. *Food Res. Int.* 137, 109708. doi: 10.1016/j.foodres.2020.109708

Saricaoglu, F. T. (2020). Application of high-pressure homogenization (HPH) to modify functional, structural and rheological properties of lentil (Lens culinaris) proteins. *Int. J. Biological Macromolecules.* 144, 760–769. doi: 10.1016/j.ijbiomac.2019.11.034

Scippa, G. S., Rocco, M., Ialicicco, M., Trupiano, D., Viscosi, V., Di Michele, M., et al. (2010). The proteome of lentil (*Lens culinaris* Medik.) seeds: Discriminating between landraces. *Electrophoresis.* 31, 497–506. doi: 10.1002/elps.200900459

Sharma, S., Singh, A., and Singh, B. (2019). Characterization of *in vitro* antioxidant activity, bioactive components, and nutrient digestibility in pigeon pea (*Cajanus cajan*) as influenced by germination time and temperature. *J. Food Biochemist.* 43, 13. doi: 10.1111/jfbc.12706

Sjöö, M., and Nilsson, L. (2017). *Starch in Food: Structure, Function and Applications, 2nd Edn*. Cambridge: Woodhead Publishing Series in Food Science, Technology and Nutrition.

Smýkal, P., Vernoud, V., Blair, M. W., Soukup, A., and Thompson, R. D. (2014). The role of the testa during development and in establishment of dormancy of the legume seed. *Front. Plant Sci.* 5, 351–370. doi: 10.3389/fpls.2014.00351

Sousa, R., Portmann, R., Dubois, S., Recio, I., and Egger, L. (2020). Protein digestion of different protein sources using the INFOGEST static digestion model. *Food Res. Int.* 130, 108996. doi: 10.1016/j.foodres.2020.108996

Subedi, M., Khazaei, H., Arganosa, G., Etukudo, E., and Vandenberg, A. (2021). Genetic stability and genotype × environment interaction analysis for seed protein content and protein yield of lentil. *Crop Sci.* 61, 342–356. doi: 10.1002/csc2.20282

Sueiro, S., Hermida, M., González, M., Lois, A., and Rodríguez-Otero, J. L. (2015). A comparison of the iso and AACC methods for determining the activity of trypsin inhibitors in soybean meal. *J. Am. Oil Chemists' Society.* 92, 1391–1397. doi: 10.1007/s11746-015-2702-8

Warsame, A. O., Michael, N., O'sullivan, D. M., and Tosi, P. (2020). Identification and quantification of major faba bean seed proteins. *J. Agri. Food Chemist.* 68, 8535–8544. doi: 10.1021/acs.jafc.0c02927

Warsame, A. O., O'sullivan, D. M., and Tosi, P. (2018). Seed storage proteins of faba bean (*Vicia faba* L.): current status and prospects for genetic improvement. *J. Agri. Food Chemist.* 66, 12617–12626. doi: 10.1021/acs.jafc.8b04992

Xu, M., Jin, Z., Simsek, S., Hall, C., Rao, J., and Chen, B. (2019). Effect of germination on the chemical composition, thermal, pasting, and moisture sorption properties of flours from chickpea, lentil, and yellow pea. *Food Chemist.* 295, 579–587. doi: 10.1016/j.foodchem.2019.05.167

Zahir, M., Fogliano, V., and Capuano, E. (2020). Effect of soybean processing on cell wall porosity and protein digestibility. *Food Funct.* 11, 285–296. doi: 10.1039/C9FO02167A

Zahir, M., Fogliano, V., and Capuano, E. (2021). Soybean germination limits the role of cell wall integrity in controlling protein physicochemical changes during cooking and improves protein digestibility. *Food Res. Int.* 143, 110254. doi: 10.1016/j.foodres.2021.110254

Zhao, W., Xue, S., Yu, Z., Ding, L., Li, J., and Liu, J. (2019). Novel ACE inhibitors derived from soybean proteins using *in silico* and *in vitro* studies. *J Food Biochemist.* 43, e12975. doi: 10.1111/jfbc.12975

Zou, Y., and Hou, X. (2017). Effects of germination conditions on peptides accumulation in soybean (*Glycine max* L. *Merr.). Acta. Alimentaria.* 46, 346–354. doi: 10.1556/066.2017.46.3.10

Zurita, J., Díaz-Rubio, M. E., and Saura-Calixto, F. (2012). Improved procedure to determine non-extractable polymeric proanthocyanidins in plant foods. *Int. J. Food Sci. Nutr.* 63, 936–939. doi: 10.3109/09637486.2012.681634

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Bautista-Expósito, Vandenberg, Peñas, Frias and Martínez-Villaluenga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 December 2021
doi: 10.3389/fpls.2021.785221

Transient Nutrient Deficiencies in Pea: Consequences on Nutrient Uptake, Remobilization, and Seed Quality

*Cécile Jacques[1], Marion Forest[1], Vincent Durey[1], Christophe Salon[1], Alain Ourry[2] and Marion Prudent[1]**

[1] Agroécologie, AgroSup Dijon, INRAE, Université de Bourgogne, University Bourgogne Franche-Comté, Dijon, France, [2] UMR 950 Ecophysiologie Végétale, Agronomie et Nutritions N, C, S, INRAE, Normandie Université, UNICAEN, Caen, France

Legume plants, such as peas, are of significant nutritional interest for both humans and animals. However, plant nutrition and thus, seed composition, depends on soil mineral nutrient availability. Understanding the impact of their deprivation on the plant mineral nutrient content, net uptake, and remobilization is of key importance but remains complex as the elements of the plant ionome are linked in intricate networks, one element deprivation impacting uptake and remobilization of other nutrients. To get a better insight into pea mineral nutrition, the transitory deprivations of 13 mineral nutrients were imposed during the vegetative growth phase. Thereafter, plants were grown under optimal mineral conditions until physiological maturity. Plant nutritional status and seed quality impacts caused by the deprivations were characterized using measurement of mineral nutrient concentration and plant biomass allocation. Our results highlight: (i) the preferential allocation of dry weight and elements to shoots at the expense of the roots under non-limiting conditions, and more particularly to the tendrils in comparison to the other shoot organs, (ii) the positive and/or negative impact of one mineral nutrient deprivation on other elements of the ionome, (iii) four different remobilization strategies for eight mineral nutrients, and (iv) possible strategies to improve seed quality *via* fine control of fertilization during a period of mineral nutrient deficiency.

Keywords: grain legume, ionome, nutrient deficiency, nutrient interactions, seed mineral composition, agroecology

OPEN ACCESS

Edited by:
Petr Smýkal,
Palacký University Olomouc, Czechia

Reviewed by:
Clarice J. Coyne,
United States Department of Agriculture (USDA), United States
Szilvia Veres,
University of Debrecen, Hungary

****Correspondence:***
Marion Prudent
marion.prudent@inrae.fr

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: *28 September 2021*
Accepted: *16 November 2021*
Published: *23 December 2021*

Citation:
Jacques C, Forest M, Durey V, Salon C, Ourry A and Prudent M (2021) Transient Nutrient Deficiencies in Pea: Consequences on Nutrient Uptake, Remobilization, and Seed Quality. Front. Plant Sci. 12:785221. doi: 10.3389/fpls.2021.785221

INTRODUCTION

Plants are sessile organisms that must constantly adapt to fluctuating local conditions. Their growth and development depend on water and mineral nutrient availability around the root system that can be modulated in time and space by soil physicochemical properties, such as rhizospheric soil pH (Verma and Minhas, 1987) that results from interactions between the chemical form of the element, root rhizodeposition, and the activity of microbial communities. Abiotic factors including extreme temperatures, anaerobic conditions, or low water availability can also impair the mineral nutrient uptake by the plant. For example, soil moisture influences mineral nutrient mobility in the soil solution (Wiersum, 1958; Barber, 1962). The decrease of nutrient mobility in the soil limits

the movement of elements from the soil solution to the rhizosphere and thus the availability of the mineral nutrients for the plant. The plant element contents have been defined as the plant ionome that can be subdivided into two categories, essential elements (or nutrients) and beneficial elements. Essential elements constitute the functional ionome and correspond to those that are needed to complete the plant lifecycle, are not replaceable by any other element to ensure a biological function, and are directly involved in plant metabolism (Arnon and Stout, 1939). Essential elements are divided into two groups, macro-nutrients, and micro-nutrients. Macro-nutrients include carbon (C), hydrogen (H), oxygen (O), nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), sulfur (S), and magnesium (Mg), while micro-nutrients include iron (Fe), copper (Cu), zinc (Zn), manganese (Mn), molybdenum (Mo), bore (B), chlorine (Cl), and nickel (Ni). These two groups of nutrients differ in the quantity needed for optimal plant growth. Indeed, macro-nutrient needs for plants represent more than 0.1% of the plant dry weight while for micro-nutrients, this represents less than 0.1% of the plant dry weight (Kirkby, 2012; Maillard et al., 2015). Beneficial elements comprise vanadium (V), cobalt (Co), or sodium (Na), and present positive effects on growth under particular growth conditions and differ among species (Kirkby, 2012; Maillard et al., 2015).

The uptake of these mineral nutrients can be modulated either *via*: (i) mechanisms triggering the root system morphological adaptations, (ii) regulations of the element transporter activities of the roots or storage of elements within plants, or (iii) rhizosphere modifications. For example, some nutrient deficiencies induce root system plasticity to improve soil prospection (López-Bucio et al., 2003; Gruber et al., 2013; Giehl and von Wiren, 2014). For instance, P or Fe deficiency increases root hair density in *Arabidopsis thaliana* (Ma et al., 2001; López-Bucio et al., 2003). N deficiency may increase primary root and lateral root length (López-Bucio et al., 2003). Mineral nutrient uptake is controlled by transporter activities which select the elements necessary for optimal plant growth, depending on their availability (Gojon et al., 2009; Tejada-Jiménez et al., 2009; Llamas et al., 2011). Soil mineral nutrients absorbed by roots can then be stored or remobilized in different plant compartments, to optimize nutrient availability for growing tissues. In this way, their storage duration during the vegetative stage depends on the need and on the availability of mineral nutrients to support the growth of new tissues (Malagoli, 2005; Abdallah et al., 2010; White, 2012) while during the reproductive stage, root activity and mineral nutrient uptake are reduced (Malagoli, 2005). Alternatively, storing the element in plant compartments such as vacuoles can reduce their overall potential toxicity for plants (Pittman, 2005; Millaleo et al., 2010).

Uptake, storage, and remobilization, all of which depend on the species, nutrient availability, and environmental conditions, are three important processes to ensure the optimal growth of plants and thus the yield and quality of seeds. Under mineral nutrient deficiency or during the reproductive phase, the exportation of elements from old tissues to growing tissues plays a crucial role in optimal growth and yield (Masclaux-Daubresse et al., 2008). The remobilization rate is specific to each element of the plant ionome. Macro-nutrients (except Ca) are rapidly transported by phloem in comparison to micro-nutrients (White, 2012). As such, it has been reported in three legume species (pea, white lupin, and narrow-leaved lupin), that N, P, and K present efficiency of mobilization of 60 to 90% from senescing organs to seeds whereas that of Mg, Zn, Mn, Fe, and Cu is lower, between 20 and 60 % (Hocking and Pate, 1977; Maillard et al., 2015). Moreover, Ca and Mn are minimally mobile in the phloem but remobilized from the root *via* xylem (Biddulph et al., 1959; White, 2012; Maillard et al., 2015). In *A. thaliana,* during seed filling, 48% of K is remobilized from leaves but less than 30% for Fe, P, S, Zn, and Cu, whereas Mg, Ca, and Mn are not remobilized (Waters and Grusak, 2008). The characterization of the mechanisms involved in these processes is therefore required to improve our knowledge of plant needs according to species, environmental conditions, and considered stages.

The elements that compose the plant ionome are linked together by a complex network extending from the soil to plant tissues (Lahner et al., 2003; Salt et al., 2008; Baxter, 2015). First, some mineral nutrients are taken up by common transporters, such as for S and Mo (Dudev and Lim, 2004), K and Na (Gassmann et al., 1996), and divalent metals (Pilon et al., 2009). This can induce competition between mineral nutrients for their uptake (Gassmann et al., 1996; Dudev and Lim, 2004; Alhendawi et al., 2005; Pilon et al., 2009). For instance, the divalent metal co-transporter IRT1 which has a high affinity for Fe is up-regulated by Fe deficiency and other cations (Ni^{2+}, Cu^{2+}, Mn^{2+}, Zn^{2+}) are then also increasingly taken up (Pilon et al., 2009). Second, some elements share similar roles in biological processes. Plant homeostasis results from interactions among mineral elements (Kirkby and Knight, 1977; Sorin et al., 2015). To maintain osmotic and acido-basic equilibrium under conditions of deficiency or overaccumulation, some mineral nutrients may be increasingly taken up under a non-optimal total ionic charge. For instance, a S deficiency induces a disequilibrium of negative ionic charge compensated for by an increase of NO_3^- and PO_4^{2-} uptake (Sorin et al., 2015). In contrast, an increase in the quantity of NO_3^- in plant tissues induces an increase in K^+, Ca^{2+}, Mg^{2+}, and Na^+ uptake to compensate for the negative charge (Kirkby and Knight, 1977). Biosynthetic pathways requiring various mineral nutrients can also be impacted by the deficiency of one of these elements. For instance, the pathway of Mo cofactor biosynthesis involves Cu and Zn (Kuper et al., 2004; Schwarz and Mendel, 2006; Llamas et al., 2011). Thus, Cu or Zn deficiency may have a negative impact on the Mo consumption in this pathway.

As such, due to different interactions between elements, the impact of an elementary deprivation on plant growth and yield is the result of a complex network and cannot be reduced to a simple deficiency of the missing element.

Thanks to the smaller size of the ionome dataset as compared with the transcriptomic and metabolomic dataset, recent studies of the ionome have provided new insights into the characterization of plant nutritional status (Baxter et al., 2008). The negative impacts of elementary deficiencies on seed quantity and quality have been recognized for years (McGrath and Zhao, 1996). For instance, B deficiency in canola plants negatively impacts plant yields (Grant and Bailey, 1993) while S deficiency reduces the accumulation of proteins rich in S in seeds

in rapeseed and pea plants (D'Hooghe et al., 2014; Henriet et al., 2019). However, the effects of transitory deficiencies, i.e., during a certain period of the crop cycle, on seed quality are not well established. In some cases, the deficiency in one element could serve to enhance the rate of other elements during vegetative growth, as previously reported for Fe deficiency and other metal cations (Pilon et al., 2009). So, it can be hypothesized that a higher metal cation in vegetative tissues may increase their remobilization to the seeds. Such impacts on seed composition can be beneficial for food quality under the context of human food deficiencies in Fe, Mg, Cu, and Zn throughout the world (Fan et al., 2008). Pea seeds are rich in Zn and Fe and can provide a solution to the most important micro-nutrient deficiencies (Amarakoon et al., 2012). Thus, in a species whose consumption is recommended by nutritionists, the characterization of seed mineral content after transient nutrient deficiencies could help in targeting the appropriate fertilization.

To extend our understanding of pea nutrition and response to mineral nutrient deprivation, several issues have been examined in this article. Firstly, elements were classified in terms of their quantity needed for plant optimal growth and their allocation into four different organs of the pea plant (roots, stems, stipules, tendrils). Secondly, we analyzed the ionome composition under nutrient deprivation to confirm interactions observed in other species and to characterize the specific interactions related to the pea plant. Moreover, this characterization allowed us to obtain an ionomic imprint of the different deprivations. Thirdly, an analysis of mature seeds enabled us to identify the effects of each transient deprivation on yield and seed quality at physiological maturity.

MATERIALS AND METHODS

Plant Growth Under Hydroponic Conditions

Pea seeds (*Pisum sativum* L. cv. Kayanne, obtained from KWS Momont, Mons-en-Pévèle, France) were calibrated, surface sterilized by exposure to 70% ethanol for 1 min, then to 0.6% sodium hypochlorite for 10 min. The seeds were then imbibed in distilled water for 2 h and pre-germinated in trays containing sand at 8% humidity for three days in the dark at 22°C, in a Fitoclima S600 germinator (Aralab, Rio de Mouro, Portugal). Seedlings were then transferred to the greenhouse in a 30 L container filled with demineralized water over 3 days, to favor radicle length and lead to homogenized seedling growth. The greenhouse environmental conditions were 21.3 ± 1.7°C during the day and 16.5 ± 1.0°C at night, with a photoperiod of 16h with artificial lighting (MACS 400W; Mazda, Dijon, France), allowing for a mean of 188.12 $\mu mol/m^2/s$. The seedlings were then transferred to 208 5 L pots (two seedlings per pot) containing a nutritive solution hereafter referred to as "Control solution, C" described in **Table 1**. After 15 days of plant growth under plethoric mineral nutrition (t0), pots were divided into 14 groups with one "Control" and 13 "Deficient" solutions (**Figure 1A**) whose compositions are detailed in **Table 1**. The 13 elementary deprived solutions used include N deficiency (N-), S deficiency (S-), P deficiency (P-), K deficiency (K-), Ca deficiency (Ca-), Cu deficiency (Cu-), Ni deficiency (Ni-), Mo deficiency (Mo-), B deficiency (B-), Zn deficiency (Zn-), Mg deficiency (Mg-), Mn deficiency (Mn-), and Fe deficiency (Fe-) (**Figure 1A**). All elements that composed each of the 14 nutrient solutions were added in excess, to avoid any potential competition for elements between the two plants of the pot. At the beginning of the plant nutrient deficiency period (t0) (**Figure 1A**), seven pots were harvested, corresponding to 14 plants. On the 201 remaining pots, a mark with attached twines identified organs formed before t0 from organs formed after t0. Three times a week, the plants were removed from their pots, weighed, and replaced in their pots. As soon as a growth cessation was observed for plants grown under a nutrient (n) deprivation, the plants in this group were harvested (td_n) along with plants from the control group (Control). The duration of the nutrient deprivation period thus depended on the nutrient that had been depleted from the solution (**Figure 1B**). When no growth cessation was observed, plants were arbitrarily harvested at td_n = 42 days after nutrient deprivation imposition. At each td_n, 6 deficient pots (12 plants) and 6 control pots (12 plants) were harvested. Their root surface was rinsed with osmotic water and plants were separated into seven samples: non-nodulated roots, stems, stipules, and tendrils formed before nutrient deprivation at t0 (hereafter referred to as stembf, stipulebf, and tendrilbf, respectively) and the stems, stipules, and tendrils developed after t0 (hereafter called stemaf, stipuleaf, and tendrilaf, respectively). At the end of the deficiency period (td_n), three pots per "Deficient" solution were filled with the non-deficient "Control" solution until the plants reached their physiological maturity stage (tm_n) (**Figure 1A**). The time to reach the physiological maturity phase depended on the deficiency considered (**Figure 1B**). At tm_n, only seeds were harvested for further analysis and counted to determine yield components.

Elemental Analysis of Plant Tissues

Following each harvest, tissue samples including non-nodulated roots, stems, stipules, tendrils formed before nutrient deprivation (at t0 and td_n) or after nutrient deprivation (at td_n), or seeds at (tm_n) were dried at 80°C for 48h, individually weighed to measure their dry weight and ground to a fine powder using MM 400 vibratory mixer mill (Retch, France). C and N concentrations were measured by Dumas procedure (Thermo electron NC 2500, Courtaboeuf, France). The other element concentrations [Sulphur (S), Phosphorus (P), Potassium (K), Calcium (Ca), Magnesium (Mg), Copper (Cu), Nickel (Ni), Manganese (Mn), Bore (B), Zinc (Zn), Iron (Fe), Vanadium (V), Cobalt (Co), Sodium (Na), Molybdenum (Mo)] were measured with a high resolution inductively coupled plasma mass spectrometer (HR ICP-MS, Element2, Thermo Fisher, Caen), following the methods previously described in Maillard et al. (2015).

Calculations and Statistical Analyses

From the elemental measurements of tissues at t0, tdn_n, and tm_n, different variables were calculated, where td_n corresponds to the sample at the end of the deprivation period for element n, and tm_n corresponds to the sample at physiological maturity for element n.

TABLE 1 | Composition of control and nutrient-deprived solutions.

	Nutrient solutions													
Salt concentration	**Control**	**N-**	**K-**	**Ca-**	**P-**	**S-**	**Mg-**	**Fe-**	**B-**	**Mn-**	**Zn-**	**Cu-**	**Mo-**	**Ni-**
$CaCO_3$	2000	3250	5000	0	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
KNO_3	1250	0	0	3750	1250	1250	1250	1250	1250	1250	1250	1250	1250	1250
$Ca(NO_3)2, 4H_2O$	1250	0	1875	0	1250	1250	1250	1250	1250	1250	1250	1250	1250	1250
KH_2PO_4	250	250	0	250	0	250	250	250	250	250	250	250	250	250
$MgSO_4$	500	500	500	500	500	0	0	500	500	500	500	500	500	500
H_3BO_3	10	10	10	10	10	10	10	10	0	10	10	10	10	10
$CuSO_4$	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0	0.7	0.7
$CoCl_2$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$NiCl_2$	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0
SiO_2	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Na_2O	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78	38.78
Na_2SO_4	0	0	0	0	0	0	0	0	0	5	3	0.7	0	0
$CaCl_2$,2 H2O	1250	1250	625	0	1250	1250	1250	1250	1250	1250	1250	1250	1250	1250
KCl	250	250	0	2000	250	0	250	250	250	250	250	250	250	250
HCl	0	0	1500	0	0	0	0	500	0	0	0	0	0	0
NaOH	0	0	0	0	0	0	0	400	0	0	0	0	0	0
KOH	0	1250	0	50	250	0	0	0	0	0	0	0	0	0
$KHSO_4$	0	0	0	0	0	0	500	0	0	0	0	0	0	0
$MnSO_4$,$4H_2O$	5	5	5	5	5	5	5	5	5	0	5	5	5	5
$ZnSO_4$,$7H_2O$	3	3	3	3	3	3	3	3	3	3	0	3	3	3
$(NH_4)6\ Mo_7O_{24}$; $4H_2O$	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0	0.7
$MgCl_2$,$6H_2O$	0	0	0	0	0	300	0	0	0	0	0	0	0	0
EDTA, (NaFe), 0.05% H_2O	200	200	200	200	200	200	200	0	200	200	200	200	200	200
H_3PO_4	0	0	250	0	0	0	0	0	0	0	0	0	0	0
NH_4NO_3	3500	0	3500	3500	3500	3500	3500	3500	3500	3500	3500	3500	3500	3500

Deprivation applied are N-, Nitrogen deprivation; K-, Potassium deprivation; Ca-, Calcium deprivation; P-, Phosphorus deprivation; S-, Sulfur deprivation; Mg-, Magnesium deprivation; Fe-, Iron deprivation; Mn-, Manganese deprivation; Mo-, Molybdenum deprivation; Zn-, Zinc deprivation; B-, Bore deprivation; Cu-, Copper deprivation; Ni-, Nickel deprivation. Salt concentration is given in $\mu mol.L^{-1}$.

The quantity (*Qty*) of element *n* in the tissue *x* (expressed in mg) was calculated as:

$$Qty_{n(x)} = [n]_x \times DW_x$$

where $[n]_x$ is the concentration of element n in the tissue x (in mg/g), and DW_x is the dry weight of the tissue x (in g).

The quantity of element *n* taken up (Up) by the plant (expressed in mg) was calculated as:

$$(Up)_n = Qty_{n(plant)}\ (td_n) - Qty_{n(plant)}\ (t0)$$

where $Qty_{n(plant)}$ (td_n) is the quantity of the element n in the plant at td_n (in mg) and $Qty_{n(plant)}$ (t0) is the quantity of the element "n" in the plant at t0 (in mg).

The percentage of element allocated to a tissue *x*, hereafter called Allocation (A) (expressed in %) was calculated as:

$$A_x = \frac{Qty_{n(x)}}{Qty_{n(plant)}} \times 100$$

where $Qty_{n(x)}$ is the quantity of the element *n* in the tissue *x* (in mg) and $Qty_{n(plant)}$ is the quantity of the element *n* in the plant (in mg).

The quantity of element *n* that was remobilized during the deficiency period (R) (expressed in mg) was calculated by mass balance (i.e., net remobilization) in the following manner:

$$Rn = Qty_{n(x)}\ (td_n) - Qty_{n(x)}\ (t0)$$

where $Qty_{n(x)}$ (td_n) is the quantity of the element *n* in the tissue *x* at td_n (in mg) and $Qty_{n(x)}$ (t0) is the quantity of the element *n* in the tissue *x* at t0 (in mg).

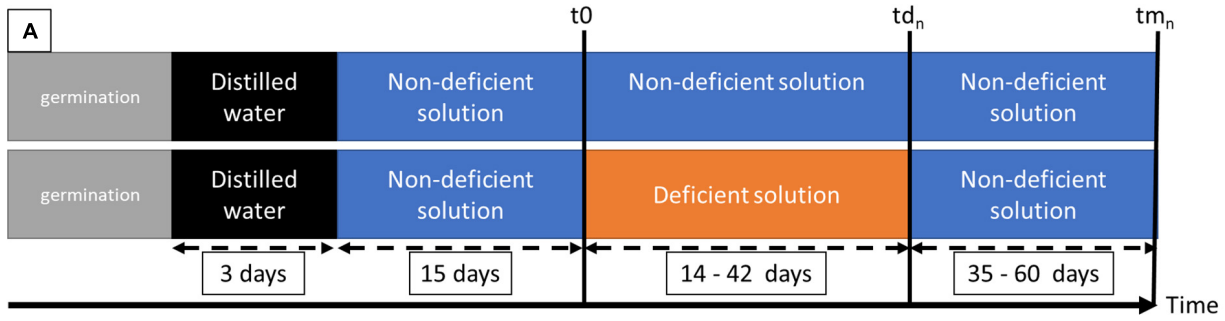
Statistical analyses were performed with R software[1]. Comparisons were performed between t0 and td_n for control and deficient plants and between control and deficient plants at td_n and tm_n. To analyze the effects of different deficiencies on measured and calculated variables, *t*-tests were carried out *via* bootstrap tests using boot.test() (package boot and MKinfer). The effect was considered significant at a *p*-value inferior to 0.05.

RESULTS

Essential and Beneficial Element Accumulation During Pea Growth

The allocation of dry matter (DW) to the shoot tissues increased progressively with time at the expense of the

[1]https://www.r-project.org/, v.3.5.2



B	Control	N-	Mg-	S-	Ca-	Mo-	P-	Fe-	K-	Cu-	Ni-	B-	Mn-	Zn-
Deprivation period (day) (leaves stage)		14 (14)	15 (15)	16 (16)	16 (16)	16 (16)	24 (20)	28 (22)	29 (22)	32 (24)	42 (29)	42 (29)	42 (29)	42 (29)
Maturity (day)	103	93	91	111	114	90	110	89	96	99	107	98	107	100

FIGURE 1 | Experimental design **(A)** showing the duration of deprivation period between t0 and td$_n$ and life cycle to maturity between t0 and tm$_n$. **(B)** Control corresponds to the solution needed for optimal growth; while solutions deficient in one element concern N, S, P, K, Ca, Mg, Cu, Ni, Mo, B, Zn, Mn, and Fe; t0 correspond to the first sample of the control plant before deprivation (7 pots of 2 plants each in control solution); td$_n$ correspond to the sample at deficiency establishment that depends on the element considered (6 pots of 2 plants each of control and deficiency solutions); tm$_n$ correspond to the sample at maturity that depends on the element considered (3 pots of 2 plants each).

allocation to the roots (**Figure 2A**). Carbon accumulation in the plant and its allocations (**Figure 2B**) to the different tissues followed approximatively the same pattern (**Supplementary Figure 1**).

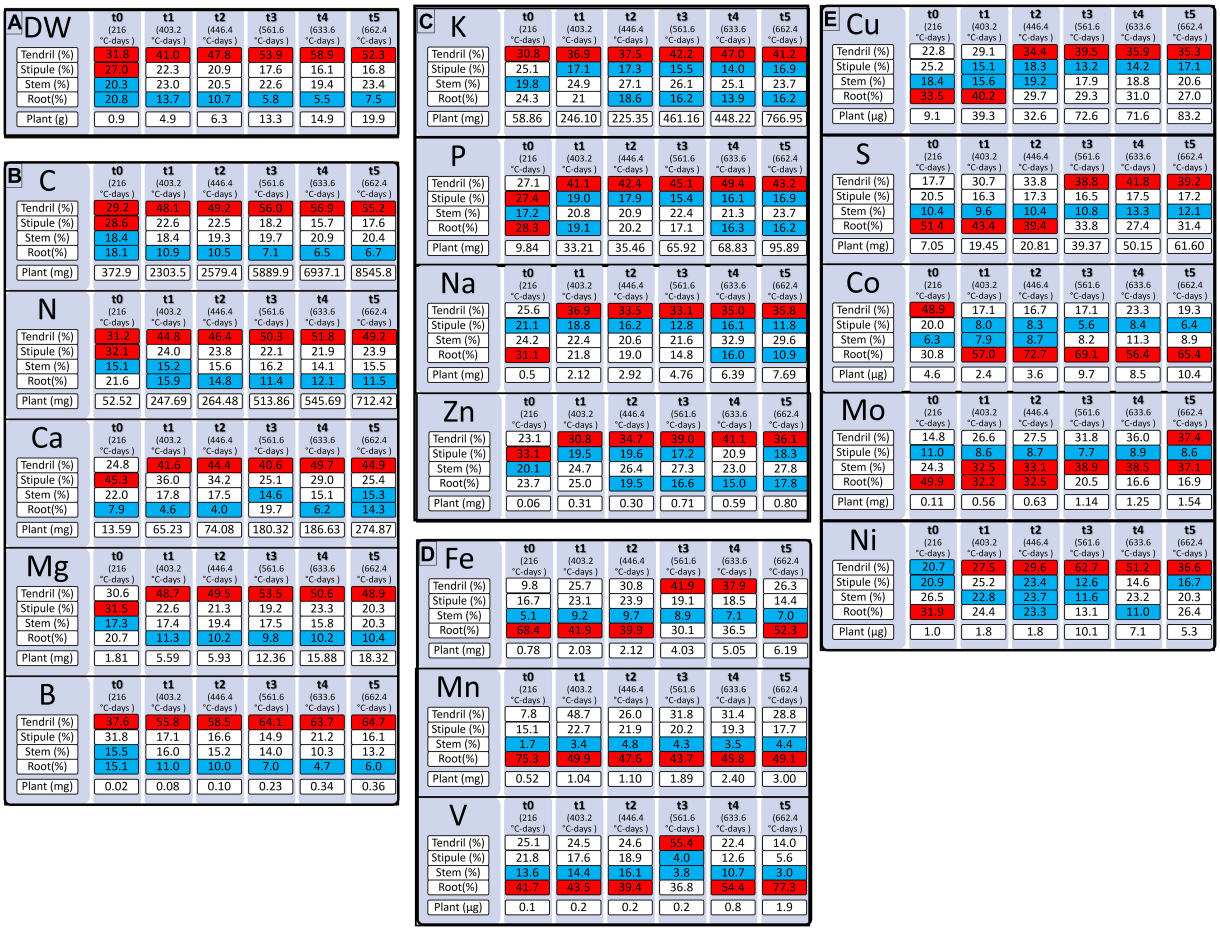
The highest macro-nutrients were N and K while amongst micro-nutrients Fe, Mn, and Mo were those with the highest accumulation in the vegetative tissues of the plant (**Figure 2**). The different elements were mostly allocated to the shoots, except for S, Fe, and Mn at t0 (15 days), Fe at t5 (32 days), Co except at t0 (15 days), and V from t4 (45 days) that were preferentially allocated to the roots. Between t1 (403°C-days) and t5 (662°C-days), all elements in the shoots that were divided into the stems, stipules, and tendrils, were mostly allocated to the tendrils except for Mo that was, for the most part, allocated to the stem. The higher quantity of element in tendrils can be explained by the preferential allocation of dry weight to this tissue during pea growth (**Figure 2A**), and even if lower concentrations of N, Mg, P, S, Ca, V, Mn, Fe, and Cu were measured in the tendrils regarding stipules (**Supplementary Figure 1** and **Supplementary Table 1**).

Finally, based on the patterns of element allocation within the plant throughout its development, three groups of elements were identified. The first group, composed of N, Ca, Mg, B, and C, was characterized by the allocation of elements preferentially to the tendrils and a lesser extent to the roots (**Figure 2B**). The second group, composed of P, K, Zn, and Na, was characterized by the allocation of elements primarily to the tendrils and to a lesser extent to the roots and/or stipules according to the age of the plant (**Figure 2C**). Finally, the third group, composed of Fe, Mn, and V, was characterized by the allocation of elements primarily to the roots and to a lesser extent to the stems (**Figure 2D**). However, S, Mo, Cu, and Co cannot be classified into one of these three groups (**Figure 2E**) as their accumulation pattern among tissues varied throughout plant growth.

Remobilization Strategies Differ Among Deprived Elements

In this study, 13 elemental deprivations were applied to pea plants at the vegetative stage. Each deprivation was continued until the cessation of growth was observed. The negative impact of macro-nutrient deprivation during vegetative growth on biomass production was observed before that of micro-element deprivation, except in the cases of Mo and Fe deficiencies (**Figure 1B**): the first deficiency was established after 14 days (14 leaves stage) of nutrient removal in the case of N deficiency, followed by Mg deficiency (15 days, 15 leaves stage), S, Ca and Mo deficiencies (16 days, 16 leaves stage), P deficiency (24 days, 20 leaves stage), Fe deficiency (28 days, 22 leaves stage), K deficiency (29 days, 22 leaves stage), and Cu deficiency (32 days, 24 leaves stage) (**Figure 1B**). For Ni, B, Mn, and Zn deprivations, no cessation of growth was observed after 42 days (29 leaves stage) of nutrient removal (**Figure 1B**). Despite the nutrient removal from the nutrient solution, significant nutrient uptakes were observed for N, P, Zn, S, Fe, B, and Ni under their elemental deprivation (**Supplementary Table 2**), probably because of the presence of trace elements in the solution. However, their uptake under deprivation was reduced by 73.3, 78.8, 76.9, 44.2, 74.6, 61.2, and 58.9%, respectively, compared with control plants. The net remobilization of these elements was thus not quantified, even if remobilization was noticeable in the cases of N and P deprivations (**Figures 3B,C**). Indeed, a significant decrease in the quantity of these two elements was observed in the shoots present before deprivation (Shootbf).

For the other nutrients, four main types of remobilizations, classified according to either a sink behavior (net increase of



A DW

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	31.8	41.0	47.8	53.9	58.9	52.3
Stipule (%)	27.0	22.3	20.9	17.6	16.1	16.8
Stem (%)	20.3	23.0	20.5	22.6	19.4	23.4
Root(%)	20.8	13.7	10.7	5.8	5.5	7.5
Plant (g)	0.9	4.9	6.3	13.3	14.9	19.9

B C

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	29.2	48.1	49.2	56.0	56.9	55.2
Stipule (%)	28.6	22.6	22.5	18.2	15.7	17.6
Stem (%)	18.4	18.4	19.3	19.7	20.9	20.4
Root(%)	18.1	10.9	10.5	7.1	6.5	6.7
Plant (mg)	372.9	2303.5	2579.4	5889.9	6937.1	8545.8

N

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	31.2	44.8	46.4	50.3	51.8	49.2
Stipule (%)	32.1	24.0	23.8	22.1	21.9	23.9
Stem (%)	15.1	15.2	15.6	16.2	14.1	15.5
Root(%)	21.6	15.9	14.8	11.4	12.1	11.5
Plant (mg)	52.52	247.69	264.48	513.86	545.69	712.42

Ca

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	24.8	41.6	44.4	40.6	49.7	44.9
Stipule (%)	45.3	36.0	34.2	25.1	29.0	25.4
Stem (%)	22.0	17.8	17.5	14.6	15.1	15.3
Root(%)	7.9	4.6	4.0	19.7	6.2	14.3
Plant (mg)	13.59	65.23	74.08	180.32	186.63	274.87

Mg

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	30.6	48.7	49.5	53.5	50.6	48.9
Stipule (%)	31.5	22.6	21.3	19.2	23.3	20.3
Stem (%)	17.3	17.4	19.4	17.5	15.8	20.3
Root(%)	20.7	11.3	10.2	9.8	10.2	10.4
Plant (mg)	1.81	5.59	5.93	12.36	15.88	18.32

B

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	37.6	55.8	58.5	64.1	63.7	64.7
Stipule (%)	31.8	17.1	16.6	14.9	21.2	16.1
Stem (%)	15.5	16.0	15.2	14.0	10.3	13.2
Root(%)	15.1	11.0	10.0	7.0	4.7	6.0
Plant (mg)	0.02	0.08	0.10	0.23	0.34	0.36

C K

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	30.8	36.9	37.5	42.2	47.0	41.2
Stipule (%)	25.1	17.1	17.3	15.5	14.0	16.9
Stem (%)	19.8	24.9	27.1	26.1	25.1	23.7
Root(%)	24.3	21	18.6	16.2	13.9	16.2
Plant (mg)	58.86	246.10	225.35	461.16	448.22	766.95

P

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	27.1	41.1	42.4	45.1	49.4	43.2
Stipule (%)	27.4	19.0	17.9	15.4	16.1	16.9
Stem (%)	17.2	20.8	20.9	22.4	21.3	23.7
Root(%)	28.3	19.1	20.2	17.1	16.3	16.2
Plant (mg)	9.84	33.21	35.46	65.92	68.83	95.89

Na

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	25.6	36.9	33.5	33.1	35.0	35.8
Stipule (%)	21.1	18.8	16.2	12.8	16.1	11.8
Stem (%)	24.2	22.4	20.6	21.6	32.9	29.6
Root(%)	31.1	21.8	19.0	14.8	16.0	10.9
Plant (mg)	0.5	2.12	2.92	4.76	6.39	7.69

Zn

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	23.1	30.8	34.7	39.0	41.1	36.1
Stipule (%)	33.1	19.5	19.6	17.2	20.9	18.3
Stem (%)	20.1	24.7	26.4	27.3	23.0	27.8
Root(%)	23.7	25.0	19.5	16.6	15.0	17.8
Plant (mg)	0.06	0.31	0.30	0.71	0.59	0.80

D Fe

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	9.8	25.7	30.8	41.9	37.9	26.3
Stipule (%)	16.7	23.1	23.9	19.1	18.5	14.4
Stem (%)	5.1	9.2	9.7	8.9	7.1	7.0
Root(%)	68.4	41.9	39.9	30.1	36.5	52.3
Plant (mg)	0.78	2.03	2.12	4.03	5.05	6.19

Mn

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	7.8	48.7	26.0	31.8	31.4	28.8
Stipule (%)	15.1	22.7	21.9	20.2	19.3	17.7
Stem (%)	1.7	3.4	4.8	4.3	3.5	4.4
Root(%)	75.3	49.9	47.6	43.7	45.8	49.1
Plant (mg)	0.52	1.04	1.10	1.89	2.40	3.00

V

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	25.1	24.5	24.6	55.4	22.4	14.0
Stipule (%)	21.8	17.6	18.9	4.0	12.6	5.6
Stem (%)	13.6	14.4	16.1	3.8	10.7	3.0
Root(%)	41.7	43.5	39.4	36.8	54.4	77.3
Plant (µg)	0.1	0.2	0.2	0.2	0.8	1.9

E Cu

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	22.8	29.1	34.4	39.5	35.9	35.3
Stipule (%)	25.2	15.1	18.3	13.2	14.2	17.1
Stem (%)	18.4	15.6	19.2	17.9	18.8	20.6
Root(%)	33.5	40.2	29.7	29.3	31.0	27.0
Plant (µg)	9.1	39.3	32.6	72.6	71.6	83.2

S

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	17.7	30.7	33.8	38.8	41.8	39.2
Stipule (%)	20.5	16.3	17.3	16.5	17.5	17.2
Stem (%)	10.4	9.6	10.4	10.8	13.3	12.1
Root(%)	51.4	43.4	39.4	33.8	27.4	31.4
Plant (mg)	7.05	19.45	20.81	39.37	50.15	61.60

Co

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	48.9	17.1	16.7	17.1	23.3	19.3
Stipule (%)	20.0	8.0	8.3	5.6	8.4	6.4
Stem (%)	6.3	7.9	8.7	8.2	11.3	8.9
Root(%)	30.8	57.0	72.7	69.1	56.4	65.4
Plant (µg)	4.6	2.4	3.6	9.7	8.5	10.4

Mo

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	14.8	26.6	27.5	31.8	36.0	37.4
Stipule (%)	11.0	8.6	8.7	7.7	8.9	8.6
Stem (%)	24.3	32.5	33.1	38.9	38.5	37.1
Root(%)	49.9	32.2	32.5	20.5	16.6	16.9
Plant (mg)	0.11	0.56	0.63	1.14	1.25	1.54

Ni

	t0 (216 °C-days)	t1 (403.2 °C-days)	t2 (446.4 °C-days)	t3 (561.6 °C-days)	t4 (633.6 °C-days)	t5 (662.4 °C-days)
Tendril (%)	20.7	27.5	29.6	62.7	51.2	36.6
Stipule (%)	20.9	25.2	23.4	12.6	14.6	16.7
Stem (%)	26.5	22.8	23.7	11.6	23.2	20.3
Root(%)	31.9	24.4	23.3	13.1	11.0	26.4
Plant (µg)	1.0	1.8	1.8	10.1	7.1	5.3

FIGURE 2 | Accumulation of dry weight [DW, **(A)**] and quantity of mineral elements **(B–E)** during pea vegetative growth from 216 to 662.4 degree-days (°C-days). Mineral elements are grouped by the quantity needed for pea plant growth. Mineral elements are N, Nitrogen; K, Potassium; Ca, Calcium, P, Phosphorus; S, Sulfur; Mg, Magnesium; Fe, Iron; Mn, Manganese; Mo, Molybdenum; Na, Sodium; Zn, Zinc; B, Bore; Cu, Copper; Ni, Nickel; V, Vanadium; Co, Cobalt. Tendril, Stipule, Stem, Root (%) correspond to the percentages of the element allocated to the organ concerning the total element quantity in the plant. "Plant" refers to the total quantity of an element in the plant in mg or µg according to the element considered. Red boxes correspond to tissues where the element is allocated in the highest quantity and blue boxes correspond to tissues where the element is the least allocated. Values (n varied from 4 to 5).

element content) or a source behavior (net decrease of element content), were observed (**Figure 3**). K was the only element remobilized from both root and shoot grown before deprivation to the shoot grown during deprivation (**Figure 3A**). Moreover, the percent of remobilization of K measured in shoot growth before deprivation was higher (91.1%) than the percent of remobilization of K measured in the root (58.9%). N and Cu were remobilized from the shoots present before deprivation to both shoots growing under deprivation and to the roots, although this measurement was biased by a residual N uptake that may have occurred during N deprivation (**Figure 3C**). Under Cu deprivation, a higher quantity of Cu was remobilized from shoots present before deprivation to shoots growing under deprivation (78%) compared with the Cu remobilized from roots to shoots growing under deprivation (22%). On the other hand, Mg and P were remobilized from the shoots present before deprivation to shoots growing under deprivation only, but not to the roots (**Figure 3D**). In addition, the percent of remobilized P is higher (79.1%) than the percent of remobilized Mg (42.8%), with approximately twice as much percent of P as the percent of Mg removed from shoots present before deprivation to shoots growing under deprivation. The last type of remobilization concerned the Ca, Mo, and Mn that were remobilized from roots to shoots with, respectively, 83, 66, and 84% of Ca, Mo, and Mn remobilized (**Figure 3B**).

Essential Element Interactions During Nutrient Deficiencies

To obtain better insight into the interactions among elements, their quantifications were performed at the end of each nutrient deprivation period and summarized in **Figure 4**. Whatever the nutrient deprivation, a lower accumulation of the deprived element was observed, except in the case of Ni deprivation, in which Ni accumulation did not change (**Figure 4** and **Supplementary Table 3**). For example, nitrate

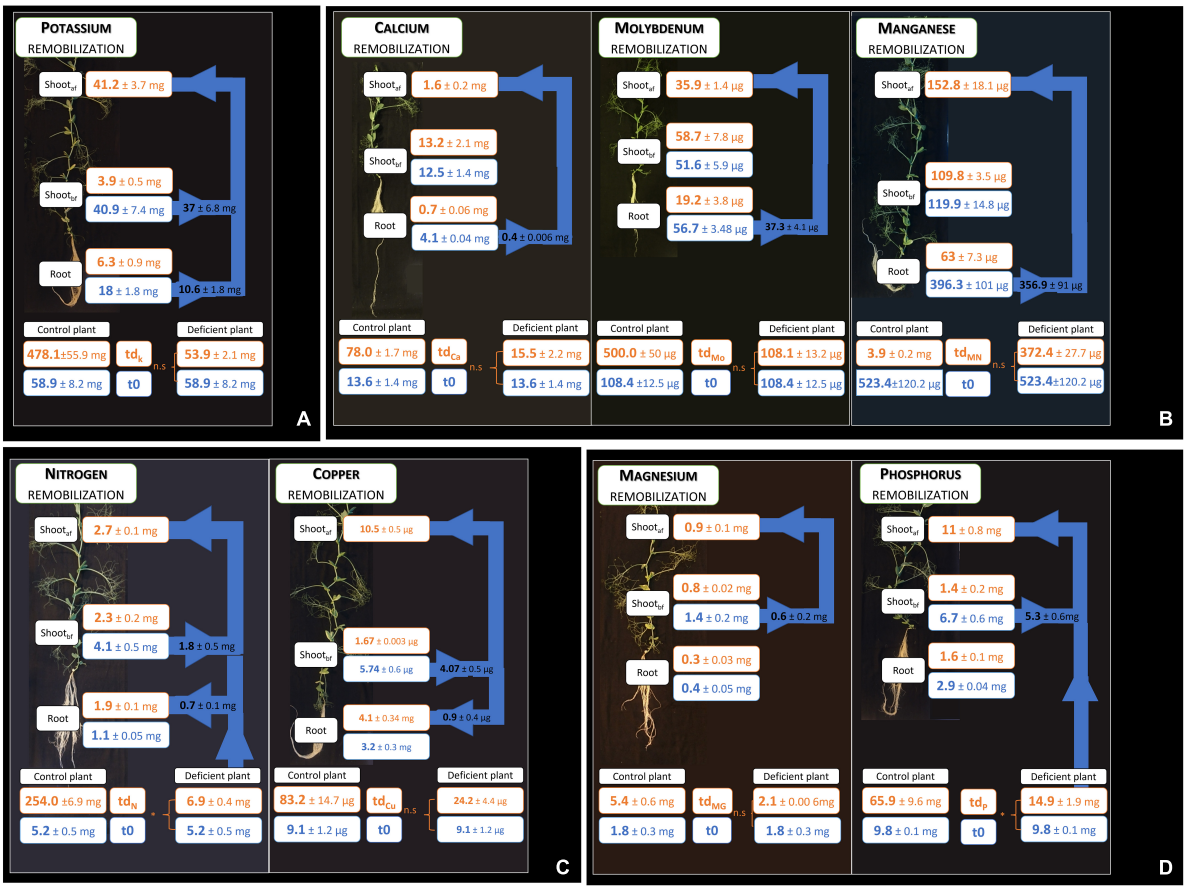


FIGURE 3 | Remobilization of elements after element deprivation. The compartments consider root and shoot growth before deprivation (shoot$_{bf}$) and shoot growth under deprivation (shoot$_{af}$); a non-significant uptake during deprivation is characterized by n.s, whereas an uptake during deprivation is characterized by *. Arrows represent the quantity of element remobilized from grown tissues to growing tissues. Four groups of remobilizations are observed in the pea according to sink and source organs **(A–D)**. Values ± SD (n varied from 4 to 5).

deprivation significantly reduced N quantity (by about 73%) in the deprived plant in comparison to the control plant. These results thus confirm the establishment of nutrient deficiency except in the case of Ni.

Overall, the deprivation of one nutrient affects the ionome composition of the plant largely (**Figure 4**) except for Ni deprivation. Mn, Zn, Cu, and B deprivation negatively impacted or have no impact on the uptake of other elements, while the other deprivations revealed both positive, negative, or neutral impacts on the accumulation of other elements (**Figure 4**). Some nutrients appeared to be reciprocally linked, i.e., deprivation of element "a" decreased the quantity of element "b" and inversely. For instance, N deprivation reduced K, Ca, and Mn accumulation in the plant, and K, Ca and Mn deprivation likewise reduced N accumulation in the plant, in comparison with the control plants. The same type of link was observed between S and (P, K, Ca, B), between P and (S, K, Ca, Mn, Fe), between K and (N, S, P, B, Mn, Zn), between Ca and (N, S, P, P, B, Cu), between Mo and (B, Zn), between Fe and P, between Zn and (P, K, Mo), between B and (S, K, Zn), between Cu and Ca, and between Mn and (N, P, K).

In contrast, some elements were antagonistically linked, i.e., the deprivation of element "a" reduced the quantity of element "b" but the deprivation of element "b" increased the quantity of element "a" in comparison to control plants. This was the case for Fe and S, B and Mg, Cu and Fe.

On the other hand, some elements were non-reciprocally linked, i.e., the deprivation of element "a" modifies the quantity of element "b", but the deprivation of element "b" does not modify the quantity of element "a". This was the case of N with P, B, Fe, Mo; K with Mo; Ca with K, Mg, Mo; P with Cu; S with Cu, Mn, Mg, N and Mo; Fe with Ca, Mn, and Mg; Mn with Ca and Mg; Mo with Fe and Mn; Zn with S, Fe, Ca, Cu, Mn, and Mg; B with Cu, Mn, and Mo.

Transitory Nutrient Deficiency Impacted Seed Yield Components and Seed Mineral Composition

In our experiment, the transitory elemental deprivation period to which the pea plants were subjected during the vegetative stage was followed by a recovery period until plant physiological

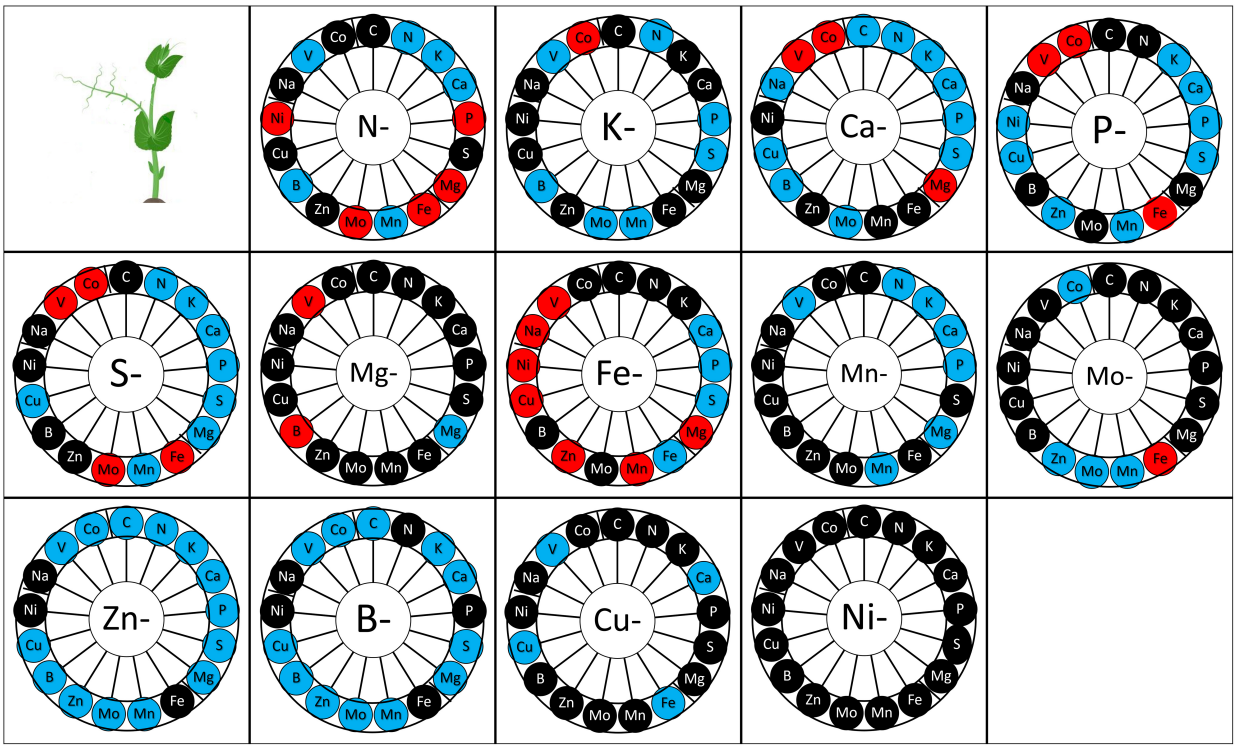


FIGURE 4 | Mineral element quantity in the plant according to the elementary deprivation considered. Deprivations include N-, Nitrogen deprivation; K-, Potassium deprivation; Ca-, Calcium deprivation; P-, Phosphorus deprivation; S-, Sulfur deprivation; Mg-, Magnesium deprivation; Fe-, Iron deprivation; Mn-, Manganese deprivation; Mo-, Molybdenum deprivation; Zn-, Zinc deprivation; B-, Bore deprivation; Cu-, Copper deprivation; Ni-, Nickel deprivation. Elements measured are C, Carbon; N, Nitrogen; K, Potassium; Ca, Calcium; P, Phosphorus; S, Sulfur; Mg, Magnesium; Fe, Iron; Mn, Manganese; Mo, Molybdenum; Na, Sodium; Zn, Zinc; B, Bore; Cu, Copper; Ni, Nickel; V, Vanadium; Co, Cobalt. Elements in red are those whose quantity increased as a result of deprivation, elements in blue are those whose quantity decreased and elements in black are those whose quantity was maintained after deprivation. Values (n varied from 4 to 5).

maturity. After ten transitory deprivations during vegetative growth, yield components such as seed dry weight, seed number, and/or weight per seed were significantly affected (**Figure 5**). A lower seed dry weight was observed for N, K, Ca, S, Mg, Fe, Mn, Mo, Zn, and Cu deficiencies (**Figure 5A**). This lower seed dry weight was linked to a lower seed number, except for Zn and Cu deprivations (**Figure 5B**) for which the lower seed dry weight was linked to a lower weight per seed (**Figure 5C**). Moreover, the more deleterious deprivation effects on yield components were found for Ca, Mg, Fe, Mn with total seed weight and a seed number reduced by at least 50% compared with the control plants.

At physiological maturity, the concentration of each element was measured in seeds (**Figure 6** and **Supplementary Table 4**). B and Ni deprivation did not impact the mineral composition of seeds. Conversely, a transient Zn or Mn deprivation increased some element concentrations in mature seeds in comparison to control plants. Interestingly, only four transitory deprivations (N, Fe, Mn, and Zn) impacted the concentration of the deprived element at maturity (**Figure 6**). In some cases, transitory deprivation enhanced the concentration of the deprived element in seeds. It was the case for N deprivation and Fe deprivation. Conversely, transitory deprivation of Zn induced a decrease in Zn quantity in seeds at maturity in comparison to control plants. The transitory deprivation that most modified the mineral composition of seeds was the Fe deprivation, increasing the N, Mg, Cu, and Fe concentrations and decreasing Mo and Ca concentrations.

DISCUSSION

In our study, the 13 mineral nutrient deprivations applied to pea plants allowed us to have a better knowledge about their uptake, their allocations in plant compartments, their remobilization between source and sink tissues, and the interactions among elements for these processes. Moreover, the effect of mineral nutrient deprivation on yield components (seed number and seed dry weight) and seed quality (mineral composition), have been characterized.

Nutrient Allocation Within Plant Under Non-stressful Conditions

Under non-limiting nutrient conditions and more broadly non-stressful conditions, the majority of plant biomass and carbon was allocated to shoot organs and notably to tendrils for this pea *afila* variety (**Figure 2**). This is congruent with non-limiting conditions concerning water and nutrients that enable the

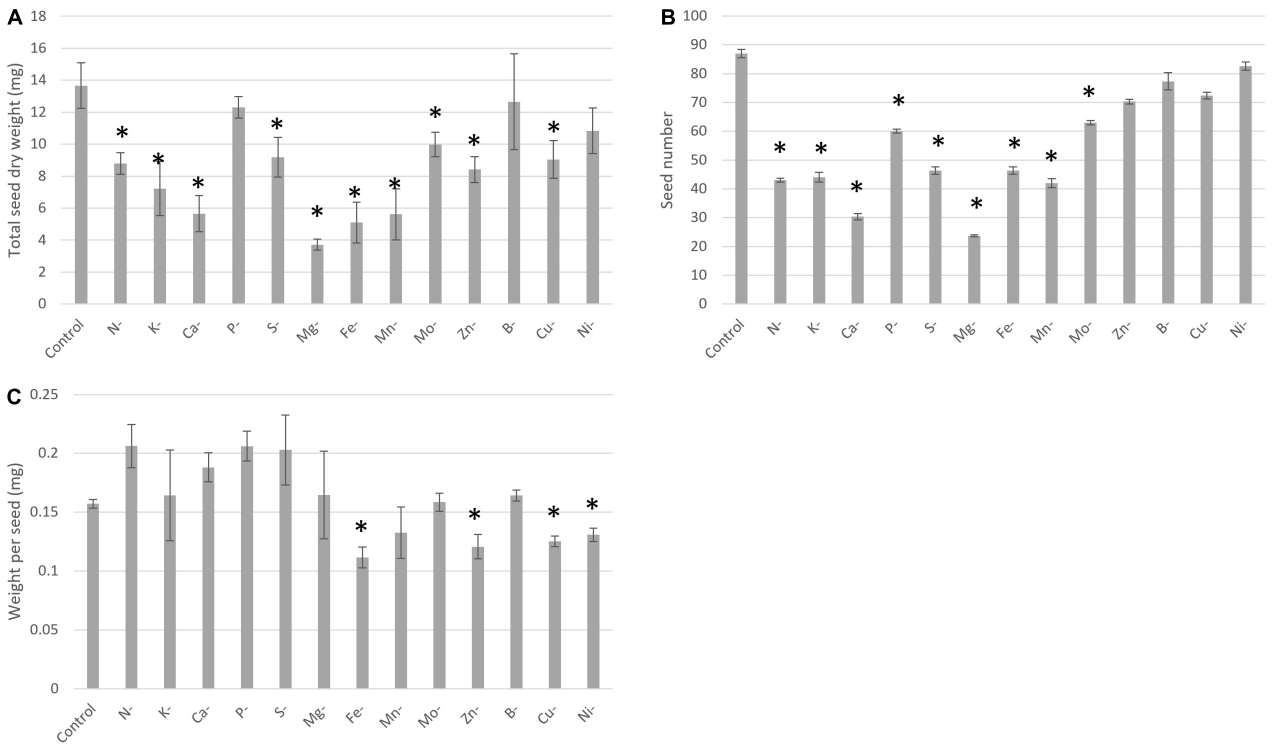


FIGURE 5 | Yield components under control and deprived conditions, seed dry weight **(A)**, seed number **(B)**, and one seed dry weight **(C)**. Conditions are: Control; N-, Nitrogen deprivation; K-, Potassium deprivation; Ca-, Calcium deprivation; P-, Phosphorus deprivation; S-, Sulfur deprivation; Mg-, Magnesium deprivation; Fe-, Iron deprivation; Mn-, Manganese deprivation; Mo-, Molybdenum deprivation; Zn-, Zinc deprivation; B-, Bore deprivation; Cu-, Copper deprivation; Ni-, Nickel deprivation. All deprivations occurred during vegetative growth for a duration that depended on the element considered. *Corresponds to deprivations that present a significant difference with the Control. Values ± SD (n = 4).

necessary nutrition for plant growth. Indeed, under non-limiting conditions, root growth is not privileged since all necessary nutrients and water are available for the root system (Ma et al., 2001; López-Bucio et al., 2003). Moreover, a higher allocation to tendrils compared with stipules may be correlated with the low surface of stipules. This can be explained by a higher xylem flux in tendrils that present a higher exchange surface than stipules, *via* a higher transpiratory rate in tendrils (Salon and Munier-Jolain, 2010). For the majority of elements (N, Mg, P, S, Ca, Fe, and Cu), the higher quantity of elements in tendrils regarding stipules was linked to higher biomass despite lower concentrations. In this way, several elements were accumulated preferentially in the tendrils, notably N and that could be linked to C accumulation in biomass production. However, some elements (S, Fe, Mn, Co, and V) were preferentially allocated to the roots at different growing stages. This was previously observed in other pea cultivars where higher proportions of Cu and Cd (Landberg and Greger, 1994), Mn and Fe (Nenova, 2008) were allocated to the root system. Another factor that may explain element partitioning in different organs is the relationship with others. Indeed, elements are linked by common physiological processes, either *via* a common transporter (Gassmann et al., 1996; Dudev and Lim, 2004; Pilon et al., 2009), or osmotic and acido-basic equilibrium maintenance (Kirkby and Knight, 1977; Sorin et al., 2015), or their role in a common metabolic pathway (Kuper et al., 2004; Schwarz and Mendel, 2006). For instance, Fe and Mn are transported by the same transporters (IRT1, Iron-Responsive Transporter 1; Ys-YSL, Yellow-Stripe 1-like; ZIP, ZRT-IRT_like Proteins) and present in the same tissue for the most (Pilon et al., 2009). In addition, K and Na present the same pattern of accumulation, an observation that may be explained by their role in potential osmotic maintenance and by some common transporter (Schachtman and Liu, 1999). Thus, allocation to different organs seems to depend on the biochemical functions of a group of nutrients and their transport, which may be common. In our study, plants were not inoculated with rhizobia and were not able to form any nodule in the root system. This leads to an N uptake that is only based on a mineral uptake of N. Because N fixation can induce changes in the pattern of nutrient uptake, promoting an alkaline nutrient uptake unlike acidic uptake pattern observed in nitrate-supplied plants (Van Beusichem, 1981), we can suppose that our results could not be directly transposed to plants grown without mineral N and under dinitrogen-fixing conditions.

Mineral allocation on different plant tissues depends on the quantity of the element taken up and on its storage in the plant (Malagoli, 2005; Gojon et al., 2009; Pilon et al., 2009; White, 2012). To enable optimal growth and development during the reproductive phase or under abiotic stresses, mineral elements are then remobilized from old tissues to growing

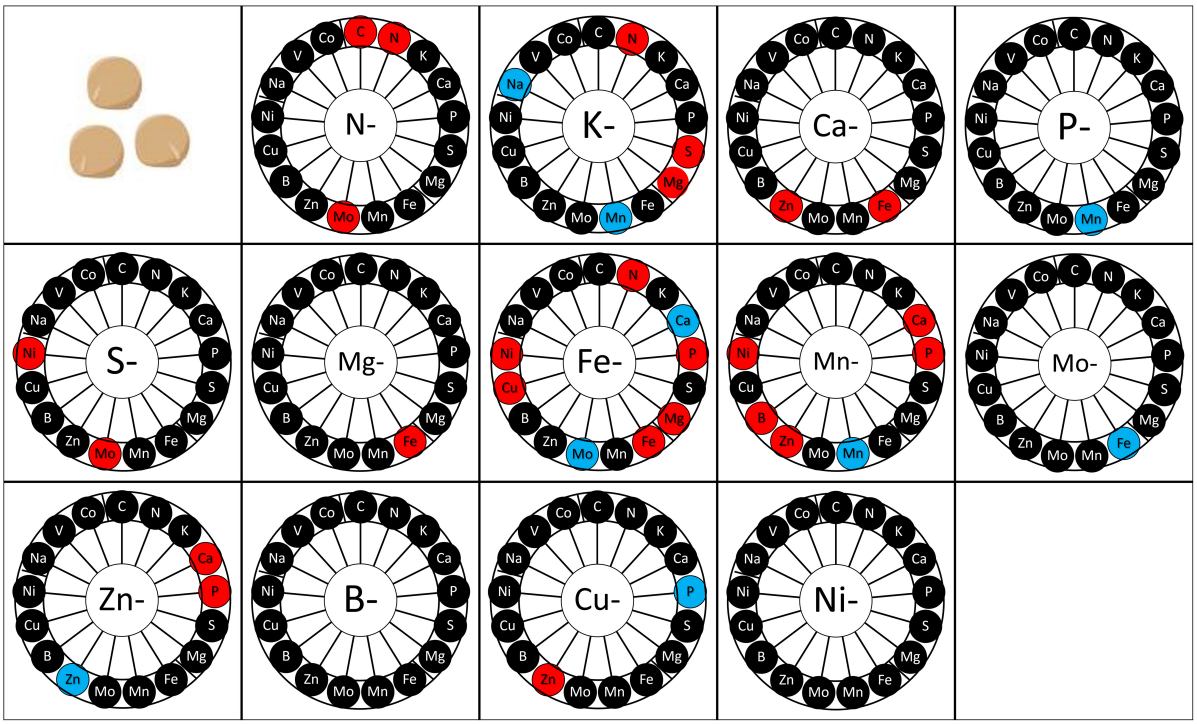


FIGURE 6 | Mineral element concentration in seeds according to the elementary deprivation considered. Deprivations included N-, Nitrogen deprivation; K-, Potassium deprivation; Ca-, Calcium deprivation; P-, Phosphorus deprivation; S-, Sulfur deprivation; Mg-, Magnesium deprivation; Fe-, Iron deprivation; Mn-, Manganese deprivation; Mo-, Molybdenum deprivation; Zn-, Zinc deprivation; B-, Bore deprivation; Cu-, Copper deprivation; Ni-, Nickel deprivation. Elements measured are C, Carbon; N, Nitrogen; K, Potassium; Ca, Calcium; P, Phosphorus; S, Sulfur; Mg, Magnesium; Fe, Iron; Mn, Manganese; Mo, Molybdenum; Na, Sodium; Zn, Zinc; B, Bore; Cu, Copper; Ni, Nickel; V, Vanadium; Co, Cobalt. Elements in red are elements whose concentration increased after deprivation, elements in blue are elements whose concentration decreased after deprivation, and elements in black are those whose concentration was maintained after deprivation. Values (n = 4).

tissues (Malagoli, 2005; Masclaux-Daubresse et al., 2008). Indeed, some stresses like water deficiency reduce mineral element availabilities in the soil for plants (Wiersum, 1958). Thus, the remobilization of mineral elements under limiting availability can allow plants to maintain their growth during a given period (Masclaux-Daubresse et al., 2008).

Four Patterns of Remobilization Identified Under Nutrient Deficiency

In our study, in the case of seven nutrient deprivations, a significant uptake was observed after the control solution was replaced by an elementary deprived solution. This observation applies to the N, P, Zn, S, Fe, B, and Ni deficiencies. In pea plants, four patterns of remobilization that depend on the source and sink organs were highlighted (**Figure 3**). Indeed, the remobilization of a deprived element may start in an old shoot and/or root and move to a young shoot and/or root. This first pattern of remobilization concerned only K, with remobilization from all organs present before K deprivation (**Figure 3A**). Indeed, K was remobilized from roots and old shoots to young shoots. In the second and third patterns, elements were remobilized from old shoots to young shoots and/or roots (**Figures 3B,C**). These two types of remobilizations concerned N, Cu, Mg, and P. It is interesting to observe the same pattern of remobilization in the cases of N and Cu in the pea plant, which could be linked to the high induction of Cu remobilization by an N deficiency (Hill and Stobbe, 1978). Moreover, as in the case of N remobilization, there is a correlation between leaf senescence and Cu remobilization (Himelblau and Amasino, 2001). Thus, under N and Cu deprivation, these two elements would be, respectively, remobilized from old shoots *via* the senescence process to meet the needs of growing shoots and roots.

On the other hand, Mg and P were remobilized from old shoots to growing shoots, but not to the roots (**Figure 3C**). However, Mg deficiency was established faster than P deficiency, after 15 and 24 days, respectively. This difference could be explained by the lower efficiency of mobilization of Mg in comparison to the efficiency of mobilization of P, yet observed for legume plants (Hocking and Pate, 1977). Lastly, the fourth strategy of remobilization that we highlighted concerned Ca, Mn, and Mo, for which remobilization was based solely on roots, not on shoots (**Figure 3D**). It could be explained by the low phloem mobility of these elements and the potential remobilization of these elements from the root *via* the xylem (Nable and Loneragan, 1984; Dayod et al., 2010; White, 2012; Maillard et al., 2015). Indeed, this process has been observed in soybeans (*Glycine max*) and in green beans (*Phaseolus vulgaris*) for Ca

(Biddulph et al., 1959; Mauk and Noodén, 1992). Moreover, the lack of Ca and Mn mobility is not species-dependent, unlike Mo. Indeed, Mo exhibits phloem mobility in different species, like wheat (*Triticum sativum*) although a lack of mobility has been demonstrated in pea plants (White, 2012; Maillard et al., 2015). Thus, the remobilization of these Ca, Mn, and Mo is nonetheless observed from roots to shoots, which suggests a xylem transport to ensure the needs of the growing shoots (Maillard et al., 2015). In this way, the lack of Mo shoot remobilization may explain the establishment of deficiency of this micro-nutrient after the same duration as that observed in Ca deficiency establishment. Thus, the establishment of mineral element deficiency depends not only on element quantity required but also on the remobilization efficiency of this element. Thus, because of the difference in remobilization efficiency under element deprivation, a specific element deprivation can be more limiting due to its lower remobilization.

Plant Nutrition: A Complex Network of Mineral Elements

Our study highlighted interactions between elements in the pea plant under 13 mineral nutrient deprivation conditions (**Figure 4**). The uptake of an element may be antagonist or synergic with other elements, depending in some cases on the acido-basic and osmotic equilibrium. To maintain the acido-basic plant equilibrium, a reduction in the quantity of an acid element may induce a reduction in a basic element (Kirkby and Knight, 1977; Sorin et al., 2015). For instance, N deprivation induces a reduction in K and Ca uptake that could compensate for the reduction in the quantity of anion (Kirkby and Knight, 1977). Moreover, the reduction of nutrient quantity induced by an elementary deprivation may be due to the involvement of the deprived element in the biological process. Indeed, in the case of Mo cofactor biosynthesis, Zn and Cu are involved in its biosynthesis. For peas, a decrease of Mo quantity is observed under Zn deprivation (Kuper et al., 2004; Schwarz and Mendel, 2006; Llamas et al., 2011). That can be due to the lower consumption of Mo for Mo cofactor biosynthesis. On the other hand, some nutrients are characterized by antagonistic uptake; this is the case for S and Mo. A deprivation of S induces an increased uptake of Mo in the pea plant as observed for rapeseed (Maillard, 2016).

Some mineral deprivation solely induced a reduction in the quantity of another element, in addition to the nutrient-deprived. This wholly negative impact on other elements concerns four micro-nutrients studied here, namely Mn, Zn, Cu, and B. However, in the cases of other nutrient deprivation, positive and negative impacts were observed on the quantities of other elements. Among these elements, Fe deprivation was the one that resulted in the most frequent increases in the quantities of other elements. This enhancement of element uptake could be explained by a higher expression of non-specific transporter induced by deprivation. Indeed, the divalent metal co-transporter IRT1 (Iron-Regulated Transporter 1) enables the uptake of Ni, Cu, Mn, Zn, and Fe and is induced by Fe deficiency (Pilon et al., 2009). In addition, Ys-YSL (Yellow-Stripe 1-like) transports Cu, Mn, Zn and Fe, and ZIP (ZRT-IRT_like Proteins) transports Mn, Zn, and Fe (Krämer et al., 2007). Thus, under Fe deprivation, competition for uptake of other cations is reduced even if the activity of the transporter is enhanced. In this way, Ni, Cu, Mn, and Zn are more easily taken up from the soil by the root regarding non-limiting Fe condition. However, Fe deprivation induces a reduction in other elements including P, S, and Ca, all of which play important roles in plant growth and development. Indeed, S is a major component of protein with 50% of the total S quantity integrated in protein in rapeseed plant (Maathuis, 2009), while P and Ca are important for membrane stability (Maathuis, 2009), and Ca is important in the regulation of Krebs cycle and osmotic equilibrium (Maathuis, 2009).

If some relations among mineral elements seem to be generic, being shared by several species and genotypes, we need to keep in mind that strong genotypic effects and genotype × environment interactions can modulate correlations among elements, as demonstrated in Arabidopsis (Baxter et al., 2012) or maize (Stich et al., 2020).

Toward Strategies to Improve Seed Quality Through Mineral Nutrition

Transitory deficiencies have an impact on plant ionome during vegetative growth that can be conserved or not during reproductive growth. Indeed, mineral nutrient deficiencies can negatively impact seed quantity and mineral composition (McGrath and Zhao, 1996). For instance, the negative impact of S deficiency has been observed in canola plants that require an important quantity of S during its crop cycle as it induced a delay of flowering and maturity and a production of smaller and poorly filled pods (Grant and Bailey, 1993). For canola, a reduction of seed yield is also observed after B deficiency, as B represents another mineral element important for its development (Nyborg and Hoyt, 1970). Moreover, like for canola, rapeseeds need an important quantity of S and present a reduction of seed quality under S deficiency (D'Hooghe et al., 2014). The reduction of rapeseed seed quality is associated with a decrease of seed viability and a reduction of accumulation of protein-rich in S. Our study also deepens our understanding of the effects of transitory deficiency on pea seed yield and composition. Indeed, our results revealed an impact of all deprivations, except deprivation in B and Ni, during vegetative growth on seed yield and/or seed composition at maturity. In pea, a negative impact of mineral deprivation was observed for N, S, P, Ca, Mg, and Fe, Mn with a reduction of seed number that was correlated with a decrease of seed biomass for N, S, Ca, Mg, and Fe. The most important deleterious deprivation impacts on seed yield were observed for Ca, Mg, Mn, and Fe deficiencies (**Figure 5**). However, the more negative impact of these deprivations on yield regarding to N deprivation in pea could be due to different durations of recovery. Indeed, the recovery period allowed for most of the plants to offset the mineral element quantity missing at the end of the deprivation period in comparison to control plants (**Figure 6**). Only four transitory deprivations revealed an impact on the concentration of the deprived element in seeds. Two cases revealed a lower concentration after Mn and Zn

deprivation; two revealed a higher concentration, i.e., N and Fe. The lower concentration of Mn and Zn could be linked either to a lower potential of remobilization of these two elements under non-limiting conditions (Sankaran and Grusak, 2014) or/and to a lower quantity of these two elements after deprivation. We can conclude that following N and Fe deprivation, the uptake and remobilization of these elements to seeds seemed to increase. In the case of transitory N deprivation, seed N and Mo concentrations increased despite a reduction of seed quantity and biomass. So, mineral deprivation can antagonistically impact seed yield and mineral composition, i.e., produce seeds richer in N and Mo even if yields are reduced. Moreover, Fe deprivation during vegetative growth induced a higher concentration of N, Mg, P, Cu, and Fe, but also of Ni. However, the increase in Ni concentration does not represent an improvement in seed quality, nor does a decrease in Mo and Ca concentration. So, an improvement of knowledge on the ionomic imprint induced by different deprivations could enable to ensure yield and seed mineral composition thanks to appropriate cultural practice. Thus, an enhancement of nutrient uptake or remobilization after a nutrient deficiency period (*via* fine control of fertilization or bio-inoculation) would improve seed quality.

CONCLUSION

The present study deepens our understanding of mineral nutrient uptake, storage, and remobilization in the pea plant. In this species, we have been able to classify elements in order of importance: N, K, Ca, P, S, Mg, Fe, Mn, Mo, Zn, Cu, and Ni. Moreover, beneficial elements are also required in different quantities, with a higher quantity of Na required compared to V and Co. Differences in the allocation of these elements observed in two leaf tissues (stipules and tendrils) have been characterized. For eight mineral nutrients, the remobilization has been characterized into four different groups that depended on relations between sink and source organs during vegetative growth. Indeed, K was remobilized from roots and shoots to shoots, N and Cu from shoots to shoots and roots, Mg and P from shoots to shoots; and Ca, Mo, and Mn from roots to shoots. The study of the pea ionome allowed us to better understand the effects of a deficiency on the nutritional status of the plant and the time needed to the establishment of deficiency in a single cultivar of pea (cv. Kayanne) grown under non-N_2-fixing conditions. Moreover, our study of the seed ionome highlighted a possible positive impact of transitory deficiency during vegetative growth on seed quality *via* the increase in the uptake of several nutrients. A detailed understanding of deprivation impacts may enable users to fine-tune the use of fertilizers when essential for optimal plant growth in soil under conditions of simple mineral nutrient deprivation, or under sub-optimal environmental conditions such as water stress.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AO, CS, and MP conceived the project. CJ, MF, VD, and MP performed the experiments. CJ analyzed the data. CJ, MF, VD, CS, AO, and MP contributed to the writing and approved the final manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was conducted through the EAUPTIC project supported by the "Fond Unique Interministériel" (3870401/1), BPIFrance (0097244/00), the Regional Council of Burgundy (0133465/00), Dijon Metropole (2018-118-20180820), and the Fonds Européen de Développement Régional (2018-6200FEO003S01889). CJ obtained a Ph.D. grant from Université de Bourgogne.

ACKNOWLEDGMENTS

We acknowledge Sylvie Girodet for CHN analysis, the PLATIN' (Plateau d'Isotopie de Normandie) core facility for the element analysis used in this study, and more particularly, Sylain Diquelou, Fanny Leroy, and Jéhane Lepoittevin. Aurélien Doria and Anne Maillard are thanked for their advice on experimental design; Jonathan Kreplak is thanked for his help in biostatistics, the greenhouse team is thanked for taking care of the plants, and Mary Bouley is acknowledged for English correction and improvement.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.785221/full#supplementary-material

REFERENCES

Abdallah, M., Dubousset, L., Meuriot, F., Etienne, P., Avice, J.-C., and Ourry, A. (2010). Effect of mineral sulphur availability on nitrogen and sulphur uptake and remobilization during the vegetative growth of Brassica napus L. *J. Exp. Bot.* 61, 2635–2646. doi: 10.1093/jxb/erq096

Alhendawi, R. A., Kirkby, E. A., and Pilbeam, D. J. (2005). Evidence that sulfur deficiency enhances molybdenum transport in xylem sap of tomato plants. *J. Plant Nutr.* 28, 1347–1353. doi: 10.1081/PLN-200067449

Amarakoon, D., Thavarajah, D., McPhee, K., and Thavarajah, P. (2012). Iron-, zinc-, and magnesium-rich field peas (Pisum sativum L.) with naturally low phytic acid: a potential food-based solution to global micronutrient malnutrition. *J. Food Compos. Anal.* 27, 8–13. doi: 10.1016/j.jfca.2012.05.007

Arnon, D. I., and Stout, P. R. (1939). The essentiality of certain elements in minute quantity for plants with special reference to copper. *Plant Physiol.* 14, 371–375. doi: 10.1104/pp.14.2.371

Barber, S. A. (1962). *Soil Nutrient Bioavailability: A Mechanistic Approach*, 2nd Edn. New York, NY: Wiley.

Baxter, I. (2015). Should we treat the ionome as a combination of individual elements, or should we be deriving novel combined traits? *J. Exp. Bot.* 66, 2127–2131. doi: 10.1093/jxb/erv040

Baxter, I. R., Vitek, O., Lahner, B., Muthukumar, B., Borghi, M., Morrissey, J., et al. (2008). The leaf ionome as a multivariable system to detect a plant's physiological status. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12081–12086. doi: 10.1073/pnas.0804175105

Baxter, I., Hermans, C., Lahner, B., Yakubova, E., Tikhonova, M., Verbruggen, N., et al. (2012). Biodiversity of mineral nutrient and trace element accumulation in *Arabidopsis thaliana*. *PLoS One* 7:e0035121. doi: 10.1371/journal.pone.0035121

Biddulph, O., Cory, R., and Biddulph, S. (1959). Translocation of calcium in the bean plant. *Plant Physiol.* 34, 512–519.doi: 10.1104/pp.34.5.512

D'Hooghe, P., Dubousset, L., Gallardo, K., Kopriva, S., Avice, J.-C., and Trouverie, J. (2014). Evidence for proteomic and metabolic adaptations associated with alterations of seed yield and quality in sulfur-limited Brassica napus L. *Mol. Cell. Proteom.* 13, 1165–1183. doi: 10.1074/mcp.M113.034215

Dayod, M., Tyerman, S. D., Leigh, R. A., and Gilliham, M. (2010). Calcium storage in plants and the implications for calcium biofortification. *Protoplasma* 247, 215–231. doi: 10.1007/s00709-010-0182-0

Dudev, T., and Lim, C. (2004). Oxyanion selectivity in sulfate and molybdate transport proteins: an ab initio/CDM study. *J. Am. Chem. Soc.* 126, 10296–10305. doi: 10.1021/ja047951n

Fan, M.-S., Zhao, F.-J., Fairweather-Tait, S. J., Poulton, P. R., Dunham, S. J., and McGrath, S. P. (2008). Evidence of decreasing mineral density in wheat grain over the last 160 years. *J. Trace Elem. Med. Biol.* 22, 315–324. doi: 10.1016/j.jtemb.2008.07.002

Gassmann, W., Rubio, F., and Schroeder, J. I. (1996). Alkali cation selectivity of the wheat root high-affinity potassium transporter HKT1. *Plant J.* 10, 869–852. doi: 10.1046/j.1365-313x.1996.10050869.x

Giehl, R. F. H., and von Wiren, N. (2014). Root nutrient foraging. *Plant Physiol.* 166, 509–517. doi: 10.1104/pp.114.245225

Gojon, A., Nacry, P., and Davidian, J.-C. (2009). Root uptake regulation: a central process for NPS homeostasis in plants. *Curr. Opin. Plant Biol.* 12, 328–338. doi: 10.1016/j.pbi.2009.04.015

Grant, C. A., and Bailey, L. D. (1993). Fertility management in canola production. *Can. J. Plant Sci.* 73, 651–670. doi: 10.4141/cjps93-087

Gruber, B. D., Giehl, R. F. H., Friedel, S., and von Wirén, N. (2013). Plasticity of the *Arabidopsis* root system under nutrient deficiencies. *Plant Physiol.* 163, 161–179. doi: 10.1104/pp.113.218453

Henriet, C., Aimé, D., Térézol, M., Kilandamoko, A., Rossin, N., Combes-Soia, L., et al. (2019). Water stress combined with sulfur deficiency in pea affects yield components but mitigates the effect of deficiency on seed globulin composition. *J. Exp. Bot.* 70, 4287–4304. doi: 10.1093/jxb/erz114

Hill, B. D., and Stobbe, E. H. (1978). Effect of light and nutrient levels on 14C-benzoylprop ethyl metabolism and growth inhibition in wild oat (Avena fatua L.)*. *Weed Res.* 18, 223–229. doi: 10.1111/j.1365-3180.1978.tb01612.x

Himelblau, E., and Amasino, R. M. (2001). Nutrients mobilized from leaves of *Arabidopsis thaliana* during leaf senescence. *J. Plant Physiol.* 158, 1317–1323. doi: 10.1078/0176-1617-00608

Hocking, P. J., and Pate, J. S. (1977). Mobilization of minerals to developing seeds of legumes. *Ann. Bot.* 41, 1259–1278. doi: 10.1093/oxfordjournals.aob.a085415

Kirkby, E. (2012). "Introduction, definition and classification of nutrients," in *Marschner's Mineral Nutrition of Higher Plants*, ed. H. Marschner (Amsterdam: Elsevier), 3–5. doi: 10.1016/B978-0-12-384905-2.00001-7.

Kirkby, E. A., and Knight, A. H. (1977). Influence of the level of nitrate nutrition on ion uptake and assimilation, organic acid accumulation, and cation-anion balance in whole tomato plants. *Plant Physiol.* 60, 349–353. doi: 10.1104/pp.60.3.349

Krämer, U., Talke, I. N., and Hanikenne, M. (2007). Transition metal transport. *FEBS Lett.* 581, 2263–2272. doi: 10.1016/j.febslet.2007.04.010

Kuper, J., Llamas, A., Hecht, H.-J., Mendel, R. R., and Schwarz, G. (2004). Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. *Nature* 430, 803–806. doi: 10.1038/nature02681

Lahner, B., Gong, J., Mahmoudian, M., Smith, E. L., Abid, K. B., Rogers, E. E., et al. (2003). Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nat. Biotechnol.* 21, 1215–1221. doi: 10.1038/nbt865

Landberg, T., and Greger, M. (1994). Influence of selenium on uptake and toxicity of copper and cadmium in pea (Pisum sativum) and wheat (Triticum aestivum). *Physiol. Plant* 90, 637–644. doi: 10.1111/j.1399-3054.1994.tb02518.x

Llamas, Á, Tejada-Jiménez, M., Fernández, E., and Galván, A. (2011). Molybdenum metabolism in the alga *Chlamydomonas* stands at the crossroad of those in *Arabidopsis* and humans. *Metallomics* 3, 578. doi: 10.1039/c1mt00032b

López-Bucio, J., Cruz-Ramırez, A., and Herrera-Estrella, L. (2003). The role of nutrient availability in regulating root architecture. *Curr. Opin. Plant Biol.* 6, 280–287. doi: 10.1016/S1369-5266(03)00035-9

Ma, J. F., Goto, S., Tamai, K., and Ichii, M. (2001). Role of root hairs and lateral roots in silicon uptake by rice. *Plant Physiol.* 127, 1773–1780. doi: 10.1104/pp.010271

Maathuis, F. J. (2009). Physiological functions of mineral macronutrients. *Curr. Opin. Plant Biol.* 12, 250–258. doi: 10.1016/j.pbi.2009.04.003

Maillard, A. (2016). *Adaptation des Végétaux à un Environnement Minéral Fluctuant?: Mécanismes Impliqués et Interactions Entre Nutriments*. Caen: Université de Caen Normandie.

Maillard, A., Diquélou, S., Billard, V., Laîné, P., Garnica, M., Prudent, M., et al. (2015). Leaf mineral nutrient remobilization during leaf senescence and modulation by nutrient deficiency. *Front. Plant Sci.* 6:317. doi: 10.3389/fpls.2015.00317

Malagoli, P. (2005). Dynamics of nitrogen uptake and mobilization in field-grown winter oilseed rape (Brassica napus) from stem extension to harvest: I. Global n flows between vegetative and reproductive tissues in relation to leaf fall and their residual N. *Ann. Bot.* 95, 853–861. doi: 10.1093/aob/mci091

Masclaux-Daubresse, C., Reisdorf-Cren, M., and Orsel, M. (2008). Leaf nitrogen remobilisation for plant development and grain filling. *Plant Biol.* 10, 23–36. doi: 10.1111/j.1438-8677.2008.00097.x

Mauk, C. S., and Noodén, L. D. (1992). Regulation of mineral redistribution in pod-bearing soybean explants. *J. Exp. Bot.* 43, 1429–1440. doi: 10.1093/jxb/43.11.1429

McGrath, S. P., and Zhao, F. J. (1996). Sulphur uptake, yield responses and the interactions between nitrogen and sulphur in winter oilseed rape (Brassica napus). *J. Agric. Sci.* 126, 53–62. doi: 10.1017/S0021859600088808

Millaleo, R., Reyes- Diaz, M., Ivanov, A. G., Mora, M. L., and Alberdi, M. (2010). Manganese as essential and toxic element for plants: transport, accumulation and resistance mechanisms. *J. Soil Sci. Plant Nutr.* 10, 470–481. doi: 10.4067/S0718-95162010000200008

Nable, R., and Loneragan, J. (1984). Translocation of manganese in subterranean clover (Trifolium subterraneum L. Cv. seaton park). II. effect of leaf senescence and of restricting supply of manganese to part of a split root system. *Funct. Plant Biol.* 11:113. doi: 10.1071/PP9840113

Nenova, V. (2008). Growth and mineral concentrations of pea plants under different salinity levels and iron supply. *Gen. Appl. Plant Physiol.* 34, 189–201.

Nyborg, M., and Hoyt, P. B. (1970). Boron deficiency in turnip rape grown on gray wooded soils. *Can. J. Soil. Sci.* 50, 87–88. doi: 10.4141/cjss70-013

Pilon, M., Cohu, C. M., Ravet, K., Abdel-Ghany, S. E., and Gaymard, F. (2009). Essential transition metal homeostasis in plants. *Curr. Opin. Plant Biol.* 12, 347–357. doi: 10.1016/j.pbi.2009.04.011

Pittman, J. K. (2005). Managing the manganese: molecular mechanisms of manganese transport and homeostasis. *New Phytol.* 167, 733–742. doi: 10.1111/j.1469-8137.2005.01453.x

Salon, C., and Munier-Jolain, N. G. (eds) (2010). *Physiology of the Pea Crop*. Boca Raton, FL: CRC Press.

Salt, D. E., Baxter, I., and Lahner, B. (2008). Ionomics and the study of the plant ionome. *Annu. Rev. Plant Biol.* 59, 709–733. doi: 10.1146/annurev.arplant.59.032607.092942

Sankaran, R. P., and Grusak, M. A. (2014). Whole shoot mineral partitioning and accumulation in pea (Pisum sativum). *Front. Plant Sci.* 5:149. doi: 10.3389/fpls.2014.00149

Schachtman, D., and Liu, W. (1999). Molecular pieces to the puzzle of the interaction between potassium and sodium uptake in plants. *Trends Plant Sci.* 4, 281–287. doi: 10.1016/S1360-1385(99)01428-4

Schwarz, G., and Mendel, R. R. (2006). Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annu. Rev. Plant Biol.* 57, 623–647. doi: 10.1146/annurev.arplant.57.032905.105437

Sorin, E., Etienne, P., Maillard, A., Zamarreño, A.-M., Garcia-Mina, J.-M., Arkoun, M., et al. (2015). Effect of sulphur deprivation on osmotic potential components and nitrogen metabolism in oilseed rape leaves: identification of a new early indicator. *J. Exp. Bot.* 66, 6175–6189. doi: 10.1093/jxb/erv321

Stich, B., Benke, A., Schmidt, M., Urbany, C., Shi, R., and Wirén, N. (2020). The maize shoot ionome: its interaction partners, predictive power, and genetic determinants. *Plant Cell Environ.* 43, 2095–2111. doi: 10.1111/pce.13823

Tejada-Jiménez, M., Galván, A., Fernández, E., and Llamas, Á (2009). Homeostasis of the micronutrients Ni, Mo and Cl with specific biochemical functions. *Curr. Opin. Plant Biol.* 12, 358–363. doi: 10.1016/j.pbi.2009.04.012

Van Beusichem, M. L. (1981). Nutrient absorption by pea plants during dinitrogen fixation. 1. comparison with nitrate nutrition. *NJAS* 29, 259–272. doi: 10.18174/njas.v29i4.16997

Verma, T. S., and Minhas, R. S. (1987). Zinc and phosphorus interaction in a wheat-maize cropping system. *Fertilizer Res.* 13, 77–86. doi: 10.1007/BF01049804

Waters, B. M., and Grusak, M. A. (2008). Whole-plant mineral partitioning throughout the life cycle in *Arabidopsis thaliana* ecotypes Columbia, Landsberg erecta, Cape Verde Islands, and the mutant line ysl1ysl3. *New Phytol.* 177, 389–405. doi: 10.1111/j.1469-8137.2007.02288.x

White, P. J. (2012). "Ion uptake mechanisms of individual cells and roots," in *Marschner's Mineral Nutrition of Higher Plants*, ed P. Marschner (Amsterdam: Elsevier), 7–47. doi: 10.1016/B978-0-12-384905-2.00002-9.

Wiersum, L. K. (1958). Density of root branching as affected by substrate and separate ions. *Acta Bot. Neerl.* 7, 174–190. doi: 10.1111/j.1438-8677.1958.tb00613.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Jacques, Forest, Durey, Salon, Ourry and Prudent. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 January 2022
doi: 10.3389/fpls.2021.739033

Genetic Correlation Between Fe and Zn Biofortification and Yield Components in a Common Bean (*Phaseolus vulgaris* L.)

*Santiago Diaz, Jose Polania†, Daniel Ariza-Suarez, Cesar Cajiao, Miguel Grajales, Bodo Raatz† and Stephen E. Beebe**

Bean Program, Crops for Health and Nutrition Area, Alliance Bioversity International – CIAT, Cali, Colombia

OPEN ACCESS

Edited by:
Jose C. Jimenez-Lopez, Estación Experimental de Zaidín, Consejo Superior de Investigaciones Científicas, Spanish National Research Council (CSIC), Spain

Reviewed by:
Dil Thavarajah, Clemson University, United States
Petr Smýkal, Palacký University Olomouc, Czechia

***Correspondence:**
Stephen E. Beebe
s.beebe@cgiar.org

†Present address:
Jose Polania, Instituto de Biotecnologia, UNAM, Cuernavaca, Mexico
Bodo Raatz, Limagrain Vegetable Seeds, la Ménitré, France

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *09 July 2021*
Accepted: *08 November 2021*
Published: *03 January 2022*

Citation:
Diaz S, Polania J, Ariza-Suarez D, Cajiao C, Grajales M, Raatz B and Beebe SE (2022) Genetic Correlation Between Fe and Zn Biofortification and Yield Components in a Common Bean (Phaseolus vulgaris L.). Front. Plant Sci. 12:739033. doi: 10.3389/fpls.2021.739033

Common bean (*Phaseolus vulgaris* L.) is the most important legume for direct human consumption worldwide. It is a rich and relatively inexpensive source of proteins and micronutrients, especially iron and zinc. Bean is a target for biofortification to develop new cultivars with high Fe/Zn levels that help to ameliorate malnutrition mainly in developing countries. A strong negative phenotypic correlation between Fe/Zn concentration and yield is usually reported, posing a significant challenge for breeders. The objective of this study was to investigate the genetic relationship between Fe/Zn. We used Quantitative Trait Loci (QTLs) mapping and Genome-Wide Association Studies (GWAS) analysis in three bi-parental populations that included biofortified parents, identifying genomic regions associated with yield and micromineral accumulation. Significant negative correlations were observed between agronomic traits (pod harvest index, PHI; pod number, PdN; seed number, SdN; 100 seed weight, 100SdW; and seed per pod, Sd/Pd) and micronutrient concentration traits (SdFe and SdZn), especially between pod harvest index (PHI) and SdFe and SdZn. PHI presented a higher correlation with SdN than PdN. Seventy-nine QTLs were identified for the three populations: 14 for SdFe, 12 for SdZn, 13 for PHI, 11 for SdN, 14 for PdN, 6 for 100SdW, and 9 for Sd/Pd. Twenty-three hotspot regions were identified in which several QTLs were co-located, of which 13 hotpots displayed QTL of opposite effect for yield components and Fe/Zn accumulation. In contrast, eight QTLs for SdFe and six QTLs for SdZn were observed that segregated independently of QTL of yield components. The selection of these QTLs will enable enhanced levels of Fe/Zn and will not affect the yield performance of new cultivars focused on biofortification.

Keywords: QTL, iron, zinc, bean, biofortification

INTRODUCTION

Common bean (*Phaseolus vulgaris* L.) is one of the most important legumes for human consumption and is cultivated worldwide, particularly in tropical and subtropical countries of Africa and America (Broughton et al., 2003). Cultivated common bean is organized in two major gene pools, Mesoamerican and Andean, which resulted from independent domestication events,

generating landraces with diverse seed quality, nutritional traits, resistance to diseases, tolerance to abiotic stress, morphologic, and agronomic traits (Schmutz et al., 2014). Bean has been documented as a highly valuable food for the human diet, an accessible and relatively inexpensive source of proteins and micronutrients, such as iron and zinc (Beebe, 2012).

Micronutrient deficiencies are among the most common global nutritional problems, affecting half of the population of the world, especially women and children in developing countries (Bouis and Welch, 2010). Fe and Zn deficiencies are among the main common causes of micromineral malnutrition worldwide. In humans, Fe is essential for several cellular mechanisms (Steinbicker and Muckenthaler, 2013). Iron Deficiency Anemia (IDA), a pathology associated with scarce iron intake and absorption, affects a third of the population of the world and has been associated with several chronic diseases (Lopez et al., 2016). Zn has a key role in regulating growth and in the functioning of the immune system (Chasapis et al., 2012), and its deficiency can contribute to infections (Hambidge, 2000).

Wide phenotypic variability has been reported in bean in terms of Fe and Zn accumulation, with seed Fe concentration from 35 to 90 mg/kg and seed Zn concentration from 20 to 54 mg/kg (Beebe et al., 2000), these ranges are higher than those of other main staple crops, such as rice, wheat, and maize (Petry et al., 2015). For this reason, common bean is one of the crops targeted for biofortification. A cup of beans supplies 25% of Fe and 15% of Zn to the recommended daily allowance in the human diet, but breeding programs have the potential to develop genotypes with two to three times more Fe and Zn concentrations (Cichy et al., 2009).

Bouis (2003) reported that biofortification breeding has multiple positive effects, such as more efficiency in nutrients, fertilizer, and irrigation demands, and can increase disease resistance and yield. However, some difficulties have been described in bean breeding for micronutrient concentration, mainly a strong negative correlation between Fe and Zn concentration and yield. Source genotypes with high iron contents (>100 mg/kg) have been identified, but most of these lines have poor agronomic performance (Raatz, 2018). For this reason, a major breeding objective is overcoming the negative correlation between the micronutrient concentrations with yield components and finding the optimum balance between these traits.

While the phenotypic variability for Fe and Zn in common bean has been utilized in breeding programs with some success, the development of genomic tools will likely be helpful to reach the goals of biofortification (Beaver and Osorno, 2009). Currently, the genetic architecture of Fe and Zn accumulation in seed in common bean has been the object of several studies (Izquierdo et al., 2018), usually using Quantitative Trait Loci (QTL) analysis to find genomic regions related to the genetic control of these traits. Nevertheless, molecular markers have not been implemented in bean breeding programs focused on mineral concentration, as parental materials employed for populations are often distant from elite breeding materials, and the QTLs found with major effects are scarce and cannot be extrapolated to other populations (Raatz, 2018).

Another factor to consider in biofortification breeding is the accompanying agronomic traits, such as drought tolerance. Hummel et al. (2018) predicted a large part of current common bean growing areas to become unsuitable for cultivation by 2050, generating reductions in yield for current varieties and potentially affecting the nutritional quality, especially the iron levels of the crop, if current climate change trends continue. Drought is the main cause of yield loss worldwide, and its incidence and duration are expected to increase due to climate change (Lauer et al., 2012). Drought stress is a principal abiotic constraint that affects beans, causing losses close to 60% in production (Beebe et al., 2008).

The objective of this study was to investigate the correlation of genetic architecture of Fe and Zn accumulation and yield components in a common bean under irrigated and drought conditions using QTL and Genome-Wide Association Studies (GWAS) analysis in three biofortified bi-parental Mesoamerican populations and to identify genomic regions associated with yield components, micromineral accumulation, and their possible relationship.

MATERIALS AND METHODS

Plant Material

In this study we used three Mesoamerican bi-parental populations of common bean (1) SCR16 × SMC40, (2) SMC44 × SCR9, and (3) SMC33 × SCR16. Parents SCR16 and SCR9 are improved Mesoamerican lines from the International Center of Agricultural Tropical (CIAT) with traits related to drought tolerance and resistance to bean common mosaic virus (*BCMV*) and bean golden yellow mosaic virus (*BGYMV*). The parents SMC40, SMC44, and SMC33 are improved Mesoamerican lines with a high accumulation of Fe and Zn in the seed. Each population consists of 100 recombinant inbred lines (RILs). Populations were advanced by single seed descent, and families from which individual plants were drawn were evaluated for iron and zinc concentration, and yield components in the F3, F4, F5, F6, F7, and F8 generations (**Table 1**).

Field Trials Design and Phenotyping

The three bi-parental populations were planted at the CIAT experimental station in Palmira, Colombia (altitude of 1,000 m.a.s.l., latitude 3° 32′ N and longitude 76° 18′ W). The experimental design used for this study was an augmented design with three blocks in which five different check lines (SER118, SER16, Tio Canela 75, SEN56, and DOR390), and each parent for each population was evenly distributed across the field. The experimental units in these trials were row plots of 2.22 m^2 for each line. The experiments were conducted in 5 years (2011, 2012, 2013, 2014, and 2015) in four irrigated trials and four drought trials, for a total of eight trials, with data taken on families as the RILs were advanced from F2.3 to F6.8 (**Table 1**). For more information about the climatic conditions see **Supplementary Figure 1**.

The seed was harvested manually by plot upon maturity (75–80 days after sowing). This was estimated when 50% of plants of

TABLE 1 | Description of trials used in the experiment.

Year	Condition	Season[1]	Trial name	Generation
2011	Drought	B	2011_Dro_B	F2.3
2012	Drought	B	2012_Dro_B	F2.4
2013	Irrigated	A	2013_Irr_A	F3.5
	Drought	B	2013_Dro_B	F3.5
2014	Irrigated	A	2014_Irr_A	F4.6
	Drought	B	2014_Dro_B	F5.7
	Irrigated	B	2014_Irr_B	F5.7
2015	Irrigated	A	2015_Irr_A	F6.8

[1]*Sowing seasons per year. A, first semester; B, second semester.*

the plot had at least one pod losing its green pigmentation. At the time of the harvest, 0.5 m per plot were collected separately to measure the number of pods per m^2 (pod number, PdN), number of seeds per m^2 (seed number, SdN), 100 seed weight (100SdW) was obtained estimating the weight of the seeds per plot and corrected to the moisture content in the seed of 14%, number of seeds per pod (Sd/Pd), and pod harvest index (PHI, %) defined as the ratio between seed weight and total pod weight, such as seed (Assefa et al., 2013). In 2011, only PHI was measured. Additional information on the description of phenotypic traits can be found in "Traits Dictionaries for Fieldbook Development."[1].

Fe and Zn Quantification

The samples to evaluate iron and zinc concentrations in the seed (SdFe and SdZn, mg/kg) were prepared according to the method described by Stangoulis and Sison (2008). In brief, we collected 5 g of seed and cleaned using a cloth dampened with high-purity water. Later, seeds were dried to 7–8% moisture content in an oven at 60°C. Finally, samples were ground with a grinding mill Retsch Mixer Mill MM 400 using grinding jars and zirconium balls (Retsch GmbH & Co KG, Germany) to avoid iron and zinc contamination. To quantify SdFe and SdZn a Energy Dispersive X-ray Fluorescence machine (EDXRF) Bruker S2 PUMA (Bruker Corporation, Billerica, MA, United States) was used. The quantification was developed according to the method described by Guild et al. (2017a). In summary, the EDXRF analysis conditions were as follows. Atmosphere: air, X-ray tube: palladium (50 W), voltage: 40 kV, current: 240 μA, peak detected: Kα, acquisition time: 30 s, tube filter: Al 500 nm, and detector: silicon drift detector. Analyses were conducted in supplied sample cups prepared as reported previously by Guild et al. (2017a). Sample cups were cleaned and prepared before each analysis to minimize cross-contamination between samples. For the calibration, we used 10 custom-made 40 mm diameter glass disks (FLUXANA GmbH & Co., Germany) with a range of nominal Fe and Zn levels to establish a non-matrix-matched calibration. SdFe and SdZn were measured in all trials except in the 2011.

Data Analysis

Statistical analysis of the phenotypic data was conducted with statistical software R (v3.3.2.). Best linear unbiased estimators and predictors (BLUEs/BLUPs) were calculated for each trait and each trial using the "lme4" Package (Bates et al., 2015). The data were modeled using the following equation:

$$y_{mij} = \mu + G_m + R_i + \varepsilon_{mij}$$

where y is a vector with the phenotypic data, μ is the overall intercept, G_m is the effect of the m-th genotype, R_i is the effect of the i-th replication, and ε_{mij} is the error term corresponding to y_{mij}. In this model, G_m term effects were treated either as fixed (to calculated BLUEs) or random (to obtain an estimate of the genetic variance and calculated BLUPs). We assumed that every random term μ and the residual ε adjust to a normal distribution with mean 0 and independent variances $\mu \sim N(0, \sigma^2_u I)$ and $\varepsilon \sim N(0, \sigma^2_\varepsilon I)$.

To determine the proportion of the genetic variance controlling for evaluated traits, broad-sense heritability (H^2) estimates were calculated using the method proposed by Cullis et al. (2006), using the equation below:

$$H^2 = 1 - \frac{\upsilon_{BLUP}}{2\sigma^2_{genotype}}$$

where υ_{BLUP} is the mean-variance of a difference of two BLUPs of genotypic effects, and $2\sigma^2_{genotype}$ is the genetic variance. The phenotypic correlation between traits measured was expressed as Pearson's correlation coefficients among BLUEs for individual trials and cross-trials means, and their significance was tested using a two-tailed t-test.

Genotyping

DNA extraction of the RILs and parents was developed on the F6.8 generation. For DNA extraction, seeds of each genotype were selected and germinated using a Calcium Sulfate 0.5-mM solution, the DNA was extracted from young leaves using liquid N^2 and the urea buffer-based extraction miniprep protocol. The genotyping was performed using the *Ape*KI-based genotyping-by-sequencing (GBS) protocol (Elshire et al., 2011), and the GBS libraries were sequenced at the HudsonAlpha Institute for Biotechnology.[2] Some lines were re-genotyped where seed quality was poor.

The mapping and variant calling processes for GBS reads are described in detail by Perea et al. (2016). In brief, the GBS reads were demultiplexed using the Next Generation Sequencing Eclipse Plugin (NGSEP; v3.3.2) (Tello et al., 2019). Adapters and low-quality bases from the raw sequencing data were trimmed using Trimmomatic (v0.36) (Bolger et al., 2014), and reads were aligned to the reference genome of *P. vulgaris* G19833 v2.1. (Schmutz et al., 2014) using Bowtie2 (v2.2.30) (Langmead and Salzberg, 2012) with default parameters. The variant calling process of RILs was performed using NGSEP following recommended parameters for GBS data (Perea et al., 2016), using the genotype calls of the parents (SMC44, SMC40, SMC33, SCR16, and SCR9) as reference. The resulting genotype matrix was filtered to remove genotype calls with quality below 40, remove markers with more than 6% heterozygote

[1]http://www.cropontologycurationtool.org/

[2]https://hudsonalpha.org/

calls, minor allele frequency (MAF) below 0.02, and remove markers in repetitive regions of the reference genome (Lobaton et al., 2018). The process to construct a genotype matrix was carried out for each population separately and grouping the populations for GWAS analysis. The final matrices contained 1,650 markers in the SCR16 × SMC40 population, 3,033 markers in the SMC44 × SCR9 population, and 481 markers in the SMC33 × SCR16 population.

Genetic Map Construction, Quantitative Trait Loci Analysis, and Genome-Wide Association Studies Analysis

For the construction of the genetic maps, the integrated tools of the IciMapping Software v4.1 (Meng et al., 2015) were used, employing the matrices above described. The genotypic data were grouped and anchored on their physical position. The default logarithm of odds (LOD) score (3.0) and rfs (0.35) were used as a maximum threshold value for grouping markers. The Kosambi genetic mapping function was chosen for genetic interval estimation using recombination frequency. The orders of markers per chromosome were set by their physical position, and distortion of segregated markers was exhibited according to segregation ratios. QTL analysis for the populations was conducted using the genetic maps of each population and the phenotypic BLUEs for each RIL. Detection of QTLs and estimation of genetic parameters for SdFe, SdZn, PHI, PdN, SdN, 100SdW, and Sd/Pd were performed using the inclusive composite interval mapping for additive effects (ICIM-ADD) procedure of the software IciMapping (v4.1) (Meng et al., 2015) with 10 cm windows and a sliding parameter of 1 cm. Significant QTL was considered by defining the LOD score at a genome-wide type I error rate of $\alpha = 0.05$ after 1,000 permutation tests for each trait. The detected QTLs having high phenotypic variation explained (PVE) (>10%) were justified as major QTL (Amanullah et al., 2021). QTL hotspots were defined as significant loci genetically linked at <5 cm between QTL.

Genome-Wide Association Studies analysis was carried out using the R package GAPIT (v3.0) (Tang et al., 2016). The association analysis was conducted using a Compression Mixed Linear Model (CMLM) approach. This model accounts for population structure using the top four principal components as fixed effects. It also accounts for random polygenic effects with a kinship matrix as variance-covariance structure, calculated using the method proposed by VanRaden (2008) implemented in the GAPIT package. Significant associations were defined when the *p*-value was equal to or smaller than the Bonferroni-corrected threshold (1.79×10^{-6}) (Johnson et al., 2010). Manhattan and quantile–quantile (Q-Q) plot graphics were made using the qqman R package (Turner, 2018).

RESULTS

Phenotyping Variation for Traits Evaluated

Five elite lines of the CIAT bean-breeding program were used to develop three Mesoamerican bi-parental populations: (1) SCR16 × SMC40, (2) SMC33 × SCR16, and (3) SMC44 × SCR9. Families in several generations and F6.8 RILs were evaluated for a range of yield components and micronutrient concentrations of iron and zinc (SdFe and SdZn) in trials managed under irrigated and drought stress conditions. The phenotypic data show high and significant phenotypic variability among RIL in all traits evaluated (**Supplementary Tables 1, 2** and **Supplementary Figure 2**). Transgressive segregation was also evident in the three populations with lower and higher values compared to parents in all traits. An effect of drought stress was not observed on SdFe, SdZn, PHI, and Sd/Pd in the populations evaluated. However, drought reduced PdN, SdN, and 100SdW by 29.5, 39.6, and 31.4%, respectively in all populations. The SdZn performance of SMC44 × SCR9 was better in most trials compared to the other populations (27.2 mg/kg on average), while SCR16 × SMC40 presented higher SdFe than the other two populations (78.3 mg/kg on average) (**Figures 1, 2** and **Supplementary Table 1**).

The trial 2015_Irr_A presented the highest Zn accumulation for the three populations (>33 mg/kg for each population). In all trials, the SMC parents (SMC40, SMC44, and SMC33) presented a better performance in SdFe and SdZn than SCR parents (SCR9 and SCR16) (**Figure 1**). Conversely, SCR parents presented higher PHI, 100SdW, and Sd/Pd than the SMC parents (**Figure 2**). The broad-sense heritabilities for PHI, 100SdW, and Sd/Pd were high for the three populations: 0.70, 0.65, and 0.73 for SCR16 × SMC40; 0.79, 0.72, and 0.67 for SMC33 × SCR16; and 0.84, 0.79, and 0.72 for SMC44 × SCR9. In contrast, the heritabilities for the SdN and PdN were lower and not stable across trials (<0.53 for PdN and <0.47 for SdN). Heritabilities for micronutrient concentrations were intermediate ranging between 0.35 and 0.65 for SdFe and 0.58 and 0.70 for SdZn. SMC44 × SCR9 presented higher heritabilities compared to the other populations in most traits evaluated except for SdZn (**Supplementary Table 1**). Overall, the phenotypic datasets from field trials present a good quality for subsequent analyses.

Pearson correlations between traits suggested two trait clusters according to the positive and generally significant correlations within clusters and negative correlations between them. On the one hand, the micronutrient concentrations of SdFe and SdZn form a group that is positively correlated among all populations (0.71–0.76). On the other hand, a second group is formed between PdN and SdN showing a strong positive correlation between them (0.75–0.82). However, PHI only shows a stable correlation with SdN and Sd/Pd (0.49–0.56 and 0.64–0.76, respectively). These agronomic traits (PHI, PdN, SdN, and Sd/Pd) show strong negative correlations with micronutrient concentrations, especially between PHI and SdFe (−0.63 to −0.69) and SdZn (−0.72 to −0.78). 100SdW only presented weak negative correlations with SdN and PdN (−0.25 and −0.27) (**Figure 3** and **Supplementary Table 3**).

Quantitative Trait Loci Analysis

Marker-trait associations were evaluated by QTL mapping using ICIM-ADD. In total, 79 QTLs for seven traits were identified for the three populations for eight individual trials: 14 for SdFe, 12 for SdZn, 13 for PHI, 11 for SdN, 14 for PdN, 6 for 100SdW, and 9 for Sd/Pd. Sixty-four QTLs of the 77 QTLs identified in this study

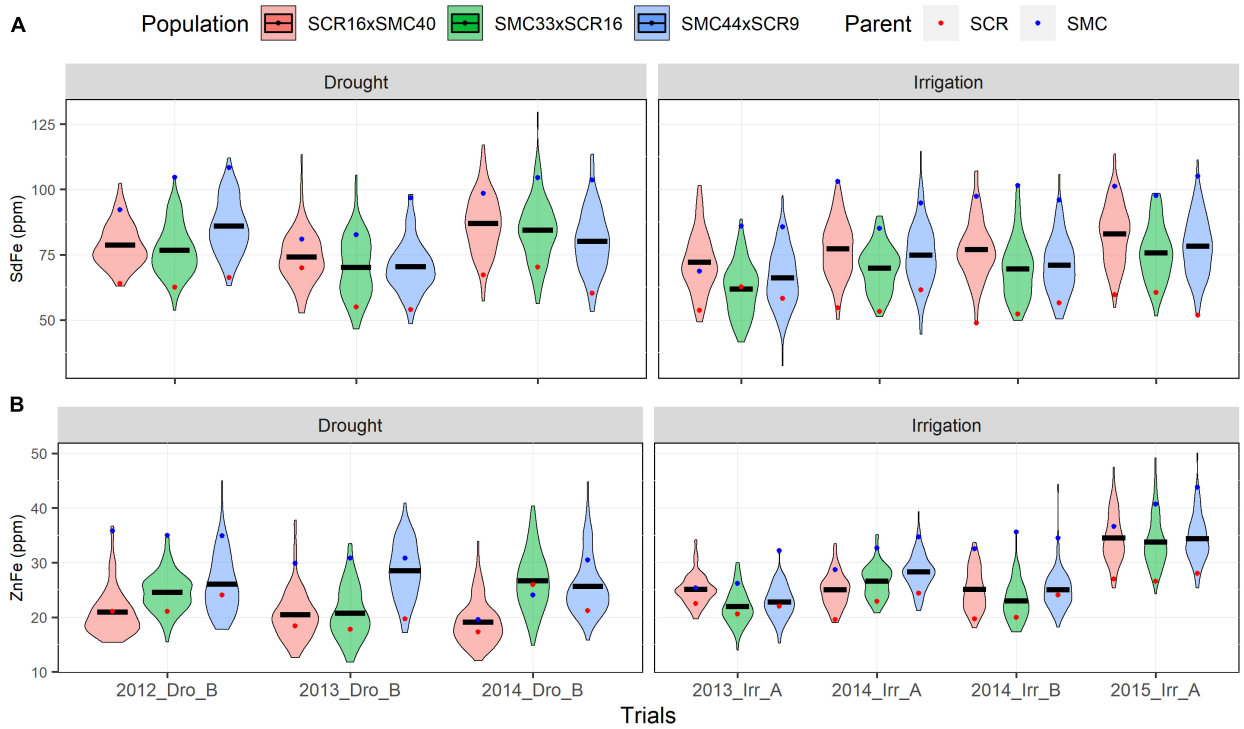


FIGURE 1 | Phenotypic distribution of micronutrients content traits. Fe accumulation (SdFe) **(A)** and Zn accumulation (SdZn) **(B)** evaluated of 2012, 2013, 2014, and 2015 under drought and irrigated conditions for the SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations.

are classified as "major QTL" (>10% PVE). Twenty-three hotspot regions were identified in which several QTLs were co-located (**Table 2** and **Supplementary Table 4**). In GWAS analysis, results show low levels of significance (**Supplementary Figures 3, 4**).

Quantitative Trait Loci for Fe accumulation were identified on seven chromosomes (Pv03, Pv05, Pv06, Pv07, Pv08, Pv09, and Pv011). Ten major QTLs were identified; SdFe6.3 in the 2013_Dro_B trial was the QTL with the highest PVE (19.47%) with an additive positive effect of 5.04 mg/kg, while SdFe6.2 had an additive effect of 6.28 mg/kg. In all QTLs found for SdFe, the additive positive effects came from SMC parents (SMC44 and SMC33) except for SdFe7.1. SdFe3.1 was the only QTL expressed in more than one trial. QTLs for SdZn were located on Pv02, Pv03, Pv04, Pv05, Pv06, Pv07, Pv08, and Pv10, of which eight were major QTLs. SdZn5.1 was the QTL that explained the greatest phenotypic variance on the trait (18.73%), with an additive effect of 2.74 mg/kg. SdZn2.3 was the only QTL with an additive positive effect from an SCR parent (1.86 mg/kg), while the other QTL for SdZn with additive positive effect came from SMC44 or SMC33. SdZn8.2 presented a higher additive effect compared with other QTL (3.57 mg/kg) (**Table 2** and **Supplementary Table 4**). SdZn10.1 was identified in two trials.

Twelve QTLs for PHI were found on Pv01, Pv02, Pv03, Pv05, Pv07, Pv08, Pv09, and Pv11. Seven of these QTLs were classified as major QTLs. PHI7.2 presented the highest PVE and highest positive additive effect (20.61% and 13.54%, respectively) in 2014_Dro_B under drought stress. All QTLs with positive additive effect for this trait originates in an SCR parent, except PHI7.3. On Pv07, three different loci controlling PHI were identified, and two of them were present in two trials. Eight major QTLs for SdN were identified with SdN8.1 as the QTL with the highest PVE (18.59%) and SdN7.1 with the highest additive effect (186.93 seeds). Eleven major QTLs for PdN were found. PdN7.1 was the QTL with the highest phenotypic variance of 17.55% and PdN4.1 had the highest additive effect of 65.33 pods. Most QTLs found for PdN with positive effect came from the SMC parent except for PdN6.1, Pdn7.2, and PdN8.1. Six QTLs for seed weight were identified on Pv06, Pv07, Pv08, and Pv11, 100SdW7.1 was the QTL with the highest PVE (16.62%) and 100SdW8.2 from SCR9 presented a high additive effect (2.44 g). The sources for all QTL with positive additive allelic effects for 100SdW were SCR parents except for 100SdW6.1. Seven major QTLs for seed per pod were identified. Sd/Pd2.1 presented the highest PVE (24.50%) and Sd/Pd7.2 showed the highest positive allelic effect (0.67 seeds per pod) (**Table 2** and **Supplementary Table 4**).

Nine QTL hotspots for yield components (PHI, SdN, PdN, SdW, and Sd/Pd) were identified. Similarly, 14 QTL hotspots for yield components and micronutrient concentration were recognized, and only one QTL hotspot was observed controlling the concentration of both Fe and Zn independently from yield components (**Supplementary Table 4**). Fourteen QTLs related with SdFe were identified, of which seven QTLs are co-localized with QTL related with Sd/Pd, SdN, PdN, 100SdW, and PHI. Similarly, of 12 QTLs identified for SdZn, eight were co-localized with QTL related to the yield components (**Table 2** and **Supplementary Table 4**). In most of these "hotpots" (Hotpots

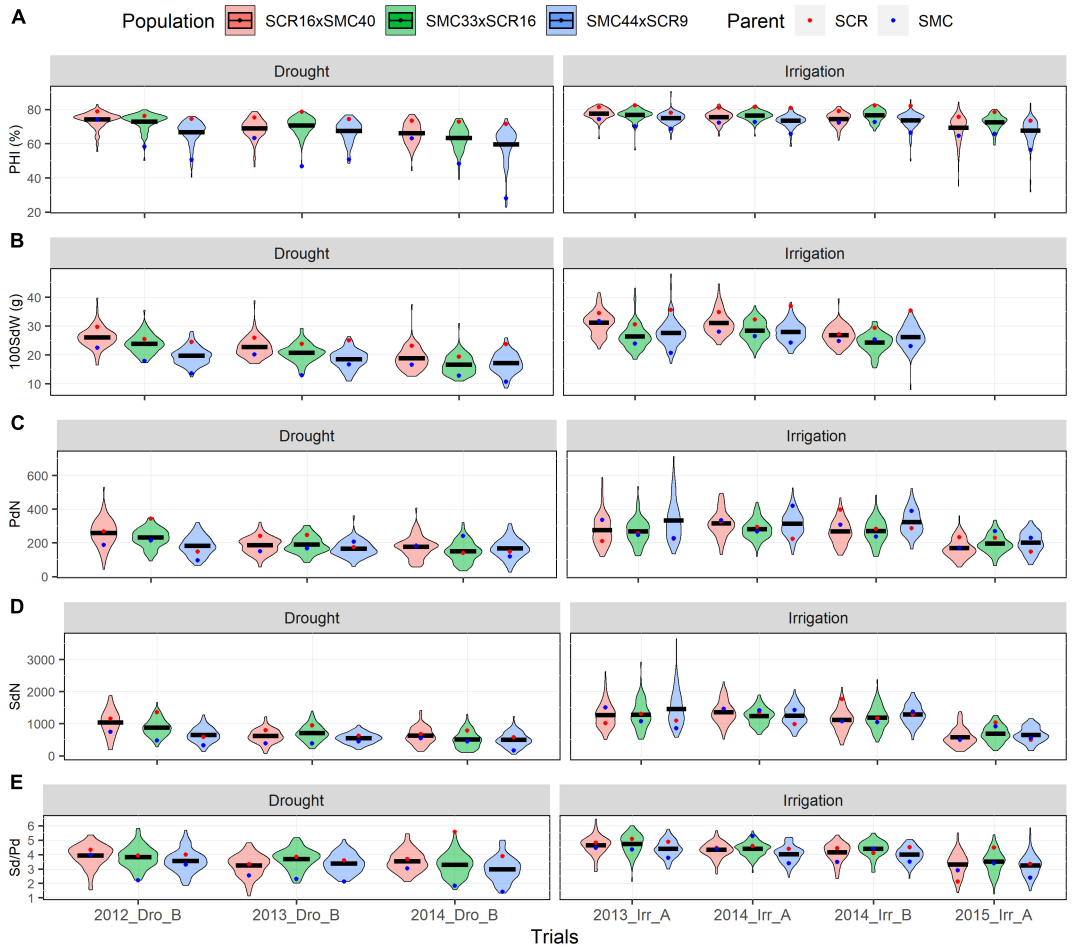


FIGURE 2 | Phenotypic distribution of agronomics traits. Pod harvest index (PHI) **(A)**, 100 seed weight (100SdW) **(B)**, pod number (PdN) **(C)**, seed number (SdN) **(D)**, seed per pod (Sd/Pd), and **(E)** evaluated of 2012, 2013, 2014, and 2015 under drought and irrigated conditions for the SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations.

4, 5, 6, 11, 12, 16, 18, 20, 21, 22, and 23), it was observed that the positive additive allelic effects for higher micronutrient concentrations were associated with the negative allelic effect for yield component traits. The high yield component QTL was usually donated by the SCR parent, while the high mineral QTL came from the SMC parent. These strong negative associations are especially visible between SdFe and SdZn with PHI and Sd/Pd, with two strong associations with 100SdW.

DISCUSSION

The populations analyzed showed ample phenotypic variability in terms of average accumulation of micronutrients: SCR16 × SMC40 (53.7–108.3 mg/kg for Fe and 17.5–36.8 mg/kg for Zn), SMC33 × SCR16 (50.0–101.6 mg/kg for Fe and 16.9–36.8 mg/kg for Zn) and SMC44 × SCR9 (49.2–107.5 mg/kg for Fe and 18.7–42.4 mg/kg for Zn) (**Supplementary Table 1**). These ranges of micronutrient concentrations are wider than those reported for other populations: Andean biparental populations AND696 × G19833 (39–79 mg/kg for Fe and 16–29 mg/kg for Zn) (Cichy et al., 2009); G21242 × G21078 (28–95 mg/kg for Fe and 17–49 mg/kg for Zn) (Blair et al., 2011); Mesoamerican biparental population G14519 × G4825 (35–97 mg/kg for Fe and 17–49 mg/kg for Zn) (Blair et al., 2010); intergene pool biparental populations DOR364 × G19833 (40–84 mg/kg for Fe and 17–42 mg/kg for Zn) (Blair et al., 2009); Cerinza × G10022 (54–100 mg/kg for Fe and 23–38 mg/kg for Zn) (Blair and Izquierdo, 2012); an African collection (23.6–78.3 mg/kg for Fe and 19–56.1 mg/kg for Zn) (Tryphone and Nchimbi-Msolla, 2010); an Andean collection (54.5–99.3 mg/kg for Fe and 21.6–39.7 mg/kg for Zn) (Katuuramu et al., 2018); a European collection (38.4–93.7 mg/kg for Fe and 18.9–43.6 mg/kg for Zn) (Caproni et al., 2020); and a multiparent advanced generation intercross (MAGIC) population (37.9–87.6 mg/kg for Fe and 18.5–39 mg/kg for Zn) (Diaz et al., 2020). Common bean has significant biofortification potential due to the wide phenotypic variability in terms of Fe and Zn concentration, producing improved lines that can accumulate >100 mg/kg of Fe and >45 mg/kg of Zn. This high accumulation capacity of

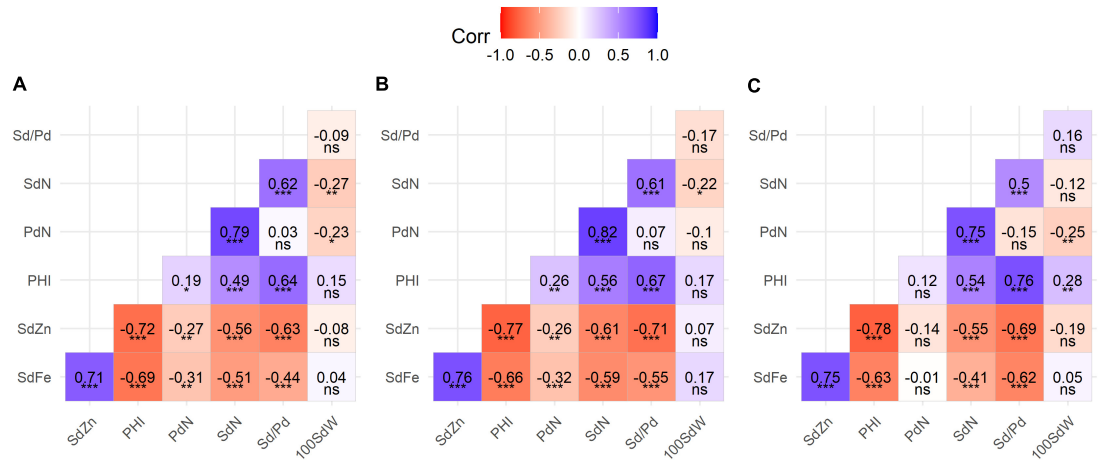


FIGURE 3 | Pearson's correlation coefficients between yield components and mineral concentration traits on three populations: **(A)** SCR16 × SMC40, **(B)** SMC33 × SCR16, and **(C)** SMC44 × SCR9. Significance of correlations indicated as ***$p < 0.0001$; **$p < 0.001$; *$p < 0.01$; ns, not significant. SdFe, seed iron; SdZn, seed zinc; PHI, pod harvest index; PdN, pod number; SdN, seed number; Sd/Pd, number of seeds per pod; 100SdW, 100 seed weight. Blue color indicates positive correlation and red color indicates negative correlation.

micronutrients exceeds that of other legume species, such as chickpea (36–91 mg/kg of Fe), pea (26–91 mg/kg of Fe), and mung bean (35–87 mg/kg of Fe), and only surpassed by lentil (37–176 mg/kg of Fe) (Shiva et al., 2016; Upadhyaya et al., 2016; Ma et al., 2017; Jha and Warkentin, 2020; Kumar and Pandey, 2020; Sab et al., 2020).

Breeding programs need fast, accurate, and cheap methods for screening large numbers of breeding lines and germplasm for genotypes with high micronutrients contents. EDXRF is an alternative to inductively coupled plasma-optical emission spectrometry (ICP-OES), the most common tool used for measuring microelements (Guild et al., 2017b). The benefits of this method compared to ICP lie in minimal sample preparation before analysis, less labor time, and cost required [∼$0.15 per sample to X-ray diffraction (XRD) method]. In addition, it is a method with a high level of prediction. Guild et al. (2017a) reported values of prediction models of 96% for Fe and 93% for Zn compared with ICP reference concentration, with an SE of prediction of 2.5 mg/kg of Fe and 1.5 mg/kg of Zn with a maximum coefficient of variation of 5% for Fe and 3.5% for Zn.

The genetic basis underlying the micronutrient concentration have been widely studied in common bean using QTL analysis, and GWAS (Gonzalez et al., 2017; Izquierdo et al., 2018; Gunjača et al., 2021). QTL mapping is an excellent strategy for gene mapping in this case because only a bi-parental population and a small set of markers are necessary a bi-parental population and a small set of markers, which are utilized to identify the genomic regions that segregate with a trait. However, this strategy is usually low in resolution since only two alleles per locus are analyzed, and genetic recombination is limited (Islam et al., 2016). Traditional QTL mapping is highly dependent on the genetic diversity of the two parents, and the effects of the detected QTLs can vary between populations. For this reason, the validation is necessary in other populations. GWAS analysis did not show any significant association. This lacks correlation to the type of population used for the biofortification analysis, as biparental populations are more suited for QTL mapping (Diaz et al., 2021). The limitations of QTL analysis can be overcome using GWAS, which can narrow down the candidate regions using natural populations. However, its use in biparental populations presents problems by low events recombination typical of this type of populations and low allelic richness required for GWAS. On the other hand, GWAS is susceptible to report false-positive associations, and its results require validation. The number of markers used in the GWAS highly affects its results (Han et al., 2018). In this study, 14 QTLs for SdFe were identified, of which 9 QTLs have been previously reported. Astudillo et al. (2013) reported two QTLs co-located respectively to SdFe3.2 and SdFe11.2 in an inter-gene pool population; the positive allele for high iron in that study came from the Andean parental (G19833). In this study, SdFe6.3 showed the highest percentage of variance explained and was the most significant QTL identified for Fe. This QTL was identified previously in an Andean biparental population (Cichy et al., 2009), in an intergene pool population (Astudillo et al., 2013), in an Andean collection (Katuuramu et al., 2018), and a Meta-QTL analysis using different populations (Izquierdo et al., 2018). Other QTLs, such as SdFe6.2, SdFe6.4, SdFe7.1, SdFe8.1, and SdFe11.1, have been identified in different populations belonging to the Andean and Mesoamerican gene-pool (Blair et al., 2010, 2011; Blair and Izquierdo, 2012; Cichy et al., 2014; Izquierdo et al., 2018; Katuuramu et al., 2018; Diaz et al., 2020). SdFe6.1 was only reported previously by Diaz et al. (2020) in a Mesoamerican Multi-parent Population and the positive allele came from a biofortified genotype MIB778. It was suggested that this allele in MIB778 may have been derived from a *Phaseolus dumosus* parent. There are no reports about SdFe3.1 but in this study, the QTL proved to be highly important for the analyzed biofortified populations.

TABLE 2 | Five most important "major QTLs" identified by inclusive composite interval mapping for additive effects (ICIM-ADD), genetic and physical position, LOD, phenotypic variation explained and additivity, mapped for SdFe, seed iron; SdZn, seed zinc; PHI, pod harvest index; PdN, pod number; SdN, seed number; Sd/Pd, number of seeds per pod in SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations evaluated in 2011, 2012, 2013, 2014, and 2015.

Trait	QTL name	Chr.	Pos. (cM)	Left marker	Right marker	LOD	PVE (%)	Add	Source
PdN	PdN2.1	2	6.0	Chr02_1904681	Chr02_1957517	3.2	13.4	27.9	SMC44
	PdN6.1	6	49.2	Chr06_28273144	Chr06_28493119	3.4	14.1	−23.1	SMC44
	PdN7.1	7	3.4	Chr07_444014	Chr07_528622	3.8	17.5	34.2	SMC33
	PdN7.2	7	14.5	Chr07_1475573	Chr07_1491651	8.5	13.2	−27.1	SMC44
	PdN8.2	8	53.0	Chr08_47670662	Chr08_51758938	3.4	14.7	28.5	SMC33
PHI	PHI2.1	2	34.4	Chr02_32651815	Chr02_33432574	5.8	17.2	−3.7	SMC33
	PHI3.1	3	7.1	Chr03_1303175	Chr03_2089482	2.9	13.2	1.8	SMC33
	PHI7.1	7	38.1	Chr07_8771144	Chr07_8994936	16.8	20.0	−11.4	SMC44
	PHI7.2	7	47.2	Chr07_21053823	Chr07_21106043	17.4	20.6	13.5	SMC44
	PHI7.3	7	80.8	Chr07_36965265	Chr07_36997111	3.7	14.4	−1.7	SCR16
Sd/Pd	Sd/Pd1.1	1	73.2	Chr01_50988725	Chr01_51029053	3.9	14.9	−0.3	SMC44
	Sd/Pd2.1	2	5.7	Chr02_1773128	Chr02_1863169	5.0	24.5	−0.4	SCR16
	Sd/Pd5.2	5	72.0	Chr05_39973668	Chr05_40035138	6.1	14.4	0.3	SCR16
	Sd/Pd7.2	7	38.1	Chr07_8771144	Chr07_8994936	8.3	14.9	−0.7	SMC44
	Sd/Pd10.1	10	35.5	Chr10_41705364	Chr10_41845732	5.6	16.7	−0.4	SMC44
SdFe	SdFe3.1	3	0.4	Chr03_128924	Chr03_178169	5.4	17.0	5.2	SMC44
	SdFe3.2	3	41.2	Chr03_32241300	Chr03_32532209	5.0	17.2	−6.3	SCR16
	SdFe6.3	6	42.6	Chr06_26259338	Chr06_26444453	10.4	19.5	5.0	SMC44
	SdFe6.4	6	52.6	Chr06_28871402	Chr06_28878273	3.0	15.4	5.7	SMC44
	SdFe7.1	7	81.5	Chr07_37050767	Chr07_37087581	3.6	14.4	4.6	SCR16
SdN	SdN2.1	2	5.7	Chr02_1773128	Chr02_1863169	3.3	15.7	−102.1	SCR16
	SdN7.1	7	3.4	Chr07_444014	Chr07_528622	3.6	14.9	177.5	SMC33
	SdN7.2	7	14.8	Chr07_1513753	Chr07_1548857	3.2	15.4	−81.2	SMC44
	SdN8.1	8	78.2	Chr08_60266587	Chr08_60277482	3.5	18.6	−109.2	SMC44
	SdN9.1	9	21.7	Chr09_20907648	Chr09_21865226	3.4	14.6	−84.5	SMC44
100SdW	100SdW6.1	6	30.0	Chr06_21622820	Chr06_21693622	3.8	13.6	−1.2	SCR16
	100SdW7.1	7	62.0	Chr07_30278406	Chr07_30493387	3.6	16.6	−1.6	SMC44
	100SdW7.2	7	74.6	Chr07_35056625	Chr07_35841130	2.9	10.1	1.6	SCR16
	100SdW8.1	8	90.0	Chr08_62043144	Chr08_62063701	4.6	11.5	−2.4	SMC44
	100SdW11.1	11	44.3	Chr11_5665676	Chr11_5713253	4.8	16.5	1.5	SCR16
SdZn	SdZn4.1	4	0.0	Chr04_22568	Chr04_55665	3.1	13.0	1.1	SMC44
	SdZn5.1	5	23.3	Chr05_3779357	Chr05_3956110	3.5	13.4	−1.4	SCR16
	SdZn5.1	5	34.3	Chr05_11706425	Chr05_11785008	5.6	18.7	2.7	SMC44
	SdZn8.2	8	92.1	Chr08_62309944	Chr08_62325409	3.5	13.0	3.6	SMC44
	SdZn10.1	10	11.7	Chr10_7899489	Chr10_8099574	7.2	18.2	−2.4	SCR16

*A complete list of significant QTL is available (**Supplementary Table 3**).*
QTLs, Quantitative Trait Loci.

For SdZn, 12 QTLs were identified, of which 6 have been reported previously. SdZn6.1 has been reported in several populations of both gene pools (Cichy et al., 2009; Blair et al., 2010; Blair and Izquierdo, 2012; Izquierdo et al., 2018; Katuuramu et al., 2018). SdZn3.1 and SdZn10.1 were only identified in an inter-gene pool bi-parental population (Astudillo et al., 2013) and an Andean Diversity Panel (Katuuramu et al., 2018). Conversely, SdZn2.3 and SdZn8.2 only were reported in a population belonging to the Mesoamerican gene pool (Freyre et al., 1998; Blair et al., 2009; Izquierdo et al., 2018; Diaz et al., 2020). Mapping certain QTL in different gene pools demonstrates the complex control over micronutrient accumulation in the common bean that can vary between the Andean and Mesoamerican genotypes. It has been reported that Andean genotypes showed higher Fe concentrations than Mesoamerican genotypes and that Mesoamerican lines presented higher Zn contents than Andean lines (Islam et al., 2002). The high degree of correspondence with other reports for both Fe and Zn lends credibility to the current results.

One of the main challenges that common bean breeders working on biofortification face is the negative correlation between iron accumulation and yield (Raatz, 2018). Genotypes with high iron accumulation have been associated with poor yield and lower yield components (fewer pods, fewer seeds, fewer seeds per pod, or poorly filled seed) (Beebe, 2020). This same tendency is reflected in the correlations presented in **Figure 3** and **Supplementary Table 3** and has been reported previously (Diaz et al., 2020). Out of 14 loci for high iron, 6 were associated with low values of yield components (SdFe6.2, SdFe6.4, SdFe7.1, SdFe9.1, SdFe11.1, and SdFe11.2). These loci

could contribute as much as 39% of the total additive value for high iron. Similarly, out of 12 loci for high Zn, 6 were associated with negative effects on yield components (SdZn2.1, SdZn2.2, SdZn2.3, SdZn6.1, SdZn7.1, and SdZn8.2), summing up to 49% total additive effect for Zn concentration. Several of these "false" loci were reported by other authors, so this has been a problem in the genetic analysis of mineral concentration for several years. On the other hand, the QTLs SdFe3.2, SdFe5.1, SdFe6.1, SdFe6.3, SdFe7.2, SdFe8.1, and SdFe11.3 were not associated with loci related with yield components, making them attractive for breeding through marker-assisted selection (MAS) because their selection would not adversely affect the yield. Another advantage of these identified QTLs is that some of these loci have been found in populations belonging to both gene pools, making them useful for improving Andean or Mesoamerican genotypes. The sum of additive effects not associated with poor yield could contribute as much as 40 mg/kg Fe. Even though it is likely not possible to maintain the full additive effects of these loci through genetic recombination, ample variability exists without affecting yield. Similarly, there are QTLs for Zn that were not co-located with QTLs related to yield components (SdZn3.1, SdZn4.1, SdZn5.1, SdZn5.2, SdZn8.1, and SdZn10.1). These QTLs can be implemented in the breeding program, which theoretically could increase Zn seed concentration by 10 mg/kg. Nonetheless, if limits exist to raising Fe and/or Zn concentration with the QTL studied here, additional sources of genetic variability for high minerals are being explored. Beebe (2020) reported that interspecific progeny derived from Phaseolus species that evolved on alkaline soils might be more receptive to raising iron levels. Initial phenotypic results with these interspecific crosses are promising but have not been studied yet to determine their genetic makeup.

The observed negative associations of seed mineral concentration with PHI and Sd/Pd, and in two cases with 100SdW, probably reflect simple factors of concentration of minerals. This may be due to a dilution effect of Fe and Zn caused by high carbohydrate translocation to the grain, a typical characteristic of genotypes with high yield and PHI. This has also been reported in other crops, such as chickpea (Diapari et al., 2014). Conversely, if seed mass is reduced, then as the iron in the pods is translocated to the seed, the concentration in the seed will be increased arithmetically. Furthermore, where selection for high iron has received priority in breeding, this will have unconsciously limited yield potential. These results reveal many of the causes behind the persistent negative correlation of yield and high minerals that have been the greatest challenge in breeding biofortified beans.

A strong phenotypic correlation between iron and zinc concentrations was observed here and has been reported previously (Blair et al., 2009; Cichy et al., 2009). Izquierdo et al. (2018) reported loci that contributed to high concentrations of both elements. They identified a MetaQTL (MQTL) denominated "MQTL_Fe&Zn_6.2," containing eight QTLs identified in previous studies, such as sources of favorable alleles for Fe and Zn in both gene pools. MQTL_Fe&Zn_6.2 displayed higher PVE (27%) compared with other MQTLs identified in the same study. However, in this work, only one locus common to SdFe and SdZn was found, and this was associated with low values of PdN. This suggests that the correlation of Fe and Zn concentration that has been observed here and elsewhere may be driven in part by poor pod formation and/or poor pod and grain filling. The hotspot 12 is the only locus in which QTLs controlling the accumulation of Fe and Zn were co-located on Pv06 (∼28–29 Mpb).

The processes of iron and zinc uptake, transport, metabolism, and storage in sinks of interest are regulated by complex genetic mechanisms, of which several genes have been identified as possible candidates regulating these processes (Connorton et al., 2017). MQTL_Fe&Zn_6.2 mentioned above is located with Hotspot 12 (SdFe6.4/SdZn6.1), which is near the gene model Phvul.006G196600 that codifies a basic helix-loop-helix DNA-binding protein (bHLH). This gene is homologous to bHLH105/ILR3, an *Arabidopsis* gene codifying a transcriptional activator of responses to Fe deficiency, that also downregulates the expression of ferritin genes, implicated in the control of Fe homeostasis (Samira et al., 2018; Tissot et al., 2019; Gao et al., 2020). FER-LIKE REGULATOR OF IRON UPTAKE (FER) is required for induction of iron mobilization genes in Arabidopsis thaliana (Bauer et al., 2007). Phvul.005G430500 is homologous to FER and is located near SdFe5.1 (∼36.8 Mpb). Phvul.011G211900 is located near SdFe11.3 and is homologous to FRD3, which has a function of root xylem loading of iron chelator and efficient uptake out of the xylem into sink cells in Arabidopsis (Green and Rogers, 2004). AtNRAMP6 is a gene codifying a multispecific vacuolar metal transporter and plays an important role in the growth of *Arabidopsis* in Fe-deficient conditions (Li et al., 2019). Phvul.005G182000 (corresponding to SdFe5.1) and Phvul.009G127900 (corresponding to SdFe9.1) are homologous to AtNRAMP6. Phvul.003G143800 is homologous of AHA2, a proton pump H+-ATPASE2 controlling acidification-reduction-transport of iron uptake in Arabidopsis (Campos, 2020). The Zrt and Irt-like protein (ZIP) is well characterized for its role in Zn and Fe transporter (Astudillo et al., 2013); Phvul.002G184200 (SdZn2.2), Phvul.008G259200 (SdZn8.2), Phvul.010G059200 (SdZn10.1), and Phvul.011G058500 (SdFe11.1) have been identified as possible candidate genes belonging to ZIP metal ion transporter family.

In terms of conventional breeding, successful cases have already been reported using these populations evaluated for obtaining varieties with high accumulation of Fe and acceptable yield in different conditions: BIO-101 in Colombia (83 mg/kg of Fe and 44 mg/kg of Zn),[3] SMR100 in Nicaragua (89 mg/kg of Fe and 35 mg/kg of Zn),[4] and INTA Bio-Apante in Nicaragua[5] were biofortified lines that used RILs of SMC40 × SCR16 population as parental materials.

An effect of drought was observed only on yield components, while Fe and Zn accumulation was not affected by the drought stress. The positive correlation between yield components has previously been reported under drought conditions (Polania et al., 2016a,b). Regarding the invariable Fe and Zn contents

[3]https://lac.harvestplus.org/en/new-biofortified-bean-varieties-bio-101-bio-107/

[4]https://inta.gob.ni/wp-content/uploads/2019/11/Brochure-Frijol-Biofortificado-HarvesPlus-2019.pdf

[5]https://lac.harvestplus.org/ometepe-biofortificados-nicaragua/

under stress conditions, Smith et al. (2019) explained that seeds have an evolutionary tendency to ensure adequate resources for germination and present buffering effects to maintain similar nutrient concentrations in drought and irrigated conditions, but the explicit mechanisms are not clear.

Pod harvest index expresses the ratio between the biomass in seed and pod and is a useful indicator of the remobilization of photosynthates to seed (Assefa et al., 2013). However, until now, it is not clear yet which of these two components have the key effect on PHI, and if there is a possible genetic relationship that explains the different behavior of PHI in each genotype. In this study, a stronger phenotypic correlation was observed between PHI with Sd/Pd and SdN than with PdN in all populations. However, the phenotypic correlation of PHI with 100SdW was low or not significant in this study (**Figure 3**), similar to results reported by Diaz et al. (2020) and Berny Mier Y Teran et al. (2019). Berny Mier Y Teran et al. (2019) suggested that sink strength due to the number of seeds is higher than sink strength based on the seed size. This could indicate that a key factor that affects PHI is the number of seeds rather than seed weight itself. In the genetic analysis, most QTLs identified for the number of seeds per pod (Sd/Pd) were co-localized with QTLs identified for SdN, and Sd/Pd has a high positive correlation with PHI. Moreover, all QTLs of PdN having a positive additive effect were associated with the SMC allele, which was parental with high micronutrient concentration but poor pod filling.

The main QTL found in this study for PHI (PHI 7.2), with the highest PVE and additive effect for the trait (**Table 2**), has been reported previously in Mesoamerican populations (Diaz et al., 2018; Berny Mier Y Teran et al., 2019) and in an Andean population (Sedlar et al., 2020), and could be used for breeding in both gene-pools. PHI1.1 and PHI8.1 have been identified previously in a Mesoamerican MAGIC population (Diaz et al., 2020). PHI1.2, PHI5.1, and PHI7.3 have been described in previous studies in Mesoamerican bi-parental populations (Asfaw et al., 2012; Diaz et al., 2018; Berny Mier Y Teran et al., 2019). PHI5.2 was reported by Sedlar et al. (2020) in an Andean bi-parental population. These QTLs seem to be independent of micronutrients content loci and could be used for MAS and theoretically could enhance PHI values by >20%, which discovers these markers very promising for breeding for drought tolerance and/or high yield. Other QTLs for other yield related-traits have also been reported, Berny Mier Y Teran et al. (2019) reported a QTL for seeds per pod near to Sd/Pd1.1. Similarly, Sedlar et al. (2020) reported a QTL near to Sd/Pd2.1, the QTL with the highest PVE. Other QTLs, such as PdN2.2, PdN4.1, SdN2.1, and SdN6.1, were reported previously in other Mesoamerican bi-parental populations (Diaz et al., 2017). Mukeshimana et al. (2014) found a QTL near to PdN8.2, this QTL was present in multiple trials.

One issue of PHI that has been discussed little is its role as a domestication trait associated with greater partitioning to grain, and whether the selection of genotypes with better pod filling was linked to other traits related to the domestication of common beans, such as pod indehiscence and determinate growth habit. PHI5.2 and PHI8.1 were co-located in the same genomic regions that controlled pod fiber or dehiscence for common bean (Parker et al., 2020). In the same way, PHI1.2 and PHI1.3 are near to the *fin* locus, a key region that controlled determinacy and photoperiod sensitivity in common bean (Kwak et al., 2012). This could indicate that PHI reflects traits that were enhanced in domestication and that loci associated with PHI could be related to domestication.

Understanding the segregation variation between the populations. Heterozygosity information in the populations is presented (**Supplementary Figure 5**). The utility of the early generation's higher variability could be implemented as sources as a pedigree follow-up. In consequence, the analysis of the use of F_3 and F_4 populations was mandatory previously for further selection cycles.

CONCLUSION

This study evaluated the phenotypic and genetic correlation between Fe/Zn accumulation and yield components using QTL analysis in three biofortified populations developed in the CIAT bean-breeding program. The results in this work validate the negative phenotypic correlation previously described between yield and Fe/Zn accumulation in phenotypic terms. In addition, we are identifying several loci controlling both traits. We hypothesize this strong correlation is a dilution effect of Fe and Zn caused by high carbohydrate translocation to grain. Conversely, loci associated with high minerals and low values of yield components should be avoided. On the other hand, we observed QTL for Fe/Zn segregating independently of the other QTL for yield components, possibly mediated by transporters with affinity to minerals. The selection of these regions will enable enhanced levels of Fe/Zn and should not affect yield performance. QTLs observed in this and previous studies suggest similar genetic control in both gene pools at some loci and these QTLs show a high potential to be implemented in a breeding program focused on biofortification.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LI31YG.

AUTHOR CONTRIBUTIONS

SB conceived the study and designed the populations. CC and MG carried out the breeding populations advancement. JP did the field trial evaluation. SD did the phenotypic data analysis and linkage analysis. BR supervised the molecular laboratory. SD and DA-S performed the genotypic data analysis. SD and SB wrote

the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This study was funded by Tropical Legumes III – Improving Livelihoods for Smallholder Farmers: Enhanced Grain Legume Productivity and Production in Sub-Saharan Africa and South Asia (OPP1114827) projects funded by the Bill & Melinda Gates Foundation. This project was supported by the CGIAR Research Program on Grain Legumes and we would like to thank USAID for their contributions, particularly in the form of the "Global Hunger and Food Security Research Strategy: Climate Resilience, Nutrition, and Policy – Feed the Future Innovation Lab for Climate Resilience in Beans (CRIB)" Project (prime award number AID-0AA-A-13-00077, sub-award number 4943-CIDT-AID-0077).

ACKNOWLEDGMENTS

We would like to thank the contributions and technical support made by the bean breeding and physiology team at the International Center for Tropical Agriculture (CIAT) for phenotyping work and population development. We are also thankful to HarvestPlus for providing support on the micromineral analysis work and the Data management team at CIAT for making the datasets used in this work available with open access.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.739033/full#supplementary-material

Supplementary Figure 1 | Precipitation, maximum, and minimum temperatures during trials at Palmira, Colombia.

Supplementary Figure 2 | Histograms of frequency distribution for Fe and Zn in seven trials of the SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations.

Supplementary Figure 3 | Genome-wide association analysis for Fe accumulation (SdFe) showing Manhattan and Q-Q plot for SMC × SCR populations with 24,785 markers in seven trials. The Bonferroni correction threshold ($p = 0.05$) is presented as a red horizontal line based on the number of markers for the population.

Supplementary Figure 4 | Genome-wide association analysis for Zn accumulation (SdZn) showing Manhattan and Q-Q plot for SMC × SCR populations with 24,785 markers in seven trials. The Bonferroni correction threshold ($p = 0.05$) is presented as a red horizontal line based on the number of markers for the population.

Supplementary Figure 5 | Heterozygosity of individuals and heterozygosity of markers in the SMC × SCR populations.

Supplementary Table 1 | Phenotypic variability and broad-sense heritability (H^2) for evaluated traits in the trials 2011, 2012, 2013, 2014, and 2015 of the SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations.

Supplementary Table 2 | Component variance, F statistic evaluated for seven trials to SdFe, seed iron; SdZn, seed zinc; PHI, pod harvest index; PdN, pod number; SdN, seed number; Sd/Pd, number of seeds per pod; 100SdW, 100 seed weight in SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations.

Supplementary Table 3 | Pearson's correlations coefficients between evaluated traits on three populations SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 for each trial and each trait. Significance of correlations indicated as $^{***}p < 0.0001$; $^{**}p < 0.001$; $^{*}p < 0.01$; ns, not significant. SdFe, seed iron; SdZn, seed zinc; PHI, pod harvest index; PdN, pod number; SdN, seed number; Sd/Pd, number of seeds per pod; 100SdW, 100 seed weight.

Supplementary Table 4 | Details of all QTL identified by inclusive composite interval mapping for additive effects (ICIM-ADD), genetic and physical position, LOD, phenotypic variation explained mapped for six traits in SCR16 × SMC40, SMC33 × SCR16, and SMC44 × SCR9 populations evaluated in 2011, 2012, 2013, 2014, and 2015.

REFERENCES

Amanullah, S., Gao, P., Agyei, B., Saroj, A., Yang, T., Liu, S., et al. (2021). Genetic linkage mapping and QTLs identification for morphology and fruit quality related traits of melon by SNP based CAPS markers. *Sci. Hortic.* 278:109849. doi: 10.1016/j.scienta.2020.109849

Asfaw, A., Blair, M., and Struik, P. (2012). Multienvironment quantitative trait loci analysis for photosynthate acquisition, accumulation, and remobilization traits in common bean under drought stress. *G3 Genes Genomes Genetics* 2, 579–595. doi: 10.1534/g3.112.002303

Assefa, T., Beebe, S., Rao, I., Cuasquer, J., Duque, M., Rivera, M., et al. (2013). Pod harvest index as a selection criterion to improve drought resistance in white pea bean. *Field Crops Res.* 148, 24–33. doi: 10.1016/j.fcr.2013.04.008

Astudillo, C., Fernandez, A., Blair, M., and Cichy, K. (2013). The phaseolus vulgaris ZIP gene family: identification, characterization, mapping, and gene expression. *Front. Plant Sci.* 4:286. doi: 10.3389/fpls.2013.00286

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using Lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bauer, P., Ling, H., and Guerinot, M. (2007). FIT, the fer-like iron deficiency induced transcription factor in *Arabidopsis*. *Plant Physiol. Biochem.* 45, 260–261. doi: 10.1016/j.plaphy.2007.03.006

Beaver, J., and Osorno, J. (2009). Achievements and limitations of contemporary common bean breeding using conventional and molecular approaches. *Euphytica* 168, 145–175. doi: 10.1007/s10681-009-9911-x

Beebe, S. (2012). Common bean breeding in the tropics. *Plant Breed. Rev.* 36, 357–426. doi: 10.1002/9781118358566

Beebe, S. (2020). Biofortification of common bean for higher iron concentration. *Front. Sustain. Food Syst.* 4:573449. doi: 10.3389/fsufs.2020.573449

Beebe, S., Gonzalez, A., and Rengifo, J. (2000). Research on trace minerals in the common bean. *Food Nutr. Bull.* 21, 387–391. doi: 10.1177/156482650002100408

Beebe, S., Rao, I., Cajiao, C., and Grajales, M. (2008). Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Sci.* 48, 582–592. doi: 10.2135/cropsci2007.07.0404

Berny Mier Y Teran, J., Konzen, E., Palkovic, A., Tsai, S., Rao, I., et al. (2019). Effect of drought stress on the genetic architecture of photosynthate allocation and remobilization in pods of common bean (*Phaseolus vulgaris* L.), a key species for food security. *BMC Plant Biol.* 19:171. doi: 10.1186/s12870-019-1774-2

Blair, M., Astudillo, C., Grusak, M., Graham, R., and Beebe, B. (2009). Inheritance of seed iron and zinc concentrations in common bean (*Phaseolus vulgaris* L.). *Mol. Breed.* 23, 197–207. doi: 10.1007/s11032-008-9225-z

Blair, M., Astudillo, C., Rengifo, J., Beebe, S., and Graham, R. (2011). QTL analyses for seed iron and zinc concentrations in an intra-genepool population of andean common beans (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 122, 511–521. doi: 10.1007/s00122-010-1465-8

Blair, M., and Izquierdo, P. (2012). Use of the advanced backcross-QTL method to transfer seed mineral accumulation nutrition traits from wild to andean cultivated common beans. *Theor. Appl. Genet.* 125, 1015–1031. doi: 10.1007/s00122-012-1891-x

Blair, M., Medina, J., Astudillo, C., Rengifo, J., Beebe, S., Machado, G., et al. (2010). QTL for seed iron and zinc concentration and content in a mesoamerican common bean (*Phaseolus vulgaris* L.) population. *Theor. Appl. Genet.* 121, 1059–1070. doi: 10.1007/s00122-010-1371-0

Bolger, A., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bouis, H. (2003). Micronutrient fortification of plants throught plant breeding: can it improve nutrition in man at low cost? *Proc. Nutr. Soc.* 62, 403–411. doi: 10.1079/pns2003262

Bouis, H., and Welch, R. (2010). Biofortification—a sustainable agricultural strategy for reducing micronutrient malnutrition in the Global South. *Crop Sci.* 50, S20–S32. doi: 10.2135/cropsci2009.09.0531

Broughton, W., Hernandez, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* Spp.) – model food legumes. *Plant Soil* 252, 55–128. doi: 10.1023/A:1024146710611

Campos, M. (2020). Metallic action! The dynamics of a tripartite iron uptake complex in *Arabidopsis* roots. *Plant Physiol.* 184, 1212–1213. doi: 10.1104/pp.20.01271

Caproni, L., Raggi, L., Talsma, E., Wenzl, P., and Negri, V. (2020). European landrace diversity for common bean biofortification: a genome-wide association Study. *Sci. Rep.* 10, 1–13. doi: 10.1038/s41598-020-76417-3

Chasapis, C., Spiliopoulou, C., Loutsidou, A., and Stefanidou, M. (2012). Zinc and human health: an update. *Arch. Toxicol.* 86, 521–534. doi: 10.1007/s00204-011-0775-1

Cichy, K., Caldas, G., Snapp, S., and Blair, M. (2009). QTL Analysis of Seed iron, zinc, and phosphorus levels in an andean bean population. *Crop Sci.* 49, 1742–1750. doi: 10.2135/cropsci2008.10.0605

Cichy, K., Fernandez, A., Kilian, A., Kelly, J., Galeano, G., Shaw, S., et al. (2014). QTL analysis of canning quality and color retention in black beans (*Phaseolus vulgaris* L.). *Mol. Breed.* 33, 139–154. doi: 10.1007/s11032-013-9940-y

Connorton, J., Balk, J., and Rodríguez-Celma, J. (2017). Iron homeostasis in plants-a brief overview. *Metallomics* 9, 813–823. doi: 10.1039/c7mt00136c

Cullis, B., Smith, A., and Coombes, N. (2006). On the design of the early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11, 381–393. doi: 10.1198/108571106X154443

Diapari, M., Sindhu, A., Bett, K., Deokar, A., Warkentin, T., Taran, B., et al. (2014). Genetic diversity and association mapping of iron and zinc concentrations in chickpea (*Cicer arietinum* L.). *Genome* 57, 459–468. doi: 10.1139/gen-2014-0108

Diaz, L., Ricaurte, J., Cajiao, C., Galeano, C., Rao, I., Beebe, S., et al. (2017). Phenotypic evaluation and QTL analysis of yield and symbiotic nitrogen fixation in a common bean population grown with two levels of phosphorus supply. *Mol. Breed.* 37:76. doi: 10.1007/s11032-017-0673-1

Diaz, L., Ricaurte, J., Tovar, E., Cajiao, C., Terán, H., Grajales, M., et al. (2018). QTL analyses for tolerance to abiotic stresses in a common bean (*Phaseolus vulgaris* L.) population. *PLoS One* 13:e0202342. doi: 10.1371/journal.pone.0202342

Diaz, S., Ariza-Suarez, D., Izquierdo, P., Lobaton, J., De La Hoz, J., Acevedo, F., et al. (2020). Genetic mapping for agronomic traits in a MAGIC population of common bean (*Phaseolus vulgaris* L.) under drought conditions. *BMC Genomics* 21:799. doi: 10.1186/s12864-020-07213-6

Diaz, S., Ariza-Suarez, D., Ramdeen, R., Aparicio, J., Arunachalam, N., Hernandez, C., et al. (2021). Genetic architecture and genomic prediction of cooking time in common bean (*Phaseolus vulgaris* L.). *Front. Plant Sci.* 11:622213. doi: 10.3389/fpls.2020.622213

Elshire, R., Glaubitz, J., Sun, Q., Poland, J., Kawamoto, K., Buckler, E., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e0019379. doi: 10.1371/journal.pone.0019379

Freyre, R., Skroch, P., Geffroy, V., Adam-Blondon, A., Shirmohamadali, A., Johnson, W., et al. (1998). Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. *Theor. Appl. Genet.* 97, 847–856. doi: 10.1007/s001220050964

Gao, F., Robe, K., Bettembourg, M., Navarro, N., Rofidal, V., Santoni, V., et al. (2020). The transcription factor BHLH121 interacts with BHLH105 (ILR3) and its closest homologs to regulate iron homeostasis in *Arabidopsis*. *Plant Cell* 32, 508–524. doi: 10.1105/tpc.19.00541

Gonzalez, A., Yuste-Lisbona, F., Fernandez-Lozano, A., Lozano, R., and Santalla, M. (2017). "Genetic mapping and QTL analysis in common bean," in *The Common Bean Genome*, eds M. P. de la Vega, M. Santalla, and F. Marsolais (Berlin: Springer), 69–108.

Green, L., and Rogers, E. (2004). FRD3 controls iron localization in *Arabidopsis*. *Plant Physiol.* 136, 2523–2531. doi: 10.1104/pp.104.045633

Guild, G., Paltridge, N., Andersson, M., and Stangoulis, J. (2017a). An energy-dispersive X-Ray fluorescence method for analysing Fe and Zn in common bean, maize and cowpea biofortification programs. *Plant Soil* 419, 457–466. doi: 10.1007/s11104-017-3352-4

Guild, G., Parkes, E., Nutti, M., Palacio-Rojas, N., and Stangoulis, J. (2017b). High-throughput measurement methodologies for developing nutrient-dense crops. *Afr. J. Food Agric. Nutr. Dev.* 17, 11941–11954. doi: 10.18697/ajfand.78.HarvestPlus06

Gunjača, J., Carović, K., Lazarević, B., Vidak, M., Petek, M., Liber, Z., et al. (2021). Genome-wide association studies of mineral content in common bean. *Front. Plant Sci.* 12:636484. doi: 10.3389/fpls.2021.636484

Hambidge, M. (2000). Zinc and health: current status and future directions: introduction. *J. Nutr.* 130(5 Suppl.), 1344–1349.

Han, K., Lee, H., Ro, N., Hur, O., Lee, J., Kwon, J., et al. (2018). QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in Capsicum. *Plant Biotechnol. J.* 16, 1546–1558. doi: 10.1111/pbi.12894

Hummel, M., Hallahan, B., Brychkova, G., Ramirez-Villegas, J., Guwela, V., Chataika, B., et al. (2018). Reduction in nutritional quality and growing area suitability of common bean under climate change induced drought stress in Africa. *Sci. Rep.* 8:16187. doi: 10.1038/s441598-018-33952-4

Islam, F., Basford, K., Jara, C., Redden, R., and Beebe, S. (2002). Seed compositional and disease resistance differences among gene pools in cultivated common bean. *Genet. Resour. Crop Evol.* 49, 285–293. doi: 10.1023/A:1015510428026

Islam, M., Thyssen, G., Jenkins, J., Zeng, L., Delhom, C., McCarty, J., et al. (2016). A MAGIC population-based genome-wide association study reveals functional association of GhRBB1_A07 gene with superior fiber quality in cotton. *BMC Genomics* 17:903. doi: 10.1186/s12864-016-3249-2

Izquierdo, P., Astudillo, C., Blair, M., Iqbal, A., Raatz, B., and Cichy, K. (2018). Meta-QTL analysis of seed iron and zinc concentration and content in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 131, 1645–1658. doi: 10.1007/s00122-018-3104-8

Jha, A., and Warkentin, D. (2020). Biofortification of pulses crops: status and future perspectives. *Plants* 9:73. doi: 10.3390/plants9010073

Johnson, R., Nelson, G., Troyer, J., Lautenberger, J., Kessing, B., Winkler, C., et al. (2010). Accounting for multiple comparisons in a Genome-Wide Association Study (GWAS). *BMC Genomics* 11:724. doi: 10.1186/1471-2164-11-724

Katuuramu, D., Hart, J., Porch, T., Grusak, M., Glahn, R., and Cichy, K. (2018). Genome-wide association analysis of nutritional composition-related traits and iron bioavailability in cooked dry beans (*Phaseolus vulgaris* L.). *Mol. Breed.* 38, 1–18. doi: 10.1007/s11032-018-0798-x

Kumar, S., and Pandey, G. (2020). Biofortification of pulses and legumes to enhance nutrition. *Heliyon* 6:e03682. doi: 10.1016/j.heliyon.2020.e03682

Kwak, M., Toro, O., Debouck, D., and Gepts, P. (2012). Multiple origins of the determinate growth habit in domesticated common bean (*Phaseolus vulgaris*). *Ann. Bot.* 110, 1573–1580. doi: 10.1093/aob/mcs207

Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lauer, J., Bijl, C., Grusak, M., Baenziger, P., Boote, K., Lingle, S., et al. (2012). The scientific grand challenges of the 21st century for the crop science society of America. *Crop Sci.* 52, 1003–1010. doi: 10.2135/cropsci2011.12.0668

Li, J., Wang, Y., Zheng, L., Li, Y., Zhou, X., Li, J., et al. (2019). The intracellular transporter AtNRAMP6 is involved in fe homeostasis in *Arabidopsis*. *Front. Plant Sci.* 10:1124. doi: 10.3389/fpls.2019.01124

Lobaton, J., Miller, T., Gil, J., Ariza, D., de la Hoz, J., Soler, A., et al. (2018). Resequencing of common bean identifies regions of inter-gene pool introgression and provides comprehensive resources for molecular breeding. *Plant Genome* 11, 1–21. doi: 10.3835/plantgenome2017.08.0068

Lopez, A., Cacoub, P., Macdougall, I., and Peyrin-Biroulet, L. (2016). Iron deficiency anaemia. *Lancet* 387, 907–916. doi: 10.1016/S0140-6736(15)60865-0

Ma, Y., Coyne, C., Grusak, M., Mazourek, M., Cheng, P., Main, D., et al. (2017). Genome-wide SNP identification, linkage map construction and QTL mapping for seed mineral concentrations and contents in pea (*Pisum sativum* L.). *BMC Plant Biol.* 17:43. doi: 10.1186/s12870-016-0956-4

Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001

Mukeshimana, G., Butare, L., Cregan, P., Blair, M., and Kelly, J. (2014). Quantitative trait loci associated with drought tolerance in common bean. *Crop Sci.* 54, 923–938. doi: 10.2135/cropsci2013.06.0427

Parker, T., Berny Mier y Teran, J., Palkovic, A., Jernstedt, J., and Gepts, P. (2020). Pod indehiscence is a domestication and aridity resilience trait in common bean. *New Phytol.* 225, 558–570. doi: 10.1111/nph.16164

Perea, C., De La Hoz, J., Cruz, D., Lobaton, J., Izquierdo, P., Quintero, J., et al. (2016). Bioinformatic analysis of genotype by sequencing (GBS) data with NGSEP. *BMC Genomics* 17(Suppl.e 5):539. doi: 10.1186/s12864-016-2827-7

Petry, N., Boy, E., Wirth, J., and Hurrell, R. (2015). Review: the potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification. *Nutrients* 7, 1144–1173. doi: 10.3390/nu7021144

Polania, J., Poschenrieder, C., Beebe, S., and Rao, I. (2016a). Effective use of water and increased dry matter partitioned to grain contribute to yield of common bean improved for drought resistance. *Front. Plant Sci.* 7:660. doi: 10.3389/fpls.2016.00660

Polania, J., Rao, I., Cajiao, C., Rivera, M., Raatz, B., and Beebe, S. (2016b). Physiological traits associated with drought resistance in andean and mesoamerican genotypes of common bean (*Phaseolus vulgaris* L.). *Euphytica* 210, 17–29. doi: 10.1007/s10681-016-1691-5

Raatz, B. (2018). Biofortification of grain legumes. *Achiev. Sustain. Cultiv. Grain Legumes* 1, 177–200. doi: 10.19103/as.2017.0023.06

Sab, S., Lokesha, R., Mannur, D. M., Somasekhar, Jadhav, K., Mallikarjuna, B. P., et al. (2020). Genome-wide SNP discovery and mapping QTLs for seed iron and zinc concentrations in chickpea (*Cicer arietinum* L.). *Front. Nutr.* 7:559120. doi: 10.3389/fnut.2020.559120

Samira, R., Li, B., Kliebenstein, D., Li, C., Davis, E., Gillikin, J., et al. (2018). The BHLH transcription factor ILR3 modulates multiple stress responses in *Arabidopsis*. *Plant Mol. Biol.* 97, 297–309. doi: 10.1007/s11103-018-0735-8

Schmutz, J., McClean, P., Mamidi, S., Wu, G., Cannon, S., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008

Sedlar, A., Zupin, M., Maras, M., Razinger, J., Šuštar-Vozlič, J., Pipan, B., et al. (2020). QTL mapping for drought responsive agronomic traits associated with physiology, phenology, and yield in an andean intra-gene pool common bean population. *Agronomy* 10:225. doi: 10.3390/agronomy10020225

Shiva, Y., Singh, U., Prasad, R., and Kaur, R. (2016). Agronomic interventions for micronutrient biofortification of pulses. *Indian J. Agron.* 61, S161–S172. doi: 10.1002/jsfa.11059

Smith, M., Veneklaass, E., Polania, J., Rao, I., Beebe, S., and Merchant, A. (2019). Field drought conditions impact yield but not nutritional quality of the seed in common bean (*Phaseolus vulgaris* L.). *PLoS One* 14:e0217099. doi: 10.1371/journal.pone.0217099

Stangoulis, J., and Sison, C. (2008). Crop sampling protocols for micronutriens analysis. *Harvest Plus Tech. Monogr. Ser.* 7, 1–20.

Steinbicker, A., and Muckenthaler, M. (2013). Out of balance-systemic iron homeostasis in iron-related disorders. *Nutrients* 5, 3034–3061. doi: 10.3390/nu5083034

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT Version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9, 1–9. doi: 10.3835/plantgenome2015.11.0120

Tello, D., Gil, J., Loaiza, C., Riascos, J., Cardozo, J., and Duitama, J. (2019). NGSEP3: accurate variant calling across species and sequencing protocols. *Bioinformatics* 35, 4716–4723. doi: 10.1093/bioinformatics/btz275

Tissot, N., Robe, K., Gao, F., Grant-Grant, S., Boucherez, J., Bellegarde, F., et al. (2019). Transcriptional integration of the responses to iron availability in *Arabidopsis* by the BHLH factor ILR3. *New Phytol.* 223, 1433–1446. doi: 10.1111/nph.15753

Tryphone, G., and Nchimbi-Msolla, S. (2010). Diversity of common bean (*Phaseolus vulgaris* l.) genotypes in iron and zinc contents under screenhouse conditions. *Afr. J. Agric. Res.* 5, 738–747. doi: 10.5897/AJAR10.304

Turner, S. (2018). Qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Sour. Softw.* 3:731. doi: 10.21105/joss.00731

Upadhyaya, H., Bajaj, D., Das, S., Kumar, V., Gowda, C. L. L., Sharma, S., et al. (2016). Genetic dissection of seed-iron and zinc concentrations in chickpea. *Sci. Rep.* 6:24050. doi: 10.1038/srep24050

VanRaden, P. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Diaz, Polania, Ariza-Suarez, Cajiao, Grajales, Raatz and Beebe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

PERSPECTIVE
published: 05 January 2022
doi: 10.3389/fpls.2021.810692

Adaptive Mechanisms Make Lupin a Choice Crop for Acidic Soils Affected by Aluminum Toxicity

Miguel A. Quiñones, M. Mercedes Lucas and José J. Pueyo**

Institute of Agricultural Sciences (ICA), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain

Almost half of the world's agricultural soils are acidic, and most of them present significant levels of aluminum (Al) contamination, with Al^{3+} as the prevailing phytotoxic species. Lupin is a protein crop that is considered as an optimal alternative to soybean cultivation in cold climates. Lupins establish symbiosis with certain soil bacteria, collectively known as rhizobia, which are capable of fixing atmospheric nitrogen. Moreover, some lupin species, especially white lupin, form cluster roots, bottlebrush-like structures specialized in the mobilization and uptake of nutrients in poor soils. Cluster roots are also induced by Al toxicity. They exude phenolic compounds and organic acids that chelate Al to form non-phytotoxic complexes in the rhizosphere and inside the root cells, where Al complexes are accumulated in the vacuole. Lupins flourish in highly acidic soils where most crops, including other legumes, are unable to grow. Some lupin response mechanisms to Al toxicity are common to other plants, but lupin presents specific tolerance mechanisms, partly as a result of the formation of cluster roots. Al-induced lupin organic acid secretion differs from P-induced secretion, and organic acid transporters functions differ from those in other legumes. Additionally, symbiotic rhizobia can contribute to Al detoxification. After revising the existing knowledge on lupin distinct Al tolerance mechanisms, we conclude that further research is required to elucidate the specific organic acid secretion and Al accumulation mechanisms in this unique legume, but definitely, white lupin arises as a choice crop for cultivation in Al-rich acidic soils in temperate climate regions.

Keywords: lupin, *Lupinus albus*, acidic soils, aluminum, cluster roots, *Bradyrhizobium*

OPEN ACCESS

Edited by:
Alfonso Clemente, Estación Experimental del Zaidín (CSIC), Spain

Reviewed by:
Ricardo A. Cabeza, University of Talca, Chile
Sergio J. Ochatt, INRA UMR 1347 Agroécologie, France

****Correspondence:***
Miguel A. Quiñones
ma.quinones@csic.es
José J. Pueyo
jj.pueyo@csic.es

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: *07 November 2021*
Accepted: *14 December 2021*
Published: *05 January 2022*

Citation:
Quiñones MA, Lucas MM and Pueyo JJ (2022) Adaptive Mechanisms Make Lupin a Choice Crop for Acidic Soils Affected by Aluminum Toxicity. Front. Plant Sci. 12:810692. doi: 10.3389/fpls.2021.810692

INTRODUCTION

White lupin (*Lupinus albus* L.) is a singular legume with increasing value both as an agronomic and a horticultural crop (Jansen, 2006). Lupins seeds present protein contents close to 40% and are among the most protein-rich plant products, making this legume an excellent protein crop and a viable alternative to soybean cultivation in cold climates (Lucas et al., 2015; De Ron et al., 2017). As a horticultural crop, white lupins play a role in ecosystem services as honey plants and annual ornamentals (Jansen, 2006). White lupin has a notable tolerance to abiotic stresses (Fernández-Pascual et al., 2007), and a considerable potential as a tool for the recovery of degraded soils and phytoremediation of toxic metal-polluted soils (Coba de la Peña and Pueyo, 2012; Quiñones et al., 2013, 2021). Lupins establish symbiosis with soil bacteria leading to the formation

of nitrogen-fixing root nodules. Moreover, several lupin species are capable of developing the so-called cluster roots (CR) (Lambers et al., 2003; Pueyo et al., 2021). P or Fe deficiency and Al toxicity, which often coexists with P deficiency in globally-distributed acidic soils, induce the formation of CR, which promote mobilization of P and other nutrients by secretion of organic acids (OAs), protons, phosphatases and phenolic compounds (Sun et al., 2016). These distinctive traits confer white lupin the competence to grow in N- and P-deprived soils, mobilizing other nutrients, which are unavailable to most plants. Unlike other legumes, lupins can grow optimally at pH > 4.0, and sub-optimally at even lower pHs (Nelson et al., 2001; Zavalin et al., 2019).

It is estimated that roughly 30% of the world's soils are acidic (Zheng, 2010). In many areas of the Pacific Ring of Fire, volcanic-originated soils are particularly acidic, Al-toxic and P-deficient Andosols that comprise approximately 50% of arable land in Chile (Redel et al., 2016). Acidic soils are also predominant in Europe, mostly in the Northern regions (**Figure 1**). Soil acidification is accentuated by some agricultural practices (Moore et al., 2001). Climate change and the fluctuations in rainfall patterns also contribute to soil acidification. While other metals are more toxic, aluminum constitutes the main problem of acidic soils, which depends on soil pH, but also on factors such as organic matter, phosphates, fluorides, sulfates or ionic strength (Carr and Ritchie, 1993). Acidic soils are usually rich in clay minerals such as iron oxides, kaolinite, montmorillonite and illite that trap P and render it inaccessible for plants (Zheng, 2010; Eriksson et al., 2016). Therefore, P deficiency and Al toxicity coexist in acidic soils severely hampering crop production. Lime and P applications are effective in the short term, they elevate the soil pH, reduce Al solubility, which is also displaced from exchangeable sites by Ca, and increase P availability. However, the soil buffering capacity can diminish the effects of the amendments. Acidification is increased by factors such as acid rain and application of ammonium-based fertilizers, and amendments need to be reapplied, which might not be sustainable in the long term (Zheng, 2010). A crop such as lupin, which tolerates Al toxicity and displays mechanisms to fix atmospheric nitrogen and mobilize P and other nutrients in nutrient-deprived acidic soils, arises as an ideal option for sustainable agriculture in temperate zones worldwide.

In plants, Al can cause cytotoxic inhibition, damages to cell structures and eventual suppression of plant growth and crop yield (Andersson, 1988). Al toxicity symptoms include disturbance of the plasma membrane, decrease in water uptake, damage to photosynthesis, and generation of reactive oxygen species (ROS) (Wei et al., 2021). Al inhibits auxin transport, induces ethylene production and reduces root growth (Chauhan et al., 2021; Wei et al., 2021). Al causes epidermal cell death, decreasing cell extensibility and conductivity (Chauhan et al., 2021). Al interferes with the uptake, transport and metabolism of multiple mineral nutrients (Andersson, 1988). It alters the expression of numerous genes, it affects the ability of DNA to replicate, and disturbs signal transduction pathways (Sade et al., 2016). Some plants have developed Al tolerance through root secretion of chelating compounds. Several transporters involved in Al resistance mechanisms belong to the multidrug and toxic compound extrusion (MATE), the aluminum-activated malate transporter (ALMT), the natural resistance associated macrophage protein (NRAMP), the ATP-binding cassette (ABC), and the aquaporin families (Wei et al., 2021). OAs and phenolic compounds bind Al and exclude it from the cell and/or sequester the resulting complexes into the vacuole or the cell wall (Sade et al., 2016). A fine regulation of the response to Al stress is achieved by the involvement of phytohormones, Ca^{2+} signaling pathways, transcription factors and miRNAs (Matsumoto, 2000).

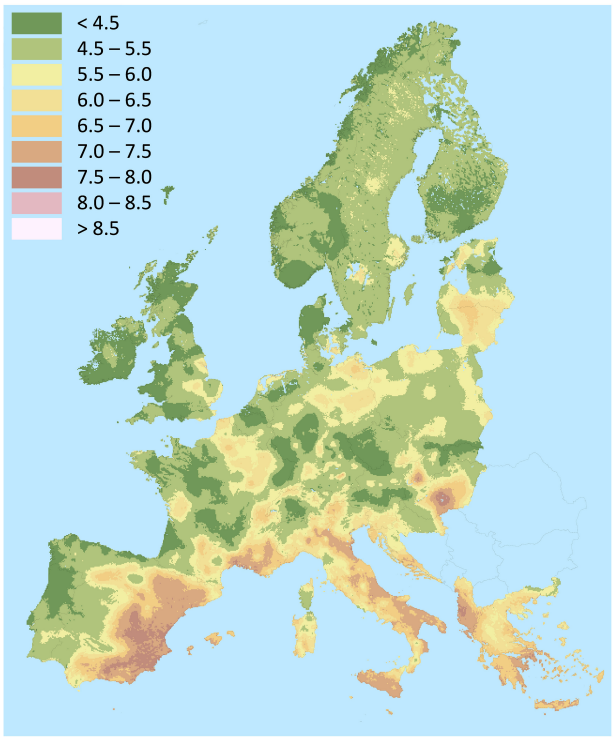


FIGURE 1 | Map of soil pH in Europe. Source: Land Resources Management Unit, Institute for Environment and Sustainability, European Soil Data Centre (ESDAC), esdac.jrc.ec.europa.eu, European Commission, Joint Research Centre (Panagos et al., 2012).

PHYSIOLOGICAL ALTERATIONS AND SPECIFIC MECHANISMS INVOLVED IN LUPIN TOLERANCE TO ALUMINUM

White lupin response to Al toxicity involves physiological and morphological modifications and mechanisms that differ from other plants, including other legumes. Due to its high tolerance, Al toxicity has been scarcely investigated in lupin, but at high enough concentrations, Al can cause a suppression of vegetative growth, with decreases in shoot and root biomasses (Hemada et al., 2020). Lupin CR exudates limit the entrance of Al into the roots through the formation of non-toxic complexes (Dakora and Phillips, 2002). Malate, citrate and oxalate are

involved in Al resistance in plants (Ma, 2000; Ma et al., 2001). Malate and citrate are secreted from root tips and regulated by specific transporters (Sharma et al., 2016; Wang et al., 2017); however, oxalate secretion transport pathways are mostly unknown (Lv et al., 2021). CR exudates consist mainly of OA anions (Dessureault-Rompré et al., 2007), protons, phenolic compounds and phosphatases (Gerke et al., 1994; Gilbert et al., 1999; Lamont, 2003) that generate changes in the rhizosphere and mobilize nutrients and toxic metals from insoluble phosphates (Dakora and Phillips, 2002). OAs act as detoxifying agents by chelating Al into non-phytotoxic forms (Dakora and Phillips, 2002). OA anion exudation by lupin CR has a strong effect on cation concentration in the rhizosphere, but this does not occur through changes in pH, but through metal complexation (Dessureault-Rompré et al., 2008), ligand exchange (Violante et al., 1991; Geelhoed et al., 1998) and dissolving organic matter (Kaiser and Zech, 1997; Beck et al., 1999). P deficiency and Al toxicity frequently coexist on acid soils, suggesting that response mechanisms might have evolved to adapt to both stresses through common regulation pathways (Sun et al., 2016).

Enzymes of the TCA cycle, the glycolysis pathway and the glyoxylate cycle are involved in OA biosynthesis (Igamberdiev and Eprintsev, 2016). The conversion of pyruvate to malate, phosphoenolpyruvate to oxaloacetate and oxaloacetate to malate are Al-induced in soybean (Dong et al., 2004). Al toxicity increases malate exudation in *Lupinus pilosus* (Ligaba et al., 2004). The activities and gene expression levels of citrate synthase, malate dehydrogenase and phosphoenolpyruvate carboxylase have been reported to increase in white lupin CR under P deficiency (Johnson et al., 1994; Uhde-Stone et al., 2003). Aconitase and malic enzyme activities are also affected by Al stress in soybean (Xu et al., 2010; Zhou et al., 2018).

Differences exist between P deficiency- and Al toxicity-induced citrate secretion (Wang et al., 2007). P deficiency induces citrate release from mature CR, while Al triggers exudation from the subapical root zones of lateral roots and from mature and senescent CR. Al-induced citrate exudation is inhibited by P limitation at the seedling stage, but is stimulated at later stages. Citrate secretion is dependent on plasma membrane H^+-ATPase in lupin under P deficit (Tomasi et al., 2009), and H^+-ATPase also participates in regulating Al-activated citrate exudation in different legumes (Shen et al., 2005; Guo et al., 2013). However, Al-activated lupin citrate secretion is independent of plasma membrane H^+-ATPase. Citrate secretion is coupled with K^+ efflux in P-deficient lupin plants, which may compensate H^+ efflux and keep the charge balance, but this Al-induced K^+ efflux is independent of Al-induced citrate exudation (Zeng et al., 2013). OAs release is mediated by the activation of transport systems and the expression of stress-induced specific genes (Ma, 2000). In white lupin, LaALMT1 contributes to malate, but not citrate release (Zhou et al., 2020). Contrary to some ALMT homologs in other species, LaALMT1 expression, which is involved in root-to-shoot translocation, is not stimulated by Al. Other LaALMT homologs, which are upregulated by Al, might have a role in lupin Al-tolerance (Zhou et al., 2020). MATE/AACT/DTX transporters might be involved in Al-stimulated citrate release in lupin (Chen and Liao, 2016; Zhou et al., 2020). Besides malic and citric, oxalic, fumaric, malonic and α-ketoglutaric acids have been identified in the rhizosphere of white lupin (Mimmo et al., 2008), and might contribute to lupin Al tolerance. In certain plants, oxalate rather than citrate or malate is involved in Al tolerance (Yang et al., 2011; Lv et al., 2021), but the regulation of oxalate secretion requires further study. Recently, a potential mechanism for Al tolerance in alfalfa has been described that involves oxalate secretion and ABA-mediated signaling (Lv et al., 2021). Flavonoids exuded by lupin roots play a key role in the establishment of symbiosis by inducing the expression of rhizobial *nod* genes (Cooper, 2007), but flavonoids secreted by CR do not differ from those secreted by regular roots (Weisskopf et al., 2006; Cesco et al., 2010). Secretion of flavonoids by CR precedes that of OAs (Cesco et al., 2010; Chen and Liao, 2016), and they can chelate the metal ions that are solubilized by OAs (Weisskopf et al., 2006; Wei et al., 2021).

Several genes involved in auxin and ethylene signaling have been shown to control both P deprivation and Al stress suggesting a P-Al signaling crosstalk (Sun et al., 2016). Phytohormones have different roles in different plants under Al stress (Rangel et al., 2007; Rademacher et al., 2011; Yang et al., 2014; Sun et al., 2016). Auxin response factors (ARF) in roots are regulated by Al-modulated miRNAs (Wang et al., 2005; Liu et al., 2007). Ethylene biosynthesis genes are induced by Al (Sun et al., 2016), and ACC synthase and ACC-oxidase activities correlate with Al-induced ethylene evolution (Sun et al., 2007; Swarup et al., 2007). Ethylene triggers auxin synthesis in the transition zone of the root apex, inhibiting root growth (Yang et al., 2014). In fact, CR regulatory components are common to regular roots (Secco et al., 2014). CR formation is initiated by the same phytohormones and transcription factors employed by simple-patterned roots (Coudert et al., 2010; Petricka et al., 2012). PIN and LAX auxin carriers (Billou et al., 2005; Grieneisen et al., 2007; Swarup et al., 2008), or Aux/IAA proteins working in combination with ARFs (Vanneste and Friml, 2009) occur in both CR and regular root tips. Root-originated auxin has also been detected in lupin CR (Meng et al., 2013). Transcription factors involved in root meristem function and CR development are also coincident (Galinha et al., 2007; Carlsbecker et al., 2010; Secco et al., 2014). A simplified representation of the molecular mechanisms in lupin CR in response to Al toxicity is presented in **Figure 2**.

ROLE OF RHIZOBIA IN LUPIN ALUMINUM TOLERANCE

Plant growth-promoting rhizobacteria (PGPR) can positively affect Al stress tolerance reducing Al uptake and accumulation, decreasing lipid peroxidation and enhancing root development through ACC deaminase and IAA production or induction of plant genes. Some rhizobia possess PGPR properties, produce exopolysaccharides and secrete OAs, phenolic compounds and sugars that can chelate Al (Kang et al., 2021). Malate-chemotaxis is a mechanism for the recruitment of beneficial rhizobacteria (Rudrappa et al., 2008). Lupin-nodulating rhizobia belong primarily to the *Bradyrhizobium* genus, although several other bacteria have been shown to form efficient symbiosis with lupins (Msaddak et al., 2021). Rhizobial entry occurs at the

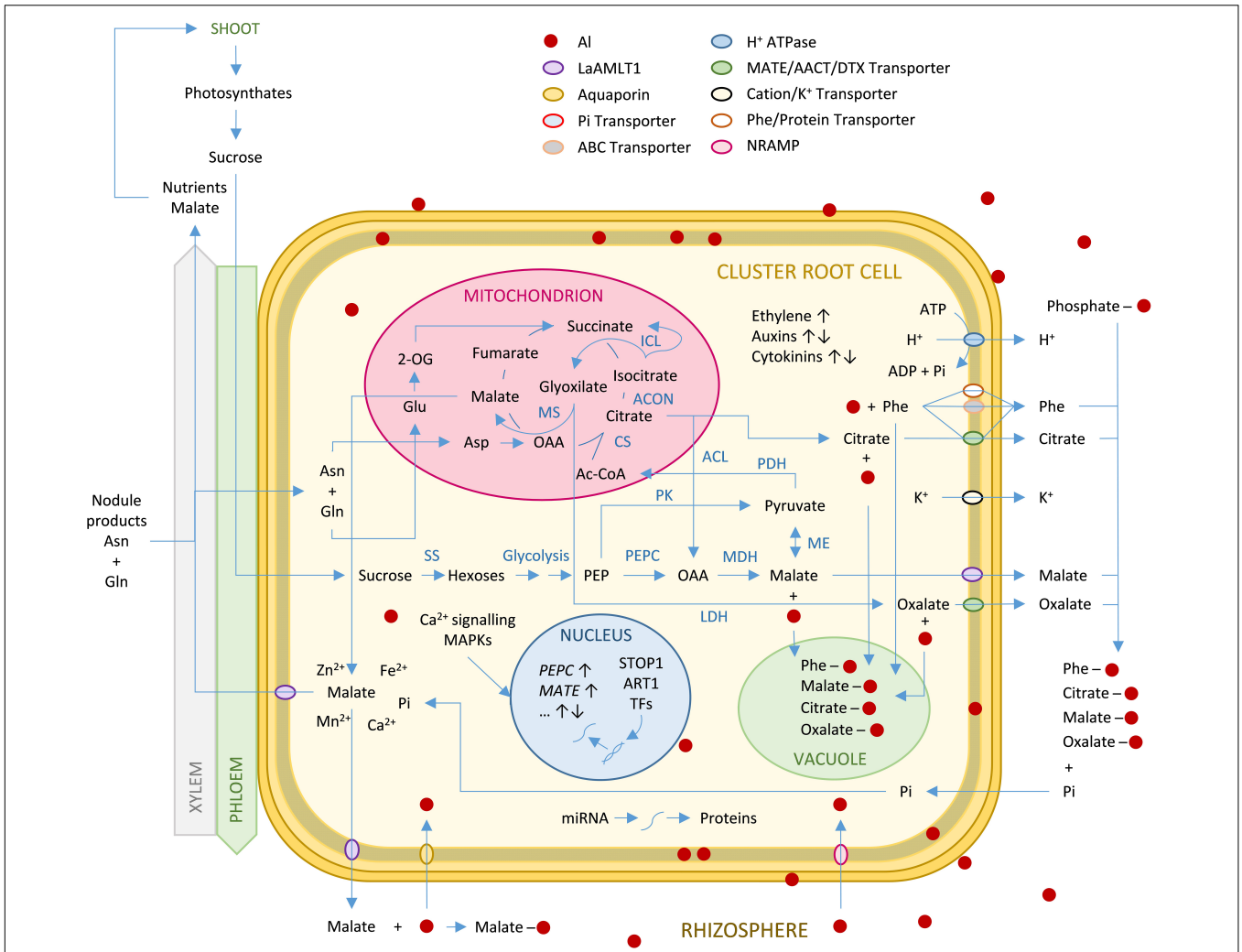


FIGURE 2 | Simplified metabolic pathways in a white lupin cluster root cell in response to Al stress. Aluminum present in the rhizosphere enters the cells primarily through aquaporins and NRAMP transporters. Through the action of Ca^{2+} signaling pathways, transcription factors and miRNAs, Al induces replication of DNA that leads to activation of the TCA cycle, the glycolysis pathway and the glyoxylate cycle. This leads to the accumulation of organic acids and phenolic compounds. Activation of specific transport systems (MATE, ALMT, ABC-transporters) leads to the exudation of these compounds into the rhizosphere. OAs and phenolic compounds chelate Al cations outside and inside the cells. Intracellular complexes are accumulated in the vacuole. Extrusion of H^+ and K^+ contribute to the maintenance of the membrane potential. Ethylene production increases, while auxin and cytokinin levels increase or decrease depending on the developmental stage and the different root zones. Abbreviations: 2-OG, 2-oxoglutarate; AACT, aluminum-activated citrate transporter; ABC, ATP-binding cassette; ACON, aconitase; AMLT, aluminum-activated malate transporter; ACL, ATP citrate lyase; ART1, Al Resistance Transcription Factor 1; Ac-CoA, Acetyl CoA; Asn, asparagine; Asp, aspartic acid; CS, citrate synthase; DTX, detoxification transporter; Gln, glutamine; Glu, glutamic acid; ICL, isocitrate lyase; LDH, lactate dehydrogenase; MAPK, mitogen-activated phosphate kinase; MATE, multidrug and toxic compound extrusion; MDH, malate dehydrogenase; ME, malic enzyme; MS, malate synthase; NRAMP, natural resistance associated macrophage protein; OAA, oxaloacetate; PDH, pyruvate dehydrogenase; PEP, phosphoenolpyruvate; PEPC, PEP carboxylase; Phe, phenolic compounds; PK, PEP kinase; SS, sucrose synthase; STOP1, Transcription factor sensitive to protein rhizotoxicity 1; TFs, transcription factors.

junction between a root hair and an adjacent epidermal cell (González-Sama et al., 2004), and rhizobia are distributed between the daughter cells (Fedorova et al., 2007; Coba de la Peña et al., 2018), while endoreduplication processes take place (González-Sama et al., 2006). Al can severely inhibit rhizobial growth and decrease nitrogenase activity (Lesueur et al., 1993; Arora et al., 2010). However, highly Al-tolerant strains have been characterized (Vargas et al., 2007; Ferreira et al., 2012; Kang et al., 2021). Strains isolated from acidic Al-contaminated soils show a high production of exopolysaccharides and a decreased outer membrane permeability (Ferreira et al., 2012). Al stress also promotes the production of Al-chelating siderophores (Roy and Chakrabartty, 2000). Overall, *Bradyrhizobium* species are more tolerant to acid and Al stress than fast-growing species (Flis et al., 1993). Several mechanisms have been proposed for acid tolerance, including changes in lipopolysaccharide composition of the outer membrane or polyamine accumulation (Chen et al., 1993; Ferreira et al., 2012; Jaiswal et al., 2018).

High exopolysaccharide production by lupin-nodulating bradyrhizobia has been proposed to prevent Hg toxicity (Arregui et al., 2021; Quiñones et al., 2021). Physical sequestration, exclusion and complexation mechanisms contribute to reduce metal availability (Rajkumar et al., 2010). Metals bind to carboxyl and hydroxyl groups on the cell surface, and to microbial secreted substances (Pulsawat et al., 2003). Many lupin species and native soil rhizobia are naturally resistant to low pH and Al toxicity, and tolerant rhizobia favor plant establishment and development (Sprent, 2009). Interestingly, white lupin appears to have developed a strategy to avoid excessive OAs microbial degradation (Weisskopf et al., 2006). The stability of carboxylates in the soil is key to maintain P acquisition and counteract Al toxicity. This is achieved by a temporal regulation of the secretion of isoflavonoids, antifungal cell wall-degrading enzymes, chitinase and glucanase prior to OAs excretion, thus reducing microorganisms population and preventing OA degradation (Neumann et al., 1999; Burzyński et al., 2000; Zhu et al., 2005).

DISCUSSION

The gradual acidification of the world's cultivated soils is predominantly driven by agricultural practices and global change. A decrease in pH leads primarily to the complexation of P, rendering it unavailable to plant roots, and to the release of phytotoxic Al cations. Alkaline amendments, such as lime, temporarily increase soil pH. However, in the long term, they might contribute to surface water eutrophication. Sustainable alternatives are needed to confront this global problem. In acidic soils, plants can prevent Al toxicity in a zone around the root tip, by avoiding its entrance into root cells mainly through the exudation of AOs (Wei et al., 2021). Lupin CRs are specialized organs composed of hundreds of small rootlets with active root tips that are able to exudate AOs and other metabolites (Pueyo et al., 2021 and references therein). We propose that CR multiply the defense mechanisms present in other plants, thus significantly increasing Al tolerance. Cultivation of lupin, a plant that is able to cope with Al toxicity and possesses the ability to fix nitrogen in symbiosis with rhizobia and the capacity to mobilize P and other various nutrients through CR exudation, is able to contribute to a sustainable management of agricultural soils affected by low pH and Al toxicity, without compromising crop production. Additionally, the elevated protein content in lupin seeds makes it an extremely valuable crop due to the increasing demand of plant protein. Accurate studies on Al content in lupin organs grown in acidic soils with different soil textures and different Al speciation must be performed to ensure the seed protein safety used for feed and food.

Lupins are cultivated worldwide, but they are a minority crop. *L. albus* readily forms CR under moderate P deficiency, Fe deficiency or Al toxicity, while some other cultivated lupins are capable of producing CR under harsher stress conditions (Pueyo et al., 2021), which implies that they probably share Al-detoxification systems. While *L. albus* is the most studied species in terms of tolerance mechanisms, other *Lupinus* species are also quite tolerant to acidic soils and Al toxicity. CR formation is triggered by specific soil conditions, but the mechanisms involved are similar to those known to control growth of regular roots. CR provide an expanded surface to interact with the rhizospheric soil. Lupins establish symbiosis with soil bacteria, mostly within the *Bradyrhizobium* genus, which contribute to Al immobilization. Al tolerance mechanisms in lupin are common to other Al-tolerant plants, but specific features derived from the root structural modifications and the interaction with symbiotic bacteria are outstanding. Here, we have analyzed the existing knowledge on the particular mechanisms that control lupin tolerance to Al. Citrate exudation induced by Al in lupin is not coupled with H^+ secretion through plasma membrane H^+-ATPase, and when Al toxicity is concomitant with P deficiency, a K^+ efflux occurs that is not caused by Al stress alone (Zeng et al., 2013). Contrary to some homologs in other plants, LaALMT1 expression is not stimulated by Al, and it is not involved in malate exudation. The transport mechanisms in Al-induced citrate secretion are not completely understood and need further investigation. This research line has not really progressed in the last few years; however, the fully white lupin sequenced genome and some novel technologies, including omics and advanced transformation techniques, provide new tools to elucidate the specific Al-triggered mechanisms that lead to citrate (and other less studied metabolites) exudation.

While different rhizobacteria can help reduce Al stress, the slow-growing bradyrhizobia that nodulate lupin have been reported to produce abundant exopolysaccharides, which immobilize metal ions. Lupin tolerance mechanisms have been investigated and deciphered to some extent; however, little is known on the accumulation of metals by lupin. Rooibos tea (*Aspalathus linearis*) is a legume that also grows in acidic, Al-rich, nutrient-poor soils (Kanu et al., 2013). Like lupin, it forms CR that exude Al-chelating molecules, and it accumulates Al–Si complexes (Kanu et al., 2013). It has been proposed as a suitable candidate for phytoremediation. Recently, white lupin has been reported as an Hg-resistant plant capable of Hg rhizosequestration, especially in CR, with null translocation of the metal to the aerial parts of the plant (Quiñones et al., 2021). It appears that CR have a role not only in elevated secretion of metal chelators, but also in the immobilization of toxic metals, a subject that certainly requires further investigation for its practical application in avoiding translocation to the aerial parts of the plant, a topic that requires further analyses, and for the elucidation of metal sequestration mechanisms in plant roots.

White lupin in symbiosis with tolerant bradyrhizobia represents a potentially powerful tool for metal-polluted soil phytoremediation and degraded landscape restoration, entailing significant environmental benefits. Moreover, lupin cultivation involves an enrichment of the soils with N through atmospheric N_2 fixation, and nutrients such as P and Fe through CR exudates-mediated solubilization that may be taken advantage of by rotating crops. In conclusion, lupin cultivation as a protein crop arises as an optimal environment-friendly alternative to exploit acidic soils in temperate zones affected by Al toxicity, where other crops might not be sufficiently productive to be economically viable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JJP conceived the work and designed the figures. JJP and MAQ wrote the first draft. MAQ, MML, and JJP contributed to the writing of the manuscript, proofread, and approved its content.

FUNDING

This work was supported by the Spanish State Research Agency (AEI), grant AGL2017-88381-R to MML and JJP. We acknowledge support of the publication fee (25%) by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

ACKNOWLEDGMENTS

We apologize to all colleagues whose work we have not cited due to space restrictions.

REFERENCES

Andersson, M. (1988). Toxicity and tolerance of aluminium in vascular plants – a literature review. *Water Air Soil Pollut.* 39, 439–462. doi: 10.1007/BF00279487

Arora, N. K., Khare, E., Singh, S., and Maheshwari, D. K. (2010). Effect of Al and heavy metals on enzymes of nitrogen metabolism of fast and slow growing rhizobia under explanta conditions. *World J. Microbiol. Biotechnol.* 26, 811–816. doi: 10.1007/s11274-009-0237-6

Arregui, G., Hipólito, P., Pallol, B., Lara-Dampier, V., García-Rodríguez, D., Varela, H. P., et al. (2021). Mercury-tolerant Ensifer medicae strains display high mercuric reductase activity and a protective effect on nitrogen fixation in *Medicago truncatula* nodules under mercury stress. *Front. Plant Sci.* 11:560768. doi: 10.3389/fpls.2020.560768

Beck, M. A., Robarge, W. P., and Buol, S. W. (1999). Phosphorus retention and release of anions and organic carbon by two Andisols. *Eur. J. Soil Sci.* 50, 157–164. doi: 10.1046/j.1365-2389.1999.00213.x

Billou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Frimi, J., et al. (2005). The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* 433, 39–44. doi: 10.1038/nature03184

Burzyński, A., Piślewska, M., and Wojtaszek, P. (2000). Active chitinases in the apoplastic fluids of healthy white lupin (*Lupinus albus* L.) plants. *Acta Physiol. Plant.* 22, 31–38. doi: 10.1007/s11738-000-0005-6

Carlsbecker, A., Lee, J. Y., Roberts, C. J., Dettmer, J., Lehesranta, S., Zhou, J., et al. (2010). Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* 465, 316–321. doi: 10.1038/nature08977

Carr, J., and Ritchie, P. (1993). Al toxicity of wheat grown in acidic subsoils in relation to soil solution properties and exchangeable cations. *Aust. J. Soil Res.* 31, 583–596. doi: 10.1071/sr9930583

Cesco, S., Neumann, G., Tomasi, N., Pinton, R., and Weisskopf, L. (2010). Release of plant-borne flavonoids into the rhizosphere and their role in plant nutrition. *Plant Soil* 329, 1–25. doi: 10.1007/s11104-009-0266-9

Chauhan, D. K., Yadav, V., Vaculík, M., Gassmann, W., Pike, S., Arif, N., et al. (2021). Aluminum toxicity and aluminum stress-induced physiological tolerance responses in higher plants. *Crit. Rev. Biotechnol.* 41, 715–730. doi: 10.1080/07388551.2021.1874282

Chen, H., Richardson, A. E., and Rolfe, B. G. (1993). Studies of the physiological and genetic basis of acid tolerance in *Rhizobium leguminosarum* biovar *trifolii*. *Appl. Environ. Microbiol.* 59, 1798–1804. doi: 10.1128/aem.59.6.1798-1804.1993

Chen, Z. C., and Liao, H. (2016). Organic acid anions: an effective defensive weapon for plants against aluminum toxicity and phosphorus deficiency in acidic soils. *J. Genet. Genomics* 43, 631–638. doi: 10.1016/j.jgg.2016.11.003

Coba de la Peña, T., and Pueyo, J. J. (2012). Legumes in the reclamation of marginal soils, from cultivar and inoculant selection to transgenic approaches. *Agron. Sustain. Dev.* 32, 65–91. doi: 10.1007/s13593-011-0024-2

Coba de la Peña, T., Fedorova, E., Pueyo, J. J., and Lucas, M. M. (2018). The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? *Front. Plant Sci.* 8:2229. doi: 10.3389/fpls.2017.02229

Cooper, J. E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *J. Appl. Microbiol.* 103, 1355–1365. doi: 10.1111/j.1365-2672.2007.03366.x

Coudert, Y., Périn, C., Courtois, B., Khong, N. G., and Gantet, P. (2010). Genetic control of root development in rice, the model cereal. *Trends Plant Sci.* 15, 219–226. doi: 10.1016/j.tplants.2010.01.008

Dakora, F. D., and Phillips, D. A. (2002). Root exudates as mediators of mineral acquisition in low-nutrient environments. *Plant Soil* 245, 35–47. doi: 10.1023/A:1020809400075

De Ron, A. M., Sparvoli, F., Pueyo, J. J., and Bazile, D. (2017). Protein crops: food and feed for the future. *Front. Plant Sci.* 8:105. doi: 10.3389/fpls.2017.00105

Dessureault-Rompré, J., Nowack, B., Schulin, R., and Luster, J. (2007). Spatial and temporal variation in organic acid anion exudation and nutrient anion uptake in the rhizosphere of *Lupinus albus* L. *Plant Soil* 301, 123–134. doi: 10.1007/s11104-007-9427-x

Dessureault-Rompré, J., Nowack, B., Schulin, R., Tercier-Waeber, M. L., and Luster, J. (2008). Metal solubility and speciation in the rhizosphere of *Lupinus albus* cluster roots. *Environ. Sci. Technol.* 42, 7146–7151. doi: 10.1021/es800167g

Dong, D., Peng, X., and Yan, X. (2004). Organic acid exudation induced by phosphorus deficiency and/or aluminium toxicity in two contrasting soybean genotypes. *Physiol. Plant.* 122, 190–199. doi: 10.1111/j.1399-3054.2004.00373.x

Eriksson, A. K., Hillier, S., Hesterberg, D., Klysubun, W., Ulén, B., and Gustafsson, J. P. (2016). Evolution of phosphorus speciation with depth in an agricultural soil profile. *Geoderma* 280, 29–37. doi: 10.1016/j.geoderma.2016.06.004

Fedorova, E. E., De Felipe, M. R., Pueyo, J. J., and Lucas, M. M. (2007). Conformation of cytoskeletal elements during the division of infected *Lupinus albus* L. nodule cells. *J. Exp. Bot.* 58, 2225–2236. doi: 10.1093/jxb/erm083

Fernández-Pascual, M., Pueyo, J. J., Felipe, M. R., Golvano, M. P., and Lucas, M. M. (2007). Singular features of the *Bradyrhizobium-Lupinus* symbiosis. *Dyn. Soil Dyn. Plant* 1, 1–16. doi: 10.1007/978-3-319-23534-9_1

Ferreira, P. A. A., Bomfeti, C. A., Soares, B. L., and de Souza Moreira, F. M. (2012). Efficient nitrogen-fixing *Rhizobium* strains isolated from amazonian soils are highly tolerant to acidity and aluminium. *World J. Microbiol. Biotechnol.* 28, 1947–1959. doi: 10.1007/s11274-011-0997-7

Flis, S. E., Glenn, A. R., and Dilworth, M. J. (1993). The interaction between aluminium and root nodule bacteria. *Soil Biol. Biochem.* 25, 403–417. doi: 10.1016/0038-0717(93)90066-K

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., et al. (2007). PLETHORA proteins as dose-dependent master regulators of *Arabidopsis* root development. *Nature* 449, 1053–1057. doi: 10.1038/nature06206

Geelhoed, J. S., Hiemstra, T., and Van Riemsdijk, W. H. (1998). Competitive interaction between phosphate and citrate on goethite. *Environ. Sci. Technol.* 32, 2119–2123. doi: 10.1021/es970908y

Gerke, J., Römer, W., and Jungk, A. (1994). The excretion of citric and malic acid by proteoid roots of *Lupinus albus* L.; effects on soil solution concentrations of phosphate, iron, and aluminum in the proteoid rhizosphere in samples of an oxisol and a luvisol. *Z. Pflanzenernährung Bodenkd.* 157, 289–294. doi: 10.1002/jpln.19941570408

Gilbert, G. A., Knight, J. D., Vance, C. P., and Allan, D. L. (1999). Acid phosphatase activity in phosphorus-deficient white lupin roots. *Plant Cell Environ.* 22, 801–810. doi: 10.1046/j.1365-3040.1999.00441.x

González-Sama, A., De La Peña, T. C., Kevei, Z., Mergaert, P., Lucas, M. M., De Felipe, M. R., et al. (2006). Nuclear DNA endoreduplication and expression of the mitotic inhibitor Ccs52 associated to determinate and lupinoid nodule organogenesis. *Mol. Plant Microbe Interact.* 19, 173–180. doi: 10.1094/MPMI-19-0173

González-Sama, A., Lucas, M. M., De Felipe, M. R., and Pueyo, J. J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytol.* 163, 371–380. doi: 10.1111/j.1469-8137.2004.01121.x

Grieneisen, V. A., Xu, J., Marée, A. F. M., Hogeweg, P., and Scheres, B. (2007). Auxin transport is sufficient to generate a maximum and gradient guiding root growth. *Nature* 449, 1008–1013. doi: 10.1038/nature06215

Guo, C. L., Chen, Q., Zhao, X. L., Chen, X. Q., Zhao, Y., Wang, L., et al. (2013). Al-enhanced expression and interaction of 14-3-3 protein and plasma membrane H^+-ATPase is related to Al-induced citrate secretion in an Al-resistant black soybean. *Plant Mol. Biol. Report.* 31, 1012–1024. doi: 10.1007/s11105-013-0569-0

Hemada, M., Mohamed, H. E., Saber, N. E.-S., and El-Ekeili, A. M. (2020). Role of salicylic acid in alleviation of aluminum effects on growth and biochemical processes in lupin (*Lupinus termis* L.) plant. *Egypt. J. Bot.* 60, 287–300. doi: 10.21608/ejbo.2019.15177.1343

Igamberdiev, A. U., and Eprintsev, A. T. (2016). Organic acids: the pools of fixed carbon involved in redox regulation and energy balance in higher plants. *Front. Plant Sci.* 7:1042. doi: 10.3389/fpls.2016.01042

Jaiswal, S. K., Naamala, J., and Dakora, F. D. (2018). Nature and mechanisms of aluminium toxicity, tolerance and amelioration in symbiotic legumes and rhizobia. *Biol. Fertil. Soils* 54, 309–318. doi: 10.1007/s00374-018-1262-0

Jansen, P. C. M. (2006). "*Lupinus albus* L," in *Record from PROTA4U*, eds M. Brink and G. Belay (Wageningen: PROTA (Plant Resources of Tropical Africa/Ressources végétales de l'Afrique Tropicale)).

Johnson, J. F., Allan, D. L., and Vance, C. P. (1994). Phosphorus stress-induced proteoid roots show altered metabolism in *Lupinus albus*. *Plant Physiol.* 104, 657–665. doi: 10.1104/pp.104.2.657

Kaiser, K., and Zech, W. (1997). Competitive sorption of dissolved organic matter fractions to soils and related mineral phases. *Soil Sci. Soc. Am. J.* 61, 64–69. doi: 10.2136/sssaj1997.03615995006100010011x

Kang, J. P., Huo, Y., Yang, D. U., and Yang, D. C. (2021). Influence of the plant growth promoting *Rhizobium panacihumi* on aluminum resistance in *Panax ginseng*. *J. Ginseng Res.* 45, 442–449. doi: 10.1016/j.jgr.2020.01.001

Kanu, S. A., Okonkwo, J. O., and Dakora, F. D. (2013). *Aspalathus linearis* (Rooibos tea) as potential phytoremediation agent: a review on tolerance mechanisms for aluminum uptake. *Environ. Rev.* 21, 85–92. doi: 10.1139/er-2012-0055

Lambers, H., Cramer, M. D., Shane, M. W., Wouterlood, M., Poot, P., and Veneklaas, E. J. (2003). Structure and functioning of cluster roots and plant responses to phosphate deficiency. *Plant Soil* 248, 9–19. doi: 10.1023/A:1025561812696

Lamont, B. (2003). Structure, ecology and physiology of root clusters – a review. *Plant Soil* 248, 1–19. doi: 10.1023/A:1022314613217

Lesueur, D., Diem, H. G., Dianda, M., and Le Roux, C. (1993). Selection of *Bradyrhizobium* strains and provenances of *Acacia mangium* and *Faidherbia albida*: relationship with their tolerance to acidity and aluminium. *Plant Soil* 149, 159–166. doi: 10.1007/BF00016605

Ligaba, A., Yamaguchi, M., Shen, H., Sasaki, T., Yamamoto, Y., and Matsumoto, H. (2004). Phosphorus deficiency enhances plasma membrane H^+-ATPase activity and citrate exudation in greater purple lupin (*Lupinus pilosus*). *Funct. Plant Biol.* 31, 1075–1083. doi: 10.1071/FP04091

Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H., and Carrington, J. C. (2007). Repression of AUXIN response factor10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J.* 52, 133–146. doi: 10.1111/j.1365-313X.2007.03218.x

Lucas, M. M., Stoddard, F. L., Annicchiarico, P., Frías, J., Martínez-Villaluenga, C., Sussmann, D., et al. (2015). The future of lupin as a protein crop in Europe. *Front. Plant Sci.* 6:705. doi: 10.3389/fpls.2015.00705

Lv, A., Wen, W., Fan, N., Su, L., Zhou, P., and An, Y. (2021). Dehydrin MsDHN1 improves aluminum tolerance of alfalfa (*Medicago sativa* L.) by affecting oxalate exudation from root tips. *Plant J.* 108, 441–458. doi: 10.1111/tpj.15451

Ma, J. F. (2000). Role of organic acids in detoxification of aluminum in higher plants. *Plant Cell Physiol.* 41, 383–390. doi: 10.1093/pcp/41.4.383

Ma, J. F., Ryan, P. R., and Delhaize, E. (2001). Aluminium tolerance in plants and the complexing role of organic acids. *Trends Plant Sci.* 6, 273–278. doi: 10.1016/S1360-1385(01)01961-6

Matsumoto, H. (2000). Cell biology of aluminum toxicity tolerance in higher plants. *Int. Rev. Cytol.* 200, 1–46. doi: 10.1016/s0074-7696(00)00001-2

Meng, Z. B., You, X. D., Suo, D., Chen, Y. L., Tang, C., Yang, J. L., et al. (2013). Root-derived auxin contributes to the phosphorus-deficiency-induced cluster-root formation in white lupin (*Lupinus albus*). *Physiol. Plant.* 148, 481–489. doi: 10.1111/j.1399-3054.2012.01715.x

Mimmo, T., Ghizzi, M., Marzadori, C., and Gessa, C. E. (2008). Organic acid extraction from rhizosphere soil: effect of field-moist, dried and frozen samples. *Plant Soil* 312, 175–184. doi: 10.1007/s11104-008-9574-8

Moore, G., Dolling, P., Porter, B., and Leonard, L. (2001). "Chemical factors affecting plant growth. Soil acidity," in *Soil Guide. A Handbook for Undersanding and Managing Agricultural Soils*, eds G. Moore, J. McFarlane, B. Purdie, and G. Wilson (Perth, WA: Department of Agriculture and Food), 127–140.

Msaddak, A., Rey, L., Imperial, J., Palacios, J. M., Mars, M., and Pueyo, J. J. (2021). Phylogenetic analyses of rhizobia isolated from nodules of *Lupinus angustifolius* in Northern Tunisia reveal *Devosia* sp. As a new microsymbiont of lupin species. *Agronomy* 11:1510. doi: 10.3390/agronomy11081510

Nelson, P., Moore, G., and Dracup, M. (2001). "Crops: soil and climatc requirements. Narrow-leafed lupins," in *Soil Guide. A Handbook for Undersanding and Managing Agricultural Soils*, eds G. Moore, J. McFarlane, B. Purdie, and G. Wilson (Perth, WA: Department of Agriculture and Food), 272–273. doi: 10.1111/nph.14094

Neumann, G., Massonneau, A., Martinoia, E., and Römheld, V. (1999). Physiological adaptations to phosphorus deficiency during proteoid root development in white lupin. *Planta* 208, 373–382. doi: 10.1007/s004250050572

Panagos, P., Van Liedekerke, M., Jones, A., and Montanarella, L. (2012). European soil data centre: response to European policy support and public data requirements. *Land Use Policy* 29, 329–338. doi: 10.1016/j.landusepol.2011.07.003

Petricka, J. J., Winter, C. M., and Benfey, P. N. (2012). Control of *Arabidopsis* root development. *Annu. Rev. Plant Biol.* 63, 563–590. doi: 10.1146/annurev-arplant-042811-105501

Pueyo, J. J., Quiñones, M. A., Coba de la Peña, T., Fedorova, E. E., and Lucas, M. M. (2021). Nitrogen and phosphorus interplay in lupin root nodules and cluster roots. *Front. Plant Sci.* 12:644218. doi: 10.3389/fpls.2021.644218

Pulsawat, W., Leksawasdi, N., Rogers, P. L., and Foster, L. J. R. (2003). Anions effects on biosorption of Mn(II) by extracellular polymeric substance (EPS) from *Rhizobium etli*. *Biotechnol. Lett.* 25, 1267–1270. doi: 10.1023/A:1025083116343

Quiñones, M. A., Fajardo, S., Fernández-Pascual, M., Lucas, M. M., and Pueyo, J. J. (2021). Nodulated white lupin plants growing in contaminated soils accumulate unusually high mercury concentrations in their nodules, roots and especially cluster roots. *Horticulturae* 7, 1–19. doi: 10.3390/horticulturae7090302

Quiñones, M. A., Ruiz-Díez, B., Fajardo, S., López-Berdonces, M. A., Higueras, P. L., and Fernández-Pascual, M. (2013). *Lupinus albus* plants acquire mercury tolerance when inoculated with an Hg-resistant *Bradyrhizobium* strain. *Plant Physiol. Biochem.* 73, 168–175. doi: 10.1016/j.plaphy.2013.09.015

Rademacher, E. H., Möller, B., Lokerse, A. S., Llavata-Peris, C. I., Van Den Berg, W., and Weijers, D. (2011). A cellular expression map of the *Arabidopsis* AUXIN response factor gene family. *Plant J.* 68, 597–606. doi: 10.1111/j.1365-313X.2011.04710.x

Rajkumar, M., Ae, N., Prasad, M. N. V., and Freitas, H. (2010). Potential of siderophore-producing bacteria for improving heavy metal phytoextraction. *Trends Biotechnol.* 28, 142–149. doi: 10.1016/j.tibtech.2009.12.002

Rangel, A. F., Rao, I. M., and Horst, W. J. (2007). Spatial aluminium sensitivity of root apices of two common bean (*Phaseolus vulgaris* L.) genotypes with contrasting aluminium resistance. *J. Exp. Bot.* 58, 3895–3904. doi: 10.1093/jxb/erm241

Redel, Y., Cartes, P., Velásquez, G., Poblete-Grant, P., Poblete-Grant, P., Bol, R., et al. (2016). Assessment of phosphorus status influenced by Al and Fe compounds in volcanic grassland soils. *J. Soil Sci. Plant Nutr.* 16, 490–506. doi: 10.4067/S0718-95162016005000041

Roy, N., and Chakrabartty, P. K. (2000). Effect of aluminum on the production of siderophore by *Rhizobium* sp. (*Cicer arietinum*). *Curr. Microbiol.* 41, 5–10. doi: 10.1007/s002840010082

Rudrappa, T., Czymmek, K. J., Paré, P. W., and Bais, H. P. (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol.* 148, 1547–1556. doi: 10.1104/pp.108.127613

Sade, H., Meriga, B., Surapu, V., Gadi, J., Sunita, M. S. L., Suravajhala, P., et al. (2016). Toxicity and tolerance of aluminum in plants: tailoring plants to suit to acid soils. *BioMetals* 29, 187–210. doi: 10.1007/s10534-016-9910-z

Secco, D., Shou, H., Whelan, J., and Berkowitz, O. (2014). RNA-seq analysis identifies an intricate regulatory network controlling cluster root development in white lupin. *BMC Genomics* 15:230. doi: 10.1186/1471-2164-15-230

Sharma, T., Dreyer, I., Kochian, L., and Piñeros, M. A. (2016). The ALMT family of organic acid transporters in plants and their involvement in detoxification and nutrient security. *Front. Plant Sci.* 7:1488. doi: 10.3389/fpls.2016.01488

Shen, H., He, L. F., Sasaki, T., Yamamoto, Y., Zheng, S. J., Ligaba, A., et al. (2005). Citrate secretion coupled with the modulation of soybean root tip under aluminum stress. Up-regulation of transcription, translation, and threonine-oriented phosphorylation of plasma membrane H^+-ATPase. *Plant Physiol.* 138, 287–296. doi: 10.1104/pp.104.900170

Sprent, J. I. (2009). *Legume Nodulation. A Global Perspective*. Chichester: Wiley-Blackwell, doi: 10.1016/j.cub.2014.01.028

Sun, L., Tian, J., Zhang, H., and Liao, H. (2016). Phytohormone regulation of root growth triggered by P deficiency or Al toxicity. *J. Exp. Bot.* 67, 3655–3664. doi: 10.1093/jxb/erw188

Sun, P., Tian, Q. Y., Zhao, M. G., Dai, X. Y., Huang, J. H., Li, L. H., et al. (2007). Aluminum-induced ethylene production is associated with inhibition of root elongation in *Lotus japonicus* L. *Plant Cell Physiol.* 48, 1229–1235. doi: 10.1093/pcp/pcm077

Swarup, K., Benková, E., Swarup, R., Casimiro, I., Péret, B., Yang, Y., et al. (2008). The auxin influx carrier LAX3 promotes lateral root emergence. *Nat. Cell Biol.* 10, 946–954. doi: 10.1038/ncb1754

Swarup, R., Perry, P., Hagenbeek, D., Van Der Straeten, D., Beemster, G. T. S., Sandberg, G., et al. (2007). Ethylene upregulates auxin biosynthesis in *Arabidopsis* seedlings to enhance inhibition of root cell elongation. *Plant Cell* 19, 2186–2196. doi: 10.1105/tpc.107.052100

Tomasi, N., Kretzschmar, T., Espen, L., Weisskopf, L., Fuglsang, A. T., Palmgren, M. G., et al. (2009). Plasma membrane H^+-ATPase-dependent citrate exudation from cluster roots of phosphate-deficient white lupin. *Plant Cell Environ.* 32, 465–475. doi: 10.1111/j.1365-3040.2009.01938.x

Uhde-Stone, C., Gilbert, G., Johnson, J. M. F., Litjens, R., Zinn, K. E., Temple, S. J., et al. (2003). Acclimation of white lupin to phosphorus deficiency involves enhanced expression of genes related to organic acid metabolism. *Plant Soil* 248, 99–116. doi: 10.1023/A:1022335519879

Vanneste, S., and Friml, J. (2009). Auxin: a trigger for change in plant development. *Cell* 136, 1005–1016. doi: 10.1016/j.cell.2009.03.001

Vargas, L. K., Lisboa, B. B., Scholles, D., Silveira, J. R. P., Jung, G. C., Granada, C. E., et al. (2007). Genetic diversity and symbiotic efficiency of black wattle-nodulating rhizobia in soils of Rio Grande do Sul State, Brazil. *Rev. Bras. Cienc. Solo* 31, 647–654. doi: 10.1590/s0100-06832007000400005

Violante, A., Colombo, C., and Buondonno, A. (1991). Competitive adsorption of phosphate and oxalate by aluminum oxides. *Soil Sci. Soc. Am. J.* 55, 65–70. doi: 10.2136/sssaj1991.03615995005500010011x

Wang, B. L., Shen, J. B., Zhang, W. H., Zhang, F. S., and Neumann, G. (2007). Citrate exudation from white lupin induced by phosphorus deficiency differs from that induced by aluminum. *New Phytol.* 176, 581–589. doi: 10.1111/j.1469-8137.2007.02206.x

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., and Chen, X. Y. (2005). Control of root cap formation by MicroRNA-targeted auxin response factors in *Arabidopsis*. *Plant Cell* 17, 2204–2216. doi: 10.1105/tpc.105.033076

Wang, J., Hou, Q., Li, P., Yang, L., Sun, X., Benedito, V. A., et al. (2017). Diverse functions of multidrug and toxin extrusion (MATE) transporters in citric acid efflux and metal homeostasis in *Medicago truncatula*. *Plant J.* 90, 79–95. doi: 10.1111/tpj.13471

Wei, Y., Han, R., Xie, Y., Jiang, C., and Yu, Y. (2021). Recent advances in understanding mechanisms of plant tolerance and response to aluminum toxicity. *Sustainability* 13:1782. doi: 10.3390/su13041782

Weisskopf, L., Tomasi, N., Santelia, D., Martinoia, E., Langlade, N. B., Tabacchi, R., et al. (2006). Isoflavonoid exudation from white lupin roots is influenced by phosphate supply, root type and cluster-root stage. *New Phytol.* 171, 657–668. doi: 10.1111/j.1469-8137.2006.01776.x

Xu, M., You, J., Hou, N., Zhang, H., Chen, G., and Yang, Z. (2010). Mitochondrial enzymes and citrate transporter contribute to the aluminium-induced citrate secretion from soybean (*Glycine max*) roots. *Funct. Plant Biol.* 37, 285–295. doi: 10.1071/FP09223

Yang, J. L., Zhu, X. F., Peng, Y. X., Zheng, C., Ming, F., and Zheng, S. J. (2011). Aluminum regulates oxalate secretion and plasma membrane H^+-ATPase activity independently in tomato roots. *Planta* 234, 281–291. doi: 10.1007/s00425-011-1402-3

Yang, Z. B., Geng, X., He, C., Zhang, F., Wang, R., Horst, W. J., et al. (2014). TAA1-regulated local auxin biosynthesis in the root-apex transition zone mediates the aluminum-induced inhibition of root growth in *Arabidopsis*. *Plant Cell* 26, 2889–2904. doi: 10.1105/tpc.114.127993

Zavalin, A. A., Sokolov, O. A., Shmyreva, N. Y., and Lukin, S. V. (2019). Legume reaction to soil acidity. *Amaz. Investig.* 8, 162–170.

Zeng, H., Feng, X., Wang, B., Zhu, Y., Shen, Q., and Xu, G. (2013). Citrate exudation induced by aluminum is independent of plasma membrane H^+-ATPase activity and coupled with potassium efflux from cluster roots of phosphorus-deficient white lupin. *Plant Soil* 366, 389–400. doi: 10.1007/s11104-012-1445-7

Zheng, S. J. (2010). Crop production on acidic soils: overcoming aluminium toxicity and phosphorus deficiency. *Ann. Bot.* 106, 183–184. doi: 10.1093/aob/mcq134

Zhou, Y., Neuhäuser, B., Neumann, G., and Ludewig, U. (2020). LaALMT1 mediates malate release from phosphorus-deficient white lupin root tips and metal root to shoot translocation. *Plant Cell Environ.* 43, 1691–1706. doi: 10.1111/pce.13762

Zhou, Y., Yang, Z., Xu, Y., Sun, H., Sun, Z., Lin, B., et al. (2018). Soybean NADP-malic enzyme functions in malate and citrate metabolism and contributes to their efflux under Al stress. *Front. Plant Sci.* 8:2246. doi: 10.3389/fpls.2017.02246

Zhu, Y., Yan, F., Zörb, C., and Schubert, S. (2005). A link between citrate and proton release by proteoid roots of white lupin (*Lupinus albus* L.) grown under phosphorus-deficient conditions? *Plant Cell Physiol.* 46, 892–901. doi: 10.1093/pcp/pci094

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Quiñones, Lucas and Pueyo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

MINI REVIEW
published: 26 January 2022
doi: 10.3389/fpls.2021.795091

A Bitter-Sweet Story: Unraveling the Genes Involved in Quinolizidine Alkaloid Synthesis in *Lupinus albus*

Claudia E. Osorio[1]† and Bradley J. Till[2†]*

[1] Instituto de Investigaciones Agropecuarias, INIA Carillanca, Temuco, Chile, [2] Veterinary Genetics Laboratory, University of California, Davis, Davis, CA, United States

OPEN ACCESS

Edited by:
Jose C. Jimenez-Lopez, Experimental Station of Zaidín, Spanish National Research Council (CSIC), Spain

Reviewed by:
Jun Murata, Suntory Foundation for Life Sciences, Japan
Giovanna Boschin, University of Milan, Italy
Magdalena Kroc, Institute of Plant Genetics, Polish Academy of Sciences (PAN), Poland
Michał Książkiewicz, Institute of Plant Genetics, Polish Academy of Sciences (PAN), Poland

***Correspondence:**
Claudia E. Osorio
claudia.osorio@inia.cl

†ORCID:
Claudia E. Osorio
orcid.org/0000-0001-6205-0730
Bradley J. Till
orcid.org/0000-0002-1300-8285

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 14 October 2021
Accepted: 30 December 2021
Published: 26 January 2022

Citation:
Osorio CE and Till BJ (2022) A Bitter-Sweet Story: Unraveling the Genes Involved in Quinolizidine Alkaloid Synthesis in Lupinus albus. Front. Plant Sci. 12:795091. doi: 10.3389/fpls.2021.795091

Alkaloids are part of a structurally diverse group of over 21,000 cyclic nitrogen-containing secondary metabolites that are found in over 20% of plant species. *Lupinus albus* are naturally containing quinolizidine alkaloid (QA) legumes, with wild accessions containing up to 11% of QA in seeds. Notwithstanding their clear advantages as a natural protecting system, lupin-breeding programs have selected against QA content without proper understanding of quinolizidine alkaloid biosynthetic pathway. This review summarizes the current status in this field, with focus on the utilization of natural mutations such as the one contained in *pauper* locus, and more recently the development of molecular markers, which along with the advent of sequencing technology, have facilitated the identification of candidate genes located in the *pauper* region. New insights for future research are provided, including the utilization of differentially expressed genes located on the *pauper locus*, as candidates for genome editing. Identification of the main genes involved in the biosynthesis of QA will enable precision breeding of low-alkaloid, high nutrition white lupin. This is important as plant based high quality protein for food and feed is an essential for sustainable agricultural productivity.

Keywords: alkaloids, quinolizidine alkaloids, white lupin, *Lupinus albus*, *pauper*

INTRODUCTION

Lupinus, part of the Fabaceae family, is a genus of plants widely distributed around the world (Clements et al., 2008). Different cultivated species are present in the Old World (*L. albus, L. angustifolius,* and *L. luteus*) as well as the New World (*L. mutabilis*) (Petterson and Mackintosh, 1994; Peterson, 2000). Despite the phenotypic differences due to environmental conditions, they are all characterized by high level of proteins, being used as human food or animal feed (Petterson and Mackintosh, 1994; Abraham et al., 2019; Sońta and Rekiel, 2019). The production of lupin seeds as an agricultural product occurs mainly in Australia but also it has an important role in agricultural productive system in parts of Europe, Africa, and South America (FAO, 2021). During 2019, the largest lupin producers were Australia (474,629 t), Russian Federation (103,792 t) and Poland (261,500 t) (FAO, 2021).

White lupin (*Lupinus albus L.*) is a legume from the Mediterranean region; its center of origin is believed to be southern Greece and Western Turkey, where its cultivation started 4,000 years ago and wild landraces still persist (Gladstones et al., 1998). White lupin (WL) is recognized as an exceptional source of protein (between 30 and 40% of the whole seed dry matter) (Bähr et al., 2014) with an adequate balance of essential amino acids, as well as an adequate protein efficiency ratio

(Sujak et al., 2006). During recent years, white lupin has gained attention for human consumption because of its levels of tocopherols, having the lowest glycemic index among consumed grains, high dietary fiber content and the absence of immunogenic epitopes causing celiac disease (Boschin et al., 2008; Boschin and Arnoldi, 2011; Fontanari et al., 2012; Bähr et al., 2014). It is also a crop with low need for phosphate fertilizers due to the presence of specialized cluster root structures and the capacity to releases phosphorus (P) from its insoluble form because of the ability to mobilize carboxylates (Lucas et al., 2015; Xu et al., 2020). Further, as a legume, application of nitrogen fertilizers can be avoided if an adequate symbiosis with *Rhyzobium* bacteria is achieved, decreasing therefore the environmental impact due to its cultivation (Pueyo et al., 2021).

However, the presence of secondary metabolites, which act as anti-nutritional compounds, in white lupin seeds limits its use. *Lupinus albus* wild varieties and landraces can accumulate up to 11% of their dry weight in the form of alkaloids, most of them belonging to the family of quinolizidine alkaloids (QA) (Rybiński et al., 2018). QA are notoriously bitter and toxic to both humans and farm animals, displaying both teratogenic and anti-cholinergic effects (Wink and Schimmer, 1999; Lourenço et al., 2002). Accordingly, traditional consumption of lupin grain involves a debittering process, which also removes a large proportion of nutrients such as soluble proteins, minerals, flavonoids, monosaccharides and sucrose from the seeds (Erbas, 2010). Despite WL importance and potential as a protein source, scarce knowledge has been generated about quinolizidine alkaloid synthesis. Cultivated WL relies on the incorporation of a natural occurring mutation at pauper locus, which decreases QA levels below the 0.02% threshold established as safe for consumption as food and feed, however, fundamental knowledge is required to maintain a reduction of QA within breeding programs.

The central objective of this review is to summarize the available resources to understand the biosynthesis of QA in *L. albus.* Identification and characterization of the genes responsible for QA biosynthesis are a challenge which breeders must pursue in order to manipulate the total amount and type present in any given genotype or commercial variety, while controlling the amount and type of QA and maintaining lupin high nutritional value. From an environmental point of view, breeding efforts could lead to optimizing selection for higher alkaloid content in the leaves (and thus resistance to pathogens and insects), while at the same time selecting for low QA in the seed (Gladstones et al., 1998). QA synthesis knowledge will provide the opportunity to ensure safe levels for human and animal consumption as well as optimal pest control, increasing yield, and decreasing the use of agrochemicals (Vishnyakova et al., 2020).

QUINOLIZIDINE ALKALOID BIOSYNTHETIC PATHWAY

Alkaloids are part of a structurally diverse group of over 21,000 cyclic nitrogen-containing secondary metabolites (Wink, 2013) that are found in over 20% of plant species (Croteau et al., 2000; De Luca and St. Pierre, 2000; Bunsupa et al., 2012b). QA are secondary metabolites that occur mostly within the Leguminosae family, but are also present in other taxa (Ohmiya et al., 1995). In lupin species, QA distinctive structure is a quinolizidine ring, which can be grouped mainly into bicyclic alkaloids, such as lupinine and its derivatives, and tetracyclic alkaloids, represented by sparteine, lupanine and hydroxilupanine (Wink, 1987). Examples exist of distinctive QA expression in different species. Isolupanine and angustifoline accumulate to high levels in *L. angustifolius* (narrow-leafed lupin). Albinine and multiflorine accumulate to high levels in *L. albus* and lupinine in *L. luteus* (Święcicki et al., 2019). QA vary in their toxicity and their deterrence against insect pests and mammals. Sparteine and lupanine appear to be the two most toxic QA to humans and laboratory animals (Allen, 1998; Petterson, 1998), with lupanine having the greatest impact on aphid survival, followed by sparteine, lupinine, 13α-hydroxylupanine and angustifoline having a moderate impact (Ridsdill-Smith et al., 2004; Philippi et al., 2015). QA toxicity against larvae from different species as well as acaricidal effect has also been reported (Hassan et al., 2019; Elma et al., 2021).

The synthesis of QA occurs through the cyclization of cadaverine, due to a L-lysine decarboxylation catalyzed by a lysine decarboxylase L/ODC (La-L/ODC) (Saito and Murakoshi, 1995; Bunsupa et al., 2012a). Cadaverine is then oxidized by a copper amine oxidase (CuAO) (Yang et al., 2017) to yield 5-aminopentanal and spontaneously cyclized to 1-piperideine Schiff base, which is a universal intermediate for the production of various Lys-derived alkaloids (Bunsupa et al., 2012b). It has been suggested that in addition to these reactions, a series of aldol-type reactions, hydrolysis, oxidative deamination and coupling gives rise to the major structural QAs [e.g., lupanine and others; (Dewick, 2002)]. The diiminium cation was proposed as an intermediate product in to yield tetracyclic alkaloids [lupanine, multiflorine, and sparteine (Fraser and Robins, 1984)]. These QA are final products, but also, can be further modified by dehydrogenation, oxygenation, hydroxylation, glycosylation, acetylation or esterification to form a wide variety of structurally related QAs (Wink and Hartmann, 1982; Saito et al., 1992, 1993; Ohmiya et al., 1995; Saito and Murakoshi, 1995; Bunsupa et al., 2012b; Boschin and Resta, 2013). Continued research will likely add more detail to the QA pathway. For example, the observation that QA can be found in high enantiomeric excess has led to the proposal that stereoselective enzyme catalysis may be involved in the QA pathway (Lichman, 2021).

GENES INVOLVED IN QUINOLIZIDINE ALKALOID BIOSYNTHESIS IN *LUPINUS*

Identification of genes involved in the QA biosynthesis has been partially achieved by identifying homologous genes in other species expressing QA, as in the case of *Lupinus angustifolius La-L/ODC* gene, which was identified as a homolog of L/*ODC* expressed in the distantly related species *S. flavescens*, *E. koreensis*, *T. chinensis*, and *B. australis* (Bunsupa et al., 2012a).

Genes encoding acyltransferase were described in *L. albus* and *L. angustifolius* (*LaHMT/HLT* and *LaAT*, *respectively*), but proof of the formation of acetylated products (13α-hydroxylupanine and 13α-hydroxymultiflorine), was only achieved for *L. albus HMT/HLT* (Saito and Murakoshi, 1995; Okada et al., 2005; Bunsupa et al., 2011).

Transcriptome experiments in different tissues of *L. angustifolius* lead to the identification of a copper amine oxidase, LaCAO (Yang et al., 2017), with cadaverine as substrate, catalyzing its transformation into 5-aminopentanal, which is then spontaneously cyclized to 1-piperideine (Yang et al., 2017). In a previous report, Okada et al. (2005) cloned and characterized an *O*-tigloytransferase from WL, an enzyme involved in the final steps of QA biosynthesis. Recently, it was proven for *Lupinus angustifolius*, that *RAP2-7* is a putative regulatory gene of QA biosynthesis/accumulation in aerial tissues (Kroc et al., 2019), with a S196R substitution being responsible for the bitter/sweet phenotype (Czepiel et al., 2021), but, however, additional studies are needed to determine the mechanism and effect on QA expression, and its role in different lupin species. In an effort to identify the missing enzymes of the QA pathway, the study of the existence of common enzymes between nicotine synthesis (monoterpene indole alkaloids, MIA; benzylisoquinole alkaloids, BIA) and lupins QA has been proposed (Frick et al., 2017). Many of these enzymes (methyltransferases, decarboxylases, oxidases, acyltransferases, cytochromes-P450, oxidoreductases, demethylases, reductases, hydroxylases, and coupling enzymes) and their encoding genes have been identified in *N. tabacum, C. roseus, C. japonica*, and *P. somniferum* (Bird et al., 2003; Dewey and Xie, 2013; Hagel and Facchini, 2013; Kilgore and Kutchan, 2016; Pan et al., 2016; Thamm et al., 2016) and it is expected that they play a role in lupins QA biosynthesis (Frick et al., 2017). *Sophora flavescens* transcriptome analysis had also identified several genes co-expressed, such as a putative *S. flavescens* L/ODC and candidate genes clustered into the same clade as L/ODC (major latex-like protein (MLP-like), a cP450, and a ripening related protein), but their function remains unknown (Han et al., 2015). Berberine bridge and berberine bridge-like enzymes catalyze oxidative reactions for the biosynthesis of BIAs (Facchini et al., 1996; Samanani et al., 2004; Kajikawa et al., 2011), possibly having similar roles in QA biosynthesis. Cytocromes-P450 have a role in hydroxylation reactions, as well as other reactions in MIA and BIA biosynthesis (Pauli and Kutchan, 1998; Thamm et al., 2016). Recently, the existence of a high number of QA biosynthesis genes controlled by a regulatory agent localized at *iucundus* locus in NLL was reported, which supports the idea that ethylene responsive transcription factor *RAP2-7* gene may control low-alkaloid phenotype in NLL, acting as a promoter of the expression of biosynthesis genes (Plewiński et al., 2019; Czepiel et al., 2021).

In addition to QAs biosynthetic genes, major *loci* controlling QA expression have been described in lupins. Cultivated lupins display lower alkaloid content than landraces, due to incorporation of "sweet" domestication genes, which were generated by natural mutation (Lin et al., 2009). Most of these mutations are recessive, such as *iucundus, esculentus*, and *depressus* in NLL, *amoenus, dulcis*, and *liber* in *L. luteus* (Lin et al., 2009). In *L. albus*, several *loci* have been reported to produce low alkaloid genotypes, with the *pauper* locus being the most effective and used worldwide in breeding programs (Gladstones, 1974; Harrison and Willliams, 1982). In *L. mutabilis*, the low alkaloid phenotype is controlled by several alleles (Clements et al., 2008). It is worth highlighting that none of the mutations identified in lupin completely eliminate QAs (Harrison and Willliams, 1982).

Pauper LOCUS

Construction of low-density linkage maps allowed identification of genomic regions involved in alkaloid biosynthesis in white lupin (Phan et al., 2007; Croxford et al., 2008; Vipin et al., 2013). There has also been development of molecular markers to identify QTLs responsible for low alkaloid content linked to these recessive loci (Yang et al., 2010; Raman et al., 2014). The development and mass use of GBS technology as a tool for breeders (Elshire et al., 2011; Annicchiarico et al., 2017), enabled progress to identify causative genes for low QA content in *L. albus*. High-density consensus maps for comparisons between *L. angustifolious* and *L. albus*, had led to the hypothesis that the *iucundus* locus responsible for low alkaloid content in NLL differs by function from *pauper* in WL (Książkiewicz et al., 2017).

Worldwide, *L. albus* breeding programs have relied mainly on the effect of *pauper locus* to produce sweet varieties, for both food and feed. Besides its importance, little is known about the *pauper* locus gene(s) with respect to their effect on alkaloid content. Earlier studies identified at least two different alleles for *pauper locus*, controlling total alkaloid content (Harrison and Willliams, 1982). The action of this recessive locus was suggested as a reduction of a common alkaloid substrate, which seems to be the ubiquitous for most lupin low alkaloid genotypes, without affecting intermediate substrates at late biosynthetic stages, when chemical differences among alkaloids are being finally specified (Harrison and Willliams, 1982).

Genetic and comparative map of *L. albus*, based on a RIL population developed from Kiev (Ukrainian cv, sweet, early flowering, anthracnose susceptible) and P27174 (Ethiopian landrace, bitter, late flowering, anthracnose resistant), allowed the discovery of 220 amplified fragment length polymorphisms and 105 gene-based markers, enabling for the first time mapping of the alkaloid locus, with flanking markers located in a region within 20 cM in both directions (Phan et al., 2007). Later, with the development of Pauper M1, a molecular marker more closely linked to the pauper locus (1.4 cM), allowed discrimination of low alkaloid content genotypes with efficiencies restricted to ∼95% for bitter and 91% for sweet non-*pauper* lines. Implementing PauperM1 required the use of sequencing gels and radioisotope primer labeling for the correct determination of alleles, which made its application restricted to authorized facilities (Lin et al., 2009). With the aid of GBS, a high-density consensus linkage map of WL genome was constructed, integrating 453 published markers with 3,597 newly developed sequence-based markers, recovering a single linkage group per every chromosome (Książkiewicz et al., 2017). This map yielded several new markers co-segregating, or closely localized to the

pauper locus than the Pauper M1 (Książkiewicz et al., 2017). In an effort to improve Pauper M1 efficacy, CAPS markers were developed (using two identified SNPs), which were substrates for restriction enzymes, *Hha*I for the bitter allele, and *Hin*fI for the sweet allele (Rychel and Książkiewicz, 2019). One of these, LAGI01_35805_F1_R1 homologous to *LaAT* (AB581532.1), different than *pauper* locus gene, showed higher efficacy than Pauper M1 (Rychel and Książkiewicz, 2019).

Recently a high quality reference *L. albus* genome allowed the study of the *pauper* genomic region, identifying several candidate genes on Chr18. This research demonstrate the existence of a gene cluster in the *pauper* locus, which comprises a 958 kb region and contains 66 genes, amongst which several are strong candidates genes encoding enzymatic activities, such as cinnamoyl-CoA reductase and acyltransferases (Hufnagel et al., 2020a). In addition to the reference genome, transcriptomic data from different organs, resequencing of 15 varieties and a pangenome dataset provide tools for further exploration of the genomics of alkaloid content[1] (Hufnagel et al., 2020b) and the complex role of the pauper locus. Analysis of the pauper marker associated with low QA was carried out in a landrace, breeding lines and cultivars of *L. albus* and QA was measured by UHPLC-HRMS. Interestingly while the marker did associate with low QA and was absent in many high QA samples, there were notable exceptions where the marker was found in high QA sample (Zafeiriou et al., 2021). Thus additional genes and regulatory elements may be important in reaching breeding objectives.

[1]https://www.whitelupin.fr/

PROSPECTIVE TECHNIQUES TO UNRAVEL BIOSYNTHESIS ALKALOID GENES IN WHITE LUPIN

The main drawback to study candidate genes and their function in WL is the recalcitrant nature of this legume to tissue culture (Nguyen et al., 2016; Aslam et al., 2020). Several attempts have been undertaken to develop *in vitro* regeneration tissue culture systems, but with limited success, which today represent a challenge in WL breeding programs (Bayliss et al., 2002; Uhde-Stone et al., 2005; Surma et al., 2013; Raza et al., 2017; Che et al., 2019; Aslam et al., 2020). It is possible, as in monocot species, that transformation is highly cultivar dependent, therefore a genetic screen for transformation aptitude in WL collections may produce a genotype amenable to transformation and thus genome editing. But, in the meantime, to overcome this situation, different reverse genetics methodologies (Till et al., 2007), such as random mutagenesis and virus induced gene silencing (VIGS) may be used to probe gene function (Gupta et al., 2013).

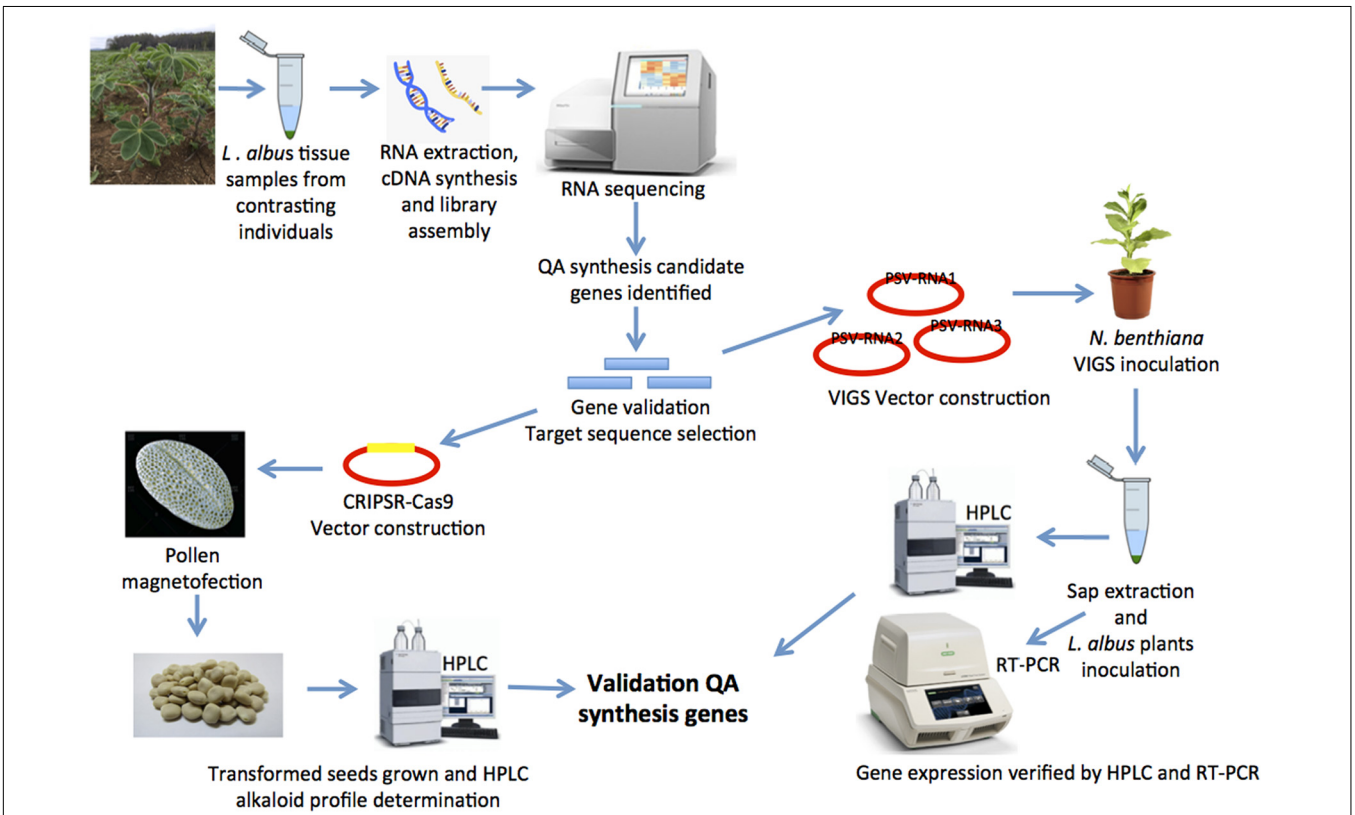


FIGURE 1 | A strategy to unravel quinolizidine alkaloid (QA) synthesis genes using virus induced gene silencing and genome editing, for functional genomics and crop improvement in *Lupinus albus*.

During the last years, the use of transient expression has facilitated gene-discovery by utilization of VIGS (Gupta et al., 2013). VIGS is an effective tool to characterize functions of candidate genes using post-transcriptional gene silencing (PTGS), which is extensively used for gene knockdowns in plants (Liu et al., 2016; Wang et al., 2016; Zhang et al., 2016). However, VIGS can also be applied as a forward genetics technique to study gene function by using cDNA libraries (Kilgore and Kutchan, 2016; Thamm et al., 2016). In white lupin, VIGS using peanut stunt virus proved to be effective tool to silence the Phytoene desaturase gene (LaPDS) (Yamagishi et al., 2015), opening possibilities for utilization of this technique to elucidate genes participating in secondary metabolite synthesis such as QA.

Advances in CRISPR-Cas technology allow fine-tuning of gene-activity and the generation larger chromosomal variation, providing a broad toolkit for gene-function analysis (Jinek et al., 2012; Cong et al., 2013; Jung and Till, 2021). Nevertheless, optimization of CRISPR technology is needed to accommodate the tissue and transformation delivery method (Char et al., 2017). Recent developments of new transformation techniques based on the utilization of functionalized nanoparticles to deliver DNA, has been proven successful in species such as cotton, sunflower and lily (Zhao et al., 2017), opening the possibilities to bypass *in vitro* regeneration in legume species, such as lupin. With the aid of magnetic fields, nanoparticles can efficiently deliver CRISPR vectors through pores present in the pollen grains, producing transformed pollen which is then used to pollinate emasculated flowers, resulting in transformed seeds (Zhao et al., 2017). The utilization of functionalized magnetic particles to deliver DNA into pollen grains, and to accelerate selection of desired individuals using speed-breeding (Watson et al., 2018; Lew et al., 2020) increases transformation efficiency and bypasses tissue culture procedures to generate plants from transformed seeds within a short period of time, broadening the possibilities for WL utilization as mayor knowledge on QA synthesis is achieved. Continued advances in nanoparticle technologies may make CRISPR-Cas approaches amenable in many recalcitrant species (Ma et al., 2021).

SUMMARY

Quinolizidine alkaloid synthesis has an important number of unresolved questions, which hinders breeding efforts worldwide in a crop with high nutritional quality such as WL. Utilization of techniques to study and manipulate genes involved in alkaloid synthesis in *L. albus* will contribute to a better understanding of the accumulation of secondary metabolites in lupin seed (**Figure 1**), contributing to the development of environmentally friendly and sustainable sources of plant protein, which are expected to be a key component of conscientious population growth. Continued efforts in white lupin breeding, leveraging knowledge gained of the genetics of QA synthesis, can have an important role in human nutrition and well being.

AUTHOR CONTRIBUTIONS

CO and BT wrote, edited and reviewed the manuscript, and approved the submitted version.

FUNDING

The authors would like to thank Instituto de Investigaciones Agropecuarias, Breeding and Plant Biotechnology Group, Plant Production Area, through grant 502610-70.

REFERENCES

Abraham, E. M., Ganopoulos, I., Madesis, P., Mavromatis, A., Mylona, P., Nianiou-Obeidat, I., et al. (2019). The use of lupin as a source of protein in animal feeding: genomic tools and breeding approaches. *Int. J. Mol. Sci.* 20:851.

Allen, J. G. (1998). "Toxins and lupinosis" in *Lupins as a Crop Plant Biology, Production and Utilization*. eds J. S. Gladstones, C. A. Atkins, and J. Hamblin (Cambridge: CAB International, University Press). 411–428.

Annicchiarico, P., Nazzicari, N., Wei, Y., Pecetti, L., and Brummer, E. C. (2017). Genotyping-by-Sequencing and its exploitation for forage and cool-season grain legume breeding. *Front. Plant Sci.* 8:679. doi: 10.3389/fpls.2017.00679

Aslam, M. M., Karanja, J. K., Zhang, Q., Lin, H., Xia, T., Akhtar, K., et al. (2020). In Vitro Regeneration Potential of White Lupin (Lupinus albus) from Cotyledonary Nodes. *Plants* 9:318. doi: 10.3390/plants9030318

Bähr, M., Fechner, A., Hasenkopf, K., Mittermaier, S., and Jahreis, G. (2014). Chemical composition of dehulled seeds of selected lupin cultivars in comparison to pea and soya bean. *LWT Food Sci. Technol.* 59, 587–590. doi: 10.1016/j.lwt.2014.05.026

Bayliss, K., Wroth, J., and Cowling, W. (2002). *Pro-embryos of Lupinus Albus Produced from Isolated Microspore Culture*. Australian: The Australian Plant Breeding Association Inc., 57–61.

Bird, D. A., Franceschi, V. R., and Facchini, P. J. (2003). A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. *Plant Cell.* 15, 2626–2635. doi: 10.1105/tpc.015396

Boschin, G., and Arnoldi, A. (2011). Legumes are valuable sources of tocopherols. *Food Chem.* 127, 1199–1203. doi: 10.1016/j.foodchem.2011.01.124

Boschin, G., D'Agostina, A., Annicchiarico, P., and Arnoldi, A. (2008). Effect of genotype and environment on fatty acid composition of Lupinus albus L. seed. *Food Chem.* 108, 600–606. doi: 10.1016/j.foodchem.2007.11.016

Boschin, G., and Resta, D. (2013). "Alkaloids Derived from Lysine: quinolizidine (a Focus on Lupin Alkaloids)" in *Natural Products: phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes*. eds G. K. Ramawat and J.-M. Mérillon (Berlin: Springer). 381–403. doi: 10.1007/978-3-642-22144-6_11

Bunsupa, S., Katayama, K., Ikeura, E., Oikawa, A., Toyooka, K., Saito, K., et al. (2012a). Lysine decarboxylase catalyzes the first step of quinolizidine alkaloid biosynthesis and coevolved with alkaloid production in leguminosae. *Plant Cell* 24, 1202–1216. doi: 10.1105/tpc.112.095885

Bunsupa, S., Yamazaki, M., and Saito, K. (2012b). Quinolizidine alkaloid biosynthesis: recent advances and future prospects. *Front. Plant Sci.* 3:239. doi: 10.3389/fpls.2012.00239

Bunsupa, S., Okada, T., Saito, K., and Yamazaki, M. (2011). An acyltransferase-like gene obtained by differential gene expression profiles of quinolizidine

alkaloid-producing and nonproducing cultivars of Lupinus angustifolius. *Plant Biotechnol.* 28, 89–94. doi: 10.5511/plantbiotechnology.10.1109b

Char, S. N., Neelakandan, A. K., Nahampun, H., Frame, B., Main, M., Spalding, M. H., et al. (2017). An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnol. J.* 15, 257–268. doi: 10.1111/pbi.12611

Che, P., Chang, S., Simon, M. K., Zhang, Z., Shaharyar, A., Ourada, J., et al. (2019). Developing a rapid and highly efficient cowpea regeneration and transformation system using embryonic axis explants. *Biorxiv* 738971. doi: 10.1101/738971

Clements, J., Prilyuk, L., Quealy, J., and Francis, G. (2008). "Interspecific crossing among the New World lupin species for Lupinus mutabilis crop improvement" in *Lupins for Health and Wealth Proceedings of the 12th International Lupin Conference*. eds J. A. Palta and J. B. Berger. (Western Australia: International Lupin Association). 14.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.

Croteau, R., Kutchan, T. M., and Lewis, N. G. (2000). Natural products (secondary metabolites). *Biochem. Mol. Biol. Plants* 24, 1250–1319. doi: 10.1128/mSystems.00186-17

Croxford, A. E., Rogers, T., Caligari, P. D., and Wilkinson, M. J. (2008). High-resolution melt analysis to identify and map sequence-tagged site anchor points onto linkage maps: a white lupin (Lupinus albus) map as an exemplar. *New Phytol.* 180, 594–607. doi: 10.1111/j.1469-8137.2008.02588.x

Czepiel, K., Krajewski, P., Wilczura, P., Bielecka, P., Święcicki, W., and Kroc, M. (2021). Expression Profiles of Alkaloid-Related Genes across the Organs of Narrow-Leafed Lupin (Lupinus angustifolius L.) and in Response to Anthracnose Infection. *Int. J. Mol. Sci.* 22:2676. doi: 10.3390/ijms22052676

De Luca, V., and St. Pierre, B. (2000). The cell and developmental biology of alkaloid biosynthesis. *Trends Plant Sci.* 5, 168–173. doi: 10.1016/s1360-1385(00)01575-2

Dewey, R. E., and Xie, J. (2013). Molecular genetics of alkaloid biosynthesis in Nicotiana tabacum. *Phytochemistry* 94, 10–27. doi: 10.1016/j.phytochem.2013.06.002

Dewick, P. M. (2002). *Medicinal Natural Products: a Biosynthetic Approach*. United States: John Wiley & Sons.

Elma, F. N., Çetín, H., Yorgancilar, M., and Acar, R. (2021). Detection of Metabolite Content in Local Bitter White Lupin Seeds (Lupinus albus L.) and Acaricidal and Insecticidal Effect of its Seed Extract. *Tarım Biliml. Derg.* 27, 407–413.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379

Erbas, M. (2010). The effects of different debittering methods on the production of lupin bean snack from bitter Lupinus albus L. seeds. *J. Food Qual.* 33, 742–757. doi: 10.1111/j.1745-4557.2010.00347.x

Facchini, P. J., Penzes, C., Johnson, A. G., and Bull, D. (1996). Molecular characterization of berberine bridge enzyme genes from opium poppy. *Plant Physiol.* 112, 1669–1677. doi: 10.1104/pp.112.4.1669

FAO (2021). *FAOSTAT. Crops and Livestock Products.* Available online at: https://www.fao.org/faostat/en/#data/QCL (accessed January 07, 2022).

Fontanari, G. G., Batistuti, J. P., da Cruz, R. J., Saldiva, P. H. N., and Arẹas, J. A. G. (2012). Cholesterol-lowering effect of whole lupin (Lupinus albus) seed and its protein isolate. *Food Chem.* 132, 1521–1526. doi: 10.1016/j.foodchem.2011.11.145

Fraser, A. M., and Robins, D. J. (1984). Incorporation of chiral [1-2 H] cadaverines into the quinolizidine alkaloids sparteine, lupanine, and angustifoline. *J. Chem. Soc. Chem. Commun.* 1477–1479. doi: 10.1039/c39840001477

Frick, K. M., Kamphuis, L. G., Siddique, K. H., Singh, K. B., and Foley, R. C. (2017). Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement. *Front. Plant Sci.* 8:87. doi: 10.3389/fpls.2017.00087

Gladstones, J. S. (1974). *Lupins of the Mediterranean region and Africa*. Perth: Department of Agriculture of Western Australia.

Gladstones, J. S., Atkins, C., and Hamblin, J. (1998). *Lupins as Crop Plants: biology, Production and Utilization*. UK: CAB international.

Gupta, B., Saha, J., Sengupta, A., and Gupta, K. (2013). Recent advances on Virus Induced Gene silencing (VIGS): plant functional genomics. *J. Plant Biochem. Physiol.* 1:e116. doi: 10.4172/2329-9029.1000e116

Hagel, J. M., and Facchini, P. J. (2013). Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant Cell Physiol.* 54, 647–672. doi: 10.1093/pcp/pct020

Han, R., Takahashi, H., Nakamura, M., Bunsupa, S., Yoshimoto, N., Yamamoto, H., et al. (2015). Transcriptome analysis of nine tissues to discover genes involved in the biosynthesis of active ingredients in Sophora flavescens. *Biol. Pharmaceut. Bull.* 38, 876–883. doi: 10.1248/bpb.b14-00834

Harrison, J., and Willliams, W. (1982). Genetical control of alkaloids in Lupinus albus. *Euphytica* 31, 357–364. doi: 10.1007/bf00021651

Hassan, M. E., Aly, N. E.-D. S., and Mikhail, M. W. (2019). Larvicidal effect of alkaloids extracted from bitter lupin seeds against mosquitoes (Culex pipiens), flies (Musca domestica) and fleas (Xenopsylla cheopis) under laboratory conditions in Egipt. *J. Egypt. Soc. Parasitol.* 49, 455–464. doi: 10.21608/jesp.2019.68192

Hufnagel, B., Marques, A., Soriano, A., Marquès, L., Divol, F., Doumas, P., et al. (2020a). High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. *Nat. Commun.* 11, 1–12. doi: 10.1038/s41467-019-14197-9

Hufnagel, B., Soriano, A., Taylor, J., Divol, F., Kroc, M., Sanders, H., et al. (2020b). Pangenome of white lupin provides insights into the diversity of the species. *Biorxiv* doi: 10.1111/pbi.13678.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jung, C., and Till, B. (2021). Mutagenesis and genome editing in crop improvement: perspectives for the global regulatory landscape. *Trends Plant Sci.* 26, 1258–1269. doi: 10.1016/j.tplants.2021.08.002

Kajikawa, M., Shoji, T., Kato, A., and Hashimoto, T. (2011). Vacuole-localized berberine bridge enzyme-like proteins are required for a late step of nicotine biosynthesis in tobacco. *Plant Physiol.* 155, 2010–2022. doi: 10.1104/pp.110.170878

Kilgore, M. B., and Kutchan, T. M. (2016). The Amaryllidaceae alkaloids: biosynthesis and methods for enzyme discovery. *Phytochem. Rev.* 15, 317–337. doi: 10.1007/s11101-015-9451-z

Kroc, M., Koczyk, G., Kamel, K. A., Czepiel, K., Fedorowicz-Strońska, O., Krajewski, P., et al. (2019). Transcriptome-derived investigation of biosynthesis of quinolizidine alkaloids in narrow-leafed lupin (Lupinus angustifolius L.) highlights candidate genes linked to iucundus locus. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-018-37701-5

Książkiewicz, M., Nazzicari, N., Nelson, M. N., Renshaw, D., Rychel, S., Ferrari, B., et al. (2017). A high-density consensus linkage map of white lupin highlights synteny with narrow-leafed lupin and provides markers tagging key agronomic traits. *Sci. Rep.* 7, 1–15. doi: 10.1038/s41598-017-15625-w

Lew, T. T. S., Park, M., Wang, Y., Gordiichuk, P., Yeap, W.-C., Mohd Rais, S. K., et al. (2020). Nanocarriers for transgene expression in pollen as a plant biotechnology tool. *ACS Mater. Lett.* 2, 1057–1066. doi: 10.1021/acsmaterialslett.0c00247

Lichman, B. R. (2021). The scaffold-forming steps of plant alkaloid biosynthesis. *Nat. Prod. Rep.* 38, 103–129. doi: 10.1039/d0np00031k

Lin, R., Renshaw, D., Luckett, D., Clements, J., Yan, G., Adhikari, K., et al. (2009). Development of a sequence-specific PCR marker linked to the gene "pauper" conferring low-alkaloids in white lupin (Lupinus albus L.) for marker assisted selection. *Mol. Breed.* 23, 153–161. doi: 10.1007/s11032-008-9222-2

Liu, N., Xie, K., Jia, Q., Zhao, J., Chen, T., Li, H., et al. (2016). Foxtail mosaic virus-induced gene silencing in monocot plants. *Plant Physiol.* 171, 1801–1807. doi: 10.1104/pp.16.00010

Lourenço, A., Máximo, P., Ferreira, L., and Pereira, M. (2002). Indolizidine and quinolizidine alkaloids structure and bioactivity. *Stud. Nat. Prod. Chem.* 27, 233–298. doi: 10.1016/s1572-5995(02)80038-2

Lucas, M. M., Stoddard, F., Annicchiarico, P., Frias, J., Martinez-Villaluenga, C., Sussmann, D., et al. (2015). The future of lupin as a protein crop in Europe. *Front. Plant Sci.* 6:705. doi: 10.3389/fpls.2015.00705

Ma, K., Li, W., Zhu, G., Sun, S., Chi, H., Yin, Y., et al. (2021). Functionalized PDA/DEX-PEI@ HA nanoparticles combined with sleeping-beauty transposons for multistage targeted delivery of CRISPR/Cas9 gene. *Biomed. Pharmacother.* 142:112061. doi: 10.1016/j.biopha.2021.112061

Nguyen, A. H., Hodgson, L. M., Erskine, W., and Barker, S. J. (2016). An approach to overcoming regeneration recalcitrance in genetic transformation of lupins and other legumes. *Plant Cell Tissue Organ Cult.* 127, 623–635. doi: 10.1007/s11240-016-1087-1

Ohmiya, S., Saito, K., and Murakoshi, I. (1995). "Lupin alkaloids" in *The Alkaloids: chemistry and Pharmacology*. ed G. A. Cordell (San Diego: Academic Press). 1–114.

Okada, T., Hirai, M. Y., Suzuki, H., Yamazaki, M., and Saito, K. (2005). Molecular characterization of a novel quinolizidine alkaloid O-tigloyltransferase: cDNA cloning, catalytic activity of recombinant protein and expression analysis in Lupinus plants. *Plant Cell Physiol.* 46, 233–244. doi: 10.1093/pcp/pci021

Pan, Q., Mustafa, N. R., Tang, K., Choi, Y. H., and Verpoorte, R. (2016). Monoterpenoid indole alkaloids biosynthesis and its regulation in Catharanthus roseus: a literature review from genes to metabolites. *Phytochem. Rev.* 15, 221–250. doi: 10.1007/s00709-011-0291-4

Pauli, H. H., and Kutchan, T. M. (1998). Molecular cloning and functional heterologous expression of two alleles encoding (S)-N-methylcoclaurine 3′-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. *Plant J.* 13, 793–801. doi: 10.1046/j.1365-313x.1998.00085.x

Peterson, D. S. (2000). The use of Lupins in Feeding Systems. *Asian Austr. J. Anim. Sci.* 13, 861–882. doi: 10.5713/ajas.2000.861

Petterson, D. (1998). "Composition and food uses of lupin" in *Lupins as Crop Plants: biology, Production, and Utilization*. eds J. Gladstones, C. Atkins, and H. Wallingford (UK: CAB International). 353–384.

Petterson, D., and Mackintosh, J. B. (1994). *The Chemical Composition and Nutritive Value of Australian Grain Legumes*. Canberra: Grains Research and Development Corp.

Phan, H. T., Ellwood, S. R., Adhikari, K., Nelson, M. N., and Oliver, R. P. (2007). The first genetic and comparative map of white lupin (Lupinus albus L.): identification of QTLs for anthracnose resistance and flowering time, and a locus for alkaloid content. *DNA Res.* 14, 59–70. doi: 10.1093/dnares/dsm009

Philippi, J., Schliephake, E., Jürgens, H. U., Jansen, G., and Ordon, F. (2015). Feeding behavior of aphids on narrow-leafed lupin (L upinus angustifolius) genotypes varying in the content of quinolizidine alkaloids. *Entomol. Exp. Appl.* 156, 37–51. doi: 10.1111/eea.12313

Plewiński, P., Książkiewicz, M., Rychel-Bielska, S., Rudy, E., and Wolko, B. (2019). Candidate Domestication-Related Genes Revealed by Expression Quantitative Trait Loci Mapping of Narrow-Leafed Lupin (Lupinus angustifolius L.). *Int. J. Mol. Sci.* 20:5670. doi: 10.3390/ijms20225670

Pueyo, J. J., Quiñones, M. A., Coba de la Peña, T., Fedorova, E. E., and Lucas, M. M. (2021). Nitrogen and Phosphorus Interplay in Lupin Root Nodules and Cluster Roots. *Front. Plant Sci.* 12:644218. doi: 10.3389/fpls.2021.644218

Raman, R., Cowley, R. B., Raman, H., and Luckett, D. J. (2014). Analyses using SSR and DArT molecular markers reveal that Ethiopian accessions of white lupin (Lupinus albus L.) represent a unique genepool. *Open J. Genet.* 04, 87–98. doi: 10.4236/ojgen.2014.42012

Raza, G., Singh, M. B., and Bhalla, P. L. (2017). In Vitro Plant Regeneration from Commercial Cultivars of Soybean. *Biomed. Res. Int.* 2017:7379693. doi: 10.1155/2017/7379693

Ridsdill-Smith, T., Edwards, O., Wang, S. F., Ghisalberti, E., and Reidy-Crofts, J. (2004). "Aphid responses to plant defensive compounds in lupins"in *Aphids in a New Millennium.* eds J. C. Simon, C. A. Dedryver, C. Rispe, and M. Hullé (France: INRA). 491–497.

Rybiński, W., Kroc, M., Święcicki, W., Wilczura, P., Kamel, K., Barzyk, P., et al. (2018). *Preliminary Estimation of Variation of Alkaloids Content in White Lupin (Lupinus albus L.) Collection*. Germany: Springer. 131–136.

Rychel, S., and Książkiewicz, M. (2019). Development of gene-based molecular markers tagging low alkaloid pauper locus in white lupin (Lupinus albus L.). *J. Appl. Genet.* 60, 269–281. doi: 10.1007/s13353-019-00508-9

Saito, K., Koike, Y., Suzuki, H., and Murakoshi, I. (1993). Biogenetic implication of lupin alkaloid biosynthesis in bitter and sweet forms of Lupinus luteus and L. albus. *Phytochemistry* 34, 1041–1044. doi: 10.1016/s0031-9422(00)90709-x

Saito, K., and Murakoshi, I. (1995). Chemistry, biochemistry and chemotaxonomy of lupin alkaloids in the Leguminosae. *Stud. Nat. Prod. Chem.* 15, 519–549. doi: 10.1016/s1572-5995(06)80142-0

Saito, K., Suzuki, H., Takamatsu, S., and Murakoshi, I. (1992). Acyltransferases for lupin alkaloids in Lupinus hirsutus. *Phytochemistry* 32, 87–91.

Samanani, N., Liscombe, D. K., and Facchini, P. J. (2004). Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. *Plant J.* 40, 302–313. doi: 10.1111/j.1365-313X.2004.02210.x

Sońta, M., and Rekiel, A. (2019). Legumes—Use for nutritional and feeding purposes. *J. Elementol.* 25, 835–849.

Sujak, A., Kotlarz, A., and Strobel, W. (2006). Compositional and nutritional evaluation of several lupin seeds. *Food Chem.* 98, 711–719.

Surma, M., Adamski, T., Swiecicki, W., Barzyk, P., Kaczmarek, Z., Kuczynska, A., et al. (2013). Preliminary results of in vitro culture of pea and lupin embryos for the reduction of generation cycles in single seed descent technique. *Acta Soc. Botanicor. Polon.* 82:3.

Święcicki, W., Czepiel, K., Wilczura, P., Barzyk, P., Kaczmarek, Z., and Kroc, M. (2019). Chromatographic fingerprinting of the Old World lupins seed alkaloids: a supplemental tool in species discrimination. *Plants* 8:548. doi: 10.3390/plants8120548

Thamm, A. M., Qu, Y., and De Luca, V. (2016). Discovery and metabolic engineering of iridoid/secoiridoid and monoterpenoid indole alkaloid biosynthesis. *Phytochem. Rev.* 15, 339–361. doi: 10.1007/s11101-016-9468-y

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S., et al. (2007). Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* 7:19. doi: 10.1186/1471-2229-7-19

Uhde-Stone, C., Liu, J., Zinn, K. E., Allan, D. L., and Vance, C. P. (2005). Transgenic proteoid roots of white lupin: a vehicle for characterizing and silencing root genes involved in adaptation to P stress. *Plant J.* 44, 840–853. doi: 10.1111/j.1365-313X.2005.02573.x

Vipin, C. A., Luckett, D. J., Harper, J. D., Ash, G. J., Kilian, A., Ellwood, S. R., et al. (2013). Construction of integrated linkage map of a recombinant inbred line population of white lupin (Lupinus albus L.). *Breed. Sci.* 63, 292–300. doi: 10.1270/jsbbs.63.292

Vishnyakova, M., Kushnareva, A., Shelenga, T., and Egorova, G. (2020). Alkaloids of narrow-leaved lupine as a factor determining alternative ways of the crop's utilization and breeding. *Vavilov J. Genet. Breed.* 24:625. doi: 10.18699/VJ20.656

Wang, R., Yang, X., Wang, N., Liu, X., Nelson, R. S., Li, W., et al. (2016). An efficient virus-induced gene silencing vector for maize functional genomics research. *Plant J.* 86, 102–115. doi: 10.1111/tpj.13142

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8

Wink, M. (1987). Physiology of the accumulation of secondary metabolites with special reference to alkaloids. *Cell Cult. Somat. Cell Genet. Plants* 4, 17–42. doi: 10.1016/b978-0-12-715004-8.50008-9

Wink, M. (2013). Evolution of secondary metabolites in legumes (Fabaceae). *South Afr. J. Bot.* 89, 164–175. doi: 10.1016/j.sajb.2013.06.006

Wink, M., and Hartmann, T. (1982). Localization of the enzymes of quinilizidine alkaloid biosynthesis in leaf chloroplasts of Lupinus polyphyllus. *Plant Physiol.* 70, 74–77. doi: 10.1104/pp.70.1.74

Wink, M., and Schimmer, O. (1999). Modes of action of defensive secondary metabolites. *Annu. Plant Rev.* 3, 17–133.

Xu, W., Zhang, Q., Yuan, W., Xu, F., Aslam, M. M., Miao, R., et al. (2020). The genome evolution and low-phosphorus adaptation in white lupin. *Nat. Commun.* 11, 1–13. doi: 10.1038/s41467-020-14891-z

Yamagishi, M., Masuta, C., Suzuki, M., and Netsu, O. (2015). Peanut stunt virus-induced gene silencing in white lupin (Lupinus albus). *Plant Biotechnol.* 15:0521.

Yang, H., Lin, R., Renshaw, D., Li, C., Adhikari, K., Thomas, G., et al. (2010). Development of sequence-specific PCR markers associated with a polygenic controlled trait for marker-assisted selection using a modified selective genotyping strategy: a case study on anthracnose disease resistance in white lupin (Lupinus albus L.). *Mol. Breed.* 25, 239–249. doi: 10.1007/s11032-009-9325-4

Yang, T., Nagy, I., Mancinotti, D., Otterbach, S. L., Andersen, T. B., Motawia, M. S., et al. (2017). Transcript profiling of a bitter variety of narrow-leafed lupin to discover alkaloid biosynthetic genes. *J. Exp. Bot.* 68, 5527–5537. doi: 10.1093/jxb/erx362

Zafeiriou, I., Polidoros, A. N., Baira, E., Kasiotis, K. M., Machera, K., and Mylona, P. V. (2021). Mediterranean White Lupin Landraces as a Valuable Genetic Reserve for Breeding. *Plants* 10:2403.

Zhang, N., Huo, W., Zhang, L., Chen, F., and Cui, D. (2016). Identification of winter-responsive proteins in bread wheat using proteomics analysis and virus-induced gene silencing (VIGS). *Mol. Cell. Proteomics* 15, 2954–2969. doi: 10.1074/mcp.m115.057232

Zhao, X., Meng, Z., Wang, Y., Chen, W., Sun, C., Cui, B., et al. (2017). Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. *Nat. Plants* 3:956.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Osorio and Till. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 March 2022
doi: 10.3389/fpls.2022.814325

Effect of Drought and Low P on Yield and Nutritional Content in Common Bean

Millicent R. Smith[1†], Eric Dinglasan[2], Erik Veneklaas[3], Jose Polania[4], Idupulapati M. Rao[4], Stephen E. Beebe[4] and Andrew Merchant[1*]*

[1] School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, Sydney, NSW, Australia,
[2] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Saint Lucia, QLD, Australia,
[3] School of Biological Sciences and Institute of Agriculture, The University of Western Australia, Crawley, WA, Australia,
[4] Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia

OPEN ACCESS

Edited by:
Sergio J. Ochatt, INRA UMR 1347 Agroécologie, France

Reviewed by:
Frédéric Marsolais, Agriculture and Agri-Food Canada (AAFC), Canada
Tendai Polite Chibarabada, WaterNet Trust, Botswana
Jorge Alberto Acosta-Gallegos, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), Mexico

***Correspondence:**
Millicent R. Smith
millicent.smith@uq.edu.au
Andrew Merchant
andrew.merchant@sydney.edu.au

†Present address:
Millicent R. Smith, School of Agriculture and Food Sciences, Faculty of Science, The University of Queensland, Gatton, QLD, Australia

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: *13 November 2021*
Accepted: *16 February 2022*
Published: *29 March 2022*

Citation:
Smith MR, Dinglasan E, Veneklaas E, Polania J, Rao IM, Beebe SE and Merchant A (2022) Effect of Drought and Low P on Yield and Nutritional Content in Common Bean. Front. Plant Sci. 13:814325. doi: 10.3389/fpls.2022.814325

Common bean (*Phaseolus vulgaris* L.) production in the tropics typically occurs in rainfed systems on marginal lands where yields are low, primarily as a consequence of drought and low phosphorus (P) availability in soil. This study aimed to investigate the physiological and chemical responses of 12 bush bean genotypes for adaptation to individual and combined stress factors of drought and low P availability. Water stress and P deficiency, both individually and combined, decreased seed weight and aboveground biomass by ∼80%. Water deficit and P deficiency decreased photosynthesis and stomatal conductance during plant development. Maximum rates of carboxylation, electron transport, and triose phosphate utilization were superior for two common bean genotypes (SEF60 and NCB226) that are better adapted to combined stress conditions of water deficit and low P compared to the commercial check (DOR390). In response to water deficit treatment, carbon isotope fractionation in the leaf tissue decreased at all developmental stages. Within the soluble leaf fraction, combined water deficit and low P, led to significant changes in the concentration of key nutrients and amino acids, whereas no impact was detected in the seed. Our results suggest that common bean genotypes have a degree of resilience in yield development, expressed in traits such as pod harvest index, and conservation of nutritional content in the seed. Further exploration of the chemical and physiological traits identified here will enhance the resilience of common bean production systems in the tropics.

Keywords: abiotic stress, legume, *Phaseolus*, phosphorus, protein, water deficit

INTRODUCTION

Common bean (*Phaseolus vulgaris* L.) is the most important grain legume for human consumption (Beebe, 2012). In tropical regions where common bean is typically cultivated, 60% of production is at risk of intermittent or terminal drought stress (Beebe et al., 2008) and 50% of the area suffers from low phosphorus (P) availability in soil (Beebe et al., 2009). Consequently, drought frequently occurs in combination with low P availability (Ho et al., 2005) in regions where producers have less capital investment for improvements (Broughton et al., 2003; Cavalieri et al., 2011). In common bean, drought has been demonstrated to significantly impact yield quantity as a consequence of

interference with pod development and seed fill (Beebe et al., 2008) and disruption of processes related to carbon partitioning (Cuellar-Ortiz et al., 2008). Low P availability also reduces the yield of common bean by impacting mainly on photosynthesis, metabolism, and carbohydrate partitioning between source and sink tissues (Nielsen et al., 2001; Hermans et al., 2006). Chemical and physiological acclimation and adaptation to the effects of abiotic stress are multifaceted and likely associated with greater overall plant efficiency allowing the plant to adapt to a range of abiotic stresses (Beebe et al., 2008). For example, physiological and chemical responses such as the remobilization of carbohydrates into reproductive tissues are a shared mechanism for resistance, at least to drought and limited P availability within common bean (Beebe et al., 2008; Cuellar-Ortiz et al., 2008). Understanding the physiological processes that underlie responses to abiotic stress and the implications of this on yield and nutritional quality is critical to identify superior common bean genotypes for production in tropical regions.

The impact of drought and low P availability on common bean yield is well described. What is less known is the impact of these resource limitations on the nutritional quality of the resulting seed. In particular, while we know that the recycling of nutrients between plant pools is essential for producing seeds with high concentrations of proteins, lipids, and starch (Bennett et al., 2011), less known is the physiological and structural adaptations to the combined effects of abiotic stress such as drought and low P on the overall nutritional quality of seed. The soluble components of leaf tissue represent the transportable pool of resources that turns over rapidly alongside less-labile components of storage and structural material, and these pools, therefore, hold potential for indicating short-term resource limitations.

To investigate the effects of both water and P deficit in source and sink tissues of common bean, a replicated study was conducted to investigate the genotypic differences in bred genotypes of common bean to an individual and combined water deficit and P deficiency on seed yield and nutritional quality. Specifically, we tested the following hypotheses: (1) variation in the magnitude of yield decrease will be observed among genotypes and treatment combinations, (2) concomitant changes in the content and concentration of nutrients (mineral nutrients, amino acids) will be observed on a background of reductions in yield due to treatment effects and (3) qualitative and quantitative changes in nutrient content and concentration in seeds will be reflected in foliar nutrient content.

MATERIALS AND METHODS

Experimental Design

Plants were established as a randomized complete block design with four treatments (main plots) and 12 genotypes as subplots with three replications. Treatments consisted of (i) low phosphorus + well-watered (LPWW), (ii) low phosphorus + water deficit (LPWD), (iii) high phosphorus + well-watered (HPWW), and (iv) high phosphorus + water deficit (HPWD). The low P treatment corresponded to a soil P application rate of 10 kg/ha, high P corresponded to a soil P application rate of 40 kg/ha (Polania et al., 2009) well-watered denotes plants watered to 80% of field capacity, and water deficit denotes suspension of irrigation at 15 days after sowing (DAS).

Plant Material

A total of 12 bush bean genotypes belonging to the Middle American gene pool previously bred by the Centro Internacional de Agricultura Tropical (CIAT) and its partners were selected for their inclusion in the trial based on their tolerance to abiotic stress or commercial availability (see **Table 1**). Genotypes selected included superior drought and low P tolerant (NCB226, SEF60, SEN56, BFS35, BFS81), drought tolerant (SEF71, RCB593), low P tolerant (SEF73, Carioca, SXB412), and commercial checks (DOR390, Tio Canela) as previously described in Beebe et al. (2008) and Polania et al. (2016c). Pedigrees of drought, and drought- and low P–adapted lines used in this study included SEA 15, which is a progeny of SEA 5, which in turn was derived from an interracial cross including races Durango and Mesoamerica of the Middle American gene pool (Terán and Singh, 2002). SEA 15 also has Apetito (G 1759), a race Durango landrace, in its pedigree (Beebe et al., 2008). Seed color and growth habits of the genotypes varied (see **Table 1**). Growth habits used in this study, as described by Singh (1981) were previously classified as, 2A: an indeterminate growth habit lacking climbing ability, 2B: an indeterminate growth habit possessing some climbing ability, and 3B: an indeterminate growth habit with long main stem guide possessing moderate climbing ability. Some of the genotypes had been characterized previously under drought conditions (see Polania et al., 2016a,c).

Experimental Site and Lysimeter Conditions

Plants were grown in soil lysimeters under a movable rainout shelter from September to December 2015, at the experimental station of the International Center for Tropical Agriculture (CIAT) in Palmira, Colombia, located at 3° 29″ N latitude, 76° 21″ W longitude, and an altitude of 965 m. Lysimeters were

TABLE 1 | Description of 12 common bean genotypes included in the lysimeter trial at CIAT grown under individual and combined treatments of low phosphorus and water deficit conditions.

Objective	Genotype	Seed color	Growth habit
Superior drought and low P tolerant	NCB226	Black	2B
	SEF60	Red	2A
	SEN56	Black	2A
	BFS35	Red	2A
	BFS81	Red	2B
Drought tolerant	SEF71	Red	2A
	RCB593	Red	2B
Low P tolerant	SEF73	Red	2B
	Carioca	Cream striped	3B
	SXB412	Cream	2B
Commercial checks	DOR390	Black	2B
	Tio Canela	Red	2A

Growth habit classification is explained in the text.

constructed from plastic transparent cylinders inserted into PVC pipes with a diameter of 20 cm and height of 120 cm (Polania et al., 2017). Lysimeters were filled with 60 kg of dry soil with the top 10 cm of the lysimeter left empty. The soil was collected from the CIAT field station in Darien, Colombia, located at 3° 53″ N latitude, 76° 31″ W longitude, and an altitude of 1,460 m. The bulk density of the soil used in the lysimeters was adjusted to 1.1 g/cm^3 to facilitate root growth and drainage. The soil is described as an Andisol and is characterized as deficient in available P (Beebe et al., 2008). Available P content was determined (Bray-II) to be 5.7 mg/kg. Available P content was determined (Bray-II) to be 7.1 mg/kg in the low P treatment and 12.5 mg/kg in the high P treatment. For the duration of the trial, mean maximum and minimum air temperatures were 31.1°C and 20.1°C, respectively, with an average relative humidity (RH) of 58.4%. The initial soil moisture content of all four treatments was at 80% field capacity. Plants in the well-watered treatment were maintained at 80% field capacity by weighing the lysimeter two times each week, irrigating the top of the lysimeter and registering soil water content. Plants in the water deficit treatment were weighed two times each week to calculate the loss of soil moisture content. At harvest, the soil moisture content in the terminal drought treatments was determined to be 39% field capacity for high P (HPWD), and 42% of field capacity for low P (LPWD). The trial was managed with weeding and spraying of insecticides and fungicides as required.

Leaf Gas Exchange Measurements

Gas exchange for each plant was measured using an LI-COR 6400 XT infra-red gas analyzer (LI-COR, Lincoln, NE, United States) for leaf-level photosynthesis, stomatal conductance, and sub-stomatal CO_2 concentration at two growth stages, flowering (DAS 32–37) and mid-pod fill (DAS 44–47). Instantaneous leaf gas exchange measurements were made between the hours of 10:00 am and 2:00 pm. A fully expanded, non-shaded leaf was chosen for measurement at each time point. For spot measurements, conditions in the chamber were set to photosynthetically active radiation (PAR) of 1,200 μmol/m^2/s, temperature maintained at 27°C and RH values were kept within the range of 65–75%. The CO_2 mole fraction of reference air in the measuring chamber was set at 400 μmol/mol for ambient measurements and 2,000 μmol/mol for maximum rate of photosynthesis (A_{max}) measurements. Spot measurements were completed within 5 days for each growth stage, with genotypes measured in order from earliest to latest flowering.

Relationships between net photosynthesis and sub-stomatal CO_2 concentration (A/c_i curves) were completed using the automated program on the LI-COR 6400 XT. Parameters were set such that c_a be stepped down from 400 to 50 μmol/mol before moving back up from 400 to 2,000 μmol/mol. Conditions in the chamber were set to a PAR of 1,200 μmol/m^2/s, leaf temperature was maintained at 27°C and RH values were kept within the range of 65–75%. A/c_i curves were carried out on three genotypes; DOR390, NCB226, and SEF60 under low and high P treatments for well-watered plants only (i.e., no combined water deficit treatment). A/c_i curves were taken at the same growth stages as mentioned earlier, flowering and mid-pod fill. Parameters that can limit maximum rates of photosynthesis; maximum rates of carboxylation (V_{cmax}), electron transport (J), and triose phosphate utilization (TPU) were derived from version 2.0 of the A/c_i curve fitting calculator described by Sharkey (2015).

Tissue Collection

Following gas exchange measurements, one fully expanded leaf not measured in the LI-COR6400 chamber was collected from each plant using a razor blade. The leaf sample was microwaved for 10 s using a conventional 900 W microwave oven to prevent metabolic activity according to the method outlined by Popp et al. (1996) and placed in an oven to dry at 65°C. At harvest, leaf area (LI-COR model LI-3000) and shoot biomass distribution (leaves, dead leaves, stems, pod walls, seed) were recorded. Shoot dry weight was determined after the shoot samples were dried in an oven at 65°C for 48 h. Harvest index and pod harvest index were determined by Equations 1, 2, respectively.

$$Harvest\ index = \frac{seed\ biomass\ dry\ weight\ at\ harvest}{aboveground\ plant\ biomass\ dry\ weight\ at\ harvest} \times 100 \quad (1)$$

$$Pod\ harvest\ index = \frac{seed\ biomass\ dry\ weight\ at\ harvest}{pod\ biomass\ dry\ weight\ at\ harvest} \times 100 \quad (2)$$

Extractions of Leaf and Seed Material

Samples of leaves and seeds were oven-dried at 65°C and ground using an oscillating matrix mill. Approximately 40 mg of ground leaf/seed sample was then weighed into a 2-mL microtube and extracted in a hot water mix according to the protocol outlined by Merchant et al. (2006). An additional 20 mg of ground seed material was placed with 1 ml of 6 M hydrochloric acid in a vacuum hydrolysis tube and digested for 24 h at 110°C. Extracts of the hot water extraction and digestion process were stored frozen at −80°C awaiting further analysis as described in the followings sections.

Analysis of Carbon and Oxygen Isotope Abundance

Determination of carbon and oxygen isotope abundance on ground bulk samples of leaves and seed was completed using a Delta V Advantage isotope ratio mass spectrometer (IRMS) (Thermo Electron) with a Conflo IV interface (ThermoFisher Scientific, Bremen, Germany).

Analysis of Plant Material for Amino Acids and Mineral Nutrients

Determination of soluble amino acids and total amino acids in extracted and digested samples, respectively, was completed using high-performance liquid chromatography (HPLC) coupled to a quadrupole time-of-flight mass spectrometer. HPLC separation was completed on an Agilent 1290 Infinity system (Agilent, Walbronn, Germany) using a Zorbax StableBond SB-CB18 column (150 mm × 2.1 mm, 3.5 μm, Agilent) including degasser, binary pump, temperature-controlled autosampler

(maintained at 4°C), and column compartment (maintained at 30°C). The mobile phase was composed of water containing 0.1% formic acid (solution A) and methanol containing 0.1% formic acid (solution B). The flow rate was 0.3 mL/min with a gradient elution of 0–100% solution B, over 23 min for positive mode, respectively. Amino acids were detected by a quadrupole-time-of-flight mass spectrometer (Agilent 6520 accurate-mass) with a dual electrospray ionization (ESI) source. The mass spectrometer was operated with a full scan in positive FT mode for amino acid analysis. ESI capillary voltage was set at 4,000 V (+) ion mode and 3,500 V (−) ion mode and fragmentor at 135 V. The liquid nebulizer was set to 30 psig and the N drying gas was set to a flow rate of 10 L/min. The drying gas temperature was maintained at 300°C. Internal reference ions were used to continuously maintain mass accuracy. Molecular ions [(M+H)+ for amino acids] were extracted from the full scan chromatograms and peak areas integrated using Agilent MassHunter Workstation software (Agilent Technologies, Santa Clara, CA, United States).

Determination of soluble and total mineral nutrients in the extracted and digested samples, respectively, was completed using an inductively coupled plasma optical emission spectrometer (Varian Vista, Agilent Technologies, Santa Clara, CA, United States). Samples were prepared with a dilution of 400 μl of supernatant in 10 ml of ultra-pure Milli-Q water. Any results lower than the detection limit of the instrument were adjusted to zero.

Statistical Analysis

Analysis was conducted in R software (R Core Team, 2021). Analysis of variance was performed wherein models were fitted using the lm function, where genotypes and treatments were treated as fixed effects. Tukey's honest significant difference (HSD) test was used for multiple comparisons of means. Correlations and principal component analysis (PCA) were conducted to determine the relationships of yield and photosynthetic parameters. All figures were generated and visualized in R using ggplot2 and complimentary grammar of graphics packages.

RESULTS

Water Deficit and Low P Availability Significantly Decrease Yield

About 97% of the cumulative variation among the different genotypes was explained by PC1 and PC2, wherein overall response to water deficit and low P availability significantly decreased yield (**Figure 1**). Overall, aboveground biomass, harvest index, pod number, seed number, and seed weight were strongly correlated with one another (r = 0.992–0.837; P = 4.44e-16–0.000), whereas pod harvest index relatively differed from the other parameters (r = 0.760–0.695; P = 4.44e-08–3.88e-10 (**Figure 1A** and **Supplementary Table 1**). Under well-watered (WW) conditions, P availability had a significant effect on all traits (aboveground biomass, P = 0.00; pod number, P = 0.00; seed number, P = 0.00; seed weight, P = 0.00) except harvest index (P = 0.23) and pod harvest index (P = 0.99) (**Figures 1B–G**). In contrast, under water deficit (WD) conditions, only pod harvest index (P = 0.001) was significantly affected by P availability (**Figure 1D**). All treatment comparisons between different yield parameters are summarized in **Supplementary Table 2**. Among yield parameters mentioned above, aboveground biomass and seed weight were strongly correlated with each other (r = 0.992; P = 0.000) (**Figure 1**). Increased seed weight was linearly proportional to high aboveground biomass, which was more prominent under well-watered (WW) conditions (HPWW/LPWW, R^2 = 0.77–0.85) (**Figure 2**). In general, superior genotypes bred for resilience to abiotic stress did perform well under all treatments compared to the commercial checks DOR390 and Tio Canela (**Figures 1A**, **2**).

Photosynthesis and Stomatal Conductance Are Primarily Impacted by Water Deficit

Leaf-level gas exchange at both ambient [CO_2] (data not shown) and maximum [CO_2] shows a significant reduction in leaf-level photosynthesis in response to WD compared to plants grown under WW conditions (**Figure 3**). Corresponding reductions in stomatal conductance (g_s) were observed with less than half the stomatal conductance for the WD treatment compared to the WW treatment (**Figure 3**). P supply impacted ambient rates of stomatal conductance under the WW treatment only (P = 0.06*)* (**Figure 3**). Ambient and maximum photosynthesis and stomatal conductance declined over development with statistically significant differences found between flowering compared to mid-pod fill (**Supplementary Table 4**). Ambient and maximum photosynthesis and stomatal conductance varied with genotype (**Figure 3**) with bred genotypes having greater instantaneous water use efficiency (higher photosynthetic capacity and lower stomatal conductance) under treatment conditions (**Figure 3**). Significant differences were also observed for maximum rates of carboxylation (V_{cmax}), electron transport (J), and triose phosphate utilization (TPU) for genotypes SEF60 and NCB226 compared to the commercial check DOR390 ($P < 0.05$), suggesting an increased capacity of superior-bred genotypes to photoassimilate under low P supply than the commercially available check (see **Supplementary Figure 1**).

Patterns in Isotope Abundance ($\delta^{13}C$ and $\delta^{18}O$) Indicate the Severity of Water Deficit

Carbon isotope abundance ($\delta^{13}C$) varied with treatment in the leaf tissue and seed with significant differences between WW and WD treatments and development stage (**Figure 4** and **Supplementary Table 5**). Within the development stage, $\delta^{13}C$ in leaf tissue significantly differs only in WW treatment (**Figure 4A**: HPWW–LPWW at flowering stage, P = 0.088; **Figure 4B**: HPWW–LPWW at mid pod fill, P = 0.013); while at harvest, $\delta^{13}C$ in seed was substantially different under both WW (HPWW—LPWW, P = 0.000) and WD (HPWD—LPWD, P = 0.000) and becoming more depleted with WW conditions compared to WD (**Figure 4C**). Overall, there was significantly

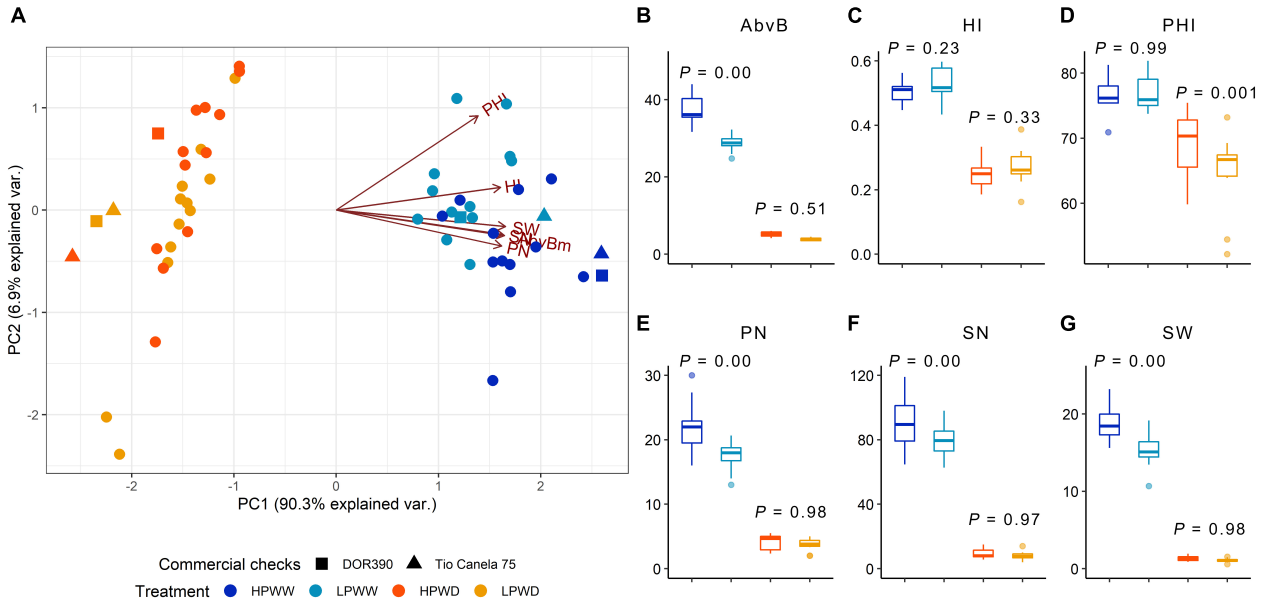


FIGURE 1 | (A) Principal component analysis of yield parameters; pod harvest index (PHI), harvest index (HI), seed weight (SW, g), seed number (SN), aboveground biomass (AbvB, g), and pod number (PN) for 12 common bean genotypes under treatments of high phosphorus well-watered (HPWW; blue), low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange). Commercial checks are illustrated by square and triangle shapes. **(B)** Aboveground biomass, **(C)** harvest index (HI), **(D)** pod harvest index (PHI), **(E)** pod number (PN), **(F)** seed number (SN), **(G)** seed weight (SW, g) for 12 common bean genotypes under treatments of high phosphorus well-watered (HPWW; blue), low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange).

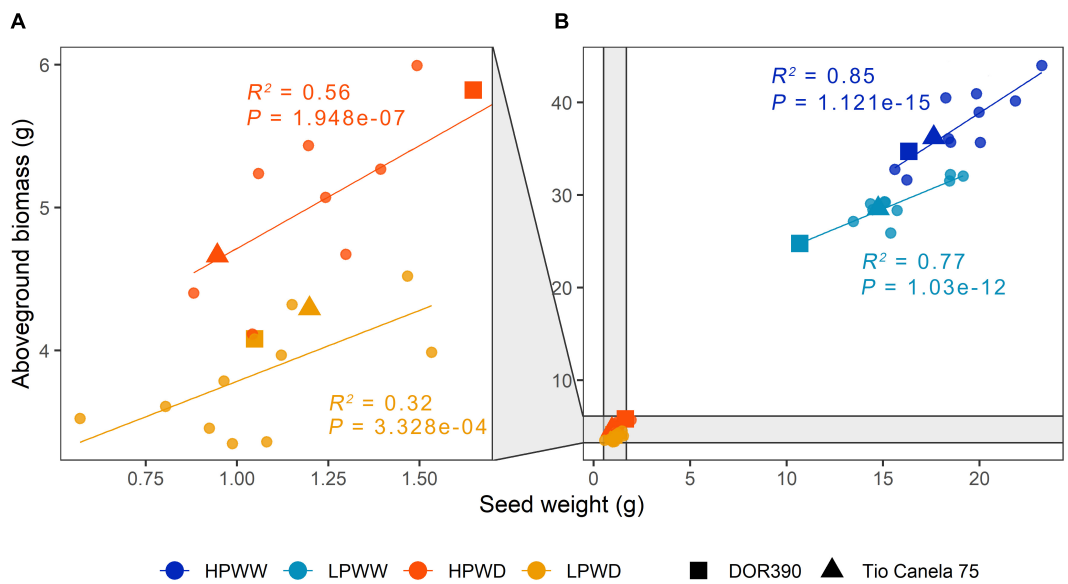


FIGURE 2 | Seed weight (g) and aboveground biomass (g) for 12 common bean lines under different treatments. **(A)** High phosphorus water deficit (HPWD; orange) and low phosphorus water deficit (LPWD; light orange). **(B)** High phosphorus well-watered (HPWW; blue) and low phosphorus well-watered (LPWW; light blue). Standard errors have been removed for clarity (n = 3). Linear regression to calculate adjusted R^2 and P values grouped according to treatment. Commercial checks are illustrated by square and triangle shapes.

(P = 0.000) lower $\delta^{13}C$ in leaf tissue at flowering (−28.65) and mid pod fill (−28.09) compared to seed at harvest (−25.35) (**Figure 4** and **Supplementary Table 5**). No variation between breeding lines and commercial checks was detected for $\delta^{13}C$ from the leaf or seed tissue (**Supplementary Table 6**). For the measured WW plants, the relationship between carbon

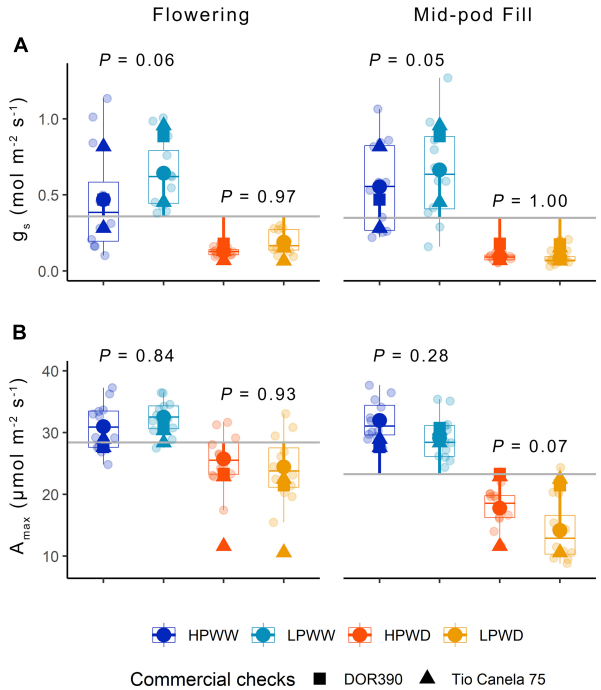


FIGURE 3 | (A) Stomatal conductance g_s (mol/m^2/s) and **(B)** A_{max} (µmol/m^2/s) for 12 common bean genotypes under treatments of low phosphorus well-watered (LPWW), high phosphorus well-watered (HPWW), low phosphorus water deficit (LPWD). and high phosphorus water deficit (HPWD) at flowering and mid-pod fill stages. *Gray horizontal lines* represent overall means across the different treatments; *segment lines* represent the mean difference of individual treatment relative to the overall means. *P* values are according to treatment means comparison calculated using HSD test at $\alpha = 0.05$. Commercial checks are illustrated by square and triangle shapes. Multiple comparisons between treatments is provided in **Supplementary Table 4**.

and oxygen isotope abundance in the leaf tissue of select genotypes SEF60, NCB226, and DOR390 was weak (HPWW $R^2 = -0.023$, $P = 0.445$; LPWW $R^2 = 0.196$, $P = 0.037$) indicating the mild effect of LP treatment on the genotypes measured (**Figure 4B**). Albeit weak, a negative relationship between $\delta^{13}C$ and $\delta^{18}O$ indicates that the dominant influence on substomatal carbon concentrations is that of limitations to maximum net photosynthetic rate (A_{max}), that is, a biochemical limitation to carboxylation imparted by low P availability rather than driven by changes in stomatal conductance.

Impact of Low P and Water Deficit on Mineral Nutrients and Amino Acids in Leaf Tissue

Within the soluble leaf extract, greater concentrations of potassium, calcium, and magnesium were present in comparison to other nutrients particularly iron and zinc (**Figure 5**). Variation in the concentration of minerals as a response to different treatments within the development stage was most evident at mid-pod fill. For example, there was significant treatment response on the concentration of phosphorus (HPWW—LPWD, $P = 0.000$) at flowering stage and calcium (HPWD—LPWW, $P = 0.000$), nitrogen (HPWW—LPWD, $P = 0.000$), and sulfur (HPWW—LPWD, $P = 0.000$) at mid-pod fill, among others (**Figure 5A**). In general, trends varied depending on the nutrient and concentration of minerals at the developmental stage (**Figures 5A,B**). The concentration of nitrogen (N) significantly declined from flowering to mid-pod fill stage for WW treatments regardless of level of P treatment (HPWW $P = 0.000$; LPWW $P = 0.000$). However, under WD treatment, there was no significant change from flowering and mid-pod fill at low P (LPWD, $P = 0.935$) but N slightly declined at high P treatment (HPWD, $P = 0.000$) (**Figures 5A,B**). The concentration of zinc significantly increased only at WD treatments at both P levels (HPWD and LPWD, $P = 0.000$) (**Figures 5A,B**). Strong differences between bred genotypes and commercial checks were only detected for zinc and iron in the soluble leaf tissue (**Supplementary Table 8**) while treatment responses do not seem to vary based on bred genotypes' tolerance to drought or low P stress (see **Table 1**).

In the soluble leaf tissue, significantly high concentrations of glutamic acid (Glu) were detected compared to all other amino acids ($P = 0.000$) at varying treatments and development stages. In addition, there was no significant change of Glu concentrations at both WW and WD and P concentrations (HPWW—LPWD, $P = 0.177$). However, Glu concentration significantly declined at mid-pod fill ($P = 0.000$) (**Figure 6**). There was a slight but not statistically significant increase in isoleucine, leucine, phenylalanine, threonine, tryptophan, and tyrosine in response to WD treatment (**Figure 6**). In general, concentrations of amino acids declined between flowering and mid-pod fill ($P = 0.000$) (**Figure 6**). Overall, there were no significant differences between bred genotypes and commercial checks across treatments and development stages, except for differences observed between low phosphorus and drought-tolerant–bred genotypes for Glu concentrations (Carioca—RCB593, $P = 0.052$).

Impact of Low P and Water Deficit on Mineral Nutrients and Amino Acids in the Seed

Nitrogen and potassium were present in higher concentrations in the digested total seed in comparison to other nutrients, particularly iron and zinc (**Figure 7A**). Significant treatment differences ($P = 0.000$), largely between WW and WD were observed for nitrogen, potassium, phosphorus, sulfur, and calcium (**Figure 7B**). Overall, there were no differences in the concentration of minerals between bred genotypes and commercial checks, except for nitrogen (**Supplementary Table 9**).

Almost all amino acids were detected in the seed tissue with the absence of only glutamic acid and tryptophan (**Figure 8A**). Statistically significant treatment differences were detected for the majority of amino acids as a consequence of the WD treatment under low P levels (HPWW—LPWD, $P = 0.000$) (**Figure 8B**). A comparison of genotype differences observed for any of the amino acids detected in the seed tissue is provided in **Supplementary Table 9**.

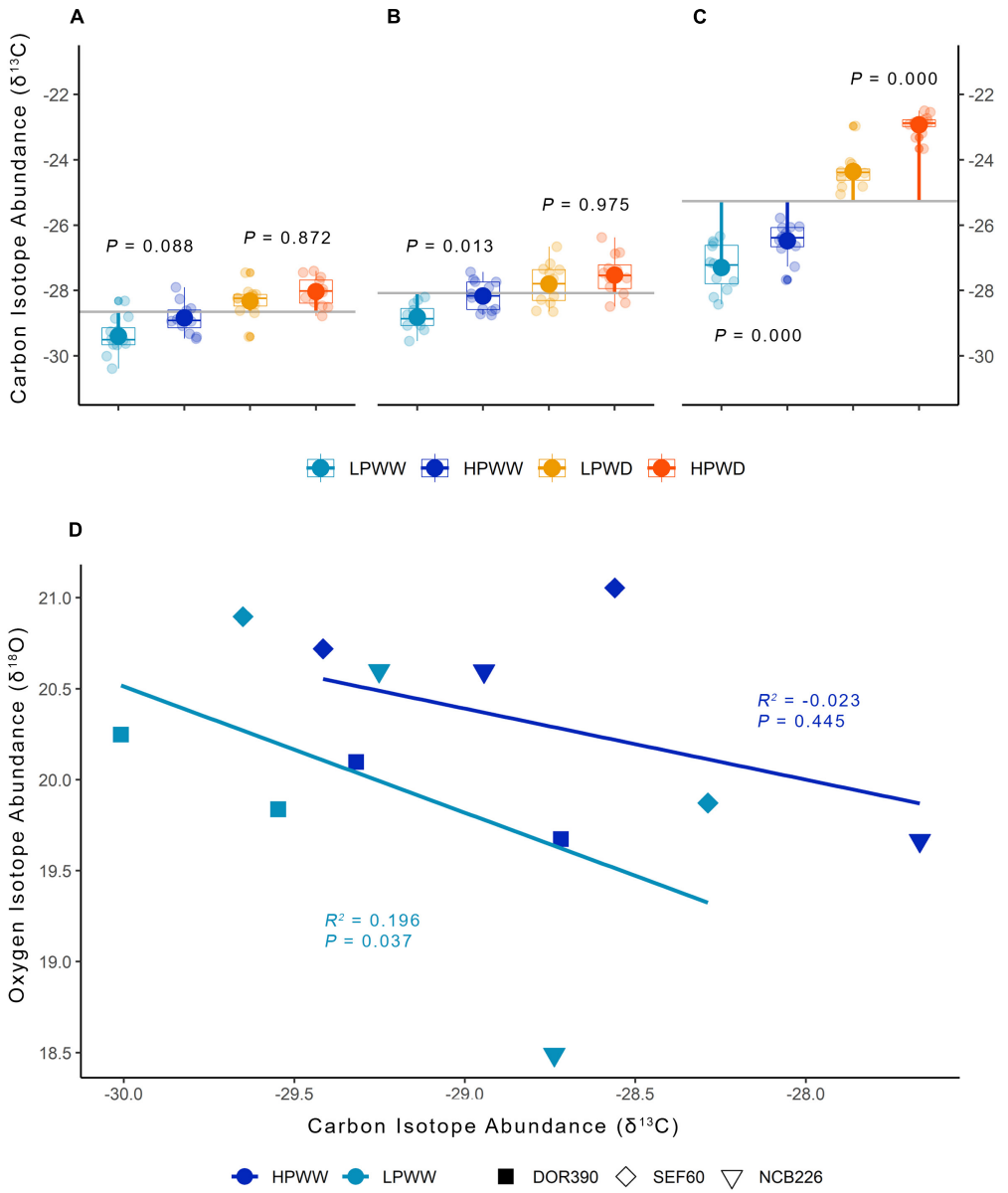


FIGURE 4 | Carbon isotope abundance ($\delta^{13}C$) for bulk leaf tissue at flowering **(A)**, mid-pod fill **(B)**, and seed at harvest **(C)** for 12 common bean genotypes under treatments of low phosphorus well-watered (light blue), high phosphorus well-watered (blue), low phosphorus water deficit (yellow-orange), and high phosphorus water deficit (orange). The *gray line* represents the overall means across treatments at different development stages while *segment lines* represent the means per treatment showing the deviations from the overall means within the development stage. *P* values are according to treatment mean comparisons within the development stage calculated using the HSD test at $\alpha = 0.05$. Multiple comparisons between treatments are provided in **Supplementary Table 5**. **(D)** Carbon isotope abundance ($\delta^{13}C$) and oxygen isotope abundance ($\delta^{18}O$) for bulk leaf samples under treatments of low phosphorus well-watered (light blue) and high phosphorus well-watered (blue) for selected lines DOR390 (*filled square*), SEF60 (*diamond*), and NCB226 (*inverted triangle*) sampled at flowering and mid-pod fill stage; mean values adjusted for treatment and development stage. Linear regression to calculate adjusted R^2 and *P* values grouped according to treatment.

DISCUSSION

Seed Yield and Aboveground Biomass Was Reduced Due to Individual and Combined Impacts of Water Deficit and Low P

Water stress and P deficiency, both individually and combined, significantly decreased seed weight and aboveground biomass by ∼80% (**Figures 1**, **2**). This is unsurprising, given the substantial evidence for both abiotic stress's influence on yield and its closely associated traits such as harvest index (Beebe et al., 2008; Smith et al., 2019). Yield loss was likely driven by reductions in aboveground biomass (i.e., less leaf area equals less photoassimilate), which in WD treatments led to a ∼75% reduction in the number of pods and seeds produced resulting in reduced yield (seed weight) (**Figure 1**). Leaf expansion is known to be highly sensitive to both water

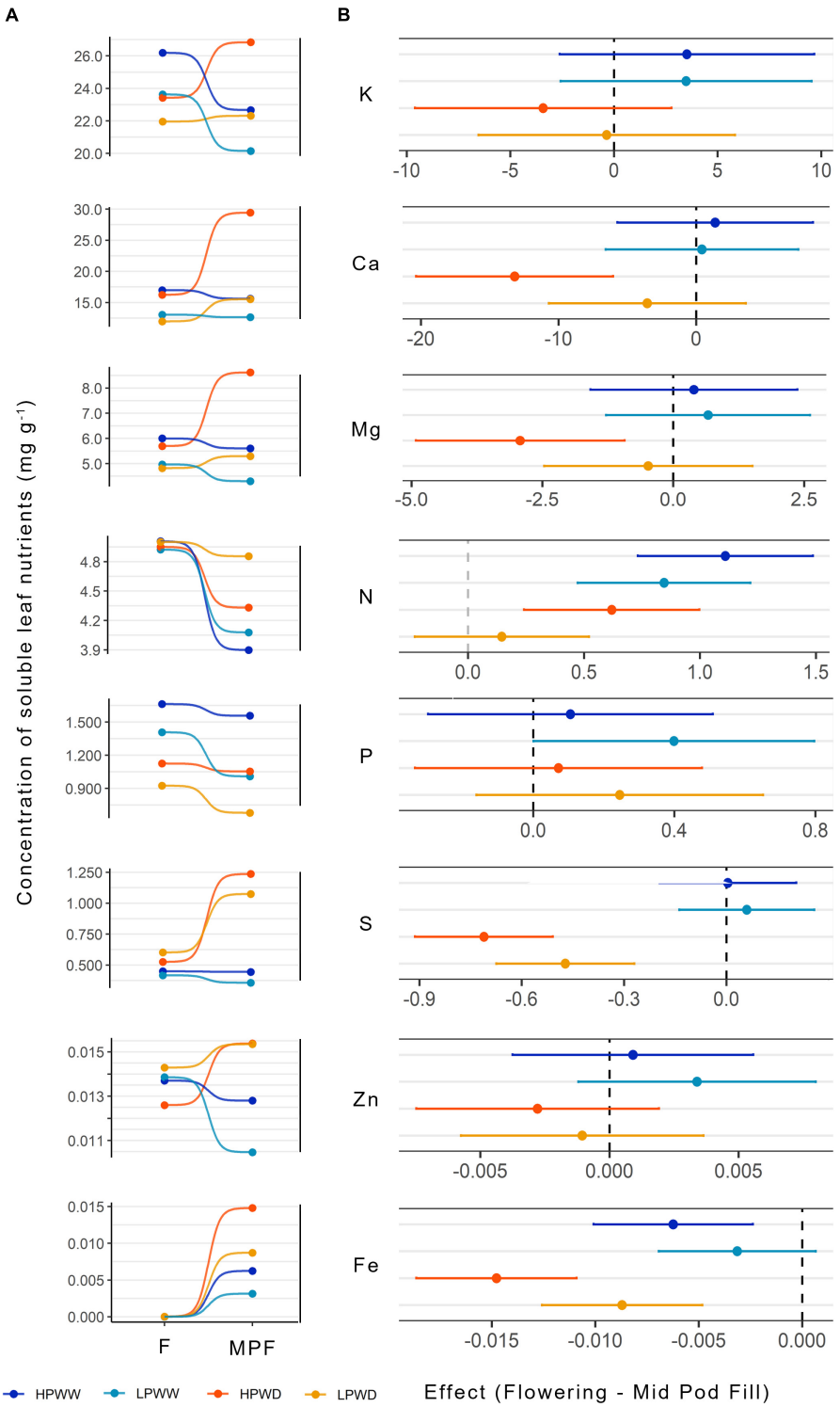


FIGURE 5 | (A) Concentration (mg/g) of soluble leaf nutrients for 12 common bean lines under treatments of high phosphorus well-watered (HPWW; blue) low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange) at flowering (F) and mid-pod fill (MPF) stage. **(B)** Effect of development stage (flowering—mid-pod fill) on the concentration of nutrients at different treatments. *Dots* represent mean value with upper and lower confidence level; *x-axis* shows the mean difference in the concentration of nutrients at four treatments; *vertical dash lines* illustrate 0 effect, where values close to 0 are not significantly different according to the HSD test at $\alpha = 0.05$. For visualization purposes, only four treatments are shown; the complete list of multiple treatments and stage comparisons is provided in **Supplementary Table 7**.

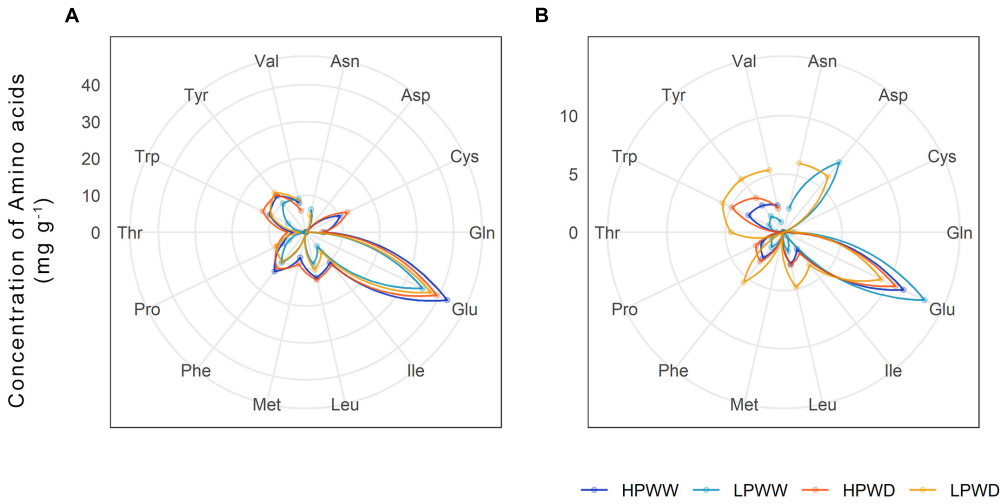


FIGURE 6 | The concentration of amino acids (mg/g) found in common bean soluble leaf fraction subjected to high phosphorus well-watered (HPWW; blue), low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange) at flowering **(A)** and mid-pod fill **(B)**. Data has been transformed for visualization purposes. Only Glu had a significant concentration level ($P = 0.000$) compared to other amino acids at all treatments at flowering and mid-pod fill.

deficit and P deficiency (Rychter and Rao, 2005). There was a significant, but not dramatic, reduction in remobilization, as indicated by lower values of pod harvest index, in WD plants compared to those in the WW treatment (**Figure 1**). Within the combined LPWD treatment, the pod harvest index was significantly lower in comparison to the combined HPWD treatment (**Figure 1**). PCA also showed the differing response of pod harvest index compared to other yield parameters suggesting that the trait is influenced by other factors beyond those measured (**Figure 1A**). This response also suggests that water and P availability play an important role in the movement of photoassimilate from the pod wall into the developing seed. Remobilization of photosynthate to seed (indicated here by measures of pod harvest index) is known to play a significant role in improved adaptation to drought and low P in common bean (Beebe et al., 2008, 2009; Assefa et al., 2013; Rao et al., 2013). However, the mechanisms that determine the movement of photoassimilate from the pod wall and into the seed are not well understood and further research is required to characterize the underlying causes of poor remobilization often described as, "lazy pod syndrome" (Beebe, 2012). More broadly, variation in seed weight and aboveground biomass associated with genotype was observed among the 12 common bean genotypes investigated (**Figure 2**). The variation in harvest index (**Figure 1C**) between the genotypes bred for adaptation to drought and low P availability and the commercial checks highlights the positive impact of breeding activities on yield. Nevertheless, further investigation is warranted to determine the mechanistic relationships of abiotic stress factors such as water deficit and P deficiency on the assimilation and transport of resources throughout yield development and the impacts this combined stress has on yield.

Water and P Deficiencies Influence Carbon Assimilation and Water Use at the Leaf and Plant Scale

Water and P deficiency significantly decreased photosynthesis and stomatal conductance over plant development (**Figure 3**). In this study, for plants grown under LPWD, aboveground biomass was significantly lower (**Figure 2**) which subsequently reduced whole-plant water use, therefore, decoupling plant growth from the treatment. For example, this may have allowed for the maintenance of photosynthesis at similar rates per unit leaf area as plants grown under the HPWD treatment which had a slightly higher aboveground biomass and hence would have required greater amounts of water at the whole plant scale (**Figures 2**, **3**). This response indicates the importance of P supply on whole-plant function *via* the regulation of source–sink dynamics (Rychter and Rao, 2005; Smith et al., 2018b).

Genotypic variation in photosynthesis and stomatal conductance was detected for all genotypes measured (**Figure 3**) and has also been observed in previous studies using common bean (Wentworth et al., 2006; Polania et al., 2016c; Rao et al., 2017). For the three genotypes, NCB226, SEF60, and DOR390, selected for A/c_i curves under the LPWW and HPWW treatments, significant treatment and genotypic differences in the biochemistry that underlies photosynthesis were only detected in the flowering stage (see **Supplementary Figure 1**). We expected greater treatment differences in biochemistry as a consequence of P deficiency due to the important role P plays in the Calvin Cycle reactions (Raines, 2003). This response highlights that the LP treatment only caused mild deficiency (also indicated by the weak negative correlations between $\delta^{13}C$ and $\delta^{18}O$, **Figure 4D**) while the WD increased in severity as development continued. Understanding how photosynthesis, along with the underlying

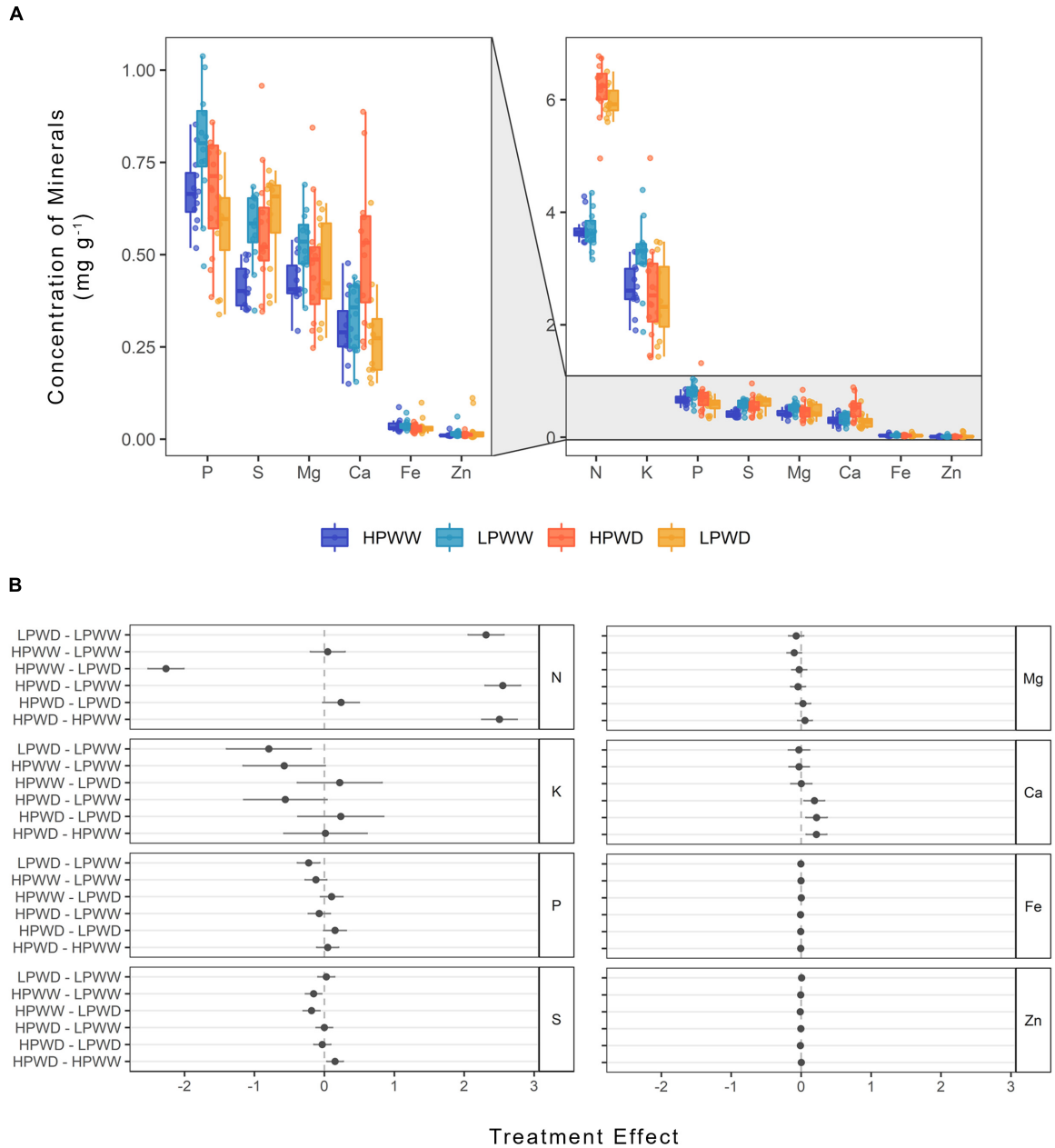


FIGURE 7 | (A) Concentration (mg/g) of nutrients found in the common bean seed subjected to high phosphorus well-watered (HPWW; blue), low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange) at harvest. Zoom in panel to the left shows nutrients that had concentrations less than 1 mg/g. **(B)** Comparison of treatment effect combination on the concentration of nutrients. *Dots* represent mean value with upper and lower confidence level; *x-axis* shows the mean difference in the concentration of nutrients between the treatments; *vertical dash lines* illustrate 0 effect, where values close to 0 are not significantly different according to the HSD test at $\alpha = 0.05$.

biochemistry, responds to multiple stresses and fluctuating environmental conditions over time is critical (see, e.g., Zhu et al., 2010).

Isotope Abundance Indicates Drought Stress in Leaf and Seed

Carbon isotope fractionation in the leaf tissue significantly decreased by approximately 2 parts per million in response to the WD treatment at all development stages (**Figure 4A**) while in the seed tissue collected at harvest, significant differences were observed between all treatments with decreases due to HP and WD treatments (individually and combined) (**Figure 4A**). These responses are similar to those detected in previous studies (Badeck et al., 2005; Smith et al., 2016, 2018a, 2019). The difference in the responses between the leaf and seed likely occurred due to heterotrophic fractionation, which may occur in part as a consequence of changing demand for metabolites in sink

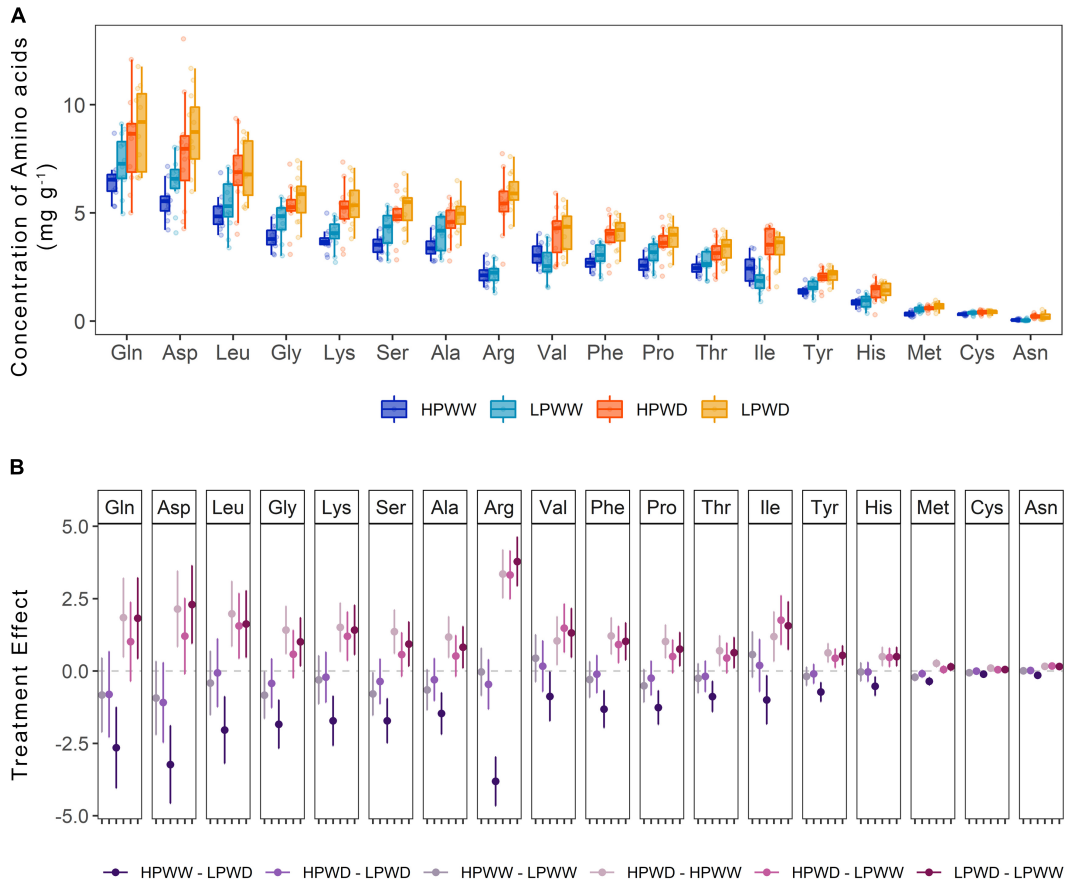


FIGURE 8 | (A) Concentration of amino acids (mg/g) found in the common bean seed tissue subjected to high phosphorus well-watered (HPWW, blue), low phosphorus well-watered (LPWW; light blue), high phosphorus water deficit (HPWD; orange), and low phosphorus water deficit (LPWD; yellow-orange) at harvest. **(B)** Comparison of treatment effect combination on the concentration of amino acids. *Dots* represent mean value with upper and lower confidence level; *y-axis* shows the mean difference in the concentration of amino acids between the treatments; *horizontal dash lines* illustrate 0 effect, where values close to 0 are not significantly different according to the HSD test at $\alpha = 0.05$.

tissues altering the carbon isotope abundance of the remaining soluble pool (see Cernusak et al., 2009; Smith et al., 2018a). This can be magnified by changes in environmental conditions (Smith et al., 2016) and in this case, resulted in the carbon isotope abundance of the seed harvested from the LP treatment significantly differing from other treatments, while in the leaf tissue the P treatment had no significant impact.

While carbon isotope abundance indicates drought stress, it does not necessarily equate directly to measures of water deficit or water use efficiency (see Seibt et al., 2008). Nevertheless, carbon isotope abundance has been proposed and utilized as a screening method to improve water use efficiency in breeding C3 crops, including that of common bean (Polania et al., 2016c). Using carbon isotope abundance from seed, common bean genotypes have previously been classified as "water spenders" and "water savers" (Polania et al., 2016c), however, in this case, no genotypic differences were detected in either leaf or seed tissue under any treatment conditions. Relationships between $\delta^{13}C$ and growth are yet to be fully described. Consideration must be made of the source of carbon for analysis, and the likelihood of the isotope abundance contained within it to reflect water use efficiency across realistic time and spatial scales. The present study did not uncover a strong relationship between $\delta^{13}C$ and $\delta^{18}O$ as expected (see Scheidegger et al., 2000; **Figure 4B**) perhaps due—at least in part—to the relatively isohydric response of common bean to such a severe water deficit. During periods of severe water deficit, changes in isotope signals are reduced because of very low gas exchange and little net assimilation of carbon. The present study highlights that the application and interpretation of carbon isotope abundance has limits imposed by plant responses at the boundaries of physiological function.

Leaf Mineral Nutrient and Amino Acid Concentrations Decline in Response to Water and P Deficiency

Within the soluble leaf tissue, concentrations of mineral nutrients and amino acids declined in response to both the individual and combined effects of WD and LP (**Figures 5**, **6**). The majority of the mineral nutrients detected within the soluble leaf

fraction at flowering and mid-pod fill decreased in response to reductions in P availability (**Figures 5A,B**). Significant variation between genotypes was detected for zinc and iron, which according to the statistical model significantly interacted with genotype and treatment (**Supplementary Table 7**). This may have been a consequence of mobilization of iron, as this has not influenced the concentration of iron present in the seed, with similar concentrations of iron found under all treatments and no significant differences between genotypes (**Figure 7B**). This finding contrasts with that of Petry et al. (2015) who reported that location and management can alter iron and zinc seed concentration; however, these findings are similar to those reported by Bulyaba et al. (2020) who found that iron and zinc concentration in the seed did not significantly vary between locations with varying soil properties or varieties. While it is known that leaves serve as a substantial source of mineral nutrients for developing seeds (Garcia and Grusak, 2015), a few studies outside the initial work of Hocking and Pate (1977) have researched concurrent changes in leaf and seed mineral content (Garcia and Grusak, 2015). Seed nutrient quality is likely regulated in part by source–sink dynamics and as such the manipulation of transport processes between the leaf tissue toward the developing reproductive tissue is a possible strategy to increase nutrient allocation within the seed (Bennett et al., 2011; Pottier et al., 2014; Garcia and Grusak, 2015; Tan et al., 2017; Smith et al., 2018b).

The concentration of most amino acids detected significantly differed with WD (**Figure 6**). Variation between genotypes was detected for amino acids at the leaf level (data not shown). In particular, the proportion of amino acids detected showed that in some genotypes, very few amino acids (aspartic acid and glutamic acid) were present in the soluble leaf fraction. The examination of the soluble leaf fraction has previously been used to infer stress responses of plants (see, e.g., Hare et al., 1998; Merchant et al., 2010; Dumschott et al., 2017). Given the variation in nutritional quality (mineral nutrients and amino acids) between each of the genotypes, further investigation may allow for the targeted improvement of the genotypes by promoting metabolites that confer resilience. While this type of monitoring could be used to reduce some of the genotypic variations and increase yield, no genotypic differences were detected for nutritional quality in the seed (**Figures 7**, **8** and **Supplementary Table 9**), and as such amelioration of reductions to nutritional quality in the seed under varying abiotic stress conditions is expected to be difficult.

Water and P Deficit Impacted Nutritional Quality Equally Among Genotypes

Water stress and P deficiency, both individually and combined, significantly decreased the concentration of most mineral nutrients detected (potassium, phosphorus, sulfur, calcium) in the seed (**Figure 7**). However, under WD conditions nitrogen concentration in the seed was found to increase by ∼50% and slightly further under the HPWD treatment (**Figures 7A,B**). This may be a consequence of increasing rates of remobilization, a well-documented response under abiotic stress conditions in common bean (Assefa et al., 2013; Beebe et al., 2013; Rao et al., 2013; Polania et al., 2016b,c). The increase in the nitrogen pool of the seed is reflected in the number of individual amino acids (**Figure 8**). Statistically significant increases (of between ∼2 and 50%) in the concentrations of most amino acids detected (alanine, arginine, asparagine, histidine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, valine) were found in the total seed in response to WD (**Figure 8B**). These results are in agreement with Gyori et al. (1998) who detected an increase in amino acids as a consequence of water stress in common bean and Nayyar et al. (2006) and Behboudian et al. (2001) who detected a similar increase of amino acids, and similar decrease of mineral nutrients, under water stress in chickpea (*Cicer arietinum* L.). Despite the relative importance of stress on the nutritive value of seeds, the mechanisms of accumulation of amino acids in the seed under stress are not well understood (Behboudian et al., 2001). Given that little attention has been directed toward the improvement of protein concentration in common beans in recent years as a result of negative correlations between protein concentration and seed yield (Beebe, 2012), further investigation into the accumulation of protein compounds under abiotic stress is required.

Nutritional quantity per seed or concentration (mineral nutrients and amino acids) did not vary in any of the genotypes measured. The lack of genotypic variation in nutritional content demonstrates that current breeding activities have maintained seed nutrient content notwithstanding successful efforts to increase seed yield quantity (Beebe et al., 2008). For instance, in this study, while superior-bred genotypes such as NCB226 maintained higher seed yield under WD and LP, compared to a commercial check such as DOR390 (**Figure 1**), and leaf-level nutrient concentration varied between the two genotypes (**Figure 5**), there were no statistically significant differences in the concentration of mineral nutrients and amino acids detected within the seed at harvest. The processes that have maintained seed nutrient concentration despite genotypic variation are not entirely clear. One hypothesis is that nutritional content may have been preserved due to the evolutionary requirements for seed germination. Given that a germinating seed requires a minimum amount of nutrients for successful germination and growth (see, e.g., White and Veneklaas, 2012) it is intuitive that the seed would reflect this need. Nevertheless, genotypic variation for nutrient content in common bean is present in the germplasm (Beebe et al., 2000; Blair et al., 2011; Blair, 2013) and has been exploited to increase the concentration of iron and zinc in common bean seed of recently released biofortified cultivars in Africa and Latin America (Haas et al., 2016; Andersson et al., 2017).

This investigation has demonstrated that water stress and P deficiency have substantial impacts on yield and yield-related parameters including aboveground biomass. Water deficit and low P led to physiological changes at the leaf level, including reductions to leaf-level photosynthesis and stomatal conductance, changes to carbon isotope abundance, and genotypic variation in the soluble pool of mineral nutrients and amino acids. Water deficit significantly increased the concentration of mineral nutrients and amino acids in the seed although no genotypic variation was detected, likely as a result of the need for viable seeds for germination. Plants produce

seeds, not for human nutrition but to sustain embryonic development at germination. Enhancing our knowledge of seed development and its resilience under realistic, natural environments containing multiple stresses has great potential to provide more informed management and breeding activities to improve the nutritional value and quantity of yield production.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

MS, EV, JP, IR, SB, and AM conceived and designed the experiment. MS, EV, JP, and IR performed data collection and site management. ED produced the data analysis and figures. MS and AM wrote the original manuscript with input from ED, EV, IR, and JP. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially supported by CGIAR Research Program on Grain Legumes through funds from the USAID. MS was the recipient of an Australian Postgraduate Award.

ACKNOWLEDGMENTS

We would like to thank the Plant Nutrition and Physiology team at the CIAT for the technical support. Metabolite samples were analyzed at Metabolomics Australia at The University of Melbourne, an NCRIS initiative under Bioplatforms Australia Pty Ltd. AM acknowledges the receipt of an Australian Research Council Fellowship FT120100200.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.814325/full#supplementary-material

Supplementary Figure 1 | A/c_i parameters V_{cmax} (μmol/m^2/s), J (μmol/m^2/s) and TPU for common bean lines DOR390, SEF60, and NCB226 under treatments of high phosphorus well-watered (HPWW; blue) and low phosphorus well-watered (LPWW; light blue) at flowering.

REFERENCES

Andersson, M. S., Saltzman, A., Virk, P. S., and Pfeiffer, W. (2017). Progress update: crop development of biofortified staple food crops under harvestplus. *Afr. J. Food Agric. Nutr. Dev.* 17, 11906–11935.

Assefa, T., Beebe, S. E., Rao, I. M., Cuasquer, J. B., Duque, M. C., Rivera, M., et al. (2013). Pod harvest index as a selection criterion to improve drought resistance in white pea bean. *Field Crops Res.* 148, 24–33. doi: 10.1016/j.fcr.2013.04.008

Badeck, F. W., Tcherkez, G., Nogues, S., Piel, C., and Ghashghaie, J. (2005). Post–photosynthetic fractionation of stable carbon isotopes between plant organs – a widespread phenomenon. *Rapid Commun. Mass Spectrom.* 19, 1381–1391. doi: 10.1002/rcm.1912

Beebe, S. (2012). "Common bean breeding in the tropics," in *Plant Breeding Reviews*, ed. J. Janick (Hoboken, NJ: John Wiley & Sons Inc).

Beebe, S. E., Rao, I. M., Blair, M. W., and Acosta-Gallegos, J. A. (2013). Phenotyping common beans for adaptation to drought. *Front. Physiol.* 4:35. doi: 10.3389/fphys.2013.00035

Beebe, S. E., Rao, I. M., Blair, M. W., and Butare, L. (2009). "Breeding for abiotic stress tolerance in common bean: present and future challenges," in *Proceedings of the 14th Australian Plant Breeding and 11th Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO) Conference*, Brisbane.

Beebe, S. E., Rao, I. M., Cajiao, C., and Grajales, M. (2008). Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Sci.* 48:582. doi: 10.2135/cropsci2007.07.0404

Beebe, S., Gonzalez, A. V., and Rengifo, J. (2000). Research on trace minerals in the common bean. *Food Nutr. Bull.* 21, 387–391.

Behboudian, M. H., Ma, Q. F., Turner, N. C., and Palta, J. A. (2001). Reactions of chickpea to water stress: yield and seed composition. *J. Sci. Food Agric.* 81, 1288–1291. doi: 10.1002/jsfa.939

Bennett, E. J., Roberts, J. A., and Wagstaff, C. (2011). The role of the pod in seed development: strategies for manipulating yield. *New Phytol.* 190, 838–853. doi: 10.1111/j.1469-8137.2011.03714.x

Blair, M. W. (2013). Mineral biofortification strategies for food staples: the example of common bean. *J. Agric. Food Chem.* 61, 8287–8294. doi: 10.1021/jf400774y

Blair, M. W., Astudillo, C., Rengifo, J., Beebe, S. E., and Graham, R. (2011). QTL analyses for seed iron and zinc concentrations in an intra–genepool population of andean common beans (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 122, 511–521. doi: 10.1007/s00122-010-1465-8

Broughton, W. J., Hernandez, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* spp.) – model food legumes. *Plant Soil* 252, 55–128. doi: 10.1023/a:1024146710611

Bulyaba, R., Winham, D. M., Lenssen, A. W., Moore, K. J., Kelly, J. D., Brick, M. A., et al. (2020). Genotype by location effects on yield and seed nutrient composition of common bean. *Agronomy* 10:347.

Cavalieri, A., Merchant, A., and van Volkenburgh, E. (2011). Why not beans? Foreword. *Funct. Plant Biol.* 38, iii–iv. doi: 10.1071/FPv38n12_FO

Cernusak, L. A., Tcherkez, G., Keitel, C., Cornwell, W. K., Santiago, L. S., Knohl, A., et al. (2009). Why are non–photosynthetic tissues generally 13C enriched compared with leaves in C3 plants? *Funct. Plant Biol.* 36, 199–213.

Cuellar-Ortiz, S. M., Arrieta-Montiel, M. D., Acosta-Gallegos, J., and Covarrubias, A. A. (2008). Relationship between carbohydrate partitioning and drought resistance in common bean. *Plant Cell Environ.* 31, 1399–1409. doi: 10.1111/j.1365-3040.2008.01853.x

Dumschott, K., Richter, A., Loescher, W., and Merchant, A. (2017). Post photosynthetic carbon partitioning to sugar alcohols and consequences for plant growth. *Phytochemistry* 144, 243–252. doi: 10.1016/j.phytochem.2017.09.019

Garcia, C. B., and Grusak, M. A. (2015). Mineral accumulation in vegetative and reproductive tissues during seed development in *Medicago truncatula*. *Front. Plant Sci.* 6:622. doi: 10.3389/fpls.2015.00622

Gyori, Z., Nemeskeri, E., and Szilagyi, S. (1998). Legumes grown under nonirrigated conditions. *J. Agric. Food Chem.* 46, 3087–3091. doi: 10.1021/jf980060c

Haas, J. D., Luna, S. V., Lung'aho, M. G., Wenger, M. J., Murray-Kolb, L. E., Beebe, S., et al. (2016). Consuming iron biofortified beans increases iron status in Rwandan women after 128 days in a randomized controlled feeding trial. *J. Nutr.* 146, 1586–1592. doi: 10.3945/jn.115.224741

Hare, P. D., Cress, W. A., and Van Staden, J. (1998). Dissecting the roles of osmolyte accumulation during stress. *Plant Cell Environ.* 21, 535–553.

Hermans, C., Hammond, J. P., White, P. J., and Verbruggen, N. (2006). How do plants respond to nutrient shortage by biomass allocation? *Trends Plant Sci.* 11, 610–617.

Ho, M. D., Rosas, J. C., Brown, K. M., and Lynch, J. P. (2005). Root architectural tradeoffs for water and phosphorus acquisition. *Funct. Plant Biol.* 32, 737–748. doi: 10.1071/Fp05043

Hocking, P. J., and Pate, J. S. (1977). Mobilization of minerals to developing seeds of legumes. *Ann. Bot.* 41, 1259–1278.

Merchant, A., Adams, M. A., Richter, A., and Popp, M. (2006). Targeted metabolite profiling provides a functional link among eucalypt taxonomy, physiology and evolution. *Phytochemistry* 67, 402–408.

Merchant, A., Peuke, A. D., Keitel, C., Macfarlane, C., Warren, C., and Adams, M. A. (2010). Phloem sap and leaf δ13C, carbohydrates and amino acid concentrations in *Eucalyptus globulus* change systematically according to flooding and water deficit treatment. *J. Exp. Bot.* 61, 1785–1793.

Nayyar, H., Kaur, S., Singh, S., and Upadhyaya, H. D. (2006). Differential sensitivity of Desi (small-seeded) and Kabuli (large-seeded) chickpea genotypes to water stress during seed filling: effects on accumulation of seed reserves and yield. *J. Sci. Food Agric.* 86, 2076–2082. doi: 10.1002/jsfa.2574

Nielsen, K. L., Eshel, A., and Lynch, J. P. (2001). The effect of phosphorus availability on the carbon economy of contrasting common bean (*Phaseolus vulgaris* L.) genotypes. *J. Exp. Bot.* 52, 329–339. doi: 10.1093/jexbot/52.355.329

Petry, N., Boy, E., Wirth, J., and Hurrell, R. (2015). The potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification. *Nutrients* 11 7, 1144–1173.

Polania, J. A., Poschenrieder, C., Beebe, S., and Rao, I. M. (2016c). Effective use of water and increased dry matter partitioned to grain contribute to yield of common bean improved for drought resistance. *Front. Plant Sci.* 7:660. doi: 10.3389/fpls.2016.00660

Polania, J., Poschenrieder, C., Rao, I., and Beebe, S. (2016a). Estimation of phenotypic variability in symbiotic nitrogen fixation ability of common bean under drought stress using N-15 natural abundance in grain. *Eur. J. Agron.* 79, 66–73. doi: 10.1016/j.eja.2016.05.014

Polania, J., Rao, I. M., Cajiao, C., Rivera, M., Raatz, B., and Beebe, S. (2016b). Physiological traits associated with drought resistance in andean and mesoamerican genotypes of common bean (*Phaseolus vulgaris* L.). *Euphytica* 210, 17–29. doi: 10.1007/s10681-016-1691-5

Polania, J., Poschenrieder, C., Rao, I. M., and Beebe, S. (2017). Root traits and their potential links to plant ideotypes to improve drought resistance in common bean. *Theor. Exp. Plant Physiol.* 29, 143–154. doi: 10.1007/s40626-017-0090-1

Polania, J., Rao, I. M., Beebe, S., and Garcia, R. (2009). Root development and distribution under drought stress in common bean (*Phaseolus vulgaris* L.) in a soil tube system. *Agron. Colomb.* 27, 25–32.

Popp, M., Lied, W., Meyer, A. J., Richter, A., Schiller, P., and Schwitte, H. (1996). Sample preservation for determination of organic compounds: microwave vs freeze drying. *J. Exp. Bot.* 47, 1469–1473.

Pottier, M., Masclaux-Daubresse, C., Yoshimoto, K., and Thomine, S. (2014). Autophagy as a possible mechanism for micronutrient remobilization from leaves to seeds. *Front. Plant Sci.* 5:11. doi: 10.3389/fpls.2014.00011

R Core Team. (2021). *R: The R Project for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raines, C. A. (2003). The calvin cycle revisited. *Photosynth. Res.* 75, 1–10.

Rao, I. M., Beebe, S. E., Polania, J., Grajales, M., Cajiao, C., Ricaurte, J., et al. (2017). Evidence for genotypic differences among elite lines of common bean in the ability to remobilize photosynthate to increase yield under drought. *J. Agric. Sci.* 155, 857–875. doi: 10.1017/S0021859616000915

Rao, I., Beebe, S., Polania, J., Ricaurte, J., Cajiao, C., Garcia, R., et al. (2013). Can tepary bean be a model for improvement of drought resistance in common bean? *Afr. Crop Sci. J.* 21, 265–281.

Rychter, A. M., and Rao, I. M. (2005). "Role of phosphorus in photosynthetic carbon metabolism," in *Handbook of Photosynthesis*, ed. M. Pessarakli (Boca Raton, FL: CRC press), 123–148.

Scheidegger, Y., Saurer, M., Bahn, M., and Siegwolf, R. (2000). Linking stable oxygen and carbon isotopes with stomatal conductance and photosynthetic capacity: a conceptual model. *Oecologia* 125, 350–357. doi: 10.1007/s004420000466

Seibt, U., Rajabi, A., Griffiths, H., and Berry, J. A. (2008). Carbon isotopes and water use efficiency: sense and sensitivity. *Oecologia* 155, 441–454. doi: 10.1007/s00442-007-0932-7

Sharkey, T. D. (2015). What gas exchange data can tell us about photosynthesis. *Plant Cell Environ.* 39, 1161–1163.

Singh, S. P. (1981). A key for identification of different growth habits of *Phaseolus vulgaris* L. *Ann. Rep. Bean Improv. Coop.* 25, 92–95.

Smith, M. R., Fuentes, D., and Merchant, A. (2018a). Chemical and isotopic markers detect water deficit and its influence on nutrient allocation in *Phaseolus vulgaris*. *Physiol. Plant.* 167, 391–403. doi: 10.1111/ppl.12899

Smith, M. R., Rao, I. M., and Merchant, A. (2018b). Source-sink relationships in crop plants and their influence on yield development and nutritional quality. *Front. Plant Sci.* 9:1889. doi: 10.3389/fpls.2018.01889

Smith, M. R., Veneklaas, E., Polania, J., Rao, I. M., Beebe, S. E., and Merchant, A. (2019). Field drought conditions impact yield but not nutritional quality of the seed in common bean (*Phaseolus vulgaris* L.). *PLoS One* 14:e0217099. doi: 10.1371/journal.pone.0217099

Smith, M., Wild, B., Richter, A., Simonin, K., and Merchant, A. (2016). Carbon isotope composition of carbohydrates and polyols in leaf and phloem sap of *Phaseolus vulgaris* L. influences predictions of plant water use efficiency. *Plant Cell Physiol.* 57, 1756–1766. doi: 10.1093/pcp/pcw099

Tan, G. Z. H., Das Bhowmik, S. S., Hoang, T. M. L., Karbaschi, M. R., Johnson, A. A. T., Williams, B., et al. (2017). Finger on the pulse: pumping iron into chickpea. *Front. plant Sci.* 8:1755. doi: 10.3389/fpls.2017.01755

Terán, H., and Singh, S. P. (2002). Comparison of sources and lines selected for drought resistance in common bean. *Crop Sci.* 42, 64–70.

Wentworth, M., Murchie, E. H., Gray, J. E., Villegas, D., Pastenes, C., Pinto, M., et al. (2006). Differential adaptation of two varieties of common bean to abiotic stress – II. *J. Exp. Bot.* 57, 699–709. doi: 10.1093/jxb/erj061

White, P. J., and Veneklaas, E. J. (2012). Nature and nurture: the importance of seed phosphorus content. *Plant Soil* 357, 1–8. doi: 10.1007/s11104-012-1128-4

Zhu, X. G., Long, S. P., and Ort, D. R. (2010). "Improving photosynthetic efficiency for greater yield," in *Annual Review of Plant Biology*, eds S. Merchant, W. R. Briggs, and D. Ort (Palo Alto, CA: Annual Reviews), 235–261.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Smith, Dinglasan, Veneklaas, Polania, Rao, Beebe and Merchant. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

REVIEW
published: 05 April 2022
doi: 10.3389/fpls.2022.869713

Protein Biofortification in Lentils (*Lens culinaris* Medik.) Toward Human Health

*Sonia Salaria[1], Jon Lucas Boatwright[1], Pushparajah Thavarajah[1], Shiv Kumar[2] and Dil Thavarajah[1]**

[1]*Plant and Environmental Sciences, Clemson University, Clemson, SC, United States,* [2]*Biodiversity and Crop Improvement Program, International Centre for Agricultural Research in the Dry Areas (ICARDA), Rabat-Institute, Rabat, Morocco*

Lentil (*Lens culinaris* Medik.) is a nutritionally dense crop with significant quantities of protein, low-digestible carbohydrates, minerals, and vitamins. The amino acid composition of lentil protein can impact human health by maintaining amino acid balance for physiological functions and preventing protein-energy malnutrition and non-communicable diseases (NCDs). Thus, enhancing lentil protein quality through genetic biofortification, i.e., conventional plant breeding and molecular technologies, is vital for the nutritional improvement of lentil crops across the globe. This review highlights variation in protein concentration and quality across *Lens* species, genetic mechanisms controlling amino acid synthesis in plants, functions of amino acids, and the effect of antinutrients on the absorption of amino acids into the human body. Successful breeding strategies in lentils and other pulses are reviewed to demonstrate robust breeding approaches for protein biofortification. Future lentil breeding approaches will include rapid germplasm selection, phenotypic evaluation, genome-wide association studies, genetic engineering, and genome editing to select sequences that improve protein concentration and quality.

Keywords: Lentil (*Lens culinaris L.*), protein, biofortification, amino acids, protein quality, food secuity

OPEN ACCESS

Edited by:
Eric Von Wettberg, University of Vermont, United States

Reviewed by:
Melike Bakir, Erciyes University, Turkey
Tadesse Sefera Gela, University of Saskatchewan, Canada

****Correspondence:***
Dil Thavarajah
dthavar@clemson.edu

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *04 February 2022*
Accepted: *14 March 2022*
Published: *05 April 2022*

Citation:
Salaria S, Boatwright JL, Thavarajah P, Kumar S and Thavarajah D (2022) Protein Biofortification in Lentils (Lens culinaris Medik.) Toward Human Health. Front. Plant Sci. 13:869713. doi: 10.3389/fpls.2022.869713

INTRODUCTION

Nutritional imbalances and deficiencies cause several malnutritional and non-communicable diseases (NCDs) in humans. A poor diet that lacks macro- and micronutrients, such as proteins, low-digestible carbohydrates (LDCs), fats, vitamins, and minerals, results in protein and micronutrient malnutrition. Low-digestible carbohydrates (LDs) are, also known as prebiotic carbohydrates, defined as 'a substrate that is selectively utilized by host microorganisms conferring a health benefit' (Gibson et al., 2017). These dietary prebiotic carbohydrates pass undigested through the upper digestive tract and are fermented by microorganisms in the colon for increased gut health. The most common human health impacts of malnutrition are stunting, intestinal health issues impairing digestion, obesity, overweight, and an increased risk of diet-related NCDs (Branca et al., 2019). Major NCDs related to poor dietary intake that threatens human life include cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes (World Health Organization, 2019). Notably, a protein-deficient diet leading to protein malnutrition has alarming consequences that affect infants, young children, and females across the globe (Semba, 2016). However, a protein-rich legume-based diet is a viable, sustainable, and healthy

option to prevent malnutrition in developing countries. Though animal proteins are extensively utilized in human diets, plant-based proteins have grown in popularity. Their demand has increased globally due to nutritional value, low carbon input, and environmental concerns (Asif et al., 2013).

Staple foods rich in macro- and micronutrients can alleviate the risk of malnutrition. Plant-based diets comprised mainly of cereals and legume staples are popular worldwide. Legume crops, including lentils (*Lens culinaris* Medikk.), have a protein concentration (20–30%) higher than cereals (10–12%) and thus have the potential to combat protein malnutrition and serve as gluten- and allergen-free protein sources. Lentil is highly nutritious, affordable and has a shorter cooking time than other pulse crops, and features high protein concentrations, low-digestible carbohydrates, minerals, vitamins, and low concentrations of phytic acid (Thavarajah et al., 2009; Kumar et al., 2015). Lentil is not a source of cholesterol, and its low-fat content makes it easier to digest than other pulse crops. Lentil proteins include both essential and non-essential amino acids but are notably low in the sulfur-containing amino acids methionine (Met) and cysteine (Cys; Khazaei et al., 2019). Biofortification is a possible approach to improve the unbalanced composition of amino acids in lentils through appropriate conventional breeding strategies and genomic selection. With increasing global protein demand, protein biofortification would justify lentils as a '*nutritional booster*' to increase global nutritional security and combat malnutrition and NCDs.

Lentil proteins are stored in the cotyledonary cells in membranous protein bodies called 'storage proteins' (Duranti and Gius, 1997). These seed proteins supply carbon (C), nitrogen (N), and sulfur (S) and compose 80% of the total protein available for germination, subsequent plant growth, and disease resistance (Khazaei et al., 2019). Storage proteins also play a defensive role against bruchids, insects of the family Bruchidae, in cowpea (*Vigna unguiculata*; Sales et al., 2000). These proteins are classified into four types: globulins (salt soluble), albumins (water-soluble), prolamins (ethanol soluble), and glutelins (acid-soluble; Osborne, 1924). Like other pulse crops, lentils are rich in globulins and albumins, whereas prolamins and glutelins are more prominent in cereals (Osborne, 1924). Globulins were the first type of storage protein reported in lentils (Osborne and Campbell, 1898) and are the principal proteins in lentils, making up ~44–70% of all storage proteins. Two subclasses of globulins, i.e., 11 s type (legumin) and 7 s type (vicilin/convicilin), were also defined (Danielson, 1950). Albumins comprise 26–61% of lentil proteins, and prolamins and glutelins only make up a small fraction (Saint-Clair, 1972; Sulieman et al., 2008).

Storage protein quantities demonstrate high variability due to the quantitative nature of the genes regulating protein synthesis in the seeds (Kumar et al., 2020). Higher genotype × environmental interactions, indicated by the moderate broad sense heritability (31.31%), is another reason for the high variation in the storage protein concentration in lentil seeds (Gautam et al., 2018). Lentil seed proteins, excluding storage proteins, also have metabolic functions. These metabolic proteins regulate numerous physiological processes in the plant, including enzymatic activity and structural and physiological functions (Scippa et al., 2010). Ultimately, lentil seed protein composition contributes to human health by providing essential amino acids necessary for metabolic processes and nutritional balance in the human body. Optimizing the plant breeding process and location sourcing may help develop better protein-enriched lentil cultivars for global plant-based protein demand. The objectives of this paper are to review the protein concentration and quality variations within the genus *Lens*, pathways and genes regulating the synthesis of amino acids, functions of amino acids for human health, and breeding strategies related to lentil protein biofortification.

LENTIL BIOFORTIFICATION

Lentil is an annual diploid (2n = 2x = 14) cool-season food legume that originated in the Middle East (Cubero, 1981). The genus *Lens* comprises *L. culinaris*, *L. ervoides*, *L. nigricans*, and *L. lamoletti*. *L. culinaris* is further divided into four taxa: *L. culinaris* ssp. *culinaris*, *L. culinaris* ssp. *orientalis*, *L. culinaris* ssp. *tomentosus*, and *L. culinaris* ssp. *odemensis* (Ferguson et al., 2000). Lens genus has been classified as primary, secondary, tertiary, and quaternary genetic pools according to the phylogeny using the Genotyping-by-sequencing (GBS). The primary gene pool contains *L. culinaris*, *L. orientalis*, and *L. tomentosus*, whereas *L. odemensis* and *L. lamoletti* are in the secondary gene pool. However, each tertiary and quaternary gene pools contain single species, *L. ervoides* and *L. nigricans*, respectively (Wong et al., 2015). Of these, only *L. culinaris* ssp. *culinaris* is domesticated and cultivated worldwide, representing crops over a 5.01 M ha area with an annual production of 6.54 M tonnes. Canada is a leading producer, contributing about 44% of the world's lentils; other major lentil-producing countries are India, the United States of America (United States), Turkey, Australia, Nepal, and Bangladesh (FAOSTAT, 2021).

Lentils are a staple food that is easily digested compared to other legumes. The biofortification of lentils could significantly fight hidden hunger and nutritional disorders. Hidden hunger is also known as micronutrient deficiency despite sufficient calorie intake (Lowe, 2021). Several breeding programs have been established worldwide that seek to biofortify lentils with protein, prebiotic carbohydrates, micronutrients, vitamins, etc. (Kumar et al., 2016a). Many lentil accessions have been screened for amino acid concentration (Iqbal et al., 2006), protein (Bhatty and Slinkard, 1979), starch (Zia-Ul-Haq et al., 2011), fatty acids (Grusak, 2009), macro- and micronutrients (Kumar et al., 2016a; Podder et al., 2020; Rasheed et al., 2020), folates (Sen Gupta et al., 2013), and antinutritional factors (Thavarajah et al., 2009, 2011). Marker-assisted breeding has also demonstrated the potential for identifying genes/quantitative trait loci (QTL) for iron (Fe) uptake (Kumar et al., 2015; Aldemir et al., 2017), Fe and Zinc (Zn) concentration (Kumar et al., 2014), and selenium (Se) concentration (Ates et al., 2016). Furthermore, the HarvestPlus Challenge program, established in 2004, was a landmark effort that increased lentil biofortification efforts worldwide. They released several lentil cultivars to economically

underprivileged global regions in Asia and Africa (Kumar et al., 2016a). Notably, numerous high Fe and Zn cultivars have been released, including Barimasur-4, -5, -6, -7, -8 and -9 in Bangladesh; Khajuraho-1, -2 and -3, Sital, Shekhar, Sisir, and Simal in Nepal; L 4704, IPL 220, Pusa Agaiti and Pusa Vaibhav in India; Idlib-2 and -3 in Syria/Lebanon; and Alemeya in Ethiopia. Smallholder farmers regularly use these biofortified lentils in Africa and Southeast Asia (Harvest Plus, 2014).

Various researchers have reported protein concentrations in current lentil cultivars in the range of 20–30% (**Table 1**). In a study (Bhatty, 1986), similar protein concentrations in wild and cultivated lentils, indicating homogeneity for protein concentration in the genus *Lens*, were identified. However, a recent study (Kumar et al., 2016b) efficiently distinguished wild species from cultivated lentils for protein concentration. In this study, *L. orientalis*, an immediate progenitor of cultivated lentils, expressed the highest average protein (24.15%) among all the wild species, followed by *L. ervoides* (22.99%). Other wild species, *L. odemensis*, and *L. nigricans* showed slightly higher average protein content (19.7 and 19.53%, respectively) than *L. culinaris*. A similar protein level was seen in *L. tomentosus* (18.75%) and cultivated lentils (18.7%). Extensive variation was observed for protein content within *L. orientalis* and *L. ervoides*, ranging from 18.3 to 27.75% and 18.9 to 32.7%, respectively. ILWL-47, an *L. ervoides* accession, had an exceptionally high protein content of about 32.7% and is, therefore, a potential candidate for protein quality improvement in lentil breeding programs (Kumar et al., 2016b). Protein subunit fraction profiling has indicated variable levels of the albumin protein fraction (APF) and globulin protein fraction (GPF) among *Lens* species, with the wild species having higher APF and GPF concentrations than the cultivated species (Bhatty, 1982). Among the evaluated wild species, *L. orientalis* and *L. ervoides* contained higher APF and GPF levels than *L. nigricans* (Bhatty, 1982).

The proportion of amino acids in lentil proteins varies across genotypes in the cultivated gene pool (**Table 2**). Met and tryptophan (Trp) represent a minor fraction among all amino acids and are thus termed limiting amino acids. Comparing lentil protein with cereal proteins indicates the good nutritional complementation between Met and lysine (Lys), but to some extent, for Trp and threonine (Thr) because cereals are rich in both Met and Trp (Bhatty, 1986). Generally, all essential amino acids except Lys are deficient in lentils, but a moderate to the high proportion of non-essential amino acids are present (Khazaei et al., 2019). Lentil proteins are also lacking in other S-containing amino acids such as Cys. The albumin fraction of lentils contains more essential amino acids than the globulin fraction (Bhatty, 1982). Recent studies also indicate that amino acids vary among distinct species of the genus *Lens*, with a spectrum of variation seen for amino acid content among *L. culinaris*, *L. orientalis*, *L. ervoides*, *L. nigricans*, and *L. odemensis*. Phenylalanine (Phe), Met, valine (Val), leucine (Leu), and isoleucine (Ile) concentrations are significantly higher in wild species than cultivated lentils (**Table 3**; Rozan et al., 2001). Similarly, the non-essential amino acid content is also higher in wild species than in *L. culinaris*. Such evidence signifies wild species are a potential source of candidate genes that can be harnessed to improve protein quality in cultivated lentils.

TABLE 1 | Genetic variation for protein concentration in cultivated lentils (*L. culinaris*).

Protein concentration (% of dry matter)	Total accessions used	Reference
24.6–30.0	23	Heuzé et al., 2021
10.5–27.1	45	Kumar et al., 2016b
21.8–27.1	14	Zaccardelli et al., 2012
25.3–29.3	35	Alghamdi et al., 2014
23.8–29.3	22	Tahir et al., 2011
24.3–30.2	4	Wang and Daun, 2006
23.9–26.3	58	Stoddard et al., 1993
25.5–28.9	24	Erskine et al., 1985

GENETIC CONTROL FOR AMINO ACID BIOSYNTHESIS IN PLANTS

The genetic mechanisms controlling seed protein concentration have similar regulation and pathways in different plants, including pulse crops. In pulse crops, genetic control of seed protein content has not been widely studied except in chickpea (*Cicer arietinum*), soybean (*Glycine max*), and pea (*Pisum sativum*). However, genetic control of seed protein content has been studied extensively in cereals (Mann et al., 2009; Olsen and Phillips, 2009; Chen et al., 2018; Borisjuk et al., 2019) and the model plant *Arabidopsis thaliana* (Jasinski et al., 2016). In chickpea, seven candidate genes that regulate seed protein concentration were identified using a

TABLE 2 | Amino acid profile of cultivated lentil genotypes (Sayeed and Njaa, 1985; Shekib et al., 1986; Kahraman, 2018).

Amino acids	Concentration (g/100 g of protein)
Ala	3.31–8.35
Arg	4.64–13.80
Asp	6.36–13.20
Cys	0.60–1.62
Glu	6.12–17.10
Gly	4.40–10.40
His	1.21–9.15
Ile	2.20–5.00
Leu	5.21–7.72
Lys	5.81–9.59
Met	0.90–2.23
Phe	3.85–7.55
Pro	3.50–5.22
Ser	4.90–6.34
Thr	1.04–4.60
Trp	0.57–1.37
Tyr	2.71–7.15
Val	4.10–5.01

Ala, alanine; Arg, arginine; Asp, aspartate/aspartic acid; Cys, cysteine; Glu, glutamate/glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.

TABLE 3 | Amino acid concentrations among different *Lens* species (Rozan et al., 2001).

Amino acids	L. culinaris	L. orientalis	L. ervoides	L. nigricans	L. odemensis
	mg amino acids/g of dry seed weight				
Ala	20.42	39.81	16.01	22.47	21.32
Arg	10.61	14.04	12.05	7.48	9.10
Asp	10.96	26.10	17.42	7.68	11.17
Cys	0.40	0.39	0.53	0.47	0.44
Glu	26.55	42.27	32.62	19.95	24.22
Gly	9.77	12.66	11.48	7.89	10.22
His	8.74	3.95	9.75	4.94	6.84
Ile	6.26	9.58	8.59	7.76	5.06
Leu	10.64	15.86	14.07	11.74	8.09
Lys	4.54	12.64	9.48	6.14	5.69
Met	1.49	1.63	1.74	1.22	1.18
Phe	6.70	10.64	9.37	9.46	5.55
Pro	11.11	11.36	11.54	10.52	8.88
Ser	11.38	15.60	14.10	8.70	11.20
Thr	5.57	7.57	6.31	4.56	5.62
Trp	NA	NA	NA	NA	NA
Tyr	6.34	7.53	6.65	6.35	5.05
Val	8.54	11.64	9.60	8.64	7.24

Ala, alanine; Arg, arginine; Asp, aspartate/aspartic acid; Cys, cysteine; Glu, glutamate/glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.

genome-wide association study of 336 *desi* and *Kabuli* accessions (Upadhyaya et al., 2016). In soybean, three QTL (qPro10a, qPro13a, and qPro17b) for protein were identified in a recombinant inbred line (RIL) population (Zhonghuang 24 × Huaxia 3) on chromosomes 10, 13, and 17, respectively (Liu et al., 2017).

Several genes regulating the seed protein concentration in soybean were found on chromosomes 15 and 20 (Patil et al., 2017). Another gene, *BIG SEEDS1* (*BS1*), controlling seed size, weight, and composition of amino acids in the protein, has been characterized in *Medicago trunculata* and soybean (Ge et al., 2016). Groups of highly coordinated genes (HCGs) controlling the aspartate family (Met, Ile, Lys, Thr, and Gly) and branched aromatic amino acid formation were also identified in *A. thaliana* (Less and Galili, 2009). These two HCGs have several genes controlling the formation of amino acids. The first group related to the aspartate family contained catabolic genes for *THA1* (Thr to Gly metabolism), *BCAT2* (Ile metabolism), *MGL* (Met catabolism), and *LKR/SDH* (Lys metabolism). However, the second group exclusively regulated Met metabolism and was termed the 'Met metabolism group.' It contained the genes *AK/HSDH1* (encoding aspartate kinase enzyme for the formation of aspartate-4-semialdehyde, the first substrate for amino acid synthesis), *CGS1* (Met synthesis), *DAPD* (Lys synthesis), *SAMS3* (Met catabolism), *BCAT3* (Ile metabolism), and *BCAT4*, *MAM1*, and *MAML* (Met catabolism). One of the two groups related to branched aromatic acids contained ten genes (*ASA1*, *ASB*, *TSA2*, *TSB1/2*, *IGPS* for Trp synthesis, *CYP79B2* for Trp catabolism, *PD* for Phe synthesis, *PAL1* and *PAL2* for Phe catabolism, and *TAT3* for tyrosine (Tyr) catabolism). In contrast, two genes (*PAL3* and *IGPS*) were reported in the second group (Less and Galili, 2009).

The genes regulating the synthesis of enzymes that mediate the formation of amino acids and their precursors have been extensively studied in plants (**Table 4**; **Figure 1**). In *A. thaliana*, glutamate is formed from precursor 2-oxoglutarate by enzymatic aminotransferases, a process that is regulated by 44 putative genes (Liepman and Olsen, 2004). Glutamate synthase production, which converts glutamine (Gln) to glutamate, is controlled by either one or two genes in the chloroplast and mitochondria (Gaufichon et al., 2016). Similarly, six genes encode Gln synthase, which converts glutamate to Gln, in *A. thaliana* (Forde and Lea, 2007). Glutamate is a precursor that synthesizes arginine (Arg) and proline (Pro) using 20 enzymes encoded by about 30 genes in *A. thaliana* (Majumdar et al., 2016). Glutamine with aspartate also forms asparagine (Asn) in plants by the transamination action of the Asn synthetase (AS) enzyme encoded by the *asnB* gene in eukaryotes (Gaufichon et al., 2010) and the *ASN* gene family (*ASN1*, *ASN2*, and *ASN3*) in *Arabidopsis* (**Table 4**; Arabidopsis Genome Initiative, 2000). A histidine (His) synthesis pathway revealed eight genes (*ATP-PRT*, *PRATP/CH*, *ProFAR-I*, *IGPS*, *IGPD*, *HPA*, *HPP*, and *HDH*) forming eight enzymes in *A. thaliana* (Rees et al., 2009). Two branched-chain amino acids, Val and Leu, form with the acetohydroxyacid synthase (AHAS) enzyme acting on pyruvate producing acetolactate. This enzyme forms the third branched-chain amino acid, Ile, by serving on a substrate formed from Thr in the pathway for 2-ketobutyrate converting Thr to Ile. A single gene encodes the AHAS enzyme in *Arabidopsis* (Singh and Shaner, 1995).

The enzyme chorismate mutase (CM) is encoded by three genes (*AtCM1*, *AtCM2*, and *AtCM3*) and is a precursor for chorismate to form prephenate for Phe and Tyr biosynthesis in plants (**Figure 1**). The formation of Trp from chorismate is regulated by three genes (*ASa1*, *ASa2*, and *ASb1*) and seven putative genes (two *Asa* and five *ASb* genes) encoding anthranilate synthase (AS) enzyme-producing anthranilate (**Table 4**). This anthranilate generates Trp using five enzymes (PAT1, PAI, IGPS, TS a, and TS b) encoded by eight genes in plants (Tzin and Galili, 2010; Parthasarathy et al., 2018). Aspartate regulates the formation of four essential amino acids, Ile, Lys, Met, and Thr, also termed aspartate-derived amino acids. Five genes encode aspartate formation enzymes in *A. thaliana* (Han et al., 2021). In C_3 plants, including lentils, two pathways are identified for serine (Ser) formation, namely photorespiratory and non-photorespiratory pathways in photosynthetic and non-photosynthetic tissues, respectively (**Figure 1**). The Ser produced in different pathways is converted into glycine (Gly) in non-photosynthetic tissues in the presence of the Ser hydroxymethyltransferase (SHM) enzyme. Ser also synthesizes Cys by following a two-step pathway in plants regulated by Ser acetyltransferase (SAT) and O-acetylserine (thiol)lyase (OASTL) enzymes encoded by five and nine genes, respectively (Howarth et al., 1997; Wirtz et al., 2004).

TABLE 4 | Genes responsible for amino acid synthesis.

Amino acid	Key precursors	Key enzymes	Genes in *Arabidopsis*	References
Glutamate	2-oxoglutarate; Glutamine	Amino transferases; glutamate synthase (GOGAT): two forms-ferredoxin (Fd) and NADH	44 putative genes; Fd form: *GLU1*, *GLU2*; NADH form: *GLT*	Liepman and Olsen, 2004; Forde and Lea, 2007
Glutamine	Glutamate	Glutamine synthase: two forms—plastidic (GS1) and cytoplasmic (GS2)	GS1 form: one gene; GS2 form: five genes	Forde and Lea, 2007; Gaufichon et al., 2016
Asparagine	Glutamine and Aspartate	Asparagine synthase	asnB gene; ASN gene family (*ASN1*, *ASN2*, *ASN3*)	Arabidopsis Genome Initiative, 2000; Gaufichon et al., 2010
Histidine	Ribose-5-phosphate	Eight enzymes	*PRATP/CH*, *ProFAR-I*, *IGPS*, *HPP*, *HDH*- single copy genes; *ATP-PRT*, *IGPD*, *HPA*-duplicated genes	Rees et al., 2009; Ingle, 2011
Leucine	Pyruvate 2-oxoisovalerate	Acetohydroxyacid synthase (AHAS) enzyme; Isopropylmalate synthase (IPMS), isopropylmalate isomerase (IPMI), and isopropylmalate dehydrogenase (IPMDH)	*AHAS* gene; IPMS: *IPMS1*, *IPMS2*, *IPMI LSU1 IPMI SSU1 IPMI SSU2*, *IPMI SSU3*, *IPMDH* gene	Singh and Shaner, 1995; Calder, 1995; Xing and Last, 2017; Knill et al., 2009
Valine	Pyruvate 2-oxoisovalerate	Acetohydroxyacid synthase (AHAS) enzyme; Amino transferase	*AHAS* gene; Single gene	Singh and Shaner, 1995; Calder, 1995; Xing and Last, 2017; Knill et al., 2009
Isoleucine	2-ketobutyrate	Acetohydroxyacid synthase (AHAS) enzyme	*AHAS* gene	Singh and Shaner, 1995
Alanine	Pyruvate and glutamate	Alanine aminotransferases	Eight genes	Parthasarathy et al., 2019
Phenylalanine Tyrosine	Chorismate Prephenate Arogenate	Chorismate mutase (CM), Prephenate aminotransferase; Phenylalanine synthesis: Arogenate dehydratase; Tyrosine synthesis: Arogenate dehydrogenase	CM: *AtCM1*, *AtCM2*, *AtCM3*, *AtPPA-AT* gene; Six genes (*ADT1*, *ADT2*, *ADT3*, *ADT4*, *ADT5*, *ADT6*)	Tzin and Galili, 2010; Dudareva et al., 2011
Tryptophan	Chorismate Anthranillite	Anthranilate synthase (AS); Anthranilate phosphoribosyltransferase (PAT1), indole-3-glycerol phosphate synthase (IGPS), tryptophan synthase alpha (TS a), phosphoribosylanthranilate isomerase (PAI), and tryptophan synthase beta (TS b)	Two genes (*TyrA1*, *TyrA2*) Three genes (*ASa1*, *ASa2*, *ASb1*) and seven putative genes (two *Asa* and five *ASb* genes); *PAT1*, *IGPS*, *TSa*, three genes (*PAI1*, *PAI2*, *PAI3*) and two genes (*TSb1 and TSb2*)	Tzin and Galili, 2010; Parthasarathy et al., 2018
Aspartate	Oxaloacetate and glutamate	Aspartate aminotransferase (AspAT); Aspartate kinase (AK)	Five genes: *AspAT1*, *AspAT2*, *AspAT3*, *AspAT4*, *AspAT4*, *AspAT5*	Han et al., 2021
Methionine, Threonine, Isoleucine Lysine	Aspartate L-aspartate-4-semialdehyde	Methionine, threonine and isoleucine synthesis: homoserine dehydrogenase (HSD); Lysine synthesis: dihydrodipicolinate synthase (DHDPS)	Five genes; Two genes; Two genes	Vauterin and Jacobs, 1994; Vauterin et al., 1999; Craciun et al., 2000; Sarrobert et al., 2000; Galili, 2011

AMINO ACIDS IMPACT HUMAN HEALTH

Amino acids are the foundational units of proteins. Structural conformations have unique chemical properties due to basic (amide) and acidic (carboxylic) chemical groups. Based on the human nutritional requirements, amino acids have been classified in several ways—essential or non-essential. Essential amino acids are indispensable because the human body cannot synthesize them; hence, appropriate concentrations in the diet are necessary (**Table 5**). Non-essential amino acids, synthesized in the human body, are also called dispensable amino acids (Reeds, 2000). However, some non-essential amino acids are considered conditionally non-essential because their abundance in the human body declines in times of stress or sickness. External sources are required to maintain necessary quantities (Fürst and Young, 2000).

The role of amino acids (individually or in combination) was first studied in rats to evaluate the necessity of Lys and Trp in food sources containing gliadin proteins. This initial study documented the adverse effects of amino acid deficiency on rats (Osborne and Mendel, 1914). Based on preliminary classical studies using model organisms (Ackroyd and Hopkins, 1916; Rose and Cox, 1924), an analogy of amino acid functions and dietary requirements in humans was first established by Rose and co-workers in 1947 (Rose et al., 1947). This study played a significant role in recognizing and classifying essential and non-essential amino acids based on their impacts on human health. Amino acids perform several crucial functions in the human body, either directly or indirectly. Amino acids have a specific role in gene expression (Oommen et al., 2005), signaling pathways for activation of immune systems (Kim et al., 2007), have nutraceutical effects for improving health status by regulating metabolic activities (Duranti, 2006), and can be used to treat genetic disorders (van Vliet et al., 2014).

Amino acids govern the epigenetic regulation of gene expression through DNA modifications. DNA modifications

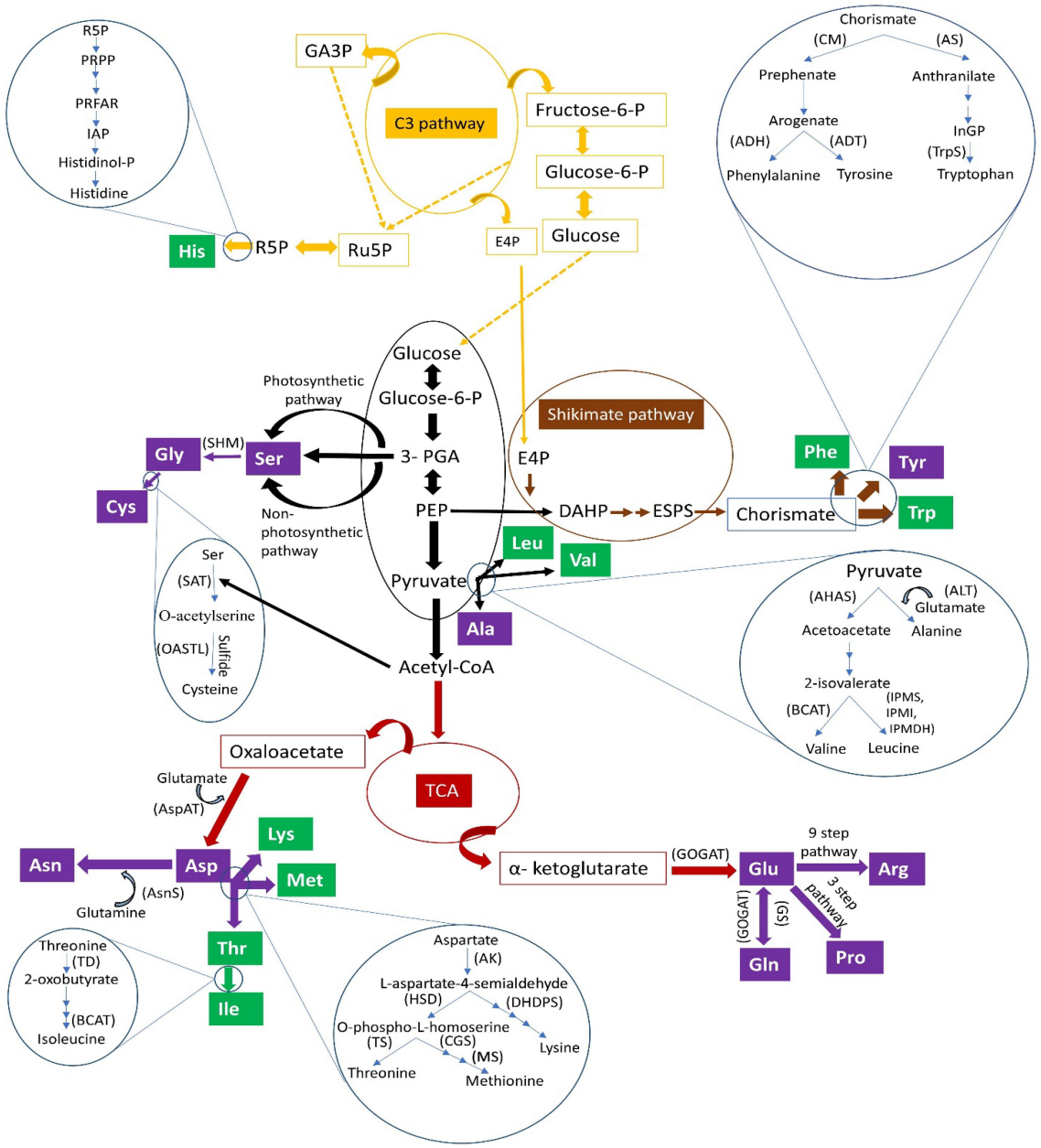


FIGURE 1 | Pathways synthesizing various essential (green boxes) and non-essential (purple boxes) amino acids. Amino acids: Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartate/aspartic acid; Cys, cysteine; Gln, glutamine; Glu, glutamate/glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine). Substrates/precursors: acetyl-CoA, acetyl-coenzyme A; DAHP, 3-deoxy-D-arabinoheptulosonate-7-phosphate; ESPS, 5-enolpyruvylshikimate-3-phosphate; E4P, erythrose 4-phosphate; fructose-6-P, fructose-6-phosphate; GA3P, glyceraldehyde 3-phosphate; glucose-6-P, glucose-6-phosphate; histidinol-P, histidinol phosphate; IAP, imidazole acetol-phosphate; PEP, phosphoenol pyruvate; 3-PGA, 3-phosphoglyceric acid; PRFAR, (*N*′-[(5′-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide) ribonucleotide); PRPP, phosphoribosyl diphosphate; R5P, ribose 5-phosphate; Ru5P, ribulose 5-phosphate. Enzymes indicated in parentheses: ADH, arogenate dehydrogenase; ADT, arogenate dehydratase; AHAS, acetohydroxyacid synthase; AK, aspartate kinase; ALT, alanine transferase; AS, anthranilate synthase; AsnS, asparagine synthetase; AspAT, aspartate aminotransferase; BCAT, branched-chain amino acid aminotransferase; CGS, cystathionine gamma synthase; CM, chorismate mutase; DHDPS, dihydrodipicolinate synthase; GOGAT, glutamate synthase; GS, glutamine synthetase; HSD, homoserine dehydrogenase; IPMS, isopropylmalate synthase; MS, methionine synthase; OASTL, O-acetylserine(thiol)lyase; SAT, serine acetyltransferase; SHM, serine hydroxymethyltransferase; TD, threonine deaminase; TrpS, tryptophan synthase; TS, threonine synthase.

such as methylation and acetylation occur due to the binding of DNA to C groups (methyl, acetyl) donated by Met, His, Ser, and Gly (Oommen et al., 2005; Kouzarides, 2007). Acetylation leads to the detachment of histones from DNA to favor its exposure-promoting transcription process. However, methylation plays a role in the reverse direction by densely packing the

TABLE 5 | Classification of amino acids based on human nutritional requirements.

Class of amino acid	Amino acids	Abbreviations
Essential		
	Histidine	His
	Isoleucine	Ile
	Leucine	Leu
	Lysine	Lys
	Methionine	Met
	Phenylalanine	Phe
	Threonine	Thr
	Tryptophan	Trp
	Valine	Val
Conditionally essential		
	Arginine	Arg
	Cysteine	Cys
	Glycine	Gly
	Glutamine	Gln
	Proline	Pro
	Tyrosine	Tyr
Non-essential		
	Alanine	Ala
	Asparagine	Asn
	Aspartate/aspartic acid	Asp
	Glutamate/glutamic acid	Glu
	Serine	Ser

DNA and encouraging gene silencing (Wu, 2010). Studies also demonstrate the role of Gln in the regulation of intestinal gene expression in rats, promoting intestinal health concerning cell growth and antioxidation activity (Wang et al., 2008). Arg supplementation in rats leads to the upregulation of gene expression, preventing oxidative stress and promoting fatty acid metabolism and glucose metabolism (McKnight et al., 2010). At the transcriptional level, amino acids regulate the activity of RNA polymerase by altering its specificity for promoters and enhancing the binding of some repressors near the non-coding sequences adjacent to the promoter region (Oommen et al., 2005). Such studies demonstrate the remarkable contribution of different amino acids in regulating gene expression.

The human immune system consists of both innate and acquired immune subsystems that regulate the response and protection of the human body upon pathogen attack (Calder, 1995). The innate immune system is a natural system that immediately activates when pathogens enter the body and can only prevent the entry and initial establishment of the pathogen. It comprises the physiological barriers, monocytes, macrophages, neutrophils, basophils, natural killer cells, mast cells, platelets, and various humoral factors (Buchanan et al., 2006). However, once the pathogen invades the innate immune system and colonizes, the acquired immune system is activated to decrease further pathogen progress. The acquired immune system consists of lymphocytes (T- and B-lymphocytes) that have immunological memory for invading pathogens (Calder, 2006). Human immune systems require a range of amino acids to produce immunoglobulins, cytokines, and other biomolecules to prevent diseases (Kim et al., 2007).

Several amino acids (branched-chain amino acids: BCAA (Leu, Ile, and Val), alanine (Ala), Gln, Ser, Pro, and Thr) regulate the proliferation of lymphocytes (Li et al., 2007). These amino acids either directly participate (Ala, Ser, and Thr) or produce signal molecules or hormones (BCAA, Gln, and Pro) to stimulate lymphocyte proliferation and create various immune responses (Li et al., 2007). Moreover, BCAAs participate in lipid metabolism (Nishimura et al., 2010) and blood glucose maintenance. In females, BCAAs also regulate blastocyst development and embryo implantation, fetal growth by hormonal secretions, stimulate mammary gland function and lactation, and increase aspartate, Gln, and glutamate synthesis (Zhang et al., 2018). Met, His, Gly, and Phe regulate the synthesis of signaling molecules controlling immune responses. Individually or in combination, these amino acids control the production of immune cell signaling molecules, leading to major immunity-boosting elements such as cytokines and antibodies (Li et al., 2007). Amino acid oxidases (AAOs) derived from L-isomers of Phe, Trp, Tyr, and Leu possess antimicrobial (Phua et al., 2012) and antitumoral functions (Lee et al., 2014).

Legumes have antinutritional compounds, including trypsin and chymotrypsin inhibitors, phytic acids, and tannins, which reduce nutrient bioavailability (Vidal-Valverde et al., 1994; Shi et al., 2017). Lentil is naturally low in phytic acid (Thavarajah et al., 2009) and contains trypsin inhibitors (3.6–7.6 units/mg protein) and tannins (1.28–3.9 mg/g; Hefnawy, 2011). Inactivity of trypsin and chymotrypsin enzymes causes difficulties in lysis proteins into small peptides and eventually affects the release of amino acids from small peptides. Tannins are phenolic inhibitors that bind to proteins *via* Lys or Met cross-links (Davis, 1981) and make insoluble complexes with carbohydrates (Reddy et al., 1985). In lentils, trypsin and chymotrypsin inhibitors and phytic acids are present in seed cotyledons, whereas tannins are concentrated mainly in the seed coat (Dueñas et al., 2002). Different food processing methods, including dehulling and cooking, are recommended to reduce these antinutritional properties (Acquah et al., 2021). Dehulling effectively reduces the tannins by removing the seed coat (Goyal et al., 2009). In pulses, other common processing treatments are soaking, hydrothermal treatments (cooking and roasting), fermentation, and irradiation (Acquah et al., 2021). Soaking reduces trypsin and chymotrypsin inhibitors, phytic acids, and tannins in lentils depending on the soaking time (Shi et al., 2017). Thermal methods are recommended for denaturing trypsin and chymotrypsin inhibitors and removing tannin in lentils (Hefnawy, 2011). Fermentation and irradiation are alternate methods to reduce antinutritional compounds (Siddhuraju et al., 2002; Maleki and Razavi, 2021) but have not been widely studied in pulses.

BREEDING APPROACHES FOR PROTEIN QUALITY IMPROVEMENT

Pulse breeding programs focus on meeting the world's food demand and ensuring global food security. The primary objectives of these breeding programs are to increase the yield by efficient selection from available germplasm, introduce hybrid lines, cross contrasting lines to exploit heterosis, develop biotic and

abiotic stress-tolerant cultivars, and induce mutations to generate novel variability with molecular and genomic techniques. Today, most conventional pulse breeding programs employ molecular markers for traits of interest. Genetic engineering technology has demonstrated remarkable potential to modify plants for specific breeding objectives. Thereby, technological advancement has broadened the scope of plant breeding to enable special-purpose breeding programs such as nutritional quality improvement programs or nutritional breeding (Kumar et al., 2020).

Conventional breeding approaches focus on improving highly heritable traits governed by a few genes. Quantitative traits with low heritability and high environmental effects, such as protein and other nutritional quality traits, do not significantly respond to selection by conventional breeding methods. In crop plants, including pulses, protein concentration negatively correlates with yield (Qureshi et al., 2013); therefore, selecting either trait negatively affects the other. For this reason, conventional approaches, such as mass selection, pedigree method, and bulk method, face challenges for protein quality improvement, but adding genetic markers into the breeding pipeline is possible. A comprehensive study comparing relative protein concentration among different lentil species identified a high protein accession, ILWL 47, belonging to *L. ervoides* (Bhatty, 1986). Lentil cultivar., IC317520, was identified as a high protein, sugar, and starch cultivar (Tripathi et al., 2019). The identified candidates can improve protein content in cultivated lentils by hybridization-based breeding methods.

Compared to selection and hybridization-based methods, mutation breeding has improved legume protein. A mutant lentil variety, NIA-MASOOR-5, with increased protein concentration, high yield, and disease resistance was created by gamma irradiation of M-85 as a parent and released in Pakistan (Ali and Shaikh, 2007). Mutation using gamma radiation has increased protein levels in mutants obtained from Chiang Mai 60, SSRSN35-19-4, and EHP 275 cultivars of soybean (Yathaputanon et al., 2009). Some high-protein and low-fiber mutants were identified from gamma ray-irradiated and ethyl methanesulfonate (EMS)-treated Himso 1563 and TS 82 cultivars in soybean (Kavithamani et al., 2010). EMS also induced beneficial mutations for protein and oil content improvement in Huayu 22 and Yueyou 45 cultivars of peanut (Chen et al., 2020). A high-yielding and high-protein chickpea mutant variety, Hyprosola or Faridpur-1, was also developed by gamma irradiation in Bangladesh (Oram et al., 1987). TAEK-SAGEL is another gamma radiation-derived, high-protein mutant variety of chickpea released in Turkey (Sağel et al., 2009). Such landmark achievements of mutation breeding in pulse crops, including lentils on a commercial scale, demonstrate the success of this method for improving quality traits.

Genomic-assisted breeding demonstrates the broad potential for improving quantitative traits, which are highly complex, controlled by many genes, and environmentally influenced (Kumar et al., 2016a). The current genomic toolbox for breeding includes genetic marker development, linkage map construction, identifying QTL and alien introgressions, candidate gene discovery, diversity analysis, genome sequencing, and pangenome construction. The use of molecular markers to gear up genomic developments in lentils for various traits has been reviewed widely (Kumar et al., 2015). Several legume crops, including dry pea (*Pisum sativum* L.), soybean, and chickpea, have been broadly investigated for use in genomic-assisted breeding to identify putative genomic regions governing seed protein concentration. The QTL mapping approach in dry pea revealed three genes regulating protein concentration using a linkage map of 207 markers (AFLP, RAPD, and STS markers; Tar'an et al., 2004). Another similar mapping study in dry pea using 204 markers (morphological, isozyme, AFLP, ISSR, STS, CAPS, and RAPD) identified genomic regions for seed protein concentration (Irzykowska and Wolko, 2004). Several other studies using genomic-assisted breeding in dry pea identified protein concentration-related genes (Tayeh et al., 2015). However, these studies are limited in the number of dry pea accessions used in each study and the genome-wide comparisons. Furthermore, a restriction-site associated DNA sequencing (RAD-seq) approach identified 47,472 SNP markers in a soybean RIL population (Liu et al., 2017), and several genes for the seed protein in soybean were found using transcriptome analysis, QTL mapping, and the genome-wide association study (GWAS) approach (Patil et al., 2017). A gene controlling seed size, weight, and composition of amino acids in total protein concentration were characterized in model legume Medicago trunculata and soybean using PCR-based markers and transcriptome profiling (Ge et al., 2016). Likewise, extensive studies in soybean have also identified several seed protein genes by exploiting genomic breeding approaches (Brummer et al., 1997; Sebolt et al., 2000; Chapman et al., 2003; Chung et al., 2003; Liang et al., 2010; Van and Mchale, 2017; Li et al., 2018; Huang et al., 2020). A high-throughput genotyping technology study identified 16,376 SNPs and revealed seven major genes for seed protein through a GWAS in 336 *desi* and *Kabuli* chickpea accessions (Upadhyaya et al., 2016). Such studies in legume crops demonstrate the success of marker-based genomic tools for improving protein concentration and quality. However, marker-based genomic-assisted studies identifying genic regions associated with seed protein content and quality have not been reported in lentils so far.

Genetic engineering technology has provided other insights to improve protein concentration in legumes. Protocols have been designed to develop transgenic lines in chickpea (Fontana et al., 1993), common bean (Russell et al., 1993), lupin (Molvig et al., 1997), peanuts (Brar et al., 1994), pea (Schroeder et al., 1993) and soybean (Hinchee et al., 1988). Several research groups have developed transgenic soybean lines with increased S-containing amino acids (Falco et al., 1995; Dinkins et al., 2001; Guo et al., 2020). Likewise, transformation studies to improve seed protein concentration in broad bean (Montamat et al., 1999), dry pea (Tegeder et al., 2007), and French bean (Tan et al., 2008) have also been reported. Recently, the genome-editing tool CRISPR/Cas 9 has emerged as a revolutionary approach to improving

staple food crops, but this approach is not widespread in pulses except in soybean.

CLOSING REMARKS

Most lentil breeding programs worldwide focus on yield improvement, disease resistance, biotic/abiotic stress tolerance, and germplasm diversity. Lentils are a nutrient-dense superfood to combat malnutrition and non-communicable diseases. As such, lentil protein quality has recently emerged as a target trait for lentil breeding programs due to the increased demand for plant-based protein. Conventional breeding is progressing for lentil crop nutritional improvement, but other genomic approaches are essential to speed up the breeding process due to the quantitative nature of these traits. Genome-wide association studies with conventional plant breeding approaches are appropriate for improving the genetic gain of quantitative traits by increasing selection accuracy through indirect selection (Rutkoski, 2019). For example, the genetic gain for lentil protein concentration can be achieved by selecting diverse parents, increasing the selection intensity, accuracy and reducing the selection cycle duration by increasing the number of generations per year. Conventional methods like pedigree, bulk, and mutation breeding can develop new breeding material using wild species, cultivars, landraces, advanced/elite breeding lines, and genetic stocks (**Figure 2**). These breeding methods will generate broadly diversified germplasm used for phenotyping and genotyping platforms to enhance selection accuracy (Xu et al., 2017). However, these conventional methods do not increase the selection intensity due to low heritability, slow progression, and visual phenotypic selection (Cobb et al., 2019). Combining genomic-assisted breeding with rapid generation methods such as single-seed descent, speed breeding, and double haploid production will enhance selection intensity and shorten the selection cycle, resulting in increased genetic gain over time (Cobb et al., 2019; **Figure 3**). Future lentil breeding efforts should focus on the rapid diversification and evaluation of lentil germplasm for protein quality through conventional breeding approaches. The development and adoption of genomic resources and tools such as genetic engineering or genome editing may also contribute to the pace of conventional breeding in lentils and eventually lead to breakthroughs in lentil protein improvement programs to ensure nutritional security and improve human health.

AUTHOR CONTRIBUTIONS

SS is a doctoral student under the supervision of DT who drafted the paper objectives, wrote the first draft, revised and edited the final version of this paper. JLB, PT, and SK edited/reviewed the final version and provided revisions and edits constructively. DT supervised SS and designed the objectives with SS, wrote parts of the paper, edited and revised the last version. All authors contributed to the article and approved the submitted version.

FUNDING

Funding support for this project was provided by the Organic Agriculture Research and Extension Initiative (OREI; award

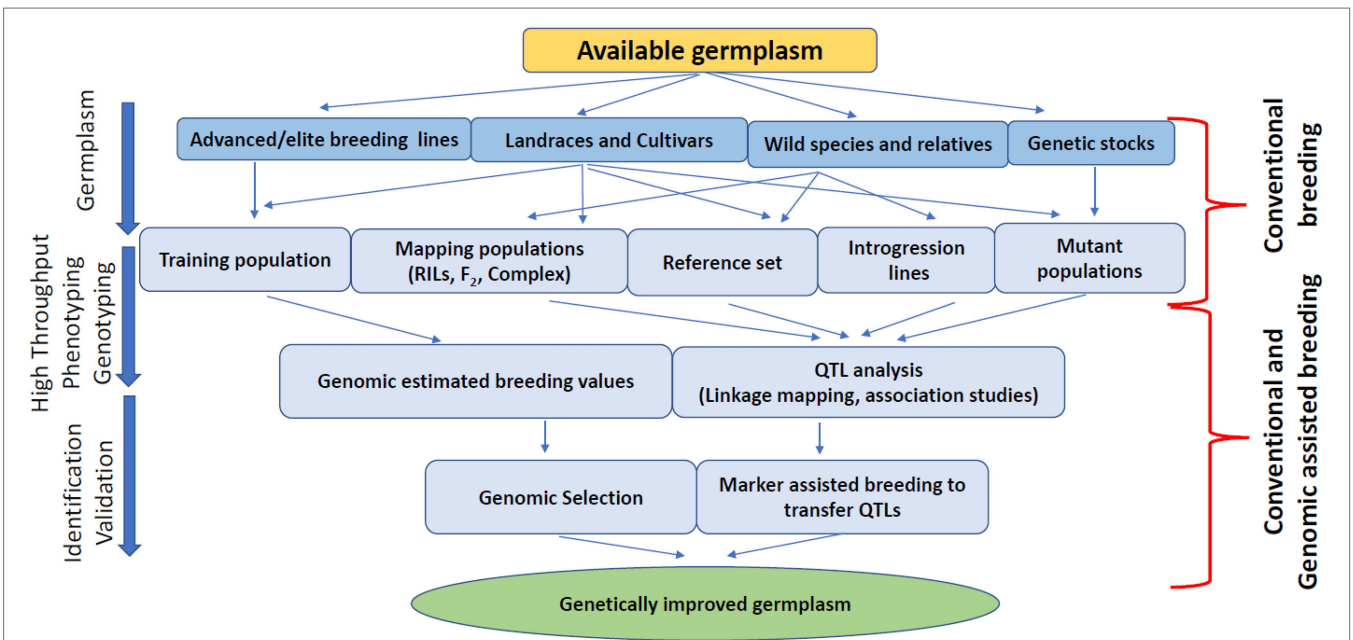


FIGURE 2 | Schematic representation of germplasm improvement for quality traits.

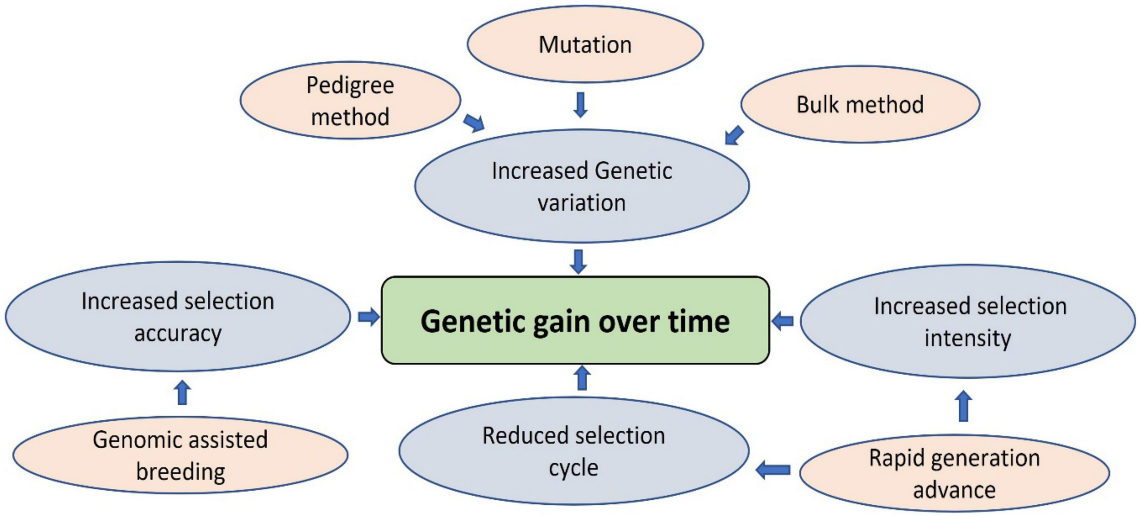


FIGURE 3 | Strategies to increase genetic gain over time.

no. 2018-51300-28431/proposal no. 2018-02799; and award no. 2021-51300-34805/proposal no. 2021-02927) of the United States Department of Agriculture, National Institute of Food and Agriculture (DT, LB), and the USDA National Institute of Food and Agriculture, [Hatch] project [1022664] (DT); the Good Food Institute (DT); and the FoodShot Global. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA.

REFERENCES

Ackroyd, H., and Hopkins, F. G. (1916). Feeding experiments with deficiencies in the amino-acid supply: arginine and histidine as possible precursors of purines. *Biochem. J.* 10, 551–576. doi: 10.1042/bj0100551

Acquah, C., Ohemeng-Boahen, G., Power, K. A., and Tosh, S. M. (2021). The effect of processing on bioactive compounds and nutritional qualities of pulses in meeting the sustainable development goal 2. *Front. Sustainable Food Syst.* 5:681662. doi: 10.3389/fsufs.2021.681662

Aldemir, S., Ateş, D., Temel, H. Y., Yağmur, B., Alsaleh, A., Kahriman, A., et al. (2017). QTLs for iron concentration in seeds of the cultivated lentil (Lens culinaris Medic.) via genotyping by sequencing. *Turk. J. Agric. For.* 41, 243–255.

Alghamdi, S. S., Khan, A. M., Ammar, M. H., El-Harty, E., Migdadi, H. M., El-Khalik, S., et al. (2014). Phenological, nutritional and molecular diversity assessment among 35 introduced lentil (Lens culinaris Medik.) genotypes grown in Saudi Arabia. *Int. J. Mol. Sci.* 15, 277–295. doi: 10.3390/ijms15010277

Ali, J. A. F., and Shaikh, N. A. (2007). Genetic exploitation of lentil through induced mutations. *Pak. J. Bot.* 39, 2379–2388.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408, 796–815. doi: 10.1038/35048692

Asif, M., Rooney, L. W., Ali, R., and Riaz, M. N. (2013). Application and opportunities of pulses in food system: a review. *Crit. Rev. Food Sci. Nutr.* 53, 1168–1179. doi: 10.1080/10408398.2011.574804

Ates, D., Sever, T., Aldemir, S., Yagmur, B., Temel, H. Y., Kaya, H. B., et al. (2016). Identification QTLs controlling genes for se uptake in lentil seeds. *PLoS One* 11:e0149210. doi: 10.1371/journal.pone.0149210

Bhatty, R. S. (1982). Albumin proteins of eight edible grain legume species. Electrophoretic patterns and amino acid composition. *J. Agric. Food Chem.* 30, 620–622. doi: 10.1021/jf00111a057

Bhatty, R. S. (1986). Protein subunits and amino acid composition of wild lentil. *Phytochemistry* 25, 641–644. doi: 10.1016/0031-9422(86)88015-3

Bhatty, R. S., and Slinkard, A. E. (1979). Composition, starch properties and protein quality of lentils. *Can. Inst. Food Sc. Technol. J.* 12, 88–92. doi: 10.1016/s0315-5463(79)73062-8

Borisjuk, N., Kishchenko, O., Eliby, S., Schramm, C., Anderson, P., Jatayev, S., et al. (2019). Genetic modification for wheat improvement: from transgenesis to genome editing. *Biomed. Res. Int.* 2019, 1–18. doi: 10.1155/2019/6216304

Branca, F., Lartey, A., Oenema, S., Aguayo, V., Stordalen, G. A., Richardson, R., et al. (2019). Transforming the food system to fight non-communicable diseases. *BMJ* 364:l296. doi: 10.1136/bmj.l296

Brar, G. S., Cohen, B. A., Vick, C. L., and Johnson, G. W. (1994). Recovery of transgenic peanut (Arachis hypogaea L.) plants from elite cultivars utilizing ACCELL® technology. *Plant J.* 55, 745–753. doi: 10.1111/j.1365-313X.1994.00745.x

Brummer, E. C., Graef, G. L., Orf, J., Wilcox, J. R., and Shoemaker, R. C. (1997). Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci.* 37, 370–378. doi: 10.2135/cropsci1997.0011183X003700020011x

Buchanan, J. T., Simpson, A. J., Aziz, R. K., Liu, G. Y., Kristian, S. A., Kotb, M., et al. (2006). DNase expression allows the pathogen group a streptococcus to escape killing in neutrophil extracellular traps. *Curr. Biol.* 16, 396–400. doi: 10.1016/j.cub.2005.12.039

Calder, P. C. (1995). Fuel utilization by cells of the immune system. *Proc. Nutr. Soc.* 54, 65–82. doi: 10.1079/pns19950038

Calder, P. C. (2006). Branched-chain amino acids and immunity. *J. Nutr.* 136, 288S–293S. doi: 10.1093/jn/136.1.288S

Chapman, A., Pantalone, V. R., Ustun, A., Allen, F. L., Landau-Ellis, D., Trigiano, R. N., et al. (2003). Quantitative trait loci for agronomic and seed quality traits in an F 2 and F 4:6 soybean population. *Euphytica* 129, 387–393. doi: 10.1023/A:1022282726117

Chen, T., Huang, L., Wang, M., Huang, Y., Zeng, R., Wang, X., et al. (2020). Ethyl methyl sulfonate-induced mutagenesis and its effects on peanut agronomic, yield and quality traits. *Agronomy* 10:655. doi: 10.3390/agronomy10050655

Chen, P., Shen, Z., Ming, L., Li, Y., Dan, W., Lou, G., et al. (2018). Genetic basis of variation in rice seed storage protein (albumin, globulin, Prolamin, and Glutelin) content revealed by genome-wide association analysis. *Front. Plant Sci.* 9:612. doi: 10.3389/fpls.2018.00612

Chung, H. L., Babka, G. L., Graef, P. E., Staswick, D. J., Lee, P. B., Cregan, R. C., et al. (2003). The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.* 43, 1053–1067. doi: 10.2135/cropsci2003.1053

Cobb, J. N., Juma, R. U., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., et al. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. *Theor. Appl. Genet.* 132, 627–645. doi: 10.1007/s00122-019-03317-0

Craciun, A., Jacobs, M., and Vauterin, M. (2000). Arabidopsis loss-of-function mutant in the lysine pathway points out complex regulation mechanisms. *FEBS Lett.* 487, 234–238. doi: 10.1016/s0014-5793(00)02303-6

Cubero, J. (1981). Origin, taxonomy and domestication. in *Lentils*, ed. Webb, H.G.C Slough, UK: Commonwealth Agricultural Bureau, 15–38.

Danielson, C. E. (1950). An electrophoretic investigation of vicilin and legumin from seeds of peas. *Acta Chem. Scand.* 4, 762–771. doi: 10.3891/acta.chem.scand.04-0762

Davis, K. R. (1981). Effect of processing on composition and Tetrahymena relative nutritive value on green and yellow peas, lentils and white pea beans. *Cereal Chem.* 58, 454–460.

Dinkins, R. D., Srinivasa Reddy, M. S., Meurer, C. A., Yan, B., Trick, H., Thibaud-Nissen, F., et al. (2001). Increased sulfur amino acids in soybean plants overexpressing the maize 15 kDa zein protein. *In Vitro Cell Dev. Biol. Plant* 37, 742–747. doi: 10.1007/s11627-001-0123-x

Dudareva, N., Maeda, H., and Yoo, H. (2011). Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. *Nat. Chem. Biol.* 7, 19–21. doi: 10.1038/nchembio.485

Dueñas, M., Hernández, T., and Estrella, I. (2002). Phenolic composition of the cotyledon and the seed coat of lentils (Lens culinaris L.). *Eur. Food Res. Technol.* 215, 478–483. doi: 10.1007/s00217-002-0603-1

Duranti, M. (2006). Grain legume proteins and nutraceutical properties. *Fitoterapia* 77, 67–82. doi: 10.1016/j.fitote.2005.11.008

Duranti, M., and Gius, C. (1997). Legume seeds: protein content and nutritional value. *Field Crop Res.* 53, 31–45. doi: 10.1016/s0378-4290(97)00021-x

Erskine, W., Williams, P. C., and Nakkoul, H. (1985). Genetic and environmental variation in the seed size, protein, yield, and cooking quality of lentils. *Field Crops Res.* 12, 153–161. doi: 10.1016/0378-4290(85)90061-9

Falco, S. C., Guida, T., Locke, M., Mauvais, J., Sanders, C., Ward, R. T., et al. (1995). Transgenic canola and soybean seeds with increased lysine. *Biotechnology* 13, 577–582. doi: 10.1038/nbt0695-577

FAOSTAT (2021). Available at: https://www.fao.org/faostat/en/#compare (Accessed November 20, 2021).

Ferguson, M. E., Maxted, N., Van Slageren, M., and Robertson, L. D. (2000). A re-assessment of the taxonomy of Lens Mill. (Leguminosae, Papilionoideae, Vicieae). *Bot. J. Linn Soc.* 133, 41–59. doi: 10.1006/bojl.1999.0319

Fontana, G. S., Santini, L., Caretto, S., Frugis, G., and Mariotti, D. (1993). Genetic transformation in the grain legume Cicer arietinum L. (chickpea). *Plant Cell Rep.* 12, 194–198. doi: 10.1007/BF00237052

Forde, B. G., and Lea, P. J. (2007). Glutamate in plants: metabolism, regulation, and signalling. *J. Exp. Bot.* 58, 2339–2358. doi: 10.1093/jxb/erm121

Fürst, P., and Young, V. R. (eds.) (2000). *Proteins, Peptides, and Amino Acids in Enteral Nutrition. Vol. 3*. Karger Medical and Scientific Publishers.

Galili, G. (2011). The aspartate-family pathway of plants. *Plant Signaling Behav.* 6, 192–195. doi: 10.4161/psb.6.2.14425

Gaufichon, L., Reisdorf-Cren, M., Rothstein, S. J., Chardon, F., and Suzuki, A. (2010). Biological functions of asparagine synthetase in plants. *Plant Sci.* 179, 141–153. doi: 10.1016/j.plantsci.2010.04.010

Gaufichon, L., Rothstein, S. J., and Suzuki, A. (2016). Asparagine metabolic pathways in Arabidopsis. *Plant Cell Physiol.* 57, 675–689. doi: 10.1093/pcp/pcv184

Gautam, N. K., Bhardwaj, R., Yadav, S., Suneja, P., Tripathi, K., and Ram, B. (2018). Identification of lentil (Lens culinaris Medik.) germplasm rich in protein and amino acids for utilization in crop improvement. *Ind. J. Genet.* 78:9. doi: 10.31742/IJGPB.78.4.9

Ge, L., Yu, J., Wang, H., Luth, D., Bai, G., Wang, K., et al. (2016). Increasing seed size and quality by manipulating BIG SEEDS1 in legume species. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12414–12419. doi: 10.1073/pnas.1611763113

Gibson, G. R., Hutkins, R., Sanders, M. E., Prescott, S. L., Reimer, R. A., Salminen, S. J., et al. (2017). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat. Rev. Gastroenterol. Hepatol.* 14, 491–502. doi: 10.1038/nrgastro.2017.75

Goyal, R. K., Vishwakarma, R. K., and Wanjari, O. D. (2009). Optimization of process parameters and mathematical modelling for dehulling of pigeonpea. *Int. J. Food Sci. Tech.* 44, 36–41. doi: 10.1111/j.1365-2621.2007.01630.x

Grusak, M. A. (2009). "Nutritional and health-beneficial quality" in *The Lentil: Botany, Production and Uses*. eds. W. Erskine, F. J. Muehlbauer, A. Sarker and B. Sharma (Cambridge, UK: CAB International), 368–390.

Guo, C., Liu, X., Chen, L., Cai, Y., Yao, W., Yuan, S., et al. (2020). Elevated methionine content in soybean seed by overexpressing maize β-zein protein. *Oil Crop Sci.* 5, 11–16. doi: 10.1016/j.ocsci.2020.03.004

Han, M., Zhang, C., Suglo, P., Sun, S., Wang, M., and Su, T. (2021). L-aspartate: an essential metabolite for plant growth and stress acclimation. *Molecules* 26:1887. doi: 10.3390/molecules26071887

Harvest Plus (2014). Biofortification progress briefs: iron and zinc lentils. *Brief* 9:19.

Hefnawy, T. H. (2011). Effect of processing methods on nutritional composition and antinutritional factors in lentils (Lens culinaris). *Ann. Agric. Sci.* 56, 57–61. doi: 10.1016/j.aoas.2011.07.001

Heuzé, V., Tran, G., Sauvant, D., Bastianelli, D., and Lebas, F. (2021). Lentil (Lens culinaris). Feedipedia, a programme by INRAE, CIRAD, AFZ and FAO. 18. Available at: https://www.feedipedia.org/node/284 (Accessed December 15, 2021).

Hinchee, M. A. W., Connor-Ward, D. V., Newell, C. A., McDonnell, R. E., Sato, S. J., Gasser, C. S., et al. (1988). Production of transgenic soybean plants using agrobacterium-mediated DNA transfer. *Bio/technology* 6, 915–922. doi: 10.1038/nbt0888-915

Howarth, J. R., Roberts, M. A., and Wray, J. L. (1997). Cysteine biosynthesis in higher plants: a new member of the Arabidopsis thaliana serine acetyltransferase small gene-family obtained by functional complementation of an *Escherichia coli* cysteine auxotroph. *Biochim. Biophys. Acta* 1350, 123–127. doi: 10.1016/s0167-4781(96)00213-8

Huang, J., Ma, Q., Cai, Z., Xia, Q., Li, S., Jia, J., et al. (2020). Identification and mapping of stable QTLs for seed oil and protein content in soybean [Glycine max (L.) Merr.]. *J. Agric. Food Chem.* 68, 6448–6460. doi: 10.1021/acs.jafc.0c01271

Ingle, R. A. (2011). Histidine biosynthesis. *The Arabidopsis Book* 2011. doi: 10.1199/tab.0141

Iqbal, A., Khalil, I. A., Ateeq, N., and Sayyar Khan, M. (2006). Nutritional quality of important food legumes. *Food Chem.* 97, 331–335. doi: 10.1016/j.foodchem.2005.05.011

Irzykowska, L., and Wolko, B. (2004). Interval mapping of QTLs controlling yield-related traits and seed protein content in Pisum sativum. *J. Appl. Genet.* 45, 297–306.

Jasinski, S., Lécureuil, A., Durandet, M., Bernard-Moulin, P., and Guerche, P. (2016). *Arabidopsis* seed content QTL mapping using high-throughput phenotyping: the assets of near infrared spectroscopy. *Front. Plant Sci.* 7:1682. doi: 10.3389/fpls.2016.01682

Kahraman, A. (2018). Sustainable food systems: relations amongst nutritional components in chickpea (Cicer arietinum L.). *Selcuk J. Agric. Food Sci.* 32, 458–461. doi: 10.15316/SJAFS.2018.123

Kavithamani, D., Kalamani, A., Vanniarajan, C., and Uma, D. (2010). Development of new vegetable soybean (*Glycine max* L. Merill) mutants with high protein and less fibre content. *Electron. J. Plant Breed.* 1, 1060–1065.

Khazaei, H., Subedi, M., Nickerson, M., Martínez-Villaluenga, C., Frias, J., and Vandenberg, A. (2019). Seed protein of lentils: current status, progress, and food applications. *Foods* 8:391. doi: 10.3390/foods8090391

Kim, S. W., Mateo, R. D., Yin, Y., and Wu, G. (2007). Functional amino acids and fatty acids for enhancing production performance of sows and piglets. *Asian-Australas. J. Anim. Sci.* 20, 295–306. doi: 10.5713/ajas.2007.295

Knill, T., Reichelt, M., Paetz, C., Gershenzon, J., and Binder, S. (2009). Arabidopsis thaliana encodes a bacterial-type heterodimeric isopropylmalate isomerase involved in both Leu biosynthesis and the met chain elongation pathway of glucosinolate formation. *Plant Mol. Biol.* 71, 227–239. doi: 10.1007/s11103-009-9519-5

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. doi: 10.1016/j.cell.2007.02.005

Kumar, S., Gupta, P., Choukri, H., and Siddique, K. H. M. (2020). "Efficient breeding of pulse crops" in *Accelerated Plant Breeding. Vol. 3*. eds. S. S. Gosal, and S. H. Wani (Cham: Springer International Publishing), 1–30.

Kumar, J., Gupta, D. S., Kumar, S., Gupta, S., and Singh, N. P. (2016a). Current knowledge on genetic biofortification in lentil. *J. Agric. Food Chem.* 64, 6383–6396. doi: 10.1021/acs.jafc.6b02171

Kumar, S., Rajendran, K., Kumar, J., Hamwieh, A., and Baum, M. (2015). Current knowledge in lentil genomics and its application for crop improvement. *Front. Plant Sci.* 6:78. doi: 10.3389/fpls.2015.00078

Kumar, H., Singh, A., Jain, N., Kumari, J., Singh, A. M., Singh, D., et al. (2014). Characterization of grain iron and zinc in lentil (Lens culinaris Medikus culinaris) and analysis of their genetic diversity using SSR markers. *Aust. J. Crop. Sci.* 8, 1005–1012.

Kumar, J., Singh, J., Kanaujia, R., and Gupta, S. (2016b). Protein content in wild and cultivated taxa of lentil (Lens culinaris ssp. culinaris Medikus). *Indian J. Genet. Plant Breed.* 76:631. doi: 10.5958/0975-6906.2016.00078.X

Lee, M. L., Fung, S., Chung, I., Pailoor, J., Cheah, S. H., and Tan, N. (2014). King cobra (Ophiophagus hannah) venom l-amino acid oxidase induces apoptosis in PC-3 cells and suppresses PC-3 solid tumor growth in a tumor xenograft mouse model. *Int. J. Med. Sci.* 11, 593–601. doi: 10.7150/ijms.8096

Less, H., and Galili, G. (2009). Coordinations between gene modules control the operation of plant amino acid metabolic networks. *BMC Syst. Biol.* 3:14. doi: 10.1186/1752-0509-3-14

Li, Y., Reif, J. C., Hong, H., Li, H., Liu, Z., Ma, Y., et al. (2018). Genome-wide association mapping of QTL underlying seed oil and protein contents of a diverse panel of soybean accessions. *Plant Sci.* 266, 95–101. doi: 10.1016/j.plantsci.2017.04.013

Li, P., Yin, Y., Li, D., Woo Kim, S., and Wu, G. (2007). Amino acids and immune function. *Br. J. Nutr.* 98, 237–252. doi: 10.1017/S000711450769936X

Liang, H., Yu, Y., Wang, S., Lian, Y., Wang, T., Wei, Y., et al. (2010). QTL mapping of Isoflavone, oil and protein contents in soybean (*Glycine max* L. Merr.). *Agric. Sci. China* 9, 1108–1116. doi: 10.1016/s1671-2927(09)60197-8

Liepman, A. H., and Olsen, L. J. (2004). Genomic analysis of aminotransferases in Arabidopsis thaliana. *Crit. Rev. Plant Sci.* 23, 73–89. doi: 10.1080/07352680490273419

Liu, N., Li, M., Hu, X., Ma, Q., Mu, Y., Tan, Z., et al. (2017). Construction of high-density genetic map and QTL mapping of yield-related and two quality traits in soybean RILs population by RAD-sequencing. *BMC Genomics* 18:466. doi: 10.1186/s12864-017-3854-8

Lowe, N. M. (2021). The global challenge of hidden hunger: perspectives from the field. *Proc. Nutr. Soc.* 80, 283–289. doi: 10.1017/S0029665121000902

Majumdar, R., Barchi, B., Turlapati, S. A., Gagne, M., Minocha, R., Long, S., et al. (2016). Glutamate, ornithine, arginine, proline, and polyamine metabolic interactions: the pathway is regulated at the post-transcriptional level. *Front. Plant Sci.* 7:78. doi: 10.3389/fpls.2016.00078

Maleki, S., and Razavi, S. H. (2021). Pulses' germination and fermentation: two bioprocessing against hypertension by releasing ACE inhibitory peptides. *Crit. Rev. Food Sci. Nutr.* 61, 2876–2893. doi: 10.1080/10408398.2020.1789551

Mann, G., Diffey, S., Cullis, B., Azanza, F., Martin, D., Kelly, A., et al. (2009). Genetic control of wheat quality: interactions between chromosomal regions determining protein content and composition, dough rheology, and sponge and dough baking properties. *Theor. Appl. Genet.* 118, 1519–1537. doi: 10.1007/s00122-009-1000-y

McKnight, J. R., Satterfield, M. C., Jobgen, W. S., Smith, S. B., Spencer, T. E., Meininger, C. J., et al. (2010). Beneficial effects of L-arginine on reducing obesity: potential mechanisms and important implications for human health. *Amino Acids* 39, 349–357. doi: 10.1007/s00726-010-0598-z

Molvig, L., Tabe, L. M., Eggum, B. O., Moore, A. E., Craig, S., Spencer, D., et al. (1997). Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (Lupinus angustifolius L.) expressing a sunflower seed albumin gene. *Proc. Natl. Acad. Sci. U. S. A* 94, 8393–839883.

Montamat, F., Maurousset, L., Tegeder, M., Frommer, W., and Delrot, S. (1999). Cloning and expression of amino acid transporters from broad bean. *Plant Mol. Biol.* 41, 259–268. doi: 10.1023/a:1006321221368

Nishimura, J., Masaki, T., Arakawa, M., Seike, M., and Yoshimatsu, H. (2010). Isoleucine prevents the accumulation of tissue triglycerides and upregulates the expression of PPARalpha and uncoupling protein in diet-induced obese mice. *J. Nutr.* 140, 496–500. doi: 10.3945/jn.109.108977

Olsen, M. S., and Phillips, R. L. (2009). "Molecular genetic improvement of protein quality in maize," in *Impacts of Agriculture on Human Health And Nutrition*. eds. I. Cakmak and R. M. Welch (EOLSS Online Publications), 1–20.

Oommen, A. M., Griffin, J. B., Sarath, G., and Zempleni, J. (2005). Roles for nutrients in epigenetic events. *J. Nutr. Biochem.* 16, 74–77. doi: 10.1016/j.jnutbio.2004.08.004

Oram, R. N., Shaikh, M. A. Q., Zaman, K. M. S., and Brown, A. H. D. (1987). Isozyme similarity and genetic differences in morphology between hyprosola, a high yielding, high protein mutant of chickpea (Cicer arietinum L.) and its parental cultivar. *Environ. Exp. Bot.* 27, 455–462. doi: 10.1016/0098-8472(87)90026-8

Osborne, T. B. (1924). *The Vegetable Proteins*. London: Longmans, Green and Co.

Osborne, T. B., and Campbell, G. F. (1898). PROTEIDS of the LENTIL.1. *J. Am. Chem. Soc.* 20, 362–375. doi: 10.1021/ja02067a007

Osborne, T. B., and Mendel, L. B. (1914). Amino-acids in nutrition and growth. *J. Biol. Chem.* 12, 484–485. doi: 10.1080/07315724.1993.10718340

Parthasarathy, A., Cross, P. J., Dobson, R. C. J., Adams, L. E., Savka, M. A., and Hudson, A. O. (2018). A three-ring circus: metabolism of the three proteogenic aromatic amino acids and their role in the health of plants and animals. *Front. Mol. Biosci.* 5:29. doi: 10.3389/fmolb.2018.00029

Parthasarathy, A., Savka, M. A., and Hudson, A. O. (2019). The synthesis and role of β-alanine in plants. *Front. Plant Sci.* 10:921. doi: 10.3389/fpls.2019.00921

Patil, G., Mian, R., Vuong, T., Pantalone, V., Song, Q., Chen, P., et al. (2017). Molecular mapping and genomics of soybean seed protein: a review and perspective for the future. *Theor. Appl. Genet.* 130, 1975–1991. doi: 10.1007/s00122-017-2955-8

Phua, C. S., Vejayan, J., Ambu, S., Ponnudurai, G., and Gorajana, A. (2012). Purification and antibacterial activities of an l-venom amino acid oxidase from king cobra (Ophiophagus hannah). *J. Venomous Anim. Toxins Incl. Trop. Dis.* 18, 198–207. doi: 10.1590/S1678-91992012000200010

Podder, R., Glahn, R. P., and Vandenberg, A. (2020). Iron- and zinc-fortified lentil (lens culinaris medik.) demonstrate enhanced and stable iron bioavailability after storage. *Front. Nutr.* 7:614812. doi: 10.3389/fnut.2020.614812

Qureshi, A., Wani, S., Lone, A., Dar, Z., Wani, S., and Nehvi, F. (2013). "Breeding for quality traits in grain legumes," in *Conventional and non-conventional interventions in crop improvement. 1st Edn*. eds. C. Malik, G. Sanghera and S. Wani (New Delhi: M D Publishers).

Rasheed, N., Maqsood, M. A., Aziz, T., and Jabbar, A. (2020). Characterizing lentil Germplasm for zinc biofortification and high grain output. *J. Soil Sci. Plant Nutr.* 20, 1336–1349. doi: 10.1007/s42729-020-00216-y

Reddy, N. R., Pierson, M. D., Sathe, S. K., and Salunkhe, D. K. (1985). Dry bean tannins: a review of nutritional implications. *J. Am. Oil Chem. Soc.* 62, 541–549. doi: 10.1007/BF02542329

Reeds, P. J. (2000). Dispensable and indispensable amino acids for humans. *J. Nutr.* 130, 1835S–1840S. doi: 10.1093/jn/130.7.1835S

Rees, J. D., Ingle, R. A., and Smith, J. A. C. (2009). Relative contributions of nine genes in the pathway of histidine biosynthesis to the control of free histidine concentrations inArabidopsis thaliana. *Plant Biotechnol. J.* 7, 499–511. doi: 10.1111/j.1467-7652.2009.00419.x

Rose, W. C., and Cox, G. J. (1924). The relation of arginine and histidine to growth. *J. Biol. Chem.* 61, 747–773. doi: 10.1016/s0021-9258(18)85123-7

Rose, W. C., Haines, W. J., and Johnson, J. E. (1947). The rôle of the amino acids in human nutrition. *J. Biol. Chem.* 146, 683–684. doi: 10.1016/s0021-9258(18)44994-0

Rozan, P., Kuo, Y., and Lambein, F. (2001). Amino acids in seeds and seedlings of the genus lens. *Phytochemistry* 58, 281–289. doi: 10.1016/s0031-9422(01)00200-x

Russell, D. R., Wallace, K. M., Bathe, J. H., Martinell, B. J., and McCabe, D. E. (1993). Stable transformation of Phaseolus vulgaris via electric-discharge mediated particle acceleration. *Plant Cell Rep.* 12, 165–169. doi: 10.1007/BF00239099

Rutkoski, J. E. (2019). A practical guide to genetic gain. *Adv. Agron.* 157, 217–249. doi: 10.1016/bs.agron.2019.05.001

Sağel, Z., Tutluer, M. I., Peşkircioğlu, H., Kantoğlu, K. Y., and Kunter, B. (2009). "The improvement of TAEK-sagel chickpea (Cicer arietinum L.) mutant variety in Turkey" in *Induced Plant Mutations in the Genomics Era*. ed. Q. Y. Shu (Rome: FAO), 319–321.

Saint-Clair, P. M. (1972). *Responses of lens Esculenta Moench to controlled environmental factors*. Wangeningen, Netherlands: Wangeningen University

Sales, M. P., Gerhardt, I. R., Grossi-de-Sá, M. F., and Xavier-Filho, J. (2000). Do legume storage proteins play a role in defending seeds against Bruchids? *Plant Physiol.* 124, 515–522. doi: 10.1104/pp.124.2.515

Sarrobert, C., Thibaud, M. C., Contard-David, P., Gineste, S., Bechtold, N., Robaglia, C., et al. (2000). Identification of an Arabidopsis thaliana mutant accumulating threonine resulting from mutation in a new dihydrodipicolinate synthase gene. *Plant J.* 24, 357–368. doi: 10.1046/j.1365-313x.2000.00884.x

Sayeed, S., and Njaa, L. R. (1985). Effect of a Bangladeshi home-cooking procedure on the amino acid content, trypsin inhibitor activity and in vitro digestibility of some legume seeds. *Qualitas plantarum. Plant Foods Human Nutr.* 354, 35, 379–388. doi: 10.1007/BF01091783

Schroeder, H. E., Schotz, A. H., Wardley-Richardson, T., Spencer, D., and Higgins, T. (1993). Transformation and regeneration of two cultivars of pea (Pisum sativum L.). *Plant Physiol.* 101, 751–757. 10.1104/pp.101.3.751

Scippa, G. S., Rocco, M., Ialicicco, M., Trupiano, D., Viscosi, V., Di Michele, M., et al. (2010). The proteome of lentil (Lens culinaris Medik.) seeds: discriminating between landraces. *Electrophoresis* 31, 497–506. doi: 10.1002/elps.200900459

Sebolt, A. M., Shoemaker, R. I., and Diers, R. W. (2000). Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Sci.* 40, 1438–1444. doi: 10.2135/cropsci2000.4051438x

Semba, R. D. (2016). The rise and fall of protein malnutrition in global health. *Ann. Nutr. Metab.* 69, 79–88. doi: 10.1159/000449175

Sen Gupta, D., Thavarajah, D., Knutson, P., Thavarajah, P., McGee, R. J., Coyne, C. J., et al. (2013). Lentils (Lens culinaris L.), a rich source of Folates. *J. Agric.Food Chem.* 61, 7794–7799. doi: 10.1021/jf401891p

Shekib, L. A. H., Zoueil, M. E., Youssef, M. M., and Mohamed, M. S. (1986). Amino acid composition and In vitro digestibility of lentil and rice proteins and their mixture (Koshary). *Food Chem.* 20, 61–67. doi: 10.1016/0308-8146(86)90167-6

Shi, L., Mu, K., Arntfield, S. D., and Nickerson, M. T. (2017). Changes in levels of enzyme inhibitors during soaking and cooking for pulses available in Canada. *J. Food Sci. Technol.* 54, 1014–1022. doi: 10.1007/s13197-017-2519-6

Siddhuraju, P., Osoniyi, O., Makkar, H. P. S., and Becker, K. (2002). Effect of soaking and ionising radiation on various antinutritional factors of seeds from different species of an unconventional legume, Sesbania and a common legume, green gram (Vigna radiata). *Food Chem.* 79, 273–281. doi: 10.1016/S0308-8146(02)00140-1

Singh, B. K., and Shaner, D. L. (1995). Biosynthesis of branched chain amino acids: from test tube to field. *Plant Cell* 7, 935–944. doi: 10.1105/tpc.7.7.935

Stoddard, F. L., Marshall, D. R., and Ali, S. M. (1993). Variability in grain protein concentration of peas and lentils grown in Australia. *Aust. J. Agric. Res.* 44:1415. doi: 10.1071/AR9931415

Sulieman, M. A., Amro, B. H., Gamaa, A. O., Mohamed, M. E. T., Elhadi, A. I. E. K., Abdullahi, H. E. T., et al. (2008). Changes in total protein digestibility, fractions content and structure during cooking of lentil cultivars. *Pak. J. Nutr.* 7, 801–805. doi: 10.3923/pjn.2008.801.805

Tahir, M., Lindeboom, N., Båga, M., Vandenberg, A., and Chibbar, R. (2011). Composition and correlation between major seed constituents in selected lentil (Lens culinaris. Medik) genotypes. *Can. J. Plant Sci.* 91, 825–835. doi: 10.4141/cjps2011-010

Tan, Q., Grennan, A. K., Pélissier, H. C., Rentsch, D., and Tegeder, M. (2008). Characterization and expression of French bean amino acid transporter PvAAP1. *Plant Sci.* 174, 348–356. doi: 10.1016/j.plantsci.2007.12.008

Tar'an, B., Warkentin, T., Somers, D. J., Miranda, D., Vandenberg, A., Blade, S., et al. (2004). Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (Pisum sativum L.). *Euphytica* 136, 297–306. doi: 10.1023/B:EUPH.0000032721.03075.a0

Tayeh, N., Aubert, G., Pilet-Nayel, M., Lejeune-Hénaut, I., Warkentin, T. D., and Burstin, J. (2015). Genomic tools in pea breeding programs: status and perspectives. *Front. Plant Sci.* 6:1037. doi: 10.3389/fpls.2015.01037

Tegeder, M., Tan, Q., Grennan, A. K., and Patrick, J. W. (2007). Amino acid transporter expression and localisation studies in pea (Pisum sativum). *Funct. Plant Biol.* 34, 1019–1028. doi: 10.1071/FP07107

Thavarajah, D., Thavarajah, P., Sarker, A., Materne, M., Vandemark, G., Shrestha, R., et al. (2011). A global survey of effects of genotype and environment on selenium concentration in lentils (Lens culinaris L.): implications for nutritional fortification strategies. *Food Chem.* 125, 72–76. doi: 10.1016/j.foodchem.2010.08.038

Thavarajah, P., Thavarajah, D., and Vandenberg, A. (2009). Low phytic acid lentils (Lens culinaris L.): a potential solution for increased micronutrient bioavailability. *J. Agric. Food Chem.* 57, 9044–9049. doi: 10.1021/jf901636p

Tripathi, K., Gore, P. G., Pandey, A., Bhardwaj, R., Singh, N., Chawla, G., et al. (2019). Seed morphology, quality traits and imbibition behaviour study of atypical lentil (Lens culinaris Medik.) from Rajasthan, India. *Genet. Resour. Crop. Evol.* 66, 697–706. doi: 10.1007/s10722-019-00745-1

Tzin, V., and Galili, G. (2010). The biosynthetic pathways for Shikimate and aromatic amino acids in Arabidopsis thaliana. *Arabidopsis Book* 8, e0132. doi: 10.1199/tab.0132

Upadhyaya, H. D., Bajaj, D., Narnoliya, L., Das, S., Kumar, V., Gowda, C. L. L., et al. (2016). Genome-wide scans for delineation of candidate genes regulating seed-protein content in chickpea. *Front. Plant Sci.* 7:302. doi: 10.3389/fpls.2016.00302

Van, K., and Mchale, L. K. (2017). Meta-analyses of QTLs associated with protein and oil contents and compositions in soybean [Glycine max (L.) Merr.] seed. *IJMS* 18:1180. doi: 10.3390/ijms18061180

van Vliet, D., Derks, T. G. J., van Rijn, M., de Groot, M. J., MacDonald, A., Heiner-Fokkema, M. R., et al. (2014). Single amino acid supplementation in aminoacidopathies: a systematic review. *Orphanet J. Rare Dis.* 9:7. doi: 10.1186/1750-1172-9-7

Vauterin, M., Frankard, V., and Jacobs, M. (1999). The Arabidopsis thaliana dhdps gene encoding dihydrodipicolinate synthase, key enzyme of lysine biosynthesis, is expressed in a cell-specific manner. *Plant Mol. Biol.* 39, 695–708. doi: 10.1023/a:1006132428623

Vauterin, M., and Jacobs, M. (1994). Isolation of a poplar and an Arabidopsis thaliana dihydrodipicolinate synthase cDNA clone. *Plant Mol. Biol.* 25, 545–550. doi: 10.1007/bf00043882

Vidal-Valverde, C., Frias, J., Estrella, I., Gorospe, M. J., Ruiz, R., and Bacon, J. (1994). Effect of processing on some antinutritional factors of lentils. *J. Agric. Food Chem.* 42, 2291–2295. doi: 10.1021/jf00046a039

Wang, J., Chen, L., Li, P., Li, X., Zhou, H., Wang, F., et al. (2008). Gene expression is altered in piglet small intestine by weaning and dietary glutamine supplementation. *J. Nutr.* 138, 1025–1032. doi: 10.1093/jn/138.6.1025

Wang, N., and Daun, J. K. (2006). Effects of variety and crude protein content on nutrients and antinutrients in lentils (Lens culinaris). *Food Chem.* 95, 493–502. doi: 10.1016/j.foodchem.2005.02.001

Wirtz, M., Droux, M., and Hell, R. (2004). O-acetylserine (thiol) lyase: an enigmatic enzyme of plant cysteine biosynthesis revisited in Arabidopsis thaliana. *J. Exp. Bot.* 55, 1785–1798. doi: 10.1093/jxb/erh201

Wong, M. M. L., Gujaria-Verma, N., Ramsay, L., Yuan, H. Y., Caron, C., Diapari, M., et al. (2015). Classification and characterization of species within the genus lens using genotyping-by-sequencing (GBS). *PLoS One* 10:e0122025. doi: 10.1371/journal.pone.0122025

World Health Organization. (2019). Non communicable diseases. Available at: https://www.who.int/gho/ncd/mortality_morbidity/en/ (Accessed December 05, 2021).

Wu, G. (2010). Functional amino acids in growth, reproduction, and health. *Adv. Nutr.* 1, 31–37. doi: 10.3945/an.110.1008

Xing, A., and Last, R. L. (2017). A regulatory hierarchy of the Arabidopsis branched-chain amino acid metabolic network. *Plant Cell* 29, 1480–1499. doi: 10.1105/tpc.17.00186

Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., et al. (2017). Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* 68, 2641–2666. doi: 10.1093/jxb/erx135

Yathaputanon, C., Bunyarut, J., Kumsueb, B., Malipan, A., and Srisombun, S.. (2009). *Protein Content in High-Protein Soybean Mutants in Thailand. Food and Agriculture Organization of the United Nations (FAO)*: Rome: FAO.

Zaccardelli, M., Lupo, F., Piergiovanni, A. R., Laghetti, G., Sonnante, G., Daminati, M. G., et al. (2012). Characterization of Italian lentil (Lens culinaris Medik.) germplasm by agronomic traits, biochemical and molecular markers genet. *Resour. Crop Evol.* 59, 727–738. doi: 10.1007/s10722-011-9714-5

Zhang, X., Wang, W., Guo, N., Zhang, Y., Bu, Y., Zhao, J., et al. (2018). Combining QTL-seq and linkage mapping to fine map a wild soybean allele characteristic of greater plant height. *BMC Genomics* 19, 226. doi: 10.1186/s12864-018-4582-4

Zia-Ul-Haq, M., Cavar, S., Qayum, M., Imran, I., and de Feo, V. (2011). Compositional studies: antioxidant and Antidiabetic activities of Capparis decidua (Forsk.) Edgew. *Int. J. Mol. Sci.* 12, 8846–8861. doi: 10.3390/ijms12128846

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Salaria, Boatwright, Thavarajah, Kumar and Thavarajah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fpls.2022.830896

Improving Association Studies and Genomic Predictions for Climbing Beans With Data From Bush Bean Populations

Beat Keller[1†], Daniel Ariza-Suarez[1,2], Ana Elisabeth Portilla-Benavides[2], Hector Fabio Buendia[2], Johan Steven Aparicio[2], Winnyfred Amongi[3], Julius Mbiu[4], Susan Nchimbi Msolla[5], Phillip Miklas[6], Timothy G. Porch[7], James Burridge[8], Clare Mukankusi[3], Bruno Studer[1] and Bodo Raatz[2†]*

[1] Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland, [2] Bean Program, International Center for Tropical Agriculture (CIAT), Cali, Colombia, [3] Bean Program, International Center for Tropical Agriculture (CIAT), Kampala, Uganda, [4] Tanzania Agricultural Research Institute (TARI), Dodoma, Tanzania, [5] Department of Crop Science and Horticulture, Sokoine University of Agriculture, Morogoro, Tanzania, [6] Department of Agriculture, Agriculture Research Service (USDA-ARS), Prosser, WA, United States, [7] Department of Agriculture, Agriculture Research Service (USDA-ARS), Tropical Agriculture Research Station, Mayaguez, PR, United States, [8] Department of Plant Science, The Pennsylvania State University, University Park, PA, United States

Common bean (*Phaseolus vulgaris* L.) has two major origins of domestication, Andean and Mesoamerican, which contribute to the high diversity of growth type, pod and seed characteristics. The climbing growth habit is associated with increased days to flowering (DF), seed iron concentration (SdFe), nitrogen fixation, and yield. However, breeding efforts in climbing beans have been limited and independent from bush type beans. To advance climbing bean breeding, we carried out genome-wide association studies and genomic predictions using 1,869 common bean lines belonging to five breeding panels representing both gene pools and all growth types. The phenotypic data were collected from 17 field trials and were complemented with 16 previously published trials. Overall, 38 significant marker-trait associations were identified for growth habit, 14 for DF, 13 for 100 seed weight, three for SdFe, and one for yield. Except for DF, the results suggest a common genetic basis for traits across all panels and growth types. Seven QTL associated with growth habits were confirmed from earlier studies and four plausible candidate genes for SdFe and 100 seed weight were newly identified. Furthermore, the genomic prediction accuracy for SdFe and yield in climbing beans improved up to 8.8% when bush-type bean lines were included in the training population. In conclusion, a large population from different gene pools and growth types across multiple breeding panels increased the power of genomic analyses and provides a solid and diverse germplasm base for genetic improvement of common bean.

Keywords: genome-wide association studies (GWAS), genomic selection, population structure, pleiotropy, growth habit, common bean (*Phaseolus vulgaris* L.), climbing and bush type bean

OPEN ACCESS

Edited by:
Eric Von Wettberg, University of Vermont, United States

Reviewed by:
Anupam Singh, University of Southern California, United States
Lorenzo Raggi, University of Perugia, Italy

***Correspondence:**
Bruno Studer
bruno.studer@usys.ethz.ch

†Present addresses:
Beat Keller, Crop Science Group, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland
Bodo Raatz, Limagrain Vegetable Seed, La Menitŕe, France

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *07 December 2021*
Accepted: *25 February 2022*
Published: *25 April 2022*

Citation:
Keller B, Ariza-Suarez D, Portilla-Benavides AE, Buendia HF, Aparicio JS, Amongi W, Mbiu J, Msolla SN, Miklas P, Porch TG, Burridge J, Mukankusi C, Studer B and Raatz B (2022) Improving Association Studies and Genomic Predictions for Climbing Beans With Data From Bush Bean Populations. Front. Plant Sci. 13:830896. doi: 10.3389/fpls.2022.830896

1. INTRODUCTION

Common bean (*Phaseolus vulgaris* L.) was domesticated about 8,000 years ago in two geographic regions, resulting in the Andean and the Mesoamerican gene pools (Gepts et al., 1986; Bitocchi et al., 2013). Within the two gene pools, several groups, including climbing and bush type beans, were identified through genetic or phenotypic characterization (Gepts et al., 1986; Rodriguez et al., 2016). Seven domestication events for the common bean were discovered by investigating a genetic locus for flowering determinacy (Kwak et al., 2008, 2012). Flowering determinacy defines the first criterion for growth type: the determinate growth type forms a reproductive terminal bud, whereas the indeterminate growth types produce a vegetative one (Singh, 1981). The second criterion describes the bush vs. climbing growth habit. Both criteria were used to distinguish four growth types: type I (determinate bush), type II (indeterminate bush), type III (indeterminate semi-climber), and type IV (indeterminate climber) (Singh, 1981). Since growth type is associated with flowering determinacy, it also affects vegetative growth and the length of the crop cycle (González et al., 2016). In recent decades, major breeding efforts have been directed toward the erect growth habit of bush type beans since this habit enables a faster cultivation cycle without staking of the plants and a single, automated harvest (Teixeira et al., 1999; Ronner et al., 2018). The joint diversity of common bean growth types may offer new insights to improve not only climbing but also bush type beans.

Although largely neglected in breeding programs, climbing beans offer three main advantages over bush type beans: first, climbing beans reach higher yield per area, with up to 5 t ha^{-1} (Rosales-Serna et al., 2004; Checa et al., 2006; Barbosa et al., 2018). Second, they have a higher symbiotic nitrogen fixation capacity, with up to 92 kg of N fixed ha^{-1} (Graham, 1981; Bliss, 1993; Checa et al., 2006; Barbosa et al., 2018). Third, they achieve higher seed iron content (SdFe), with up to 10 mg/100 g (Blair et al., 2010; Blair, 2013; Petry et al., 2015; Mukamuhirwa and Rurangwa, 2018). Indeed, the production of climbing beans can be more profitable, and they are preferentially adopted in higher and/or drought prone regions by small holder farmers in Uganda and Rwanda (Ronner et al., 2018; Katungi et al., 2019). However, these advantages come with the cost of staking the plants and a longer vegetative period, partly due to the indeterminate growth type (White et al., 1992).

To shorten the cultivation cycle, Kornegay et al. (1992) suggested crossing type I and II bean lines with type IV lines to achieve an increase in yield while selecting against climbing ability. However, the climbing growth habit is tightly linked to plant development and productivity. The relation of days to flowering (DF), vegetative growth, and plant production was investigated in two Andean (type I) x Mesoamerican (type IV) recombinant inbred line populations (González et al., 2016). In those mixed climbing and bush type populations, the QTL containing the *PvTFL1y* flowering gene explained 32% of the variation for DF, 66% for vegetative growth (length of the main stem), and 19% for the rate of plant production, including traits such as yield and seed weight (González et al., 2016). In general, more days to physiological maturity (DPM) resulted in an increased yield of lines among and across the different growth types (White and Izquierdo, 1989; Keller et al., 2020). However, within growth type I and type II, DPM was not related to yield in near-isogenic lines (White et al., 1992). This suggests that the relationship between yield, DF, and DPM can be partially uncoupled.

The higher SdFe in climbing beans is promising to combat iron deficiency in human nutrition, which causes anemia, increases morbidity, and leads to economic losses (Boccio and Iyengar, 2003). About 30% of the global population suffers from anemia, especially women and children in developing countries (Black et al., 2008; Stein, 2010). Increasing SdFe in legumes is a possible avenue to improve nutritional quality in the human diet (Petry et al., 2015; Rehman et al., 2019). In the last years, a few biofortified lines with higher SdFe were successfully released (HarvestPlus, 2022). Iron biofortified beans showed higher phytic acid concentrations, which decreased the relative but not absolute iron absorption (Petry et al., 2014). However, the SdFe of climbing beans has not been investigated intensively and SdFe is negatively correlated with yield (Kelly and Bornowski, 2018). Such tradeoffs, including the longer DPM, which is associated with the climbing habit, need to be taken into account when improving climbing beans.

Efficient breeding for multiple traits requires detailed knowledge about the underlying genetic architecture of the target traits and their correlations with other key characteristics. The genetic architecture of traits can be investigated by genome-wide association studies (GWAS) and genomic prediction models (Crossa et al., 2017; Cortes et al., 2021). For common bean, GWAS have been carried out successfully and QTL for various traits were tagged with molecular markers (Miklas et al., 2006; Wu et al., 2020). Recently, genomic predictions were evaluated for agronomic traits in an elite Andean breeding panel (Keller et al., 2020). This study revealed the prediction abilities (PAs) for the genomic estimated breeding values (GEBV) based on genomic data only, as proposed by Meuwissen et al. (2001). This allows breeders to efficiently select superior lines before they enter field trials (Crossa et al., 2017). In general, PA is increased for more heritable traits and in lines, which are closely related to the training population (TP). In order to deal with different populations or breeding panels, multivariate models were suggested to account for population structure (Lehermeier et al., 2015). Following another approach, the TP can be optimized for genetic relatedness to improve the accuracy of the GEBV (Akdemir and Isidro-Sánchez, 2019; Sarinelli et al., 2019). These approaches have great potential to improve PAs, given the increasing availability of genotypic and phenotypic data (Spindel and McCouch, 2016). The efficient use of prediction models based on appropriate TPs and available data sets is a key factor in the development of new and adapted climbing bean cultivars.

In this study, comprehensive genetic analyses across all bean growth types and the two gene pools were carried out using 1,869 lines belonging to five breeding panels. We hypothesized that (*i*) climbing and bush type beans show, despite specific growth habit loci, overall a high genetic similarity, that (*ii*) combined analysis will increase the power of GWAS as well as the genomic predictions across all growth types and that (*iii*)

predictions for climbing beans with an optimized TP including bush beans will outperform predictions which were based on the climbing bean growth types only. Hence, we evaluated whether breeding programs can supplement their trial data with already existing data to improve GWAS and genomics predictions. The overall objective was to provide molecular markers and genomic prediction models which can be used to speed up the selection of new bean lines across all growth types and gene pools.

2. MATERIALS AND METHODS

2.1. Germplasm

The germplasm used in this study consisted of five bean breeding panels across all growth types: the newly composed climbing bean panel (VEC), the Andean diversity panel (ADP), the panel of progeny from two-way crosses between five Andean and five Mesoamerican parents (AxM), the Mesoamerican introgression panel (MIP), and the elite Andean breeding panel (VEF), whereas the latter four are bush type bean panels known from previous studies.

2.1.1. Climbing Bean Panel

The VEC comprised climbing bean lines of growth types III and IV. The lines were selected for grain quality, commercial seed type, disease resistance, SdFe, seed zinc concentration (SdZn), and agronomic performance. They represent the genetic variation used in climbing bean breeding at the International Center for Tropical Agriculture (CIAT). The VEC was composed mainly of lines from the Andean gene pool, but it also included a line of the Mesoamerican gene pool (G2333) and a few lines of admixed origin. In total, the VEC was comprised of 290 lines including twelve breeding groups, four genebank accessions, and six cultivars (**Supplementary Table 1**).

Field trials including all VEC lines were conducted in Colombia, Uganda, and Tanzania to collect phenotypic data for this study.

2.1.2. Bush Type Bean Panels

Four bush type breeding panels were used in this study: (*i*) the ADP consisting of 352 Andean bush cultivars and breeding lines from public and private breeding programs as described by Cichy et al. (2015). The ADP genotyping-by-sequencing (GBS) data was available *via* the ARS Feed-the-Future Grain Legumes Project (arsftfbean.uprm.edu/bean/?p=472); (*ii*) the AxM as described by Mayor Duran et al. (2016) and Mayor Duran (2016); (*iii*) the MIP consisting of Mesoamerican breeding lines with interspecific introgressions from *P. acutifolious, P. dumosus, and P. coccineus* in their pedigrees and their parental lines as described by Diaz et al. (2021); (*iv*) the VEF as described by Keller et al. (2020).

For the VEF, MIP, and AxM, phenotypic and genotypic data of 605, 217 and 200 lines were available, respectively. For the ADP, field trials were conducted in Mozambique, Tanzania, and in the United States to collect phenotypic data for this study.

2.2. Phenotyping

Agronomic traits were evaluated in the VEC and ADP as previously described by Keller et al. (2020). Briefly, DF represents the days from planting until 50% of the plants in the plot had at least one open flower. The seed yield per plot was normalized to a moisture content of 14% and extrapolated to yield per hectare. The weight of 100 seeds (100SdW) was measured separately. The growth type was assessed according to the four categories described by Singh (1981). The growth habit described climbing ability and differentiated only between bush (types I and II) and climbing types (type III and IV). Further traits such as DPM, pod harvest index (PHI), SdFe, SdZn, and canning quality were phenotyped only for the VEC. Analogous to DF, DPM represents the days until 50% of the pods in one plot had lost their green pigmentation. For the PHI, the seed dry weight of 20 pods at harvest was divided by the corresponding pod dry weight. The SdFe and SdZn were assessed on dried and ground seeds as described by Stangoulis and Sison (2008). The SdFe and SdZn content was then quantified by the X-ray fluorescence method using the X-Supreme 8000 instrument (Oxford Instruments, UK) (Guild et al., 2017). The canning quality was assessed by a trained sensory panel at Michigan State University as described by Cichy et al. (2014). Briefly, the beans were soaked in distilled water, canned at 100°C, sealed, and stored for 2 weeks. Upon opening, the canning quality was assessed by a trained consumer panel and expressed as an overall score from 1 (unacceptable appearance) to 5 (excellent appearance). The rated criteria included color, bean splitting, free starch clumps, and brine clarity after cooking (Cichy et al., 2014).

2.2.1. Field Trials for Climbing Beans

The field trials for the VEC were carried out at five locations in three countries: Darién (3°53′31″N 76°31′00″W, altitude of 1,491 m a.s.l.), Palmira (3°30′03.0″N 76°21′03.5″W, altitude 965 m a.s.l.), and Popayán (2°25′39″N 76°37′17″W, altitude of 1,750 m a.s.l.) in Colombia; Kawanda (0°24′49″N 32°31′59″E, altitude of 1,190 m a.s.l.) in Uganda; and Kagera (1°24′56.5″S 31°46′48.8″E, altitude of 1,320 m a.s.l.) in Tanzania (**Supplementary Table 2**). Each line was arranged in an alpha lattice design with three replicates.

Regarding the field trials in Colombia, the experimental units consisted of one row with a row-to-row distance of 0.95 m and with seven seeds sown manually per meter row length. Row length per plot differed between locations with 2.5 m, 2.2 m, and 2.0 m used in Darién, Palmira, and Popayán, respectively (**Supplementary Table 2**). Climbing beans (mainly type IV) required a trellis to support the plant. Wooden poles of 3 m height were distributed in the field in squares of 5 x 5 m. Wires were spanned between the poles at a height of 2.3 m. Each bean plant climbed on a string hanging from the wire. Eight strings per plot were deployed. Plant protection was carried out when needed using good agricultural practices.

The soil type in Darién and Popayán was an Inceptisol (Typic Dystrandept) with about 70 g/kg and 140 g/kg of soil organic matter, respectively, and a pH of around 5 (Barbosa et al., 2018). The soil type in Palmira was Mollisol with a pH between 7.0 and 7.5.

In Uganda and Tanzania, the plots were planted in single 3.0 m rows with about 30 seeds planted per row and at 0.6 m between the rows (**Supplementary Table 2**). Granular N:P:K (nitrogen,

P_20_5 and K_2O) fertilizer was manually applied at 125 kg/ha. Trials were manually weeded two times. Climbing bean plants were staked with 2 to 3 m long wooden poles. Five poles were used per plot.

2.2.2. Field Trials for Bush Beans

The field trials for the ADP were carried out at five locations in three countries: Arusha (3°21′41″S 36°37′34″E, altitude of 1,387 m a.s.l.), Morogoro (6°51′14.2″S 37°39′27.6″E, altitude of 526 m a.s.l.), and Mbeya (8°54′52″S 33°31′05″E, altitude of 1,780 m a.s.l.) in Tanzania; Chokwe (24°30′04″S 33°00′09″E, altitude of 35 m a.s.l.) in Mozambique; and Wilcox (32°01′44″N 109°41′27″W, altitude of 1,321 m a.s.l.) in the United States (**Supplementary Table 3**).

The trials were conducted in two to three replications in a randomized complete block design (**Supplementary Table 3**). The plot size differed between trials with lengths of 5–8 m, 1–0.5 m spacing between rows, and 1–4 rows. The plot sizes, soil types, and growth conditions of each trial are described in **Supplementary Table 3**. Irrigated trials were watered about two times a month and drought trials were rain fed. The growing season of one trial (MZCH15D_heat) coincided with the high temperature season at the site. The trials in Tanzania and in the United States were not fertilized. In the trials in Mozambique, single superphosphate and ammonium sulfate were applied at about 30 kg P/ha and 6.3 kg N/ha, respectively.

2.2.3. Available Phenotypic Data From Previous Studies

Of the 290 VEC lines, 43 were phenotypically characterized in two trials in a previous study (Barbosa et al., 2018). Phenotypic data were also publicly available for the AxM (Mayor Duran, 2016), the MIP (Diaz et al., 2021), and the VEF (Keller et al., 2020) from ten, one and three trials, respectively (**Supplementary Table 4**).

2.3. Genotyping

GBS was conducted as previously described by Nay et al. (2019). Briefly, DNA was extracted from leaf tissue and digested with the ApeKI restriction enzyme as described in Elshire et al. (2011). The sequencing was performed on the Illumina HiSeq 2500 platform at the HudsonAlpha Genome Sequencing Center (Huntsville, AL, United States). The sequence reads were demultiplexed using NGSEP (v3.3.0) (Tello et al., 2019), trimmed using Trimmomatic (v0.36) (Bolger et al., 2014), and aligned to the reference genome of *P. vulgaris* G19833 v2.1 (Schmutz et al., 2014) using Bowtie2 (v2.2.30) (Langmead and Salzberg, 2012). The variant calling was carried out using NGSEP, filtering SNPs with a genotype quality below 40, minor allele frequency (MAF) below 0.05, and removing SNPs with less than 60% of genotype calls, which yielded a matrix with 20% of missing data. The imputation of the missing data was performed with Beagle v.5.0 (Browning et al., 2018) using 100 as an effective population size and using the genetic map reported by Diaz et al. (2019). The SNP calling was carried out once on the VEC separately and once on all five panels together. All VEC lines were genotyped together with 55 additional climbing lines from a previous study (Barbosa et al., 2018).

2.4. Data Analysis

2.4.1. Phenotypic Data

Best linear unbiased estimators (BLUEs) were extracted from phenotypic data in two stages. In the first stage, the field data were corrected for spatial effects using the SpATS R package, setting row and column as random effects (Rodríguez-Álvarez et al., 2018). The number of plants harvested was binned (binwidth = 5) and added as a random effect in the spatial analysis. The plots with residuals bigger than ±3 times the SD were treated as outliers and removed iteratively as described in Keller et al. (2020). The BLUEs were extracted for each line in each trial from the SpATS model (first-stage BLUEs) from all the data sets, except for the ADP. In the ADP lines, the first-stage BLUEs were extracted using replicate (block) as a factor for the fixed effects since the row and column information was not available for these randomized complete block design trials. In the second stage, second-stage BLUEs for each line (L_i) were calculated across all trials using the following model:

$$y_{ij} = L_i + E_j + \varepsilon_{ij} \quad (1)$$

where y_{ij} is the first-stage BLUE of the i^{th} line in the j^{th} trial, L_i is the fixed effect for each line, E_j is the fixed effect for each trial, and ε_{ij} the error term. The inverse of the squared standard error (SE) of the mean was used as a weight, i.e., $\varepsilon \sim N(0, R)$; $R = \bigoplus_{j=0}^{n} R_j$ where $R_j = diag[(SE_{ij})^2]$, and (SE_{ij}) is the standard error of a mean of the i^{th} line in the j^{th} trial (Möhring and Piepho, 2009). In addition, BLUEs for yield were scaled (mean = 0, SD=1 resulting in Yd_scaled) for each panel to compare relative differences between the lines beyond panel and growth type. Kernel density estimates of the BLUEs were calculated using a Gaussian kernel with 1/30 bandwidth of the data range. The estimates were drawn as a smoothed histogram with the integral equal to one using the ggplot R package (Wickham, 2016). The significance of the genotype-by-environment interaction (GxE) was tested by comparing model 1 with and without an interaction term EL_{ij} between j^{th} trial and i^{th} line using the likelihood-ratio test. The test statistic was compared with a chi-square value with one degree of freedom and the p-value was adjusted following the work by Self and Liang (1987).

2.4.2. Linkage Disequilibrium and Population Structure

Pairwise measures of linkage disequilibrium (LD) were calculated for each population in sliding windows of 100 markers. The LD measures were corrected for kinship in the population (r_V^2) as implemented in the R package LDcorSV (v1.3.2) (Mangin et al., 2012). The LD decay was estimated regressing the pairwise r_V^2 values on the physical distance of their markers using the locally estimated scatterplot smoothing implemented in the R function "loess" (v4.1.0), with a span value of 0.5.

The phylogenetic tree was constructed using the GGTREE R package on the hierarchically clustered SNP matrix (Yu et al., 2017). The population structure was assessed on the SNP matrix

using principal components analysis (PCA) implemented in the FactoMineR R package (Lê et al., 2008). The correlation of the supplemental phenotypic traits (second-stage BLUEs) with the principal components (PCs) of the SNP matrix was calculated using the same FactoMineR package (Lê et al., 2008).

2.4.3. Genome-Wide Association Studies

To carry out GWAS, the Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) algorithm implemented in GAPIT was used (Huang et al., 2019). The first five PCs were used to correct for population structure. The imputed SNP matrix was used for GWAS. Identified QTL tagged by SNP markers were labeled with "Trait_Chr_Postion" as QTL ID, whereas yield was abbreviated with Yd and growth habit with GH. The SNP position derived from the reference genome G19833 v2.1 (Schmutz et al., 2014) was in Mbp rounded to two digits.

The network of significant marker-trait associations was visualized similarly as suggested in Fang et al. (2017) using the ggnetwork R package (Briatte et al., 2020). The distance between the connected nodes represents the LD between the two SNPs calculated as the squared Pearson correlation coefficient. The SNPs were connected when LD >0.25. The haplotypes associated with one trait were assembled using the identified SNPs in the indicated region for that trait. In case there were less than six SNPs selected, these SNPs were directly assembled to haplotypes. When more than five SNPs were selected, groups of haplotypes were constructed by hierarchical clustering of all lines based on those SNPs using the stats R package. The optimal number of clusters was determined using the average silhouette width implemented in the factoextra R package (Kassambara and Mundt, 2020).

2.5. Genomic Prediction

The GEBV were estimated using Bayesian generalized linear regression (BGLR) implemented in the BGLR R package (Pérez and de los Campos, 2014). For the factorial models, the BGLR extension for the Multiple-Trait Model (MTM) was used (http://quantgen.github.io/MTM/vignette.html). The Gibbs sampler ran with 20,000 iterations of which the first 10,000 were burned-in and the remaining were thinned by factor 5. Three different model approaches were tested.

2.5.1. Genotype Model Among and Across Trials

For each trait, the phenotypes adjusted per trial (y_{ij}) were modeled as the sum of the GEBV for each line (g_i) estimated based on SNP marker information, i.e., the additive relationship matrix (K), a fixed effect for the trial (E_j), and an error term $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$. The following linear model was used:

$$y_{ij} = g_i + E_j + \varepsilon_{ij} \quad (2)$$

with $g \sim N(0, K\sigma^2_g)$ and K was calculated as the normalized cross product of the SNP matrix using the rrBLUP package (Endelman, 2011; Endelman and Jannink, 2012). The y_{ij} were calculated either among all trials (using first-stage BLUEs), for all trials separately (using first-stage BLUEs), or across all trials (using second-stage BLUEs). In the case of the genotype model using all trials separately, j represents always the same environment. The same applied to the genotype model when using second-stage BLUEs across all trials.

2.5.2. GxE Model

In the GxE model, the GEBV for each trait were estimated for each environment (i.e., location) by adding an effect for the interaction between the j^{th} environment and i^{th} GEBV (gE_{ij}) to the model (2). This resulted in the following model:

$$y_{ij} = g_i + E_j + gE_{ij} + \varepsilon_{ij} \quad (3)$$

with $gE \sim N(0, I \otimes K\sigma^2_{gE})$, where I is the identity matrix for the environments and $\otimes$ denotes the Kronecker product. This means no correlations between environments were considered.

2.5.3. Factor Analysis Model

In the factor analysis (FA) model, the GEBV for each line in each environment (g_{ij}) were estimated for each trait using SNP marker information and the covariance of the phenotypes (y_{ij}) between the trials. The following equation was used:

$$y_{ij} = g_{ij} + \varepsilon_{ij} \quad (4)$$

with $g \sim N_j(0, G \otimes K\sigma^2_g)$, where G represents a covariance matrix of phenotypes between trials calculated as $G = BB^T + \Psi$, where B is a matrix of loadings (regressions of the original random effects into common factors) and Ψ is a diagonal matrix whose non-null entries give the variances of factors that are trait-specific. The model residuals were assumed to follow a multivariate normal distribution $\varepsilon \sim N_j(0, R_\varepsilon \otimes I_n)$, where R_ε is a covariance matrix of model residuals and I_n represents an n-dimensional identity matrix, where n is the number of phenotypes per environment. Three common factors were selected.

2.5.4. Training Population Optimization

For each VEC line in the validation set, the 50 closest related lines from all five panels were added to the TP (TP optimized). Genetic relationships were calculated as the cophenetic distances of the hierarchically clustered lines based on the SNP matrix, i.e., as the height of the dendrogram where the two branches including the considered lines join each other. This means that the TP could consist of between 50 lines (in case the lines to be tested would have all the same 50 closest relatives) and 2,500 lines (if they would not have the same closest relatives). The number of related lines (50) was chosen arbitrarily but seemed reasonable for a panel of 1,500 lines with some population structure. The optimized TP was compared to a TP containing only VEC lines (TP VEC) and a TP containing all lines from all five panels (TP extended).

2.5.5. Cross-Validation

The cross-validation was done 100 times, randomly splitting the dataset into training and validation set, i.e., for each cross-validation step, 70% of the lines were selected for the training and 30% for the validation set. The lines in the training and

validation sets were referred to as TP and new lines, respectively. The Pearson correlation coefficient (r) between predicted and observed values was calculated at each cross-validation step to assess the PA. The prediction accuracy (PAcc) is defined as the quotient of PA and the square root of heritability.

2.5.6. Genomic BLUPs Without Cross-Validation

The genomic r is defined as the Pearson correlation coefficient of modeled vs. observed values when all lines were used in the TP to fit the model. The genomic r was derived from the predicted genomic BLUPs without cross-validation.

3. RESULTS

The traits DF, 100SdW, SdFe, and yield were analyzed in 27, 28, 6, and 33 field trials, respectively, using a total of 1,869 lines belonging to five different breeding panels (**Supplementary Figure 1**). For this study, six and eleven field trials were newly evaluated, adding data for the VEC and ADP, respectively. The number of evaluated lines per trial was 290 for the VEC (**Supplementary Table 2**) and ranged between 41 and 268 for the ADP (**Supplementary Table 3**).

3.1. Phenotypes of Climbing Beans

The lines of the VEC showed different phenotypic distributions for the traits DF, 100SdW, SdFe, and yield among all field trials (**Figure 1A**). Especially, strong environmental effects were observed for yield. The trials carried out in the lowlands and in the warmer climates (Palmira, Kawanda, and Kagera) had less yield compared to the remaining trials in the high-altitude locations Darién and Popayán. Phenotypic variation among and across trials was also observed for additional traits, including important breeding targets such as canning quality and PHI (**Supplementary Figure 2**). The phenotypic correlations between trials were evaluated using the first-stage BLUEs: positive correlations were revealed across the trials for DF, 100SdW, and SdFe, whereas yield of the Pal19D and TzKg19D trials were mainly negatively correlated to the other trials (**Figure 1B**). The likelihood-ratio test confirmed significant GxE (p value < 0.001) for all traits, especially for yield, where the variance component for GxE was higher than for the lines (**Supplementary Table 5**).

Across all trials, the correlations between traits were evaluated using the second-stage BLUEs. Positive correlations were revealed between the traits SdFe, SdZn, DPM, and DF (**Supplementary Figure 3**). However, the correlations were negative between yield, nitrogen use efficiency, and 100SdW. Additionally, PHI was negatively correlated to SdFe and SdZn. Since yield showed strong GxE, the correlations to other traits differed across trials, e.g., the yield was negatively correlated with DF in the Pop15B and Pal19D trials but positively correlated with DF in the remaining trials (**Supplementary Figure 4**). In summary, strong GxE was observed for yield, while the remaining traits showed moderate GxE among the different environments and trials.

3.2. Comparing Phenotypes of the Five Breeding Panels

The phenotypic distribution of DF, 100SdW, SdFe, and yield were compared among all five panels across all trials (**Figure 2A**). The climbing beans of the VEC showed on average 28%, 21%, and 67% higher DF, SdFe, and yield than the bush type panels, respectively. The trait 100SdW depended primarily on the gene pools, showing lower values in the Mesoamerican MIP and the AxM, consisting of inter gene pool crosses. The observation in the VEC, that SdFe correlated negatively and DF and 100SdW positively to yield, was also true in the four bush bean panels (**Supplementary Figure 4**).

3.3. Structure and Diversity of the Five Breeding Panels

In the joint analysis of all five panels, 14,913 SNP markers distributed over the whole genome were kept from the raw 169,087 SNPs after filtering for genotype quality calls, MAF, and missingness. The dendrogram of the 1,869 lines showed grouping into Mesoamerican (represented by the MIP) and Andean origin, whereas the AxM and part of the VEC formed an admixture branch (**Figure 2B**). These genetic groups were also visible in the PCA analysis of all lines: the first PC clearly grouped the lines according to their Andean and Mesoamerican origin spreading the admixed AxM lines in-between (**Figure 2C**). In agreement, the first PC was highly correlated with 100SdW ($r = -0.67$) which differentiated the two gene pools (**Figures 2A,C**). The second PC explained variation for growth type, separating mostly the VEC lines from the others, and was correlated to DF ($r = 0.58$, **Figure 2C**). The first and second PC explained 28.9 and 3.4% of the genetic variance, respectively. The third PC, explaining 2.4% of the variance, captured variation mainly between the AxM and MIP, whereas the fourth and fifth PC showed the smallest variation for the VEF (**Supplementary Figure 5**). Finally, the sixth PC showed no clear pattern among the panels. Therefore, five first PCs were included as fixed effects for GWAS. The LD decay observed for all the breeding panels was faster for the combined panel than for the separate panels, enabling higher detection power for GWAS (**Figure 2D**). In general, the genetic diversity was bigger between the two gene pools than between the growth types.

3.4. Genome-Wide Association Studies Across All Panels

Carrying out GWAS using the second-stage BLUEs across all trials and breeding panels, a total of 69 significant marker-trait associations were identified below the 1% Bonferroni corrected significance level (**Figure 3A** and **Table 1**). The observed p-value distribution showed a clear deviation of the identified significant SNPs from the expected uniform distribution if no genetic linkage was present (**Figure 3B**). A p-value inflation was visible mainly in growth habit and 100 SdW. Most striking was the region at around 45 Mbp on Chr 1 associated with significant effects for growth habit, DF, DPM, and 100 SdW (**Supplementary Figure 6**). Furthermore, different QTL for SdFe and scaled yield across all panels were identified.

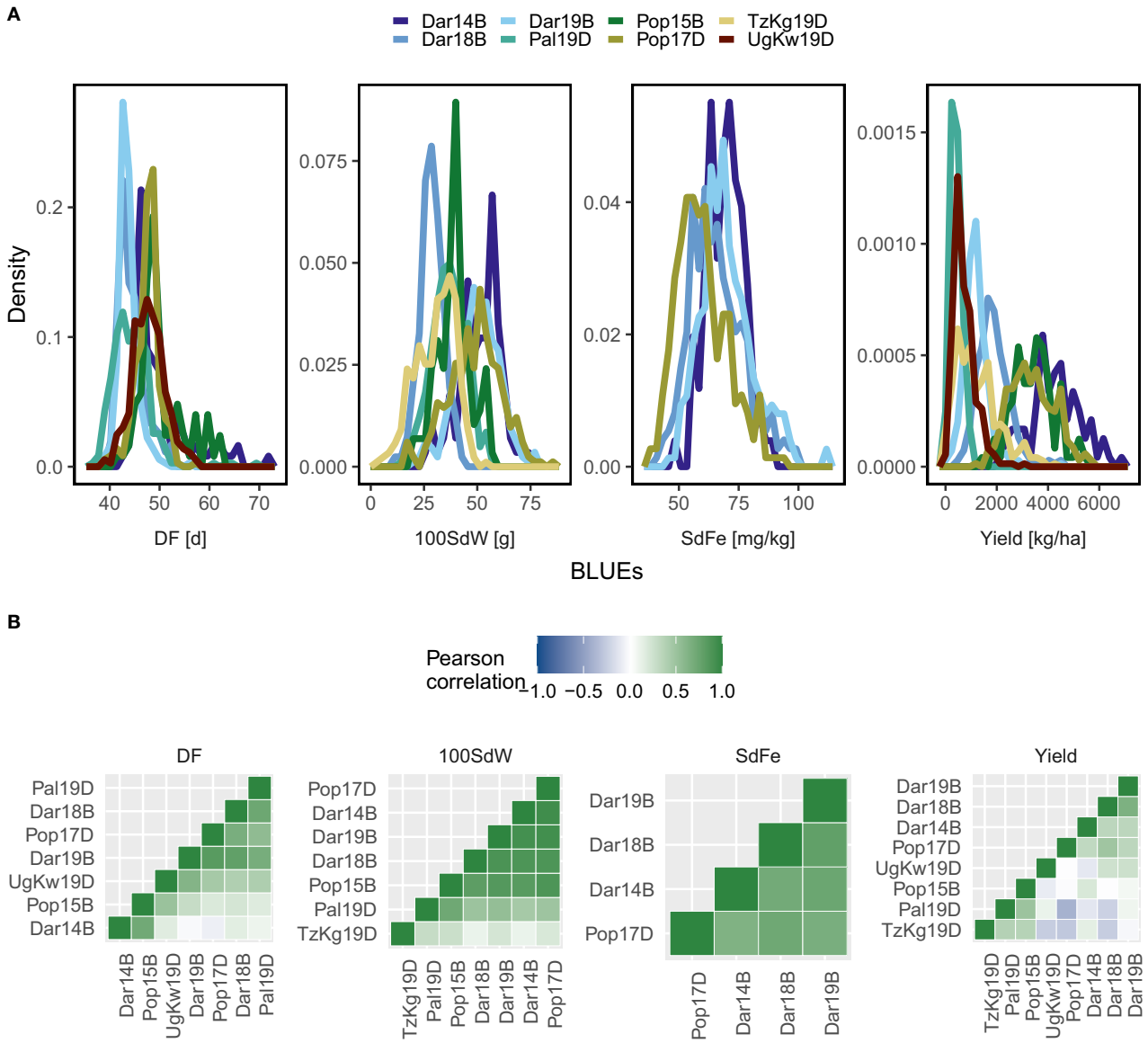


FIGURE 1 | Phenotypes of the climbing bean panel (VEC). **(A)** Density diagrams for days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield of up to 290 VEC lines in eight trials are shown. Best linear unbiased predictors (BLUEs) were calculated from each trial corrected for spatial effects in the field. **(B)** Pearson correlation coefficients were calculated across all trials for each trait based on the BLUEs. Trials were abbreviated based on the location Darién (Dar), Palmira (Pal), Popayán (Pop) in Colombia, Kagera in Tanzania (TzKg), or Kawanda in Uganda (UgKw), the year and the planting season (sequentially A to D). For a detailed description of each trial refer to **Supplementary Table 2**.

3.4.1. Growth Habit and Pleiotropy

For growth habit, 38 significant marker-trait associations were detected (**Table 1**). These QTL determined growth habit, e.g., bush (type I and II) from climbing beans (type III and IV) or they distinguished growth determinacy, e.g., separated determinate types I from indeterminate type II, III, and IV (**Supplementary Figure 7A**). The SNPs associated with growth habits on Chr 1, 3, and 6 (GH_1_43.71, GH_3_1.29, and GH_6_23.87) differentiated mainly the determinate growth type I from the other three types. In contrast, two QTL on Chr 4 (GH_4_1.42 and GH_4_40.05) and on QTL on Chr 5 (GH_5_0.74) differentiated bush from climbing types. The minor SNP variant on Chr 2 (GH_2_40.05) was the only one exclusively associated with the two climbing growth types (type III and IV). However, this association is to interpret cautiously since this SNP variant was rare (MAF = 0.05).

The region at the end of Chr 1 showed significant pleiotropic effects on different traits (**Supplementary Figure 6**). In that region, significant SNPs were identified not only for growth

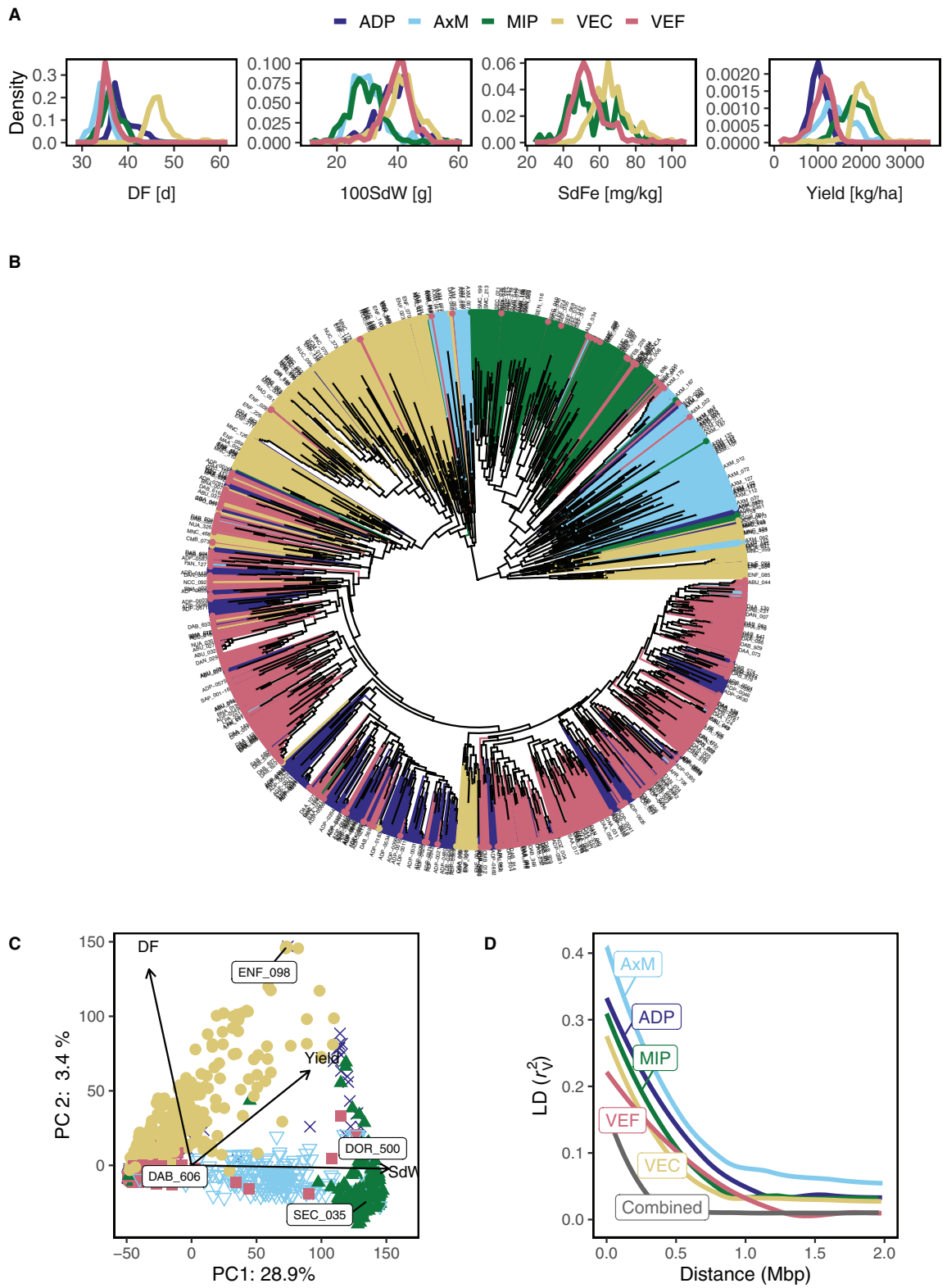


FIGURE 2 | Phenotypic and genotypic characterization of five common bean breeding panels including lines with bush and climbing growth habit originating from the Andean and Mesoamerican gene pools. **(A)** Density diagrams of best linear unbiased estimators among five breeding bean panels are shown for days to

(Continued)

FIGURE 2 | flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield. **(B)** Dendrogram of 1,869 lines characterized by 14,913 SNPs shows the hierarchical relationships between lines and panels [following the same color code for panels as in **(A)**]. **(C)** Principal components (PC) 1 and 2 visualize the genetic similarity across all five breeding panels. The arrows show quantitative supplementary phenotypic traits. Their cosines indicate the correlation with PC axes and their length approximate the SD of the variable. The extreme lines on the PC 1 axis are labeled. **(D)** Linkage disequilibrium (LD) decay is shown for all panels separately and combined. The LD was calculated in sliding windows of 100 markers and corrected for kinship in the population (r^2_V).

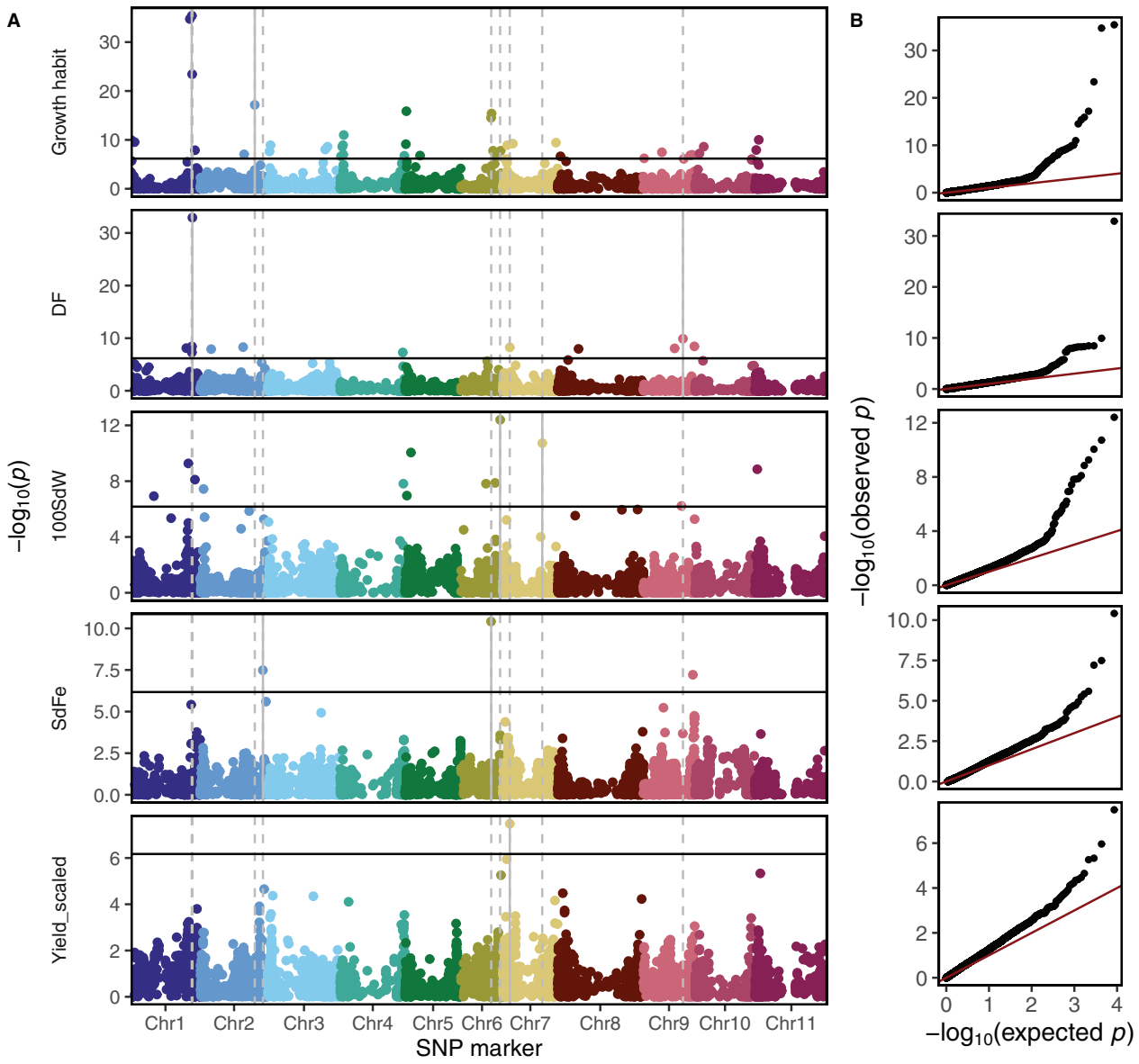


FIGURE 3 | Genome-wide association studies among five breeding panels differing in growth habits. **(A)** The Manhattan plots show the genetic associations with climbing growth habit, days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield. Seed yield was scaled among panels to allow comparison between them. The horizontal black lines show the Bonferroni corrected significance threshold at the 1% level. The vertical lines indicate the position of the two most significant markers for each trait. **(B)** Quantile distribution plots show the deviation of expected to observed *p*-values of SNP to trait associations for each trait.

habits but also for DF, DPM, and 100SdW. Additionally, the significant SNPs for growth habit on Chr 3 and 6 showed a tendency toward pleiotropic effects as observed for the QTL on Chr 1 (**Supplementary Figure 7B**). Interestingly, the QTL on Chr 4 (GH_4_1.42 and GH_4_40.05) showed again a different pattern than the others: these SNPs increased 100SdW

with increasing DF while, surprisingly, yield (scaled across populations) decreased. In summary, the 38 SNPs significantly associated with growth habit determined the four different growth types in different proportions, whereas only a few of these SNPs expressed pleiotropic effects.

3.4.2. SNP Effects Among Breeding Panels

In general, the significant marker-trait associations identified in the whole population of 1,869 lines showed effects also in the panels separately (**Figure 4**). An exception was the QTL DF_9_29.38 which revealed significant associations only in the VEF and ADP, the two Andean panels. An interesting breeding target for yield is Yd_7_4.86, expressing a clear effect in all five panels. This QTL was physically linked to DF_7_4.82 and in LD with several other QTL on different chromosomes (**Figure 5A**). Finally, two QTL for SdFe (SdFe_2_46.07 and SdFe_6_22.37) showed an effect in all three panels which had data for SdFe (**Figure 4**). Except for DF, major QTL for 100SdW, SdFe, and yield were identified showing effects in all five breeding panels.

3.4.3. Haplotype Effects Across and Among Growth Types

For traits with complex genetic architecture, single SNPs poorly explain the variance caused by the associated genetic region. Therefore, haplotypes were constructed on Chr 1 between 43.71 and 45.47 Mbp for SNPs significantly associated either with DF or growth type. The haplotypes for DF on Chr 1 explained 44.6% of the variance for DF across all growth types (**Supplementary Figure 8A**). In contrast, the best SNP for DF on Chr 1 explained only 8% of the variance. The haplotypes for growth type on Chr 1 differentiated type I from the other types (**Supplementary Figure 8B**). The first haplotype ("101") was almost exclusively associated with the determinate growth type I. The second haplotype ("001") was mainly associated with growth types II and IV. The remaining three haplotypes were associated with climbing growth habit (types III and IV). An important breeding goal is to shorten DF in all growth types while maintaining other agronomic traits. Therefore, all SNPs in the region from 44.60 to 45.47 Mbp on Chr 1 were clustered, resulting in ten distinct haplotypes according to the average silhouette width (**Figure 5B**). As expected, the haplotypes showed strong effects on DF and explained almost 50% of the variation for DF. However, these haplotype effects were not consistent among the growth types, explaining 3.7, 36.9, 0.0, and 5.3% of the DF variation in growth types I, II, III, and IV, respectively. In addition, the effect on the remaining traits varied substantially across the haplotypes. Since only a few SNPs expressed significant pleiotropic effects, some trade-offs between traits could be removed by breaking the LD of QTL on different chromosomes (**Figure 5A**). In summary, the QTL on Chr 1 between 44.60 to 45.47 Mbp controlled major processes across the growth types but showed minor effects among them. Furthermore, the QTL exhibited varying pleiotropic effects on SdFe, 100SdW, and yield which can be decreased by breaking the LD between this and further QTL (e.g., Yd_7_4.86).

3.5. GWAS Within the Climbing Bean Germplasm

For the genetic analyses within the climbing bean germplasm, a total of 15,589 SNPs were identified in the VEC. The population structure was moderate with PC1 and PC2 explaining 19.1 and 5.8 % of the genetic variance, respectively (**Supplementary Figure 9**). In total, 22 significant marker-trait associations were identified in the VEC (**Supplementary Table 6**). Two QTL for PHI and DF were identified on Chr 5 in proximity at 38.67 and 39.34 Mbp, respectively, indicating tight linkage (**Supplementary Figure 10A**). Several SNPs significantly associated with SdFe and SdZn were identified on Chr 2, 4, 6, 7, and 10. Furthermore, a QTL for canning quality was identified on Chr 7 at 2.67 Mbp. No clear *p*-value inflation or deflation was observed compared to the expected *p*-value distribution (**Supplementary Figure 10B**). In summary, further SNPs were identified in the VEC separately to specifically improve climbing beans.

3.6. Genomic Prediction

The GEBV for new VEC climbing bean lines (validation set) were calculated either with parts of the VEC or with parts of all five panels as TP.

3.6.1. Genomic Prediction Among Environments

The PAs for new VEC lines within each trial and across all trials differed between the traits as well as the used model approaches (**Figure 6**). In general, PAs followed the heritability and the genomic *r* calculated using all available lines as TP. PAs for yield reached about $r \approx 0.5$ in the high-altitude locations Darién and Popayán, where climbing beans are better adapted. The PAs for yield in other locations were lower. The trials Dar14B and Pop15B showed lower PAs with higher variability because these trials comprised only 100 lines.

On average, the FA model showed the best performance for DF. The genotype model for single trials performed best for 100SdW, SdFe, and yield. The FA and GxE model reached slightly lower average PAs for SdFe and yield, respectively. The strong GxE for yield was reflected in the different marker effects among the locations (**Supplementary Figure 11**). The PAcc reached the highest values for 100SdW using the genotype model among trials (77.4%), DF using the FA model (67.4%), SdFe using the genotype model for single trials (63.5%), and yield using the genotype model for single trials(72.5%, **Supplementary Table 7**). In summary, the PAs differed among models and traits, i.e., when predicting for a single trial, the genotype model performed best, except for DF. When predicting for multiple years in one location, the FA model was promising for DF and SdFe while for yield and 100SdW, the GxE models performed best.

3.6.2. Genomic Prediction Across Environments With Optimization of the Training Population

To increase PAs for new VEC lines, three different approaches were tested: when only VEC lines were in the TP (TP VEC), when all lines of the five panels were in the TP (TP extended), or when distantly related lines were excluded (TP optimized;

TABLE 1 | Significant marker-trait associations across five bean breeding panels below the 1% significance level according to genome-wide association studies.

QTL ID	Trait	Chr	Pos	*p*-value	Marker	MAF
GH_1_0.53	Growth habit	1	526025	1.18E-10	Pv2.1_01_526025_A/T	0.24
GH_1_2.85	Growth habit	1	2850955	2.78E-10	Pv2.1_01_2850955 C/G	0.38
GH_1_43.71	Growth habit	1	43707108	1.93E-35	Pv2.1_01_43707108 A/G	0.41
GH_1_45.04	Growth habit	1	45044047	3.90E-36	Pv2.1_01_45044047 C/T	0.13
GH_1_45.37	Growth habit	1	45374662	4.01E-24	Pv2.1_01_45374662 T/C	0.08
GH_1_47.44	Growth habit	1	47439968	1.23E-08	Pv2.1_01_47439968 G/C	0.16
GH_2_32.21	Growth habit	2	32208182	9.73E-08	Pv2.1_02_32208182 C/G	0.32
GH_2_40.05	Growth habit	2	40045968	6.88E-18	Pv2.1_02_40045968 C/A	0.05
GH_3_1.29	Growth habit	3	1289075	2.35E-08	Pv2.1_03_1289075 T/A	0.45
GH_3_1.92	Growth habit	3	1921492	1.22E-09	Pv2.1_03_1921492 A/G	0.03
GH_3_42.49	Growth habit	3	42488118	9.24E-09	Pv2.1_03_42488118 G/C	0.3
GH_3_44.06	Growth habit	3	44056005	2.62E-09	Pv2.1_03_44056005 A/T	0.12
GH_4_0.45	Growth habit	4	448766	2.45E-07	Pv2.1_04_448766 A/T	0.03
GH_4_1.42	Growth habit	4	1421295	2.52E-09	Pv2.1_04_1421295 A/G	0.47
GH_4_1.92	Growth habit	4	1919859	1.29E-09	Pv2.1_04_1919859 A/G	0.26
GH_4_2.16	Growth habit	4	2164168	9.82E-08	Pv2.1_04_2164168 T/C	0.08
GH_4_2.56	Growth habit	4	2559941	1.04E-11	Pv2.1_04_2559941 A/T	0.25
GH_4_47.17	Growth habit	4	47174835	1.95E-07	Pv2.1_04_47174835 T/A	0.26
GH_5_0.39	Growth habit	5	394641	7.38E-10	Pv2.1_05_394641 T/G	0.13
GH_5_0.74	Growth habit	5	739798	1.33E-16	Pv2.1_05_739798 G/A	0.28
GH_5_10.7	Growth habit	5	10696009	1.63E-07	Pv2.1_05_10696009 G/C	0.02
GH_6_22.3	Growth habit	6	22301003	3.03E-15	Pv2.1_06_22301003 A/G	0.03
GH_6_22.51	Growth habit	6	22508433	3.94E-16	Pv2.1_06_22508433 T/G	0.16
GH_6_23.87	Growth habit	6	23868436	1.66E-08	Pv2.1_06_23868436 G/C	0.37
GH_6_26.05	Growth habit	6	26054074	2.23E-07	Pv2.1_06_26054074 C/T	0.1
GH_6_29.43	Growth habit	6	29429040	1.53E-08	Pv2.1_06_29429040 G/A	0.02
GH_7_3.05	Growth habit	7	3047903	1.18E-09	Pv2.1_07_3047903 G/A	0.28
GH_7_7.15	Growth habit	7	7150019	6.01E-10	Pv2.1_07_7150019 C/T	0.1
GH_7_38.99	Growth habit	7	38987037	3.73E-10	Pv2.1_07_38987037 A/G	0.42
GH_8_2.11	Growth habit	8	2107245	2.42E-07	Pv2.1_08_2107245 C/T	0.23
GH_9_0.86	Growth habit	9	860918	6.16E-07	Pv2.1_09_860918 G/A	0.17
GH_9_13.92	Growth habit	9	13924731	3.14E-08	Pv2.1_09_13924731 T/C	0.49
GH_9_34.74	Growth habit	9	34742076	1.48E-07	Pv2.1_09_34742076 A/G	0.23
GH_9_36.11	Growth habit	9	36110952	1.03E-07	Pv2.1_09_36110952 A/T	0.01
GH_10_3.19	Growth habit	10	3191949	6.26E-08	Pv2.1_10_3191949 G/T	0.21
GH_10_6.5	Growth habit	10	6495216	2.67E-09	Pv2.1_10_6495216 G/T	0.27
GH_11_1.01	Growth habit	11	1010422	1.14E-08	Pv2.1_11_1010422 T/C	0.49
GH_11_2.78	Growth habit	11	2775768	9.58E-11	Pv2.1_11_2775768 T/G	0.16
DF_1_41.08	DF	1	41082526	7.72E-09	Pv2.1_01_41082526 G/A	0.22
DF_1_44.6	DF	1	44604072	1.88E-08	Pv2.1_01_44604072 A/C	0.35
DF_1_44.93	DF	1	44927394	5.64E-09	Pv2.1_01_44927394 C/T	0.25
DF_1_45.04	DF	1	45044047	5.11E-08	Pv2.1_01_45044047 C/T	0.13
DF_1_45.23	DF	1	45233651	3.51E-09	Pv2.1_01_45233651 A/G	0.08
DF_1_45.47	DF	1	45469012	1.21E-33	Pv2.1_01_45469012 G/A	0.1
DF_2_7.9	DF	2	7900892	1.22E-08	Pv2.1_02_7900892 A/C	0.26
DF_2_31.53	DF	2	31525478	4.99E-09	Pv2.1_02_31525478 A/C	0.16
DF_4_45.98	DF	4	45975326	5.51E-08	Pv2.1_04_45975326 T/C	0.38
DF_7_4.82	DF	7	4824001	5.29E-09	Pv2.1_07_4824001 C/G	0.26
DF_8_15.48	DF	8	15481886	1.12E-08	Pv2.1_08_15481886 T/A	0.01

(Continued)

TABLE 1 | Continued

QTL ID	Trait	Chr	Pos	*p*-value	Marker	MAF
DF_9_23.12	DF	9	23116201	9.41E-09	Pv2.1_09_23116201 C/T	0.28
DF_9_29.38	DF	9	29378513	1.25E-10	Pv2.1_09_29378513 C/T	0.07
DF_9_37.83	DF	9	37825106	3.55E-09	Pv2.1_09_37825106 T/G	0.21
SdW_1_17.16	100SdW	1	17164005	1.19E-07	Pv2.1_01_17164005 A/G	0.16
SdW_1_42.52	100SdW	1	42515535	5.49E-10	Pv2.1_01_42515535 A/G	0.07
SdW_1_47.26	100SdW	1	47260087	7.66E-09	Pv2.1_01_47260087 A/G	0.28
SdW_2_2.23	100SdW	2	2234763	3.64E-08	Pv2.1_02_2234763 T/G	0.45
SdW_4_46.54	100SdW	4	46540184	1.51E-08	Pv2.1_04_46540184 C/G	0.36
SdW_5_1.06	100SdW	5	1062441	1.09E-07	Pv2.1_05_1062441 C/A	0.44
SdW_5_4.07	100SdW	5	4065359	8.94E-11	Pv2.1_05_4065359 C/T	0.28
SdW_6_18.46	100SdW	6	18456447	1.48E-08	Pv2.1_06_18456447 G/A	0.2
SdW_6_25.25	100SdW	6	25248245	1.34E-08	Pv2.1_06_25248245 G/C	0.21
SdW_6_28.9	100SdW	6	28898246	3.91E-13	Pv2.1_06_28898246 G/A	0.16
SdW_7_28.77	100SdW	7	28765036	1.85E-11	Pv2.1_07_28765036 C/T	0.02
SdW_9_28.29	100SdW	9	28287352	6.16E-07	Pv2.1_09_28287352 G/A	0.12
SdW_11_1.55	100SdW	11	1547189	1.39E-09	Pv2.1_11_1547189 C/T	0.12
SdFe_2_46.07	SdFe	2	46068228	3.20E-08	Pv2.1_02_46068228 G/C	0.19
SdFe_6_22.37	SdFe	6	22365971	3.95E-11	Pv2.1_06_22365971 A/C	0.32
SdFe_9_36 8	SdFe	9	36804490	6.26E-08	Pv2.1_09_36804490 C/T	0.16
Yd_7_4.86	Yield scaled	7	4856975	3.28E-08	Pv2.1_07_4856975 C/T	0.23

The QTL ID, trait, chromosome (Chr), physical position (Pos) in bp, association strength (p value), name of the marker, and minor allele frequency (MAF) of the significantly associated SNPs is reported. The panels included bush and climbing growth habits and were evaluated for days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield (Yield_scaled, scaled for each panel).

Supplementary Figure 12). The optimized TP increased PAs for DF, SdFe, and yield (scaled among panels) when adding bush type lines of other panels and reached a PAcc of 66.8, 66.6, and 22.7% corresponding to a 0.7, 1.8, and 8.8% increase on the averaged PA, respectively (**Figure 6**). Regarding 100SdW, the TP optimization and extension decreased PAs slightly compared to the TP with only VEC lines. In summary, in complex traits such as SdFe and yield, the PA can be improved by adding related lines from other panels which are not in the TP even though they were tested in different trials.

4. DISCUSSION

Based on the largest assembly of phenotypic and genotypic common bean data, we showed increased PAs for important traits of climbing bean by the addition of related bush type beans from other trials to the TP. In addition, the extended pool of lines, including 1,869 genotypes from distinct breeding panels, was useful to predict growth type and to increase power in the detection of QTL using GWAS (Spindel and McCouch, 2016). Hence, this comprehensive study provides a solid basis to harness the large genetic diversity of common bean germplasm and to implement marker-assisted and genomic selection strategies for more efficient climbing bean breeding.

4.1. QTL Across Breeding Panels

For all studied traits, QTL with clear effects on the phenotypes in all five breeding panels were detected (**Figure 4**). This diverse joint group of panels with fast LD decay enabled the identification of SNPs tightly linked to the causal loci while controlling for population structure (Sul et al., 2018; Huang et al., 2019). On the one hand, several QTL for growth habit and DF were confirmed from previous studies. On the other hand, especially for SdFe and 100SdW, new QTL and candidate genes were identified.

4.1.1. Previously Described and New QTL for DF and Growth Habit

Considering significant marker-trait associations less than 1 Mbp away from previously reported positions, QTL for DF and growth habit were confirmed on Chr 1, 4, 9, and 11. In addition, on Chr 6, 7, and 8, significantly associated SNPs were mapped to a distance of 1 to 4 Mbp from previously reported QTL. The terminal flowering gene *PvTFL1y* (*fin* locus identified as Phvul.001G189200 on Chr 1 at 44.85 Mbp) (Norton, 1915; Koinange et al., 1996; Kwak et al., 2008; Repinski et al., 2012; González et al., 2016) was confirmed by DF_1_44.60 (**Table 1**). The *phytochrome A* gene (*Ppd* locus identified as Phvul.001G221100 on Chr 1 at 47.64 Mbp) conferring photoperiod sensitivity (Coyne and Schuster, 1974; Gu et al., 1998; Kamfwa et al., 2015; Weller et al., 2019) was tightly linked to GH_1_47.43. In agreement, the haplotype "101" constructed in the region from 43.71 to 45.37 Mbp almost exclusively differentiated between determinate and indeterminate growth types (**Supplementary Figure 8B**). Thus, multiple SNPs are required

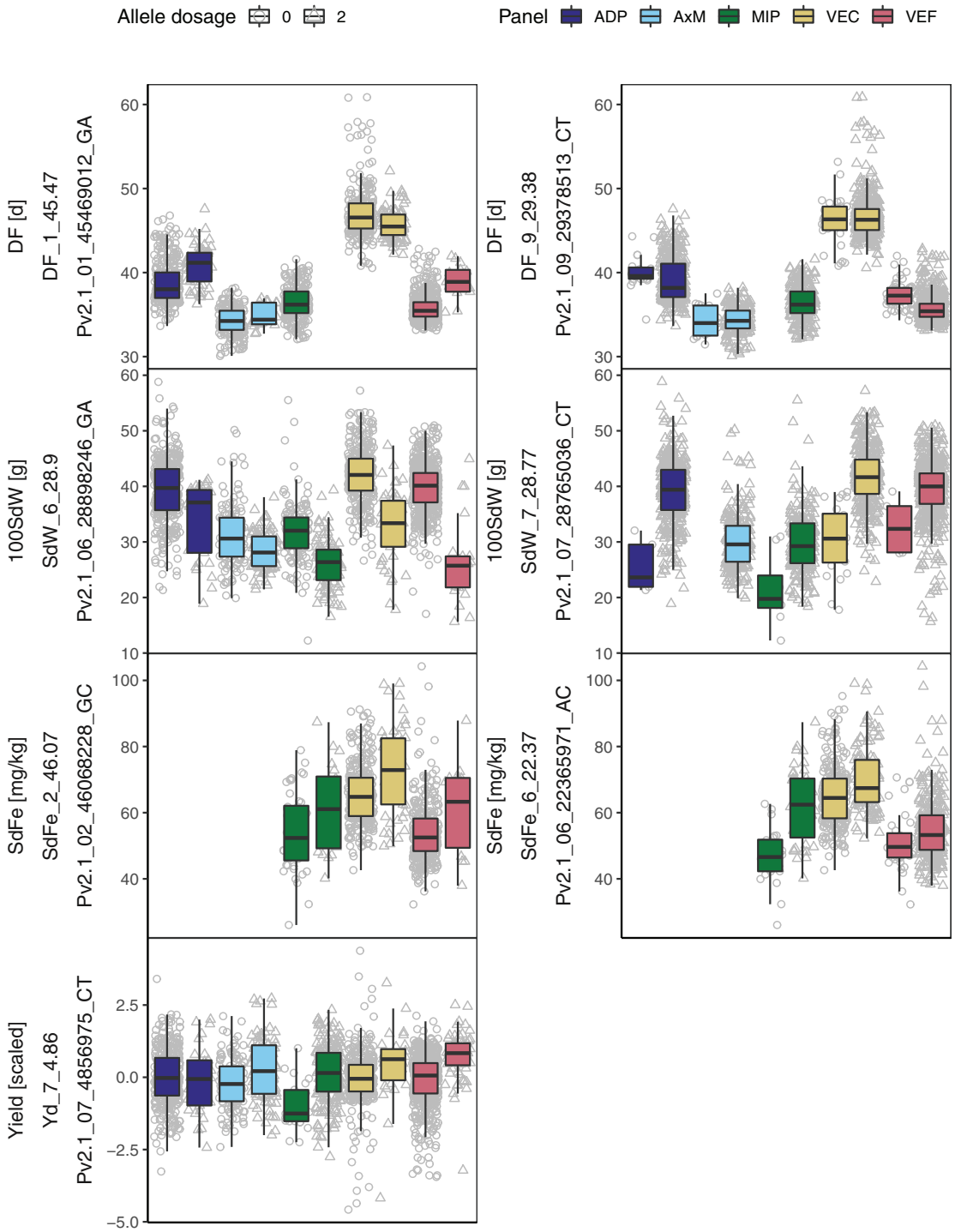


FIGURE 4 | Boxplots for allele dosage effect (0 or 2 alternative alleles) of the most significant SNPs associated with days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield for five breeding panels. The panels included lines with bush and climbing growth habits originating from the Andean and Mesoamerican gene pools. The trait, the name of the associated marker and the QTL ID is given.

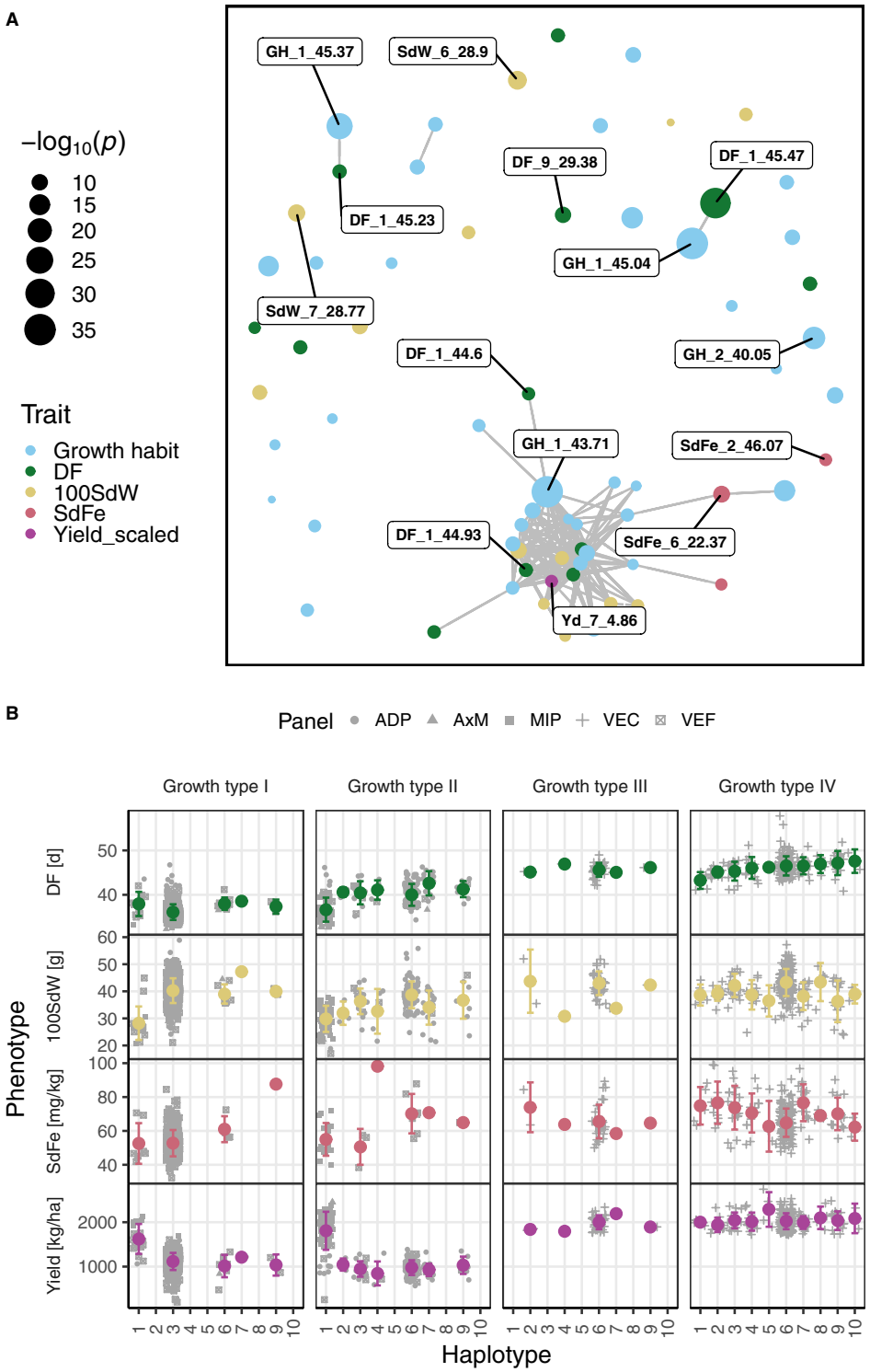


FIGURE 5 | Significant marker-trait associations were analyzed for growth habit, days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield scaled among the five breeding panels (Yield_scaled). **(A)** A network of the SNPs significantly (below the 1% significance level) associated with each of the four traits forming clusters according to their linkage disequilibrium is shown. Each dot represents a significant SNP and its size the associated $-log_{10}$ *p*-value. The SNPs on Chr 1 between 43.71 and 45.47 as well as the two most significant SNPs per trait are labeled with the QTL ID. **(B)** Haplotypes including all SNPs between 44.60 and 45.47 Mbp on chromosome 1 were constructed using hierarchical clustering. The averaged SNP effects of the haplotypes were evaluated in all traits among all growth types, i.e., growth type I (determinate bush type), type II (indeterminate bush), type III (determinate climber), and type IV (indeterminate climber). The error bars show the SD.

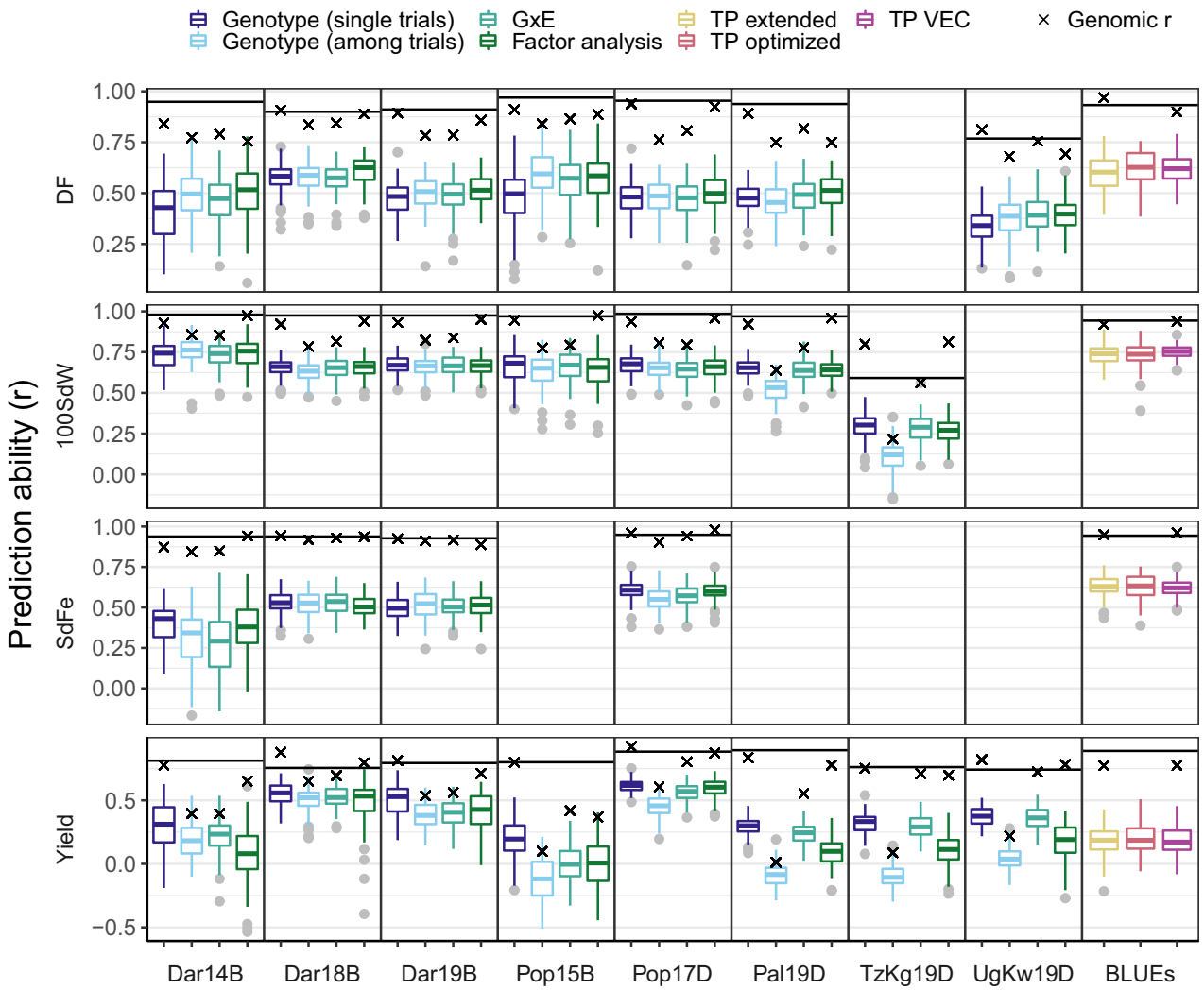


FIGURE 6 | Prediction abilities for different traits and models for new lines from the climbing bean panel (VEC). The first-stage best linear unbiased estimators (BLUEs) for each trial or second-stage BLUEs across trials were used. The first-stage BLUEs models were tested accounting for either genotypic effects only, i.e., for each trial separately (single trials) or all together (among trials), genotypic x environment interaction (GxE), or correlation between trials (Factor analysis). The second-stage BLUEs were used for models which take into account genotypic effects based on the VEC (TP VEC), on all five panels (TP extended), and all five panels with optimization of the training population (TP optimized). The predicted traits were days to flowering (DF), 100 seed weight (100SdW), seed iron concentration (SdFe), and seed yield. Seed yield was scaled among panels for the models based on second-stage BLUEs to allow comparison between the different growth habits. The horizontal line is the square root of heritability indicating the heritable variance of the trait in each trial. Trials were abbreviated based on the location Darién (Dar), Palmira (Pal), Popayán (Pop) in Colombia, Kagera in Tanzania (TzKg), or Kawanda in Uganda (UgKw), the year, and the planting season (sequentially A to D). For a detailed description of each trial refer to **Supplementary Table 2**.

to determine growth type including the allelic version of the *PvTFL1y* gene.

In the region of the two identified SNPs for growth habit on Chr 4 (GH_4_1.42 and GH_4_1.92), a QTL associated with climbing ability and plant height was previously reported (linked to Pvctt001 marker at 0.51 Mbp) (Checa and Blair, 2008). Furthermore, in proximity to GH_6_29.43, a QTL for DF was previously reported at 31.6 Mbp (Raggi et al., 2019). The terminal flowering gene *PvTFL1z* (Phvul.007G229300 a homolog of *PvTFL1y*) was located on Chr 7 at 35.31 Mbp (Kwak et al., 2008). In our study, GH_7_38.99 was detected proximal to PvTFL1z. On the upper arm of Chr 8 at 4.9 Mbp, another QTL for DF was reported previously (Raggi et al., 2019). Similarly, we detected GH_8_2.11 at less than 3 Mbp distance. A second fin locus (*fin'*) on Chr 9 was tagged by molecular markers at 13.39 Mbp (de Campos et al., 2011) and at around 20 cM (González et al., 2016). This *fin'* locus is probably tagged by GH_9_13.92. A QTL for DF on Chr 11 at around 9 cM reported in Bhakta et al. (2017) was linked to GH_11_1.01 (at around 8.8 cM). A new QTL for growth habit, GH_2_40.05 was identified with one SNP variant

exclusively associated to climbing beans, however, with a low MAF of 5%. Newly identified SNPs for growth habit with lower LOD scores have to be interpreted carefully due to the observed *p*-value inflation, indicating remaining population structure. In conclusion, the joint analysis consisting of diverse common bean populations showed high detection power for DF and growth habit QTL. Four to potentially seven QTL known from previous studies and several new QTL for DF and growth habit were identified and tagged by tightly linked SNP markers.

4.1.2. QTL for SdFe and Their Independence From Growth Habit

Four meta-QTL were previously reported for SdFe on Chr 1 (between 43.3 and 48.5 Mbp), on Chr 6 (between 28.2 and 29.5 Mbp), on Chr 9 (between 11.7 and 13.5 Mbp), and Chr 11 (between 2.3 and 5.3 Mbp) (Izquierdo et al., 2018). All four QTL fall right onto or next to QTL for growth habit identified in the current study. Since climbing beans in general exhibit a higher SdFe (Blair et al., 2010; Petry et al., 2015), the previously reported QTL is probably confounded with population structure. The effects of these QTL on SdFe were also detected in our analysis but the associations did not exceed the 5% significance threshold (**Supplementary Figure 7B**). A QTL for SdZn was recently reported on Chr 1 at 49.37 Mbp in European landraces next to *PvTFL1y*, suggesting a genetic linkage between DF and SdZn (Caproni et al., 2020). From the three major QTL for SdFe detected in our study, two were identified for the first time and SdFe_6_22.37 was previously reported in proximity at 22.8 Mbp (Diaz et al., 2020). The identified SNPs for SdFe on Chr 2 and 6 showed an effect in all three evaluated breeding panels and are thus independent of the growth habit (**Figure 4**). The QTL at the end of Chr 2 (SdFe_2_46.07) is of particular interest because it was linked to a QTL for DPM in this study and a pleiotropic QTL affecting DF, DPM, 100SdW, and yield in the VEF (Keller et al., 2020). In conclusion, various QTL for SdFe reported previously seemed to be confounded with growth habit, while our joint analysis allowed us to detect new SNPs associated with SdFe across different breeding panels.

4.1.3. Candidate Gene Identification for SdFe and 100SdW

Four new candidate genes were identified for SdFe and 100SdW: on Chr 6 at 22.18 Mbp, in a distance of less than 200,000 bp from the SNP most significantly associated to SdFe (SdFe_6_22.37), the Phvul.006G113100 gene was annotated as a homologous to a ferric-chelate reductase, reported to be involved in iron uptake from the soil (Robinson et al., 1999; Wu et al., 2005; Jeong et al., 2008; Asard et al., 2013). In proximity, less than 125,000 bp away from SdFe_2_46.07 and SdFe_9_36.80, the genes Phvul.002G292900 and Phvul.009G247600, respectively, were annotated. These two genes putatively express *Atox1*-related copper transport proteins, which are involved in copper and iron homeostasis (Himelblau et al., 1998; Puig et al., 2007). A QTL for copper and iron uptake was shown previously to have close genetic linkage (Waters and Grusak, 2008). Regarding 100SdW, a putative asparagine synthetase (Phvul.006G188400) on Chr 6 at 28.87 Mbp, less than 30,000 bp away from SdW_6_28.90, showed major effects in all breeding panels (**Figure 4**). The asparagine synthetase remobilizes nitrogen from sources to sinks and was reported to increase seed weight and soluble protein content in Arabidopsis seed (Gaufichon et al., 2016, 2017). In conclusion, for the major QTL for SdFe and 100SdW, plausible candidate genes were identified whose putative functions remain to be further validated.

4.2. QTL Detected Within Breeding Panels

Several QTL were identified only in the VEC, e.g., a pleiotropic QTL for DF and PHI on Chr 5 between 38.68 and 39.34 Mbp. The QTL for PHI differed from QTL previously identified in bi-parental bush type populations (Mukeshimana et al., 2014; Diaz et al., 2018). The PHI was weakly and positively correlated to DF and yield (**Supplementary Figure 3**). This pleiotropic effect of the identified QTL might be related to the time from flowering until harvest which could affect DF and PHI. Regarding DF, the major QTL tagged by DF_1_44.60 was mapped more closely to *PvTFL1y* in the current joint analysis than in a separate analysis of the ADP and VEF at 48.34 and 49.72 Mbp, respectively (Kamfwa et al., 2015; Keller et al., 2020). In agreement, the LD decay of the combined panel was faster than that of the separate panels. Furthermore, we concluded that the genetic control of flowering is different on the single SNP level (**Figure 4**) and haplotype level among the growth types (**Figure 5B**). The major QTL for DF on Chr 1 showed no effect within growth type III lines. One possible explanation is that DF within the climbing growth habit is regulated by additional genes or growth habit specific alleles. Additionally, haplotype 1 showed increased yield for growth types I and II while the effect on DF differed. It demonstrates that DF and yield are loosely linked as suggested earlier by White et al. (1992). Similarly, the QTL for SdFe on Chr 2 at 2.91 and 48.86 Mbp were valid only for the VEC and were not detected in the joint analysis suggesting some specific alleles were present only in the climbing bean panel.

4.3. Genomic Predictions Across and Within Breeding Panels

Using only VEC lines as TP in the different modeling approaches, the FA model performed well for DF and SdFe, showing the importance of covariance between trials for those traits (**Figure 6**). Such FA models work well for different environments, where some lines were already tested, were previously reported for sweet cherry (*Prunus avium* L.) (Hardner et al., 2019). The PAcc for yield of the genotype model using the first- and second-stage BLUEs among all environments were lower in comparison to the FA model and GxE model which took environmental differences into account. In agreement, the high GxE was visible in the correlations between the trials which reached values from –0.41 to 0.64 (**Figure 1B**). Those interactions were additionally reflected in the relatively high PA of the GxE model in yield and in the differing marker effects among locations (**Figure 6** and **Supplementary Figure 11**). Comparing different traits, the correlations between yield and DPM in bush type beans altered when changing from low- to high-altitude locations (Diaz et al., 2018). Therefore, the single-trial genotype model and the GxE model, which account for

location effects such as high- and lowland, worked best for the predictions of yield. Similarly, high GxE for yield was shown before in the VEF under different environmental conditions (Keller et al., 2020). Across all trials, the TP optimization turned out to be a promising strategy to predict the performance of new climbing bean lines. For complex traits such as SdFe and yield, optimizing the TP with related lines from different panels tested under different conditions improved the PA by 1.8 and 8.8%, respectively. Especially for smaller breeding programs or new breeding panels, it is, therefore, advisable to optimize the TP even when the added lines were tested in different field trials.

5. CONCLUSION

The benefits of introducing genes conferring climbing ability into new breeding lines are limited because growth habits are fixed in specific production systems and their modification would have pleiotropic effects, most likely affecting traits like DF and SdFe negatively. The most significant QTL for growth habit at the end of Chr 1 was pleiotropic. However, this QTL was in LD with several other QTL which can be selected separately since they were located on different chromosomes (**Figure 5A**). In addition, this QTL had no effect on common beans of the growth type III. Regarding other QTL, we detected stable SdFe and yield QTL which showed effects in all tested panels without significant pleiotropic effects. The identified markers were validated in very diverse germplasm including all growth types and both gene pools. The resulting fast LD decay allowed mapping of QTL more precisely than was achieved in separated panels. Genomic prediction models were established across populations enabling the selection and hybridization of the best lines of all populations to combine favorable alleles. Especially for yield, GxE needs to be modeled when breeding for different environments. The large common bean diversity presented in this study was used to identify markers across populations and to establish and improve prediction models. The joint population will provide a basis to exploit this genetic diversity and will contribute to the quick and targeted development of new (climbing bean) lines.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://doi.org/10.7910/DVN/RLAWYN.

AUTHOR CONTRIBUTIONS

BR and BS conceived the study. CM, AP-B, HB, WA, SM, JM, TP, JB, and PM conducted the experiments and provided the phenotypic data. BK, DA-S, and JA performed modeling, data analysis, and interpretation. DA-S prepared the genotypic data. BK drafted the manuscript, which was improved by DA-S, BS, PM, TP, and BR. All authors approved to the manuscript submission.

FUNDING

The authors would like to thank the COOP Research Program on Sustainability in Food Value Chains of the ETH Zurich World Food System Center and the ETH Zurich Foundation for supporting this project. The COOP Research Program was supported by the COOP Sustainability Fund. The work was additionally founded by Tropical Legumes III-Improving Livelihoods for Smallholder Farmers: Enhanced Grain Legume Productivity and Production in Sub-Saharan Africa and South Asia (OPP1114827), and by the AVISA-Accelerated varietal improvement and seed delivery of legumes and cereals in Africa (OPP1198373) projects funded by the Bill & Melinda Gates Foundation. The Norman Borlaug Cooperative Research Initiative (NBCRI) Grain Legumes Project of the US Agency for International Development, Project No. 0210-22310-004-96R supported the Tanzanian trials in Arusha and Mbeya. We would like to thank USAID for their contributions through the CGIAR Research Program on Grain Legumes and Dryland Cereals. Open access publication is kindly covered by ETH Zurich.

ACKNOWLEDGMENTS

We very much thank the CIAT bean teams for supporting and carrying out the experiments. Special thanks to Brenda Nakyanzi, Eninka Mndolwa, and others for managing the field trials. Many thanks to Lucy Milena Diaz Martinez and Eliana Macea for DNA extraction and library preparation. We further acknowledge Prof. Achim Walter from ETH Zurich for sharing infrastructure with the Molecular Plant Breeding group.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.830896/full#supplementary-material

REFERENCES

Akdemir, D., and Isidro-Sánchez, J. (2019). Design of training populations for selective phenotyping in genomic prediction. *Sci. Rep.* 9, 1446. doi: 10.1038/s41598-018-38081-6

Asard, H., Barbaro, R., Trost, P., and Bérczi, A. (2013). Cytochromes b561: ascorbate-mediated trans-membrane electron transport. *Antioxidants Redox Signal.* 19, 1026–1035. doi: 10.1089/ars.2012.5065

Barbosa, N., Portilla, E., Buendia, H. F., Raatz, B., Beebe, S., and Rao, I. (2018). Genotypic differences in symbiotic nitrogen fixation ability and seed yield of climbing bean. *Plant Soil.* 428, 223–239. doi: 10.1007/s11104-018-3665-y

Bhakta, M. S., Gezan, S. A., Michelangeli, J. A. C., Carvalho, M., Zhang, L., Jones, J. W., et al. (2017). A predictive model for time-to-flowering in the common bean based on QTL and environmental variables. *G3* 7, 3901–3912. doi: 10.1534/g3.117.300229

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., et al. (2013). Molecular analysis of the parallel domestication of the common bean (Phaseolus vulgaris) in Mesoamerica and the Andes. *New Phytol.* 197, 300–313. doi: 10.1111/j.1469-8137.2012.04377.x

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., Onis, M., d., et al. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* 371, 243–260. doi: 10.1016/S0140-6736(07)61690-0

Blair, M. W. (2013). Mineral biofortification strategies for food staples: the example of common bean. *J. Agric. Food Chem*. 61, 8287–8294. doi: 10.1021/jf400774y

Blair, M. W., Prieto, S., Diaz, L. M., Buendia, H. F., and Cardona, C. (2010). Linkage disequilibrium at the APA insecticidal seed protein locus of common bean (*Phaseolus vulgaris* L.). *BMC Plant Biol*. 10, 15. doi: 10.1186/1471-2229-10-79

Bliss, F. A. (1993). "Breeding common bean for improved biological nitrogen fixation," in *Enhancement of Biological Nitrogen Fixation of Common Bean in Latin America: Results from an FAO/IAEA Co-ordinated Research Programme, 1986–1991, Developments in Plant and Soil Sciences*, eds F. A. Bliss and G. Hardarson (Dordrecht: Springer Netherlands), 71–79.

Boccio, J. R., and Iyengar, V. (2003). Iron deficiency. *Biol. Trace Elem. Res*. 94, 1–31. doi: 10.1385/BTER:94:1:1

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Briatte, F., Bojanowski, M., Canouil, M., Charlop-Powers, Z., Fisher, J. C., Johnson, K., et al. (2020). ggnetwork: Geometries to Plot Networks with 'ggplot2'. Available online at: https://CRAN.R-project.org/package=ggnetwork

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet*. 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015

Caproni, L., Raggi, L., Talsma, E. F., Wenzl, P., and Negri, V. (2020). European landrace diversity for common bean biofortification: a genome-wide association study. *Sci. Rep*. 10, 19775. doi: 10.1038/s41598-020-76417-3

Checa, O., Ceballos, H., and Blair, M. W. (2006). Generation means analysis of climbing ability in common bean (*Phaseolus vulgaris* L.). *J. Heredity* 97, 456–465. doi: 10.1093/jhered/esl025

Checa, O. E., and Blair, M. W. (2008). Mapping QTL for climbing ability and component traits in common bean (*Phaseolus vulgaris* L.). *Mol. Breed*. 22, 201–215. doi: 10.1007/s11032-008-9167-5

Cichy, K. A., Fernandez, A., Kilian, A., Kelly, J. D., Galeano, C. H., Shaw, S., et al. (2014). QTL analysis of canning quality and color retention in black beans (*Phaseolus vulgaris* L.). *Mol. Breed*. 33, 139–154. doi: 10.1007/s11032-013-9940-y

Cichy, K. A., Porch, T. G., Beaver, J. S., Cregan, P., Fourie, D., Glahn, R. P., et al. (2015). A Phaseolus vulgaris diversity panel for andean bean improvement. *Crop Sci*. 55, 2149–2160. doi: 10.2135/cropsci2014.09.0653

Cortes, L. T., Zhang, Z., and Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *Plant Genome* 14, e20077. doi: 10.1002/tpg2.20077

Coyne, D. P., and Schuster, M. L. (1974). Inheritance and linkage relations of reaction to *Xanthomonas phaseoli* (E. F. Smith) Dowson (common blight), stage of plant development and plant habit in *Phaseolus vulgaris* L. *Euphytica* 23, 195–204. doi: 10.1007/BF00035858

Crossa, J., Perez-Rodriguez, P., Cuevas, J., Montesinos-Lopez, O., Jarquin, D., de los Campos, G., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci*. 22, 961–975. doi: 10.1016/j.tplants.2017.08.011

de Campos, T., Oblessuc, P. R., Sforça, D. A., Cardoso, J. M. K., Baroni, R. M., de Sousa, A. C. B., et al. (2011). Inheritance of growth habit detected by genetic linkage analysis using microsatellites in the common bean (*Phaseolus vulgaris* L.). *Mol. Breed*. 27, 549–560. doi: 10.1007/s11032-010-9453-x

Diaz, L. M., Ricaurte, J., Tovar, E., Cajiao, C., Terán, H., Grajales, M., et al. (2018). QTL analyses for tolerance to abiotic stresses in a common bean (*Phaseolus vulgaris* L.) population. *PLoS ONE* 13, e0202342. doi: 10.1371/journal.pone.0202342

Diaz, S., Ariza-Suarez, D., Izquierdo, P., Lobaton, J. D., de la Hoz, J., Acevedo, F., et al. (2019). Replication Data for: Genetic mapping for agronomic traits in a MAGIC population of common bean (*Phaseolus vulgaris* L.) under drought conditions'. *BMC Genomics*. 21, 799. doi: 10.7910/DVN/JR4X4C

Diaz, S., Ariza-Suarez, D., Ramdeen, R., Aparicio, J., Arunachalam, N., Hernandez, C., et al. (2021). Genetic architecture and genomic prediction of cooking time in common bean (*Phaseolus vulgaris* L.). *Front. Plant Sci*. 11, 622213. doi: 10.3389/fpls.2020.622213

Diaz, S., Ariza-Suarez, D., Izquierdo, P., Lobaton, J. D., de la Hoz, J., Acevedo, F., et al. (2020). Replication data for: genetic mapping for agronomic traits in a MAGIC population of common bean (*Phaseolus vulgaris* L.) under drought conditions. *BMC Genomics* 21, 799. doi: 10.1186/s12864-020-07213-6

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379. doi: 10.1371/journal.pone.0019379

Endelman, J. B. (2011). Ridge Regression and other kernels for genomic selection with r package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Endelman, J. B., and Jannink, J.-L. (2012). Shrinkage estimation of the realized relationship matrix. *G3* 2, 1405–1413. doi: 10.1534/g3.112.004259

Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., et al. (2017). Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol*. 18, 161. doi: 10.1186/s13059-017-1289-9

Gaufichon, L., Marmagne, A., Belcram, K., Yoneyama, T., Sakakibara, Y., Hase, T., et al. (2017). ASN1-encoded asparagine synthetase in floral organs contributes to nitrogen filling in Arabidopsis seeds. *Plant J*. 91, 371–393. doi: 10.1111/tpj.13567

Gaufichon, L., Rothstein, S. J., and Suzuki, A. (2016). Asparagine metabolic pathways in Arabidopsis. *Plant Cell Physiol*. 57, 675–689. doi: 10.1093/pcp/pcv184

Gepts, P., Osborn, T. C., Rashka, K., and Bliss, F. A. (1986). Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. *Econ. Bot*. 40, 451–468. doi: 10.1007/BF02859659

González, A. M., Yuste-Lisbona, F. J., Saburido, S., Bretones, S., De Ron, A. M., Lozano, R., et al. (2016). Major contribution of flowering time and vegetative growth to plant production in common bean as deduced from a comparative genetic mapping. *Front. Plant Sci*. 7, 1940. doi: 10.3389/fpls.2016.01940

Graham, P. H. (1981). Some problems of nodulation and symbiotic nitrogen fixation in *Phaseolus vulgaris* L.: a review. *Field Crops Res*. 4, 93–112. doi: 10.1016/0378-4290(81)90060-5

Gu, W., Zhu, J., Wallace, D. H., Singh, S. P., and Weeden, N. F. (1998). Analysis of genes controlling photoperiod sensitivity in common bean using DNA markers. *Euphytica* 102, 125–132. doi: 10.1023/A:1018340514388

Guild, G. E., Paltridge, N. G., Andersson, M. S., and Stangoulis, J. C. R. (2017). An energy-dispersive X-ray fluorescence method for analysing Fe and Zn in common bean, maize and cowpea biofortification programs. *Plant Soil*. 419, 457–466. doi: 10.1007/s11104-017-3352-4

Hardner, C. M., Hayes, B. J., Kumar, S., Vanderzande, S., Cai, L., Piaskowski, J., et al. (2019). Prediction of genetic value for sweet cherry fruit maturity among environments using a 6K SNP array. *Hortic. Res*. 6, 1–15. doi: 10.1038/s41438-018-0081-7

HarvestPlus (2022). *BIOFORTIFICATION A food-systems approach to ensuring healthy diets globally*. Technical brief, HarvestPlus and the World Food Programme.

Himelblau, E., Mira, H., Lin, S.-J., Cizewski Culotta, V., Peñarrubia, L., and Amasino, R. M. (1998). Identification of a functional homolog of the yeast copper homeostasis gene ATX1 from Arabidopsis. *Plant Physiol*. 117, 1227–1234. doi: 10.1104/pp.117.4.1227

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8, 1–12. doi: 10.1093/gigascience/giy154

Izquierdo, P., Astudillo, C., Blair, M. W., Iqbal, A. M., Raatz, B., and Cichy, K. A. (2018). Meta-QTL analysis of seed iron and zinc concentration and content in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet*. 131, 1645–1658. doi: 10.1007/s00122-018-3104-8

Jeong, J., Cohu, C., Kerkeb, L., Pilon, M., Connolly, E. L., and Guerinot, M. L. (2008). Chloroplast Fe(III) chelate reductase activity is essential for seedling viability under iron limiting conditions. *Proc. Natl. Acad. Sci. U.S.A*. 105, 10619–10624. doi: 10.1073/pnas.0708367105

Kamfwa, K., Cichy, K. A., and Kelly, J. D. (2015). Genome-wide association study of agronomic traits in common bean. *Plant Genome* 8, 1–12. doi: 10.3835/plantgenome2014.09.0059

Kassambara, A., and Mundt, F. (2020). factoextra: extract and visualize the results of multivariate data analyses. Available online at: https://CRAN.R-project.org/package=factoextra

Katungi, E., Larochelle, C., Mugabo, J., and Buruchara, R. (2019). Climbing bean as a solution to increase productivity in land-constrained environments: evidence from Rwanda. *Outlook Agric.* 48, 28–36. doi: 10.1177/00307270188 13698

Keller, B., Ariza-Suarez, D., de la Hoz, J., Aparicio, J. S., Portilla-Benavides, A. E., Buendia, H. F., et al. (2020). Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Front. Plant Sci.* 11, 1001. doi: 10.3389/fpls.2020.01001

Kelly, J. D., and Bornowski, N. (2018). "Marker-assisted breeding for economic traits in common bean," in *Biotechnologies of Crop Improvement, Volume 3: Genomic Approaches*, eds S. S. Gosal and S. H. Wani (Cham: Springer International Publishing), 211–238.

Koinange, E. M. K., Singh, S. P., and Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Sci.* 36, 1037–1045. doi: 10.2135/cropsci1996.0011183X003600040037x

Kornegay, J., White, J. W., and de la Cruz, O. O. (1992). Growth habit and gene pool effects on inheritance of yield in common bean. *Euphytica* 62, 171–180. doi: 10.1007/BF00041751

Kwak, M., Toro, O., Debouck, D. G., and Gepts, P. (2012). Multiple origins of the determinate growth habit in domesticated common bean (*Phaseolus vulgaris*). *Ann. Bot.* 110, 1573–1580. doi: 10.1093/aob/mcs207

Kwak, M., Velasco, D., and Gepts, P. (2008). Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (*Phaseolus vulgaris*). *J. Heredity* 99, 283–291. doi: 10.1093/jhered/esn.005

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an r package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01

Lehermeier, C., Schön, C.-C., and Campos, G. d. l. (2015). Assessment of genetic heterogeneity in structured plant populations using multivariate whole-genome regression models. *Genetics* 201, 323–337. doi: 10.1534/genetics.115.177394

Mangin, B., Siberchicot, A., Nicolas, S., Doligez, A., This, P., and Cierco-Ayrolles, C. (2012). Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. *Heredity* 108, 285–291. doi: 10.1038/hdy.2011.73

Mayor Duran, V. M. (2016). *Identificación de QTLs de frijol común (Phaseolus vulgaris) asociados a tolerancia a sequía* (Ph.D. thesis). Universidad Nacional de Colombia, Palmira.

Mayor Duran, V. M., Raatz, B., and Blair, M. W. (2016). Desarrollo de líneas de frijol (*Phaseolus vulgaris* L.) tolerante a sequía a partir de cruces inter acervo con genotipos procedentes de diferentes orígenes (Mesoamericano y Andino). *Acta Agronómica* 65, 431–438. doi: 10.15446/acag.v65n4.48680

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Miklas, P. N., Kelly, J. D., Beebe, S. E., and Blair, M. W. (2006). Common bean breeding for resistance against biotic and abiotic stresses: from classical to MAS breeding. *Euphytica* 147, 105–131. doi: 10.1007/s10681-006-4600-5

Möhring, J., and Piepho, H.-P. (2009). Comparison of weighting in two-stage analysis of plant breeding trials. *Crop Sci.* 49, 1977–1988. doi: 10.2135/cropsci2009.02.0083

Mukamuhirwa, F., and Rurangwa, E. (2018). Evaluation for high iron and zinc content among selected climbing bean genotypes in rwanda. *Adv. Crop Sci. Technol.* 6, 1–6. doi: 10.4172/2329-8863.1000344

Mukeshimana, G., Butare, L., Cregan, P. B., Blair, M. W., and Kelly, J. D. (2014). Quantitative trait loci associated with drought tolerance in common bean. *Crop Sci.* 54, 923–938. doi: 10.2135/cropsci2013. 06.0427

Nay, M. M., Mukankusi, C. M., Studer, B., and Raatz, B. (2019). Haplotypes at the Phg-2 locus are determining pathotype-specificity of angular leaf spot resistance in common bean. *Front. Plant Sci.* 10, 1–11. doi: 10.3389/fpls.2019.01126

Norton, J. (1915). *Inheritance of habit in the common bean* (Masters Theses). 1911-February 2014.

Pérez, P., and de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198, 483. doi: 10.1534/genetics.114.164442

Petry, N., Boy, E., Wirth, J. P., and Hurrell, R. F. (2015). Review: the potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification. *Nutrients* 7, 1144–1173. doi: 10.3390/nu7021144

Petry, N., Egli, I., Gahutu, J. B., Tugirimana, P. L., Boy, E., and Hurrell, R. (2014). Phytic acid concentration influences iron bioavailability from biofortified beans in Rwandese women with low iron status. *J. Nutr.* 144, 1681–1687. doi: 10.3945/jn.114.192989

Puig, S., Andrés-Colás, N., García–Molina, A., and Peñarrubia, L. (2007). Copper and iron homeostasis in Arabidopsis: responses to metal deficiencies, interactions and biotechnological applications. *Plant, Cell Environ.* 30, 271–290. doi: 10.1111/j.1365-3040.2007.01642.x

Raggi, L., Caproni, L., Carboni, A., and Negri, V. (2019). Genome-wide association study reveals candidate genes for flowering time variation in common bean (*Phaseolus vulgaris* L.). *Front. Plant Sci.* 10, 962. doi: 10.3389/fpls.2019.00962

Rehman, H. M., Cooper, J. W., Lam, H.-M., and Yang, S. H. (2019). Legume biofortification is an underexploited strategy for combatting hidden hunger. *Plant Cell Environ.* 42, 52–70. doi: 10.1111/pce.13368

Repinski, S. L., Kwak, M., and Gepts, P. (2012). The common bean growth habit gene PvTFL1y is a functional homolog of Arabidopsis TFL1. *Theor. Appl. Genet.* 124, 1539–1547. doi: 10.1007/s00122-012-1808-8

Robinson, N. J., Procter, C. M., Connolly, E. L., and Guerinot, M. L. (1999). A ferric-chelate reductase for iron uptake from soils. *Nature* 397, 694–697. doi: 10.1038/17800

Rodriguez, M., Rau, D., Bitocchi, E., Bellucci, E., Biagetti, E., Carboni, A., et al. (2016). Landscape genetics, adaptive diversity and population structure in Phaseolus vulgaris. *New Phytol.* 209, 1781–1794. doi: 10.1111/nph.13713

Rodríguez-Álvarez, M. X., Boer, M. P., van Eeuwijk, F. A., and Eilers, P. H. (2018). Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Stat.* 23, 52–71. doi: 10.1016/j.spasta.2017.10.003

Ronner, E., Descheemaeker, K., Marinus, W., Almekinders, C. J. M., Ebanyat, P., and Giller, K. E. (2018). How do climbing beans fit in farming systems of the eastern highlands of Uganda? Understanding opportunities and constraints at farm level. *Agric. Syst.* 165, 97–110. doi: 10.1016/j.agsy.2018.05.014

Rosales-Serna, R., Kohashi-Shibata, J., Acosta-Gallegos, J. A., Trejo-López, C., Ortiz-Cereceres, J., and Kelly, J. D. (2004). Biomass distribution, maturity acceleration and yield in drought-stressed common bean cultivars. *Field Crops Res.* 85, 203–211. doi: 10.1016/S0378-4290(03)00161-8

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., et al. (2019). Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theor. Appl. Genet.* 132, 1247–1261. doi: 10.1007/s00122-019-03276-6

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008

Self, S. G., and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.* 82, 605–610. doi: 10.1080/01621459.1987.10478472

Singh, S. P. (1981). *A key for identification of different growth habits of Phaseolus vulgaris L.* Technical report, International Center for Tropical Agriculture (CIAT), Cali, Colombia.

Spindel, J. E., and McCouch, S. R. (2016). When more is better: how data sharing would accelerate genomic selection of crop plants. *New Phytol.* 212, 814–826. doi: 10.1111/nph.14174

Stangoulis, J., and Sison, C. (2008). Crop sampling protocols for micronutrient analysis. *Harvest Plus Tech. Monogr. Ser.* 7, 1–20.

Stein, A. J. (2010). Global impacts of human mineral malnutrition. *Plant Soil.* 335, 133–154. doi: 10.1007/s11104-009-0228-2

Sul, J. H., Martin, L. S., and Eskin, E. (2018). Population structure in genetic studies: confounding factors and mixed models. *PLoS Genet.* 14, e1007309. doi: 10.1371/journal.pgen.1007309

Teixeira, F. F., Ramalho, M. A. P., and Abreu, N. D. F. B. (1999). Genetic control of plant architecture in the common bean (*Phaseolus vulgaris* L.). *Genet Mol Biol.* 22, 577–582. doi: 10.1590/S1415-47571999000400019

Tello, D., Gil, J., Loaiza, C. D., Riascos, J. J., Cardozo, N., and Duitama, J. (2019). NGSEP3: accurate variant calling across species and sequencing protocols. *Bioinformatics* 35, 4716–4723. doi: 10.1093/bioinformatics/btz275

Waters, B. M., and Grusak, M. A. (2008). Quantitative trait locus mapping for seed mineral concentrations in two *Arabidopsis thaliana*

recombinant inbred populations. *New Phytol.* 179, 1033–1047. doi: 10.1111/j.1469-8137.2008.02544.x

Weller, J. L., Vander Schoor, J. K., Perez-Wright, E. C., Hecht, V., González, A. M., Capel, C., et al. (2019). Parallel origins of photoperiod adaptation following dual domestications of common bean. *J. Exp. Bot.* 70, 1209–1219. doi: 10.1093/jxb/ery455

White, J., Kornegay, J., Castillo, J., Molano, C., Cajiao, C., and Tejada, G. (1992). Effect of growth habit on yield of large-seeded bush cultivars of common bean. *Field Crops Res.* 29, 151–161. doi: 10.1016/0378-4290(92)90084-M

White, J. W., and Izquierdo, J. (1989). "Dry bean: physiology of yield potential and stress tolerance," in *Common Beans: Research for Crop Improvement* (Santiago: CAB International).

Wickham, H. (2016). "Toolbox," in *ggplot2: Elegant Graphics for Data Analysis, Use R*!, ed H. Wickham (Cham: Springer International Publishing), 33–74.

Wu, H., Li, L., Du, J., Yuan, Y., Cheng, X., and Ling, H.-Q. (2005). Molecular and biochemical characterization of the Fe(III) chelate reductase gene family in *Arabidopsis thaliana*. *Plant Cell Physiol.* 46, 1505–1514. doi: 10.1093/pcp/pci163

Wu, J., Wang, L., Fu, J., Chen, J., Wei, S., Zhang, S., et al. (2020). Resequencing of 683 common bean genotypes identifies yield component trait associations across a north–south cline. *Nat. Genet.* 52, 118–125. doi: 10.1038/s41588-019-0546-0

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evolut.* 8, 28–36. doi: 10.1111/2041-210X.12628

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Keller, Ariza-Suarez, Portilla-Benavides, Buendia, Aparicio, Amongi, Mbiu, Msolla, Miklas, Porch, Burridge, Mukankusi, Studer and Raatz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 26 May 2022
doi: 10.3389/fpls.2022.892077

Genetic Mapping of the *Gmpgl3* Mutant Reveals the Function of GmTic110a in Soybean Chloroplast Development

*Hui Yu[1], Qiushi Wang[1], Zhirui Zhang[1,2], Tao Wu[1,2], Xinjing Yang[1,2], Xiaobin Zhu[1,2], Yongheng Ye[1,2], Jiantian Leng[1], Suxin Yang[1] and Xianzhong Feng[1,2,3]**

[1]Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun, China, [2]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China, [3]Zhejiang Lab, Hangzhou, China

The generation of oxygen and organic matter in plants mainly depends on photosynthesis, which directly affects plant growth and development. The chloroplast is the main organelle in which photosynthesis occurs. In this study, a *Glycine max* pale green leaf 3-1 (*Gmpgl3-1*) mutant was isolated from the soybean mutagenized population. The *Gmpgl3-1* mutant presented with decreased chlorophyll contents, reduced chloroplast stroma thylakoids, reduced yields, and decreased numbers of pods per plant. Bulked segregant analysis (BSA) together with map-based cloning revealed a single-nucleotide non-synonymous mutation at the 341st nucleotide of the first exon of the chloroplast development-related *GmTic110a* gene. The phenotype of the knockout plants was the same as that of the mutant. The *GmTic110a* gene was highly expressed in the leaves at various developmental stages, and its protein was localized to the inner chloroplast membrane. Split luciferase complementation assays and coimmunoprecipitation (co-IP) experiments revealed that GmTic110a interacted with GmTic20, GmTic40a, and GmTic40b in tobacco leaves. These results indicated that the *GmTic110a* gene plays an important role in chloroplast development.

Keywords: soybean, *Gmpgl3* mutant, gene mapping, GmTic110a, GmTic20, GmTic40a/b

OPEN ACCESS

Edited by:
Petr Smýkal,
Palacký University Olomouc, Czechia

Reviewed by:
Jean-David Rochaix,
Université de Genève, Switzerland
Shinji Masuda,
Tokyo Institute of Technology, Japan

***Correspondence:**
Xianzhong Feng
fengxianzhong@iga.ac.cn

Specialty section:
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: *08 March 2022*
Accepted: *09 May 2022*
Published: *26 May 2022*

Citation:
Yu H, Wang Q, Zhang Z, Wu T, Yang X, Zhu X, Ye Y, Leng J, Yang S and Feng X (2022) Genetic Mapping of the Gmpgl3 Mutant Reveals the Function of GmTic110a in Soybean Chloroplast Development. Front. Plant Sci. 13:892077. doi: 10.3389/fpls.2022.892077

INTRODUCTION

Plant leaves are the most important tissues for photosynthesis. Chlorophyll is an important pigment involved in photosynthesis in chloroplasts (Waters and Langdale, 2009), and the development of plant chloroplasts positively correlates with the chlorophyll content in leaves (Davis and Fajer, 1979; Wang et al., 2003) and leaf photosynthesis rates (Peng et al., 2008). Mutations in chlorophyll synthesis-related genes can directly or indirectly affect chlorophyll biosynthesis or degradation pathways, leading to the loss of plant chlorophyll, thereby affecting the photosynthesis of plants and causing yellow leaves, albinism, striped leaf spots, purple–brown patches, or other characteristics of chlorophyll-deficient mutants (Awan et al., 1980). To date, researchers have studied chlorophyll-deficient mutants of *Arabidopsis thaliana* (Carol et al., 1999), tobacco (Okabe and Straub, 1977), corn (Lonosky et al., 2004), rice (Ki-Hong et al., 2003), soybean (Stockinger and Walling, 1994),

pea (Highkin et al., 1969), wheat (Cao et al., 2006), barley (Preiss and Thornber, 1995) and other plant species. Chlorophyll-deficient mutants usually present with a decreased photosynthesis rate and reduced yields, and death can occur in severe cases.

Plant chloroplasts synthesize important amino acids through photosynthesis and are the main sources of energy for plant cells. Chloroplasts play an important role in plant growth and cellular metabolism (Tiller and Bock, 2014). The transport of substances in and out of chloroplasts depends on the translocon at the outer envelope membrane of chloroplasts (TOC) and the translocon at the inner envelope membrane of chloroplasts (TIC; Soll, 2004; Jocelyn and Paul, 2005; Jarvis, 2008). The TOC and TIC form a complex to facilitate this process. The chloroplast transport proteins on the outer chloroplast membrane identified to date include Toc159 (Waegemann and Soil, 1991; Schnell et al., 1994), Toc34 (Kessler et al., 1994; Schnell et al., 1994), Toc75 (Tranel et al., 1995; Sveshnikova et al., 2000) and Toc64 (Sohrt and Soll, 2000; Becker et al., 2004). Recent extensive studies have significantly updated our understanding of the components and mechanisms of the chloroplast translocon machinery (Kikuchi et al., 2013, 2018; Nakai, 2015, 2018, 2020). These studies have significantly revised the long-accepted "classical" model for chloroplast protein import: In the classical model, Tic110 (Ishida and Terakura, 1987; Inaba et al., 2005), Tic40 (Stahl et al., 1999; Chou et al., 2003), Tic20 (Kouranov and Schnell, 1997; Kasmati et al., 2011), and Tic21 (Vitale et al., 2015) are the main components, but they are not found in the translocon proposed by Nakai (2015, 2018, 2020). In the revised model, the 1-megadalton TIC complex consists of Tic214 (ycf1; de Vries et al., 2015; Bölter and Soll, 2017), Tic100 (Oshima et al., 1987; Ramundo et al., 2020), Tic56 (Köhler et al., 2015, 2016), Tic20 (Kikuchi et al., 2009), and Tic21 (Kouranov and Schnell, 1997; Teng et al., 2006), which functionally and physically cooperate with the ATP-driven import motor YCF2/FTSHI complex (Kikuchi et al., 2013, 2018; Thomson et al., 2020).

Tic110 is an important chloroplast inner membrane protein (Schnell et al., 1994; Inaba et al., 2005; Balsera et al., 2009). It has been reported that the Tic110 protein interacts with several molecular chaperones, such as Hsp93 and Hsp70, to form an inner membrane transport channel scaffold that ensures the successful import of various proteins into the chloroplast to perform cell biological functions (Inaba et al., 2003). Tic110 proteins interact with Tic32 proteins to perform redox functions and to regulate Ca^{2+} homeostasis in the chloroplast (Hormann et al., 2004). In addition, Tsai et al. (2013) reported that Tic110 is most likely a scaffolding component important for protein–protein interactions to recruit other translocon components and chaperones in the stroma (Tsai et al., 2013). In Arabidopsis and soybean, a lack of Tic110 blocks the transport of the inner and outer chloroplast membranes, affecting the development of chloroplasts and resulting in yellow leaves (Inaba et al., 2005; Sandhu et al., 2016). Tic110 assists in the formation of a scaffold for the assembly of the ATP-dependent import motor in the stroma (Richardson and Schnell, 2020). Soybean is an important source of grain and oil. This species is also the main source of high-quality protein for human diets and animal feed. As such, soybean occupies an important position in grain production worldwide. Obstruction of plant chloroplast development could lead to yellow leaves, which severely affects photosynthesis and plant yield; in severe cases, this results in dwarf-type plants or even plants that fail to produce harvestable yields. Therefore, it is of great scientific importance to study the regulatory molecular mechanisms of chloroplast membrane transport proteins.

In this study, we report the characterization of a *Glycine max* pale green leaf mutant (*Gmpgl3-1*). The chloroplast development-related gene *GmTic110a* encodes a chloroplast inner membrane protein. *Gmpgl3-1*, *Gmpgl3-2,* and *Gmpgl3-3* are allelic mutants of the *GmTic110a* gene. The biological function of the *GmTic110a* gene was preliminarily analyzed in this research. We used split luciferase complementation and coimmunoprecipitation (co-IP) analyses, and the results indicated that GmTic110a proteins can interact with other GmTic proteins in tobacco. Our research lays a theoretical foundation for studies of the molecular mechanism underlying soybean chloroplast development.

MATERIALS AND METHODS

Plant Material

All plants used in this study were grown at the Changchun Agricultural Station, Northeast Institute of Geography and Agroecology, Chinese Academy of Science (CAS), China. A *Gmpgl3-1* mutant was isolated from an M_2 population induced by ethylmethanesulfonate (EMS). The *Gmpgl3-1* mutant was backcrossed to Williams 82 five times from 2014 to 2018 to purify the genetic background of the *Gmpgl3-1* mutant. For protoplast isolation, Arabidopsis seeds of the Columbia ecotype (Col-0) were surface sterilized, vernalized, and then sown on 1/2-strength Murashige and Skoog (MS) media until the seedlings reached the four-leaf stage. Then, the seedlings were grown in pots containing peat moss and vermiculite (1/1, v/v) in a growth chamber under 150 μmol m^{-2} s^{-1} irradiance and a 14 h dark/10 h light photoperiod at 25°C, and the relative humidity was maintained at 60–75%. Leaves were collected from 3- to 4-week-old seedlings for transfection assays.

Mapping of *GmTic110a via* Bulked Segregant Analysis

Three F_2 populations derived from a cross between the *Gmpgl3-1* mutant and the Chinese soybean cultivar Hedou 12 were used to map the *GmTic110a* gene. DNA from 50 F_2 individuals with the *Gmpgl3-1* mutant phenotype and 50 F_2 individuals with the wild-type phenotype were bulked

Abbreviations: TOC, Translocon at the Outer Envelope Membrane of Chloroplasts; TIC, Translocon at the Inner Envelope Membrane of Chloroplasts; EMS, Ethyl Methanesulfonate; BSA, Bulked Segregation Analysis; SNP, Single-Nucleotide Polymorphism; NCBI, National Center for Biotechnology Information; MEGA 7.0, Molecular Evolutionary Genetics Analysis Software, Version 7.0; MEME, Multiple Expectation Maximization for Motif Elicitation; Pn, Photosynthetic Rate; Gs, Stomatal Conductance; Ci, Intercellular CO_2 Concentration; Tr, Transpiration Rate; F0, Initial Fluorescence; Fm, Maximal Fluorescence; Fv, Variable Fluorescence; cDNA, Complementary DNA; qRT–PCR, Real-Time Quantitative PCR.

into mutant and wild-type pools, respectively. Insertion–deletion (INDEL) markers for preliminary mapping were used according to a previously described method (Song et al., 2015). New molecular markers for fine mapping were generated; these markers are listed in **Supplementary Table S4**. The candidate genomic regions were identified *via* BSA of the F_2 population at a depth of approximately 30× using an Illumina HiSeq 2000 device (Illumina Inc., San Diego, CA, United States). The Genome Analysis Toolkit (GATK, version 3.8) was used to detect single-nucleotide polymorphisms (SNPs; Mckenna et al., 2010). Genomic regions in which Δ(SNP index) was >0.5 were selected as candidate regions.

Database Searching and Phylogenetic Analysis

GmTic110a homologs were identified by querying the GmTic110a sequence in the NCBI[1] database *via* the BLASTP program. Multiple sequence alignments were performed using ClustalX version 2.0 (Larkin et al., 2007) and were manually corrected. The obtained sequence was used as input to construct an unrooted phylogenetic tree with the neighbor-joining algorithm *via* the Molecular Evolutionary Genetics Analysis version 7.0 (MEGA 7.0) phylogenetic program (Sudhir et al., 2016). Bootstrap analysis was performed using 1,000 replicates. The protein motifs of GmTic110a-like genes were subsequently profiled by Multiple Expectation maximization for Motif Elicitation (MEME; Bailey et al., 2009).

Determination of Pigment Contents and Chlorophyll Fluorescence Analysis

To determine pigment contents, leaves of 21-day-old *Gmpgl3-1* mutants and Williams 82 plants were collected and measured according to a previously reported procedure (Gregor and Marsálek, 2004). The pigment contents were calculated according to the following formulas: chlorophyll a = 13.95*A665-6.88*A649; chlorophyll b = 24.96*A649-7.32*A665; and carotenoids = (1,000*A470-2.05*Ca-114.8*Cb)/245. The photosynthesis rate (Pn), stomatal conductance (Gs), intercellular CO_2 concentration (Mckenna et al.), and transpiration rate (Tr) of the leaves were measured using an LI-6400 photosynthesis system (LI-COR, Lincoln, NE, United States; Yamori et al., 2011), and the initial fluorescence (F0), maximal fluorescence (Fm), and variable fluorescence (Fv) values were measured using an OS-30p chlorophyll fluorometer (Opti-Sciences, Hudson, NY, United States). The maximum quantum yield of photosystem II (Fv/Fm) and the maximum photochemical yield of PSII (Fv/F0) were calculated as previously described (Genty et al., 1989). The plants were dark-adapted for 30 min before measurement. All the measurements involved the use of ten plants and were performed from 11:00 am to noon during the beginning of the flowering period. The operation of the machine and subsequent analysis were performed according to the manufacturer's instructions.

Transmission Electron Microscopy Analysis

Williams 82 and *Gmpgl3-1* mutant plants grown for 21 days were selected. The leaves were cut into rectangular pieces that were approximately 2 mm*1 mm, and the plant materials were vacuum fixed in 2.5% glutaraldehyde solution with 0.2 mol of phosphate buffer. The samples were postfixed for 3 h in 1% osmium tetroxide at 4°C. The samples were then treated according to previously described methods (Kowalewska et al., 2016). Ultrathin sections were obtained using an MT-X (RMC, Tucson, AZ, United States) ultramicrotome and stained with uranyl acetate for 20 min followed by lead citrate for 10 min. Observations of the samples and recording of images were performed using a Hitachi H-7650 electron microscope (Tokyo, Japan).

CRISPR/Cas9 Vector Construction and Soybean Transformation

To obtain *GmTic110a*-knockout plants, the CRISPR/Cas9 gene editing system for targeted genome modification of plants was used (Shan et al., 2013). Several 20-nt single-guide RNAs (sgRNAs) highly specific for Cas9 target sites were identified using the web-based tool CRISPR-P version 2.0[2] (Lu et al., 2017). A pair of 24-bp long oligonucleotides (5′-GATTGCGGCGGCTGGATACGGCCT-3′ and 5′-AAACAGGCCGTATCCAGCCGCCGC-3′) specific to the *GmTic110a* sequence were annealed and cloned into a modified *VK005-04-soU6-2-GmUbi3* knockout expression vector (Du et al., 2016). The resulting recombinant plasmid (*VK005-GmTic110a*) was introduced into *Agrobacterium tumefaciens EHA105*, which was then used to transform Williams 82 cotyledonary explants (Zhao et al., 2016). Three independent *GmTic110a*-knockout transgenic plants were obtained for further phenotypic analysis.

Analysis of the Expression Profile of the *GmTic110a* Gene

New leaves at the VE (emergence) stage; stem tips, stems, and roots at the V1 (first unrolled trifoliate leaf) stage; leaves and flowers at the R1 (beginning bloom) stage; leaves and flowers at the R2 (full bloom) stage; and leaves at the R3 (beginning of pod development) stage were collected. Total RNA was subsequently extracted using TRIzol reagent (Tiangen, Lot 118,721; China) according to the manufacturer's instructions. The integrity of the RNA was determined through agarose gel electrophoresis, and complementary DNA (cDNA) was synthesized using 5 μg of RNA with oligo$(dT)_{18}$ primers and Moloney murine leukemia virus (M-MLV) reverse transcriptase (TransGen Lot N31204; China) according to the manufacturer's protocol. Relative transcript levels of *GmTic110a* were analyzed through real-time quantitative PCR (qRT–PCR) on an Mx3005P instrument (Stratagene, La Jolla, CA, United States) in conjunction with SYBR Green Master Mix (Genstar Lot 9 BC01; China). The PCR parameters were 95°C for 30 s (1 cycle), 95°C for 5 s, and 60°C for 20 s (40 cycles), which was followed by a

[1]http://www.ncbi.nlm.nih.gov/

[2]http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR

melting curve analysis at 95°C for 60 s, 55°C for 30 s, and 95°C for 30 s. The internal control gene *GmActin11* (*Glyma.18G290800*) was used for normalization of the transcript levels of *GmTic110a* in the samples (Hu et al., 2009). The relative fold differences were calculated *via* the $2^{-\Delta\Delta Ct}$ method. Three independent biological replicates were used to confirm the expression profiles. The specific primer pairs used are listed in **Supplementary Table S4**.

Subcellular Localization Analysis

We next sought to determine the subcellular localization of GmTic110a, GmTic110a^{G114A}, Gmtic110a^{T805S}, GmTic110a^{CR1}, GmTic110a^{CR2}, and GmTic110a^{CR3} from the knockout transgenic strains. For this analysis, the full-length cDNA sequence and the mutated and knockout sequences of *GmTic110a* were cloned into pUC19-GFP (Zheng et al., 2017), and the resulting recombinant plasmid was transiently introduced into Arabidopsis (Col-0) protoplasts using 20% polyethylene glycol (Yu et al., 2015). The fluorescence signals were visualized using a Nikon C2 confocal laser scanning microscope (Japan) under a 488 nm excitation wavelength and 495–540 nm emission wavelengths to determine the subcellular localization of GmTic110a, GmTic110a^{G114A}, Gmtic110a^{T805S}, GmTic110a^{CR1}, GmTic110a^{CR2}, and GmTic110a^{CR3}. Chloroplast autofluorescence was detected at wavelengths of 488 nm (excitation) and 680–700 nm (emission). Image processing was performed with ImageJ.[3] The specific primer pairs used are listed in **Supplementary Table S4**.

Luciferase Complementation Assays

Luciferase complementation assays were performed as described previously (Chen et al., 2008), with minor modifications. The coding DNA sequences (CDSs) of *GmTic110a*, *GmTic20*, *GmTic40a*, and *GmTic40b* were cloned into either a *pCAMBIA1300-NLuc* or a *pCAMBIA1300-CLuc* vector. *pCAMBIA1300-GmTic110a-NLuc*, *pCAMBIA1300-GmTic20-CLuc*, *pCAMBIA1300-GmTic40a-CLuc*, and *pCAMBIA1300-GmTic40b-CLuc* in various combinations were transferred into *Nicotiana benthamiana* leaves by *A. tumefaciens*-mediated transformation. Agrobacterium cells with N-Luc and C-Luc vectors were resuspended in infiltration buffer (pH 5.6; 10 mm 2-(N-morpholino) ethanesulfonic acid, 10 mm $MgCl_2$, and 150 mm acetosyringone) to reach an optimal optical density at 600 nm in the range of 0.9 to 1. After 3 h of incubation at room temperature, the suspensions were infiltrated into the leaves of 4-week-old *N. benthamiana* plants, which were then cultivated for 2 days at 23°C. To inject tobacco leaves with 1 mmol of luciferin (115144–35-9, GoldBio) for measurements of luciferase activity, the leaves were maintained in the dark for 5 min. Images were captured using a chemiluminescence image analysis system (4600SF, Tanon). The sequences of the primers used are listed in **Supplementary Table S4**.

[3]http://rsb.info.nih.gov/ij/

Co-IP Assays

Co-IP assays were performed as described previously (Zhou et al., 2015; Liu et al., 2020; Wang et al., 2020), with minor modifications. The CDSs of *GmTic110a*, *GmTic20*, *GmTic40a*, and *GmTic40b* were cloned into either a *pCAMBIA1300-FLAG* or *pCAMBIA1300-HA* vector, resulting in *pCAMBIA1300-GmTic110a-FLAG*, *pCAMBIA1300-GmTic20-HA*, *pCAMBIA1300-GmTic40a-HA*, and *pCAMBIA1300-GmTic40b-HA* vectors. To measure protein–protein interactions, *A. tumefaciens* strain *EHA105* containing pairs of these constructs together with *pCAMBIA1300-GmTic110a-FLAG*, *pCAMBIA1300-GmTic20-HA*, *pCAMBIA1300-GmTic40a-HA*, and *pCAMBIA1300-GmTic40b-HA* were coinfiltrated into the leaves of 4-week-old *N. benthamiana* plants. Samples (1 g each) were then collected at 3 days after infiltration, ground in liquid nitrogen and homogenized in 1.5 ml of extraction buffer (50 mm Tris–HCl [pH 7.5], 150 mm NaCl, 1 mm EDTA [pH 8.0], 0.2% [v/v] Triton X-100, 20% [v/v] glycerol, and 1× protease inhibitor cocktail [pH 7.5]). The lysates were incubated at 4°C for 30 min and subsequently centrifuged at 15,000 × g for 30 min at 4°C. After instantaneous centrifugation, the supernatants were added to 500-µl suspensions of anti-DDDDK-tag-FLAG magnetic beads (No. M185-11R, Medical and Biological Laboratories), incubated at 4°C for 4 h, and then washed 4 times with extraction buffer. The proteins were eluted from the beads with 30 µl of 1 × Protein Loading Buffer, boiled for 5 min, and then centrifuged at 8,000 × g for 1 min at room temperature. The supernatants were electrophoretically separated *via* 10% SDS–PAGE and transferred to a nitrocellulose membrane (No. q0600003, GE Healthcare Life Sciences). Immunoblots were performed using an anti-FLAG antibody (1:5000; No. M180-5, Medical and Biological Laboratories) for probing *pCAMBIA1300-GmTic110a-FLAG* and an anti-HA antibody (1/5000, No. M180-3, Medical and Biological Laboratories) for probing *pCAMBIA1300-GmTic20-HA*, *pCAMBIA1300-GmTic40a-HA*, or *pCAMBIA1300-GmTic40b-HA*. The sequences of the primers used are listed in **Supplementary Table S4**.

RESULTS

Phenotypic Characterization of the Chloroplast Development-Related Mutant *Gmpgl3*

Compared with the wild-type Williams 82, the *Gmpgl3-1* mutant showed a pale green leaf phenotype from the seedling stage to the mature stage (**Figure 1A**). Moreover, on the basis of their phenotypes and genotypes, we identified two allelic mutants named *Gmpgl3-2* and *Gmpgl3-3* from within the mutant library. The pale green leaf phenotype was observed for the *Gmpgl3-2* mutant (**Figure 1A**), while the pale green leaf phenotype was not observed for the *Gmpgl3-3* mutant (**Supplementary Figure S1**). We analyzed mature Williams 82, *Gmpgl3-1,* and *Gmpgl3-2* plants separately. The results showed that the overall heights of the *Gmpgl3-1* and *Gmpgl3-2* mutants were reduced by 7 and 2.5%, respectively, compared with

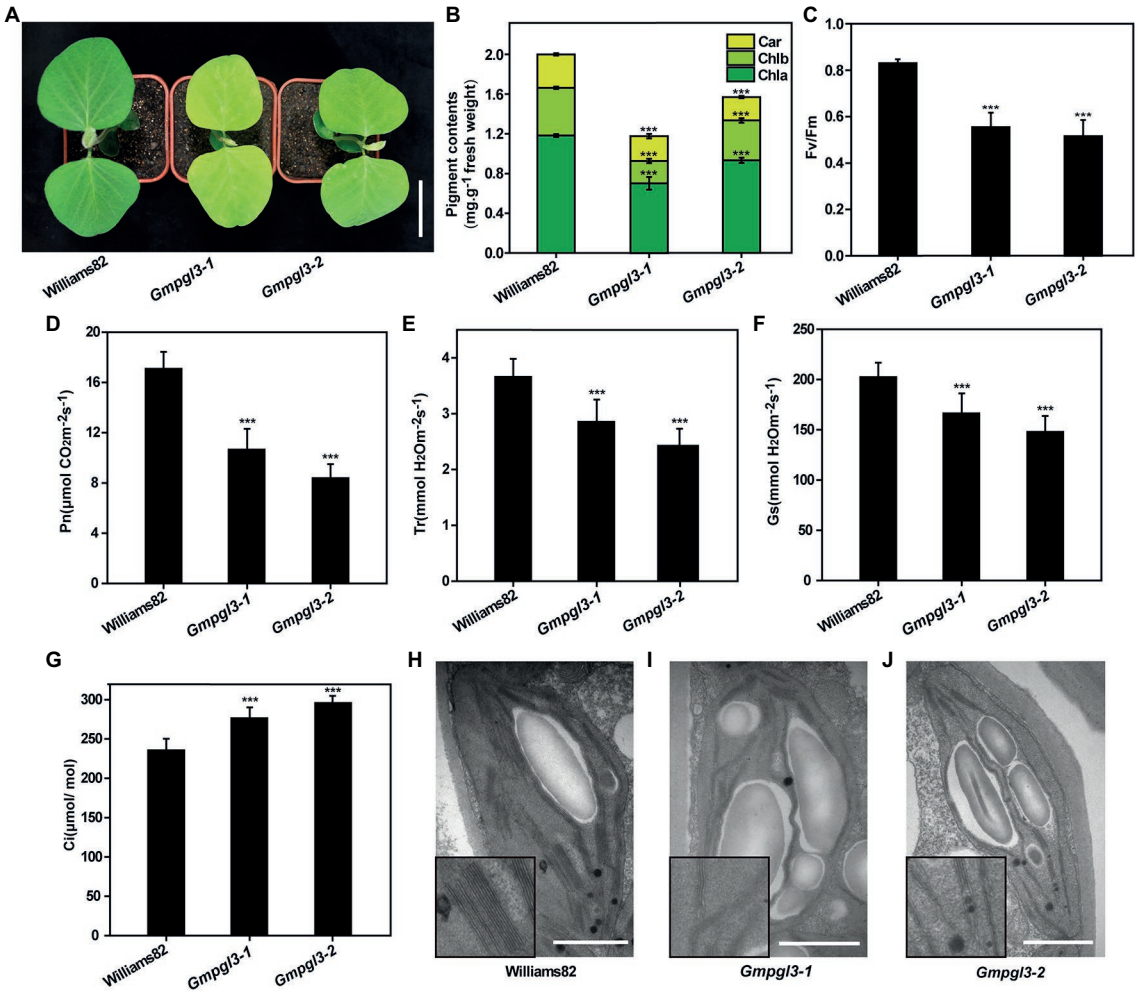


FIGURE 1 | (A) Phenotypes of the Williams 82, *Gmpgl3-1,* and *Gmpgl3-2* mutants. Scale bars, 5 cm. **(B–G)** Pigment contents and photosynthesis parameters of the Williams 82, *Gmpgl3-1,* and *Gmpgl3-2* mutants. **(B)** Chlorophyll a (Chl a), chlorophyll b (Chl b) and carotenoid (Car). **(C)** Chlorophyll fluorescence parameter (Fv/FM). **(D)** Photosynthetic rate (Pn). **(E)** Transpiration rate (Tr). **(F)** Stomatal conductance (Gs). **(G)** Intercellular CO_2 concentration (Ci). **(H)** Chloroplast structure in Williams 82. **(I,J)** Chloroplast structure in the *Gmpgl3-1* and *Gmpgl3-2* mutants. Scale bars, 1 μm. *** represents significant differences compared with the control (Williams 82) at $p < 0.001$, and the error bars represent standard deviations.

Williams 82, while the number of nodes was reduced by 17.3%. Moreover, compared with Williams 82, the number of pods per plant for the *Gmpgl3-1* and *Gmpgl3-2* mutants was reduced by 30.8 and 39.5%, respectively; the number of grains per plant for the *Gmpgl3-1* and *Gmpgl3-2* mutants was reduced by 39.5 and 39.2%, respectively; the grain weight per plant for the *Gmpgl3-1* and *Gmpgl3-2* mutants was reduced by 50.8 and 48.2%, respectively; and the 100-seed weight for the *Gmpgl3-1* and *Gmpgl3-2* mutants was reduced by 18.1 and 14.8%, respectively (**Supplementary Table S1**).

Because the *Gmpgl3-1* and *Gmpgl3-2* mutants exhibited a pale green leaf phenotype throughout the growth period, the chlorophyll and carotenoid contents of Williams 82, *Gmpgl3-1*, and *Gmpgl3-2* leaves were measured spectrophotometrically, and the results showed that both chlorophyll a and chlorophyll b contents in the *Gmpgl3-1* and *Gmpgl3-2* mutants were significantly lower than those in the wild type (**Figure 1B**). This finding indicates that the *GmTic110a* gene mutation may affect the stability of Chl a/b. The chlorophyll content of the *Gmtic110a* mutant was reduced by 44.2%. The chloroplast fluorescence Fv/Fm values of *Gmpgl3-1* and *Gmpgl3-2* were significantly lower than those of Williams 82, which showed that the photosynthetic efficiency of *Gmpgl3-1* and *Gmpgl3-2* was lower than that of Williams 82 (**Figure 1C**). The photosynthetic rate (Pn) of the *Gmpgl3-1* and *Gmpgl3-2* mutants was reduced by 37.7 and 50.9%, respectively, compared with that of Williams 82. In addition, compared with those of the wild type, the stomatal conductance (Gs) of the *Gmpgl3-1* and *Gmpgl3-2* mutants decreased by 22 and 33.7%, respectively, the transpiration rate (Tr) of the *Gmpgl3-1* and *Gmpgl3-2* mutants decreased by 17.9 and 26.9%, respectively, and the intercellular CO_2 concentration of the *Gmpgl3-1* and *Gmpgl3-2* mutants increased by 17.3 and 25.6%, respectively (**Figures 1D–G**). Taken together, these results showed that the

decrease in the chlorophyll content in the *Gmpgl3-1* and *Gmpgl3-2* mutants significantly affected the photosynthetic ability of the leaves of those plants.

Compared with that in the wild-type (Williams 82) chloroplasts, the number of starch grains in the chloroplasts of the *Gmpgl3-1* and *Gmpgl3-2* mutants increased, and the basal thylakoids became thinner (**Figures 1H–J**). Because there are photosynthetic pigment components on the thylakoid membrane and because photosynthesis mainly occurs within thylakoids, both the chlorophyll a and chlorophyll b contents in the mutants were significantly reduced, which led to thinning of the basal thylakoid membrane and a reduction in the photosynthesis rate.

Genetic Mapping of the *Gmpgl3-1* Mutation Locus From the F_2 Population

The *Gmpgl3-1* mutant was crossed with Hedou 12 to generate a segregating population for mapping the *Gmtic110a* gene. The F_2 segregating population comprised 537 plants: 423 wild-type plants and 114 mutant plants. A 3:1 segregation ratio was observed for the three F_2 segregating populations ($\chi^2 = 1.89$, $df = 1$, $p = 0.17$; **Supplementary Table S2**), indicating that the *Gmpgl3-1* mutant is the result of a single recessive gene. The F_2 population was used to identify the *Gmpgl3-1* locus. A total of 60 INDEL markers covering all 20 chromosomes were used for mapping, and the mapping results showed that *Gmpgl3-1* was restricted to a 2-Mb region (41–43 Mb) on chromosome 02 (**Figures 2A,B**). To finely map the *Gmpgl3-1* locus, we developed 7 INDEL markers, that is, MOL3067, MOL4032, MOL3069, MOL3071, MOL3073, MOL2733, and MOL0699; the *Gmpgl3-1* locus was further narrowed down to a 0.44-Mb region between 41.79 Mb and 42.23 Mb on chromosome 02, which harbors 18 annotated genes (**Figure 2C**). To identify the causal mutation, the DNA from 40 F_2 individuals carrying the *Gmpgl3-1* mutation under homozygous conditions and the DNA of 50 F_2 individuals exhibiting the wild-type phenotype were pooled into a *Gmpgl3-1* bulk and a Williams 82 bulk for further BSA. The *Gmpgl3-1* mutant was resequenced to a depth of approximately 30× using an Illumina HiSeq 2000 device. We identified a G_{-341} to A_{-341} transition in the first exon of *Glyma.02G233700* (**Figure 2D**), which caused a non-synonymous substitution of Gly_{-114} to Asp_{-114} in the predicted protein. No other mutations were discovered among the 18 genes in the candidate *Gmpgl3-1* genomic region (**Figure 2C**; **Supplementary Table S3**). However, a single-base mutation (A–2413 to T–2413) was identified in the 14th exon of the *GmTic110a* gene of the *Gmpgl3-2* mutant (**Figure 2D**). The expression level of *Glyma.02G233700* decreased in the *Gmpgl3-1* and *Gmpgl3-2* mutants (**Supplementary Figure S2**), suggesting that *Glyma.02G233700* is the *GmTic110a* gene.

GmTic110a Encodes an Inner Chloroplast Membrane Protein

BLAST searches revealed that GmTic110a, which encodes a chloroplast inner membrane protein, is highly homologous to the soybean GmTic110b protein (Glyma.14G201500; 98.4% amino acid similarity) and the Arabidopsis Tic110 protein (At1G06950; 84.3% amino acid similarity). In Medicago, the gene with the highest homology to *GmTic110a* is Medtr5g074690, whose sequence is 90.5% identical to that of *GmTic110a* and 86.4% identical to that of Medtr3g466170. Both of their proteins are 994 and 985 amino acids in length. Amino acid sequence analysis resulted in the generation of a phylogenetic tree composed of the *GmTic110a* homologous gene and its homologs from dicotyledonous plant species (*G. max*, *A. thaliana*, and *Medicago sativa*), monocotyledonous plant species (*Oryza sativa*, *Zea mays*, and *Sorghum bicolor*), *Selaginella tamariscina*, and *Physcomitrella patens* (**Figure 3C**). The results suggest that two GmTic110 homologs are evolutionarily conserved among plant species and share a common genomic structure in the observed plant species. The results of MEME analysis showed that *GmTic110a* contains 12 conserved motifs (**Figure 3D**). The mutation site of the *Gmpgl3-1* mutant is located within the conserved TM2 domain, the mutation site of the *Gmpgl3-2* mutant is located within the conserved chaperone-binding (co) domain, and the *Gmpgl3-3* mutant site is located within the terminal TM1 domain (**Figures 3A,B**).

Using qRT–PCR, we examined the *GmTic110a* expression patterns in new leaves at the VE stage; tips, stems, and roots at the V1 stage; leaves and flowers at the R1 stage; leaves and flowers at the R2 stage; and leaves at the R3 stage. *GmTic110a* was slightly expressed in the tips and stems and had low expression in the roots and flowers. The highest expression levels were detected in the leaves at various stages (new leaves at the VE stage, leaves at the R1 stage, leaves at the R2 stage, and leaves at the R3 stage; **Figure 3E**), indicating that *GmTic110a* may play an important role in leaf development and regulatory processes. These results also explained why the *GmTic110a* mutation severely affected leaf growth at various stages.

CRISPR/Cas9-Mediated *GmTic110a* Gene Editing of Transgenic Plants

To confirm whether *Gmpgl3-1* was the *GmTic110a* gene, loss-of-function transgenic lines were generated by inducing mutations in the *GmTic110a* gene using the CRISPR/Cas9 system. The resulting CRISPR/Cas9-induced mutations in three separate mutants led to the development of a *GmTic110*-specific mutant phenotype (**Figure 4A**). The results showed that the phenotypes of *GmTic110a*^{CR1}, *GmTic110a*^{CR2}, and *GmTic110a*^{CR3} were the same as those of *Gmpgl3-1* and *Gmpgl3-2*. The *GmTic110a*^{CR1} mutant contains a 7-bp substitution corresponding to the CDS of the *GmTic110a* gene from the 261st bp to the 273rd bp. The *GmTic110a*^{CR2} mutant had a 2-bp deletion from the 271st to the 272nd bp, and the *GmTic110a*^{CR3} mutant had a 2-bp deletion at 269th bp of the CDS of *GmTic110a* and a 1-bp substitution (Gly to Ala) at the 272nd bp (**Figure 4B**). These data suggested that the *GmTic110a*^{CR1}, *GmTic110a*^{CR2} and *GmTic110a*^{CR3} mutants have strong alleles and that complete loss of *Gmpgl3* function strongly influences soybean development (**Supplementary Table S1**). Moreover, the expression levels of *Glyma.02G233700* decreased in *GmTic110a*^{CR1}, *GmTic110a*^{CR2},

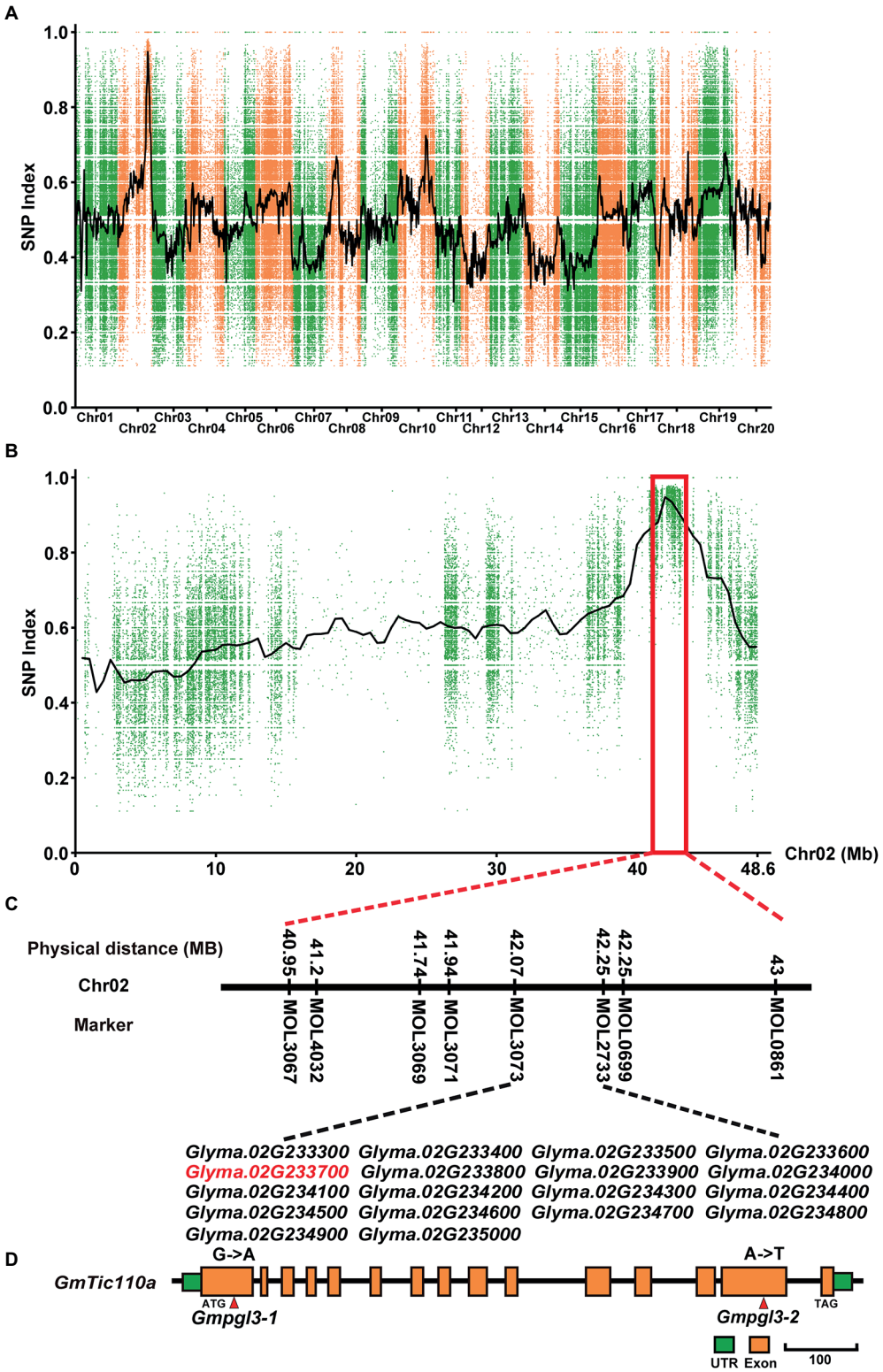


FIGURE 2 | Map-based cloning of the GmTic110a locus. **(A)** SNP index plot of all chromosomes of the F_2 plants. **(B)** SNP index plots of chromosome 02 of the *Gmpgl3a* mutant from the F_2 population. **(C)** Physical position of the *GmTic110a* candidate gene. **(D)** Schematic diagram showing the structure of *GmTic110a*. The red lines indicate mutation sites within the *GmTic110a* gene in the two mutant lines.

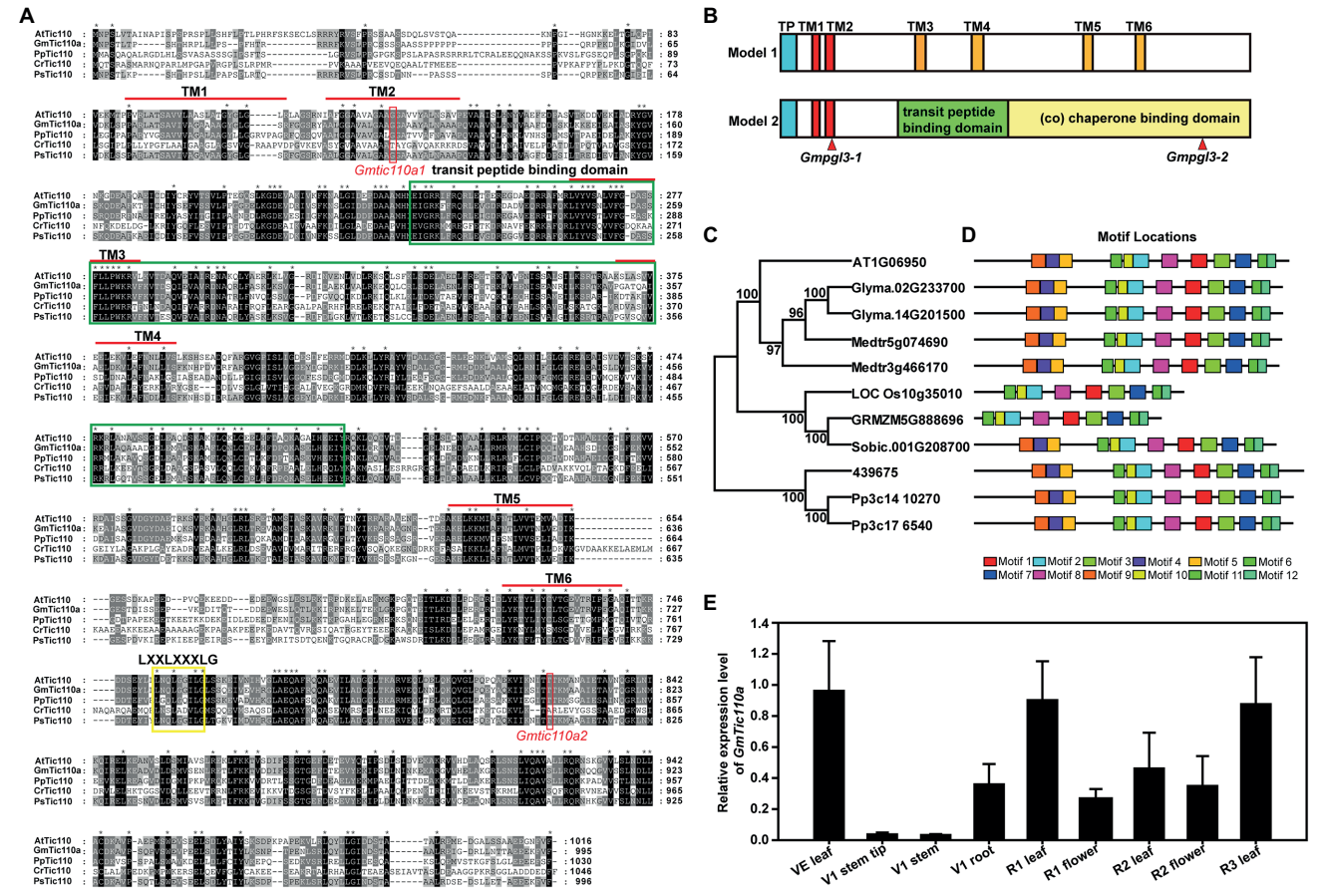


FIGURE 3 | (A) Multiple sequence alignments of the Tic110 protein in *Arabidopsis thaliana* (At), *Glycine max* (Gm), *Physcomitrella patens* (Pp), *Chlamydomonas reinhardtii* (Cr), and *Pisum sativum* (Ps). The dark region represents identical amino acids, and the grey region represents similar amino acids. **(B)** Schematic diagram representations of the two structural models of the Tic110 protein. In terms of the locations of the proposed TM domains, the red boxes represent TM1 and TM2, the orange boxes represent TM3 to TM6, the blue boxes represent transit peptides (TPs), the green box represents TP binding, and the yellow box represents a co-domain. **(C)** Phylogenetic trees based on the multiple sequence alignments of the Tic110 proteins. Bootstrap values from 1,000 replicates are indicated at each node. **(D)** Conserved motifs of Tic110 proteins in *G. max*, *A. thaliana*, *Medicago sativa*, *Oryza sativa*, *Zea mays*, *Sorghum bicolor*, *Selaginella tamariscina*, and *P. patens* were identified using the MEME search tool. Different motifs (1–12) are represented by boxes with different colors. **(E)** Tissue-specific expression profiles were determined *via* qRT–PCR.

and *GmTic110a*^{CR3} (**Figure 4C**). These results further suggested that *Glyma.02G233700* is the *GmTic110a* gene.

The GmTic110a Protein Localizes to the Chloroplast Inner Membrane

To confirm the subcellular localization of GmTic110a, the colocalization of green fluorescent proteins (*GmTic110a*-GFP) and the AtPIC1-mCherry marker protein (localization to the inner envelope of chloroplasts; Duy et al., 2007) was analyzed. As shown in **Supplementary Figure S3**, the *GmTic110a* protein targeted the inner chloroplast membrane (**Figure 5A**). At the same time, a single transformation of an empty vector that contained GFP was used as a positive control, and the results revealed that the fluorescent signal of the empty vector was detected throughout the entire cell protoplast (**Figure 5G**). The results were consistent with the bioinformatics predictions of the subcellular localization of *GmTic110a* (**Supplementary Table S5**). To determine the effects of *Gmpgl3* mutations on protein localization, we transiently transformed PUC19-GFP-GmTic110a^{G114A}, PUC19-GFP-Gmtic110a^{T805S}, PUC19-GFP-GmTic110a^{CR1}, PUC19-GFP-GmTic110a^{CR2}, and PUC19-GFP-GmTic110a^{CR3} into Arabidopsis protoplast cells. The results showed that proteins resulting from a mutated or knocked out *GmTic110a* gene were located in the inner chloroplast membrane, but the GFP fluorescence signal remained diffuse (**Figures 5B–F**), indicating that *GmTic110a* mutation or knockout altered the structure of the *GmTic110a* protein, thereby affecting the subcellular localization of the protein.

GmTic110a Interacts With GmTic20, GmTic40a, and GmTic40b

Previous studies have shown that AtTic110 mediates transport across the inner membrane *via* interactions with AtTic40 and other proteins (Kovacheva et al., 2005; Chiu and Li, 2008;

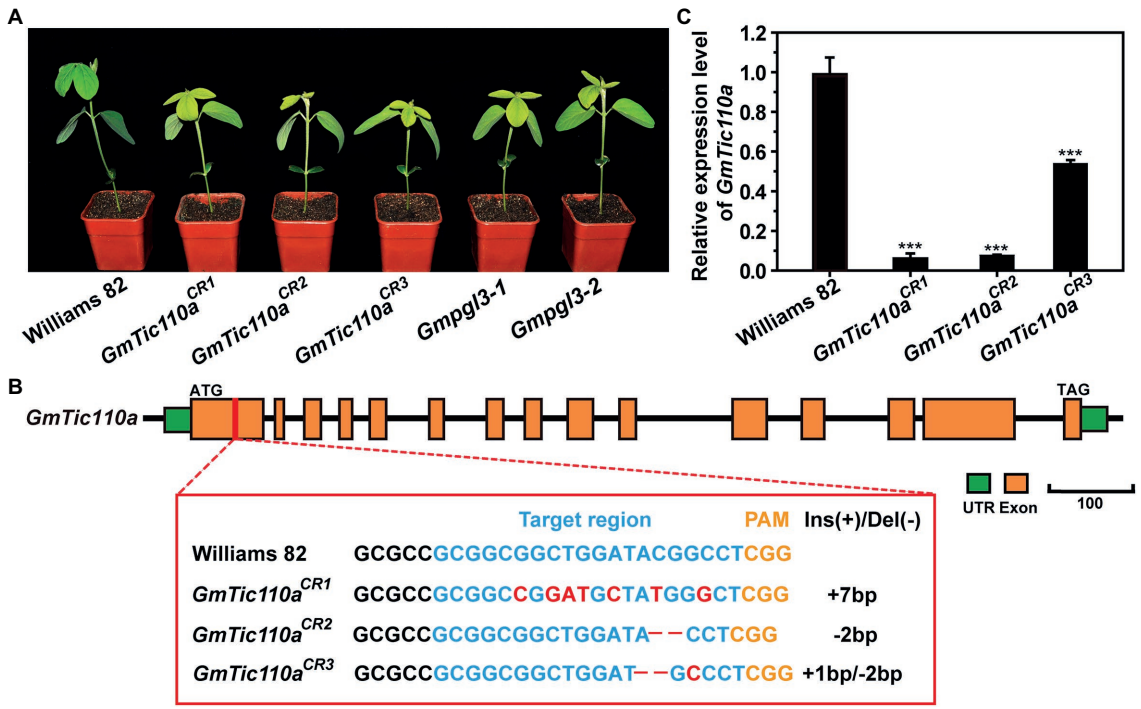


FIGURE 4 | (A) Phenotypes of Williams 82, the *Gmpgl3* mutant, and CRISPR/Cas9-edited plants (*GmTic110a*^{CR1}, *GmTic110a*^{CR,2} and *GmTic110a*^{CR3}). Scale bar = 1 cm. **(B)** sgRNA target sequences of Williams 82, *GmTic110a*^{CR1}, *GmTic110a*^{CR2}, and *GmTic110a*^{CR3}. The sgRNA target sequence is shown in blue letters, and the protospacer-adjacent motif (PAM) site is shown in yellow letters. The red letters indicate a single-base substitution. **–** indicates a deletion of the corresponding nucleotide. **(C)** Relative expression of the *GmTic110a* gene in unifoliate leaves of Williams 82, the *Gmpgl3* mutant, and CRISPR/Cas9-edited plants. The asterisks indicate statistically significant differences, as determined by Student's *t*-test (***, $p < 0.001$), and the error bars represent the standard deviations.

Yuan et al., 2021). Because the GmTic40 protein contains a highly conserved transmembrane (TM) motif, we investigated whether GmTic40 functions in conjunction with AtTic40. We used split luciferase complementation assays to confirm whether *GmTic110a* proteins could interact with GmTic20, GmTic40a, and GmTic40b (**Figures 6A,C,E**). The interactions of *GmTic110a* with GmTic20, GmTic40a, and GmTic40b were verified by co-IP analyses. As shown in **Figure 6**, we transiently coexpressed *GmTic110a* with GmTic20, GmTic40a, and GmTic40b in *N. benthamiana* leaves. Total proteins were isolated, after which they and anti-FLAG magnetic beads were incubated together to immunoprecipitate anti-FLAG. The results showed that GmTic110a, GmTic20, GmTic40a, and GmTic40b were present in the immunoprecipitate (**Figures 6B–F**), indicating that *GmTic110a* could interact with GmTic20, GmTic40a, and GmTic40b *in vivo*.

DISCUSSION

Currently, there are two hypotheses concerning the Arabidopsis Tic110 protein structure. Hypothesis 1 proposes that the Tic110 protein is composed of 6 TM domains (TM1, TM2, TM3, TM4, TM5, and TM6), while hypothesis 2 proposes that the Tic110 protein consists of 2 TM domains, one transit peptide-binding domain and 1 co-domain. We compared the amino acid sequences of the Tic110 proteins of *A. thaliana*, *G. max*, *P. patens*, *C. reinhardtii*, and *P. sativum* and identified six TM domains, transit peptide-binding domain domains, and co-domains (**Figure 3A**). By analyzing the structural site of the *Gmpgl3* mutant, we confirmed that the mutation site of the *Gmpgl3-1* mutant lies in the TM domain of TM2 (position 113 aa), the mutation site of the *Gmpgl3-2* mutant lies in the co-domain (amino acid position 805 aa), and the mutation site of the *Gmpgl3-3* mutant lies in the TM domain of TM1 (amino acid position 94 aa; **Figure 3A**). Phenotypic observations revealed that both the *Gmpgl3-1* and *Gmpgl3-2* mutants showed a pale green leaf phenotype (**Figure 1A**); *Gmpgl3-3* did not display a pale green leaf phenotype (**Supplementary Figure S1**). We speculated that the site variation of the *Gmpgl3-1* mutant affected the TM domain of TM2, indicating that it is the key site of protein function. In contrast, the mutation site of the *Gmpgl3-3* mutant is an amino acid at the end of the TM domain of TM1, and the mutation did not cause phenotypic abnormalities, indicating that this site is not a key site of the TM domain. The phenotype of the *Gmpgl3-2* mutant is similar to that of *Gmpgl3-1*, as both show a pale green leaf phenotype. The gene mutation site of *Gmpgl3-2* is in the co-domain. Mutations in this site may affect the *GmTic110a* gene and its function, which may impact the development of chloroplasts. The *GmTic110a* protein was knocked out by CRISPR/Cas9 technology, and the resulting three plants with point mutations

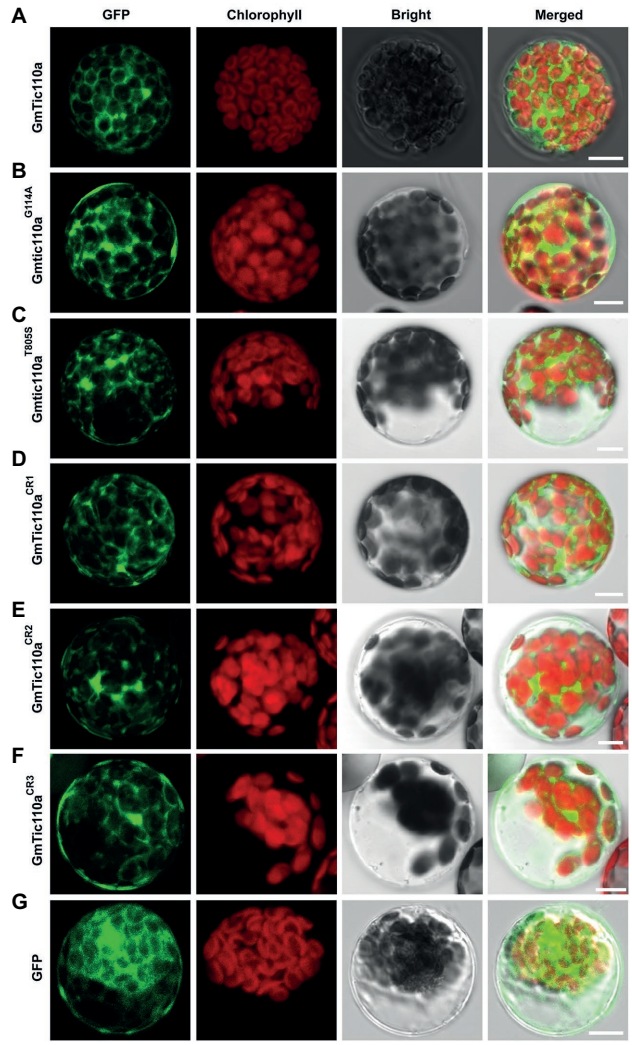


FIGURE 5 | Subcellular localization of GmTic110a, GmTic110a^{G114A}, Gmtic110a^{T805S}, GmTic110a^{CR1}, GmTic110a^{CR2}, and GmTic110a^{CR3}. **(A-G)** Transient expression of GFP-GmTic110a, GFP-GmTic110a^{G114A}, GFP-Gmtic110a^{T805S}, GFP-GmTic110a^{CR1}, GFP-GmTic110a^{CR2}, GFP-GmTic110a^{CR3} and GFP in Arabidopsis protoplasts. GFP, GFP fluorescence; Chlorophyll, chlorophyll autofluorescence; Bright, bright field. Merged, merged image of GFP fluorescence, chlorophyll autofluorescence and bright field images. Scale bars = 10 μm.

or deletions in the TM1 domain all showed similar phenotypes, which further confirmed that the phenotype was caused by a mutation in this gene. The knockout experiment using CRISPR/Cas9 technology also showed that the protein structural change due to the mutation is the main cause of the pale green leaf phenotype.

A protein must be in a suitable subcellular location to perform its function. Therefore, studying protein subcellular location is highly important for understanding protein function. This study used the online tool WoLF PSORT (Horton et al., 2006) and Target-P 1.1 Server (Emanuelsson et al., 1999) to predict the subcellular location of *GmTic110a*, which showed that the N-terminal domain contains chloroplast transit peptides, indicating that TIC110a is located on chloroplasts (**Supplementary Table S5**). To confirm whether the GmTic110a protein is located on the chloroplast, we extracted protoplasts from *A. thaliana* and transiently expressed GmTic110a-GFP constructs. The results showed that GmTic110a is located mainly on the chloroplast membrane (**Figure 5A**), which is consistent with the results predicted by the online tool. These results also indicated that the soybean GmTic110a protein is on the inner membrane of the chloroplast and is responsible for the TM transport of proteins into the chloroplast to perform normal functions. It was reported that the Arabidopsis *tic110* mutant exhibits pale green leaves and an albino phenotype. Trypsin digestion was used to prove that the Arabidopsis TIC110 fusion protein is located on the chloroplast membrane (Duy et al., 2007; Yuan et al., 2021). In this study, we confirmed that Tic110 proteins of soybean and Arabidopsis have the same subcellular localization. To confirm that the function of the GmTic110a protein is related to its specific membrane localization, we compared the GmTic110a protein localization patterns between mutant, deletion, and wild-type *GmTic110a* proteins. Point mutations or deletions of the GmTic110a protein alter the normal localization of the protein, and the protein is diffusely distributed throughout the chloroplast membrane (**Figures 5B–F**). Mutation or deletion of the TM1 domain of GmTic110a affected protein localization and normal function in the cells. The mutation or deletion position occurs after the leader peptide (amino acids 33–56 aa) in TM1, in the TM2 domain or in the co-domain, which does not affect the positioning of the *GmTic110a* protein in the chloroplast membrane. Point mutations or deletions in these conserved domains may affect the TM positioning or interactions with the protein, thereby affecting the chloroplast membrane localization and function of the GmTic110a protein.

Tic20, Tic110, and Tic40 are considered components of the TIC import machinery in the chloroplast; however, Kikuchi et al. (2013) reported that only Tic56, Tic100, and Tic214 were isolated from the 1-megadalton complex when using a tagged form of Tic20. Ramundo et al. (2020) also demonstrated that the TIC complex contains Tic20, Tic56, Tic100, and Tic214 by combining transcriptomic, biochemical, and genetic tools in the green alga Chlamydomonas, indicating that the complex is widely conserved among photosynthetic organisms (Ramundo et al., 2020). This result conflicts with our finding of the *GmTic110a* interaction with GmTic20 in the above studies, even though a similar interaction was also reported previously (Kouranov et al., 1998; Chen et al., 2002; Inaba et al., 2003). We suspected that these results were more likely caused by method limitations, and further physical interaction experiments will clear this confusion in the future. It is still unclear how Tic110 and Tic40 interact with the 1-megadalton complex, as they might be recruited to coordinate chaperone functions during later stages and/or are only required for the import of some preproteins (Lee and Hwang, 2018; Nakai, 2018; Thomson et al., 2020). Lee et al. reported that there were no differences in the import of preprotein *via* the wild-type transit peptide between

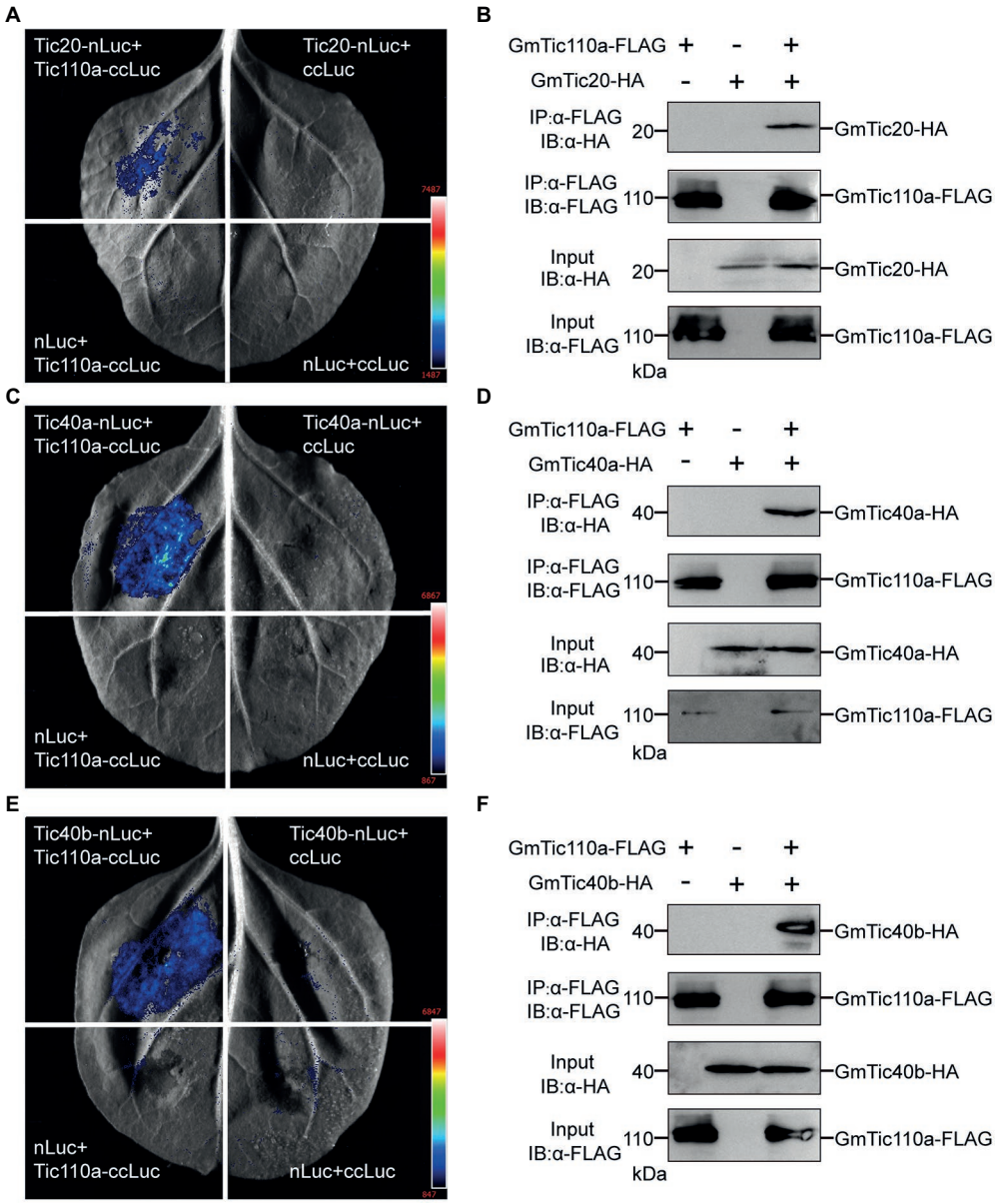


FIGURE 6 | *GmTic110a* interacts with GmTic20, GmTic40a, and GmTic40b. **(A,C,E)** Luciferase complementation assay showing that *GmTic110a* interacts with GmTic20, GmTic40a, and GmTic40b in *Nicotiana benthamiana*. Luciferase activity was detected 3 days after injection. **(B,D,F)** Interactions with GmTic20, GmTic40a, and GmTic40b in *N. benthamiana* according to a Co-IP assay. Immunoblots of the total protein extracts (20% input) and the immunoprecipitation product were performed using an anti-HA antibody (a-HA) or an anti-FLAG antibody (a-FLAG), respectively.

tic40 and wild-type protoplasts of *Arabidopsis thaliana*, while the import of N-terminal mutants of the RbcS protein (RbcS-nt) was dependent on Tic40; however, HA (hemagglutinin)-tagged Tic40 showed an intermediate form present in the stroma of *tic40* protoplasts (Lee and Hwang, 2019). In this study, we determined that GmTic110a interacted with GmTic20, GmTic40a, and GmTic40b in tobacco leaves (**Figure 6**); however, GmTic110a may also interact with other unknown partners. Thus, in the future, additional biochemical experiments will be performed to evaluate whether there is a direct interaction between them.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

XF, SY, and HY designed the research. HY, QW, ZZ, TW, and XY performed the experiments. XZ, YY, and JL analyzed

the data. HY, QW, and XF wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant nos. 31700213 and U21A20215), Science and Technology Development Plan Project of Jilin Province of China (grant nos. 20210302005NC), and Zhejiang Lab (grant nos. 2021PE0AC04).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.892077/full#supplementary-material

REFERENCES

Awan, M. A., Konzak, C. F., Rutger, J. N., and Nilan, R. A. (1980). Mutagenic effects of sodium azide in rice. *Crop Sci.* 20, 663–668. doi: 10.2135/cropsci1980.0011183X002000050030x

Bailey, T. L., Mikael, B., Buske, F. A., Martin, F., Grant, C. E., Luca, C., et al. (2009). MEME suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Balsera, M., Goetze, T. A., Kovács-Bogdán, E., Schürmann, P., Wagner, R., Buchanan, B. B., et al. (2009). Characterization of Tic110, a channel-forming protein at the inner envelope membrane of chloroplasts, unveils a response to Ca^{2+} and a stromal regulatory disulfide bridge. *J. Biol. Chem.* 284, 2603–2616. doi: 10.1074/jbc.M807134200

Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., et al. (2004). Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. *Mol. Biol. Cell* 15, 5130–5144. doi: 10.1091/mbc.E04-05-0405

Bölter, B., and Soll, J. (2017). Ycf1/Tic214 is not essential for the accumulation of plastid proteins. *Mol. Plant* 10, 219–221. doi: 10.1016/j.molp.2016.10.012

Cao, L., Wang, H., Sun, D., and Feng, Y. (2006). Photosynthesis and chlorophyll fluorescence characters of xantha wheat mutants. *Acta Botan. Boreali-Occiden. Sin.* 26, 2083–2087. doi: 10.3321/j.issn:1000-4025.2006.10.020

Carol, P., Stevenson, D., Bisanz, C., Breitenbach, J., Sandmann, G., Mache, R., et al. (1999). Mutations in the Arabidopsis gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. *Plant Cell* 11, 57–68. doi: 10.1105/tpc.11.1.57

Chen, X., Smith, M. D., Fitzpatrick, L., and Schnell, D. J. (2002). In vivo analysis of the role of at Tic20 in protein import into chloroplasts. *Plant Cell* 14, 641–654. doi: 10.1105/tpc.010336

Chen, H., Zou, Y., Shang, Y., Lin, H., Wang, Y., Cai, R., et al. (2008). Firefly luciferase complementation imaging assay for protein-protein interactions in plants. *Plant Physiol.* 146, 368–376. doi: 10.2307/40065847

Chiu, C. C., and Li, H. M. (2008). Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. *Plant J.* 56, 793–801. doi: 10.1111/j.1365-313X.2008.03638.x

Chou, M. L., Fitzpatrick, L. M., Tu, S. L., Budziszewski, G., Potter-Lewis, S., Akita, M., et al. (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J.* 22, 2970–2980. doi: 10.1093/emboj/cdg281

Davis, M. S., and Fajer, A. F. (1979). Ligated chlorophyll cation radicals: their function in photosystem II of plant photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 76, 4170–4174. doi: 10.1073/pnas.76.9.4170

de Vries, J., Sousa, F. L., Bölter, B., Soll, J., and Gould, K. (2007). YCF1: a green TIC?. *Plant Cell* 27, 1827–1833. doi: 10.1105/tpc.114.135541

Du, H. Y., Zeng, X. R., Zhao, M., Cui, X. P., Wang, Q., Yang, H., et al. (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *J. Biotechnol.* 217, 90–97. doi: 10.1016/j.jbiotec.2015.11.005

Duy, D., Wanner, G., Meda, A. R., Von Wiren, N., Soll, J., and Philippar, K. (2007). PIC1, an ancient permease in Arabidopsis chloroplasts, mediates iron transport. *Plant Cell* 19, 986–1006. doi: 10.1105/tpc.106.047407

Emanuelsson, O., Nielsen, H., and Von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984. doi: 10.1110/ps.8.5.978

Genty, B., Briantais, J. M., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochim. Biophys. Acta Gen. Subj.* 990, 87–92. doi: 10.1016/S0304-4165(89)80016-9

Gregor, J., and Marsálek, B. (2004). Freshwater phytoplankton quantification by chlorophyll a: a comparative study of in vitro, in vivo and in situ methods. *Water Res.* 38, 517–522. doi: 10.1016/j.watres.2003.10.033

Highkin, H. R., Boardman, N. K., and Goodchild, D. J. (1969). Photosynthetic studies on a pea-mutant deficient in chlorophyll. *Plant Physiol.* 44, 1310–1320. doi: 10.1104/pp.44.9.1310

Hormann, F., Kuchler, M., Sveshnikov, D., Oppermann, U., Li, Y., and Soll, J. (2004). Tic32, an essential component in chloroplast biogenesis. *J. Biol. Chem.* 279, 34756–34762. doi: 10.1074/jbc.M402817200

Horton, P., Park, K. J., Obayashi, T., and Nakai, K. (2006). Protein subcellular localisation prediction with WoLF PSORT. In *Proceedings of the 4th Asia-Pacific Bioinformatics Conference 2006.*

Hu, R., Fan, C., Li, H., Zhang, Q., and Fu, Y. F. (2009). Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Mol. Biol.* 10:93. doi: 10.1186/1471-2199-10-93

Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F., et al. (2005). Arabidopsis Tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* 17, 1482–1496. doi: 10.1105/tpc.105.030700

Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F., and Schnell, D. J. (2003). AtTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *J. Biol. Chem.* 278, 38617–38627. doi: 10.1074/jbc.M306367200

Ishida, H., and Terakura, K. (1987). Lattice dynamics of TiC(110) and TiC(111) surfaces. *Phys. Rev. B Condens. Matter* 36, 4403–4406. doi: 10.1103/physrevb.36.4403

Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* 179, 257–285. doi: 10.2307/25150450

Jocelyn, B., and Paul, J. (2005). Recognition and envelope translocation of chloroplast preproteins. *J. Exp. Bot.* 56, 2287–2320. doi: 10.1093/jxb/eri243

Kasmati, A. R., Töpel, M., Patel, R., Murtaza, G., and Jarvis, P. (2011). Molecular and genetic analyses of Tic20 homologues in Arabidopsis thaliana chloroplasts. *Plant J.* 66, 877–889. doi: 10.1111/j.1365-313X.2011.04551.x

Kessler, F., Blobel, G., Patel, H. A., and Schnell, D. J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* 266, 1035–1039. doi: 10.1126/science.7973656

Ki-Hong, J., Junghe, H., Choong-Hwan, R., Youngju, C., Yong-Yoon, C., Akio, M., et al. (2003). Characterization of a rice chlorophyll-deficient mutant using the T-DNA gene-trap system. *Plant Cell Physiol.* 44, 463–472. doi: 10.1093/pcp/pcg064

Kikuchi, S., Asakura, Y., Imai, M., Nakahira, Y., Kotani, Y., Hashiguchi, Y., et al. (2018). A Ycf2-FtsHi heteromeric AAA-ATPase complex is required for chloroplast protein import. *Plant Cell* 30, 2677–2703. doi: 10.1105/tpc.18.00357

Kikuchi, S., Be Dard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., et al. (2013). Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339, 571–574. doi: 10.1126/science.1229262

Kikuchi, S., Oishi, M., Hirabayashi, Y., Dong, W. L., Hwang, I., and Nakai, M. (2009). A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* 21, 1781–1797. doi: 10.1105/tpc.108.063552

Köhler, D., Helm, S., Agne, B., and Baginsky, S. (2016). Importance of translocon subunit Tic56 for rRNA processing and chloroplast ribosome assembly. *Plant Physiol.* 172, 2429–2444. doi: 10.1104/pp.16.01393

Köhler, D., Montandon, C., Hause, G., Majovsky, P., Kessler, F., Baginsky, S., et al. (2015). Characterization of chloroplast protein import without Tic56, a component of the 1-megadalton translocon at the inner envelope membrane of chloroplasts. *Plant Physiol.* 167, 972–990. doi: 10.1104/pp.114.255562

Kouranov, A., Chen, X., Fuks, B., and Schnell, D. J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J. Cell Biol.* 143, 991–1002. doi: 10.1083/jcb.143.4.991

Kouranov, A., and Schnell, D. J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. *J. Cell Biol.* 139, 1677–1685. doi: 10.1083/jcb.139.7.1677

Kovacheva, S., Bédard, J., Patel, R., Dudley, P., Twell, D., Ríos, G., et al. (2005). In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J.* 41, 412–428. doi: 10.1111/j.1365-313X.2004.02307.x

Kowalewska, M., Mazur, R., Suski, S., Garstka, M., and Mostowska, A. (2016). Three-dimensional visualization of the tubular-lamellar transformation of the internal plastid membrane network during runner bean chloroplast biogenesis. *Plant Cell* 28, 875–891. doi: 10.1105/tpc.15.01053

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Lee, D. W., and Hwang, I. (2018). Evolution and design principles of the diverse chloroplast transit peptides. *Mol. Cells* 41, 161–167.

Lee, D. W., and Hwang, I. (2019). Protein import into chloroplasts via the Tic40-dependent and independent pathways depends on the amino acid composition of the transit peptide. *Biochem. Biophys. Res. Commun.* 518, 66–71. doi: 10.1016/j.bbrc.2019.08.009

Liu, Z., Li, N., Zhang, Y., and Li, Y. (2020). Transcriptional repression of GIF1 by the KIX-PPD-MYC repressor complex controls seed size in Arabidopsis. *Nat. Commun.* 11:1846. doi: 10.1038/s41467-020-15603-3

Lonosky, P. M., Zhang, X., Honavar, V. G., Dobbs, D. L., Fu, A., and Rodermel, S. R. (2004). A proteomic analysis of maize chloroplast biogenesis. *Plant Physiol.* 134, 560–574. doi: 10.1104/pp.103.032003

Lu, X., Lai, Y. C., Du, W. G., Man, W. Q., Chen, S. Y., Zhang, J. S., et al. (2017). A PP2C-1 allele underlying a quantitative trait locus enhances soybean 100-seed weight. *Mol. Plant* 10, 670–684. doi: 10.1016/j.molp.2017.03.006

Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Nakai, M. (2015). The TIC complex uncovered: The alternative view on the molecular mechanism of protein translocation across the inner envelope membrane of chloroplasts. *Biochim. Biophys. Acta* 1847, 957–967. doi: 10.1016/j.bbabio.2015.02.011

Nakai, M. (2018). New perspectives on chloroplast protein import. *Plant Cell Physiol.* 59, 1111–1119. doi: 10.1093/pcp/pcy083

Nakai, M. (2020). Reply: The revised model for chloroplast protein import. *Plant Cell* 32, 543–546. doi: 10.1105/tpc.19.00821

Okabe, K., and Straub, S. J. (1977). Genetic characterization and high efficiency photosynthesis of an Aurea mutant of tobacco. *Plant Physiol.* 60, 150–156. doi: 10.2307/4264941

Oshima, C., Aizawa, T., Wuttig, M., Souda, R., Otani, S., Ishizawa, Y., et al. (1987). Surface-phonon dispersion curves of TiC(100). *Phys. Rev. B Condens. Matter* 36, 7510–7517. doi: 10.1103/physrevb.36.7510

Peng, S., Khush, G. S., Virk, P., Tang, Q., and Zou, Y. (2008). Progress in ideotype breeding toincrease rice yield potential. *Field Crops Res.* 108, 32–38. doi: 10.1016/j.fcr.2008.04.001

Preiss, S., and Thornber, J. P. (1995). Stability of the apoproteins of light-harvesting complex I and II during biogenesis of thylakoids in the chlorophyll b-less barley mutant chlorina f2. *Plant Physiol.* 107, 709–717. doi: 10.1104/pp.107.3.709

Ramundo, S., Asakura, Y., Salomé, P. A., Strenkert, D., Boone, M., Mackinder, L. C. M., et al. (2020). Coexpressed subunits of dual genetic origin define a conserved supercomplex mediating essential protein import into chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.* 117, 32739–32749. doi: 10.1073/pnas.2014294117

Richardson, L. G. L., and Schnell, D. J. (2020). Origins, function, and regulation of the TOC-TIC general protein import machinery of plastids. *J. Exp. Bot.* 71, 1226–1238. doi: 10.1093/jxb/erz517

Sandhu, D., Atkinson, T., Noll, A., Johnson, C., Espinosa, K., Boelter, J., et al. (2016). Soybean proteins GmTic110 and GmPsbP are crucial for chloroplast development and function. *Plant Sci.* 252, 76–87. doi: 10.1016/j.plantsci.2016.07.006

Schnell, D. J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. *Science* 266, 1007–1012. doi: 10.1126/science.7973649

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Sohrt, K., and Soll, J. (2000). Toc64, a new component of the protein translocon of chloroplasts. *J. Cell Biol.* 148, 1213–1222. doi: 10.1083/jcb.148.6.1213

Soll, J. (2004). Protein import into chloroplasts. *Mol. Cell. Biol.* 5, 529–535. doi: 10.1016/S1369-5266(02)00296-0

Song, X., Wei, H., Cheng, W., Yang, S., Zhao, Y., Li, X., et al. (2015). Development of INDEL markers for genetic mapping based on whole genome Resequencing in soybean. *G three* 5, 2793–2799. doi: 10.1534/g3.115.022780

Stahl, T., Glockmann, C., Soll, J., and Heins, L. (1999). Tic40, a new "old" subunit of the chloroplast protein import translocon. *J. Biol. Chem.* 274, 37467–37472. doi: 10.1074/jbc.274.52.37467

Stockinger, E. J., and Walling, L. L. (1994). A chlorophyll a/b-binding protein gene from soybean (*Glycine max* [L.] *Merr.*). *Plant Physiol.* 104, 1475–1476. doi: 10.1104/pp.104.4.1475

Sudhir, K., Glen, S., and Koichiro, T. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Sveshnikova, N., Grimm, R., Soll, J., and Schleiff, E. (2000). Topology studies of the chloroplast protein import channel Toc75. *Biol. Chem.* 381, 687–693. doi: 10.1515/BC.2000.089

Teng, Y. S., Su, Y. S., Chen, L. J., Lee, Y. J., and Li, H. H. M. (2006). Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. *Plant Cell* 18, 2247–2257. doi: 10.1105/tpc.106.044305

Thomson, S. M., Pulido, P., and Jarvis, R. P. (2020). Protein import into chloroplasts and its regulation by the ubiquitin-proteasome system. *Biochem. Soc. Trans.* 48, 71–82. doi: 10.1042/BST20190274

Tiller, N., and Bock, R. (2014). The translational apparatus of plastids and its role in plant development. *Mol. Plant* 7, 1105–1120. doi: 10.1093/mp/ssu022

Tranel, P. J., Froehlich, J., Goyal, A., and Keegstra, K. (1995). A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. *EMBO J.* 14, 2436–2446. doi: 10.1002/j.1460-2075.1995.tb07241.x

Tsai, J. Y., Chu, C. C., Yeh, Y. H., Chen, L. J., Li, H. M., and Hsiao, C. D. (2013). Structural characterizations of the chloroplast translocon protein Tic110. *Plant J.* 75, 847–857. doi: 10.1111/tpj.12249

Vitale, A., Schnell, D. J., Raikhel, N. V., and Chrispeels, M. J. (2015). "Protein sorting and vesicle traffic," in *Biochemistry & Molecular Biology of Plants. 2nd Edn*. eds. B. B. Buchanan, W. Gruissem and R. L. Jones. (Chichester: John Wiley & Sons, Ltd.), 151–190.

Waegemann, K., and Soil, J. (1991). Characterization of the protein import apparatus in isolated outer envelopes of chloroplasts. *Plant J.* 1, 149–158. doi: 10.1111/j.1365-313X.1991.00149.x

Wang, B., Lan, T., Wu, W. R., and Li, W. M. (2003). Mapping of QTLs controlling chlorophyll content in rice. *Acta Genet. Sin.* 30, 1127–1132. doi: 10.1016/S0891-0618(02)00103-5

Wang, D., Liang, X., Bao, Y., Yang, S., Zhang, X., Yu, H., et al. (2020). A malectin-like receptor kinase regulates cell death and pattern-triggered immunity in soybean. *EMBO Rep.* 21:e50442. doi: 10.15252/embr.202050442

Waters, M. T., and Langdale, J. A. (2009). The making of a chloroplast. *EMBO J.* 28, 2861–2873. doi: 10.1038/emboj.2009.264.

Yamori, W., Takahashi, S., Makino, A., Price, G. D., Badger, M. R., and Von Caemmerer, S. (2011). The roles of ATP synthase and the cytochrome b6/f complexes in limiting chloroplast electron transport and determining photosynthetic capacity. *Plant Physiol.* 155, 956–962. doi: 10.1104/pp.110.168435

Yu, H., Jiang, W., Liu, Q., Zhang, H., Piao, M., Chen, Z., et al. (2015). Expression pattern and subcellular localization of the ovate protein family in rice. *PLoS One* 10:e0118966. doi: 10.1371/journal.pone.0118966

Yuan, H., Pawlowski, E. G., Yang, Y., Sun, T., Thannhauser, T. W., Mazourek, M., et al. (2021). Arabidopsis ORANGE protein regulates plastid pre-protein import through interacting with tic proteins. *J. Exp. Bot.* 72, 1059–1072. doi: 10.1093/jxb/eraa528

Zhao, B., Dai, A., Wei, H., Yang, S., Wang, B., Jiang, N., et al. (2016). Arabidopsis KLU homologue GmCYP78A72 regulates seed size in soybean. *Plant Mol. Biol.* 90, 33–47. doi: 10.1007/s11103-015-0392-0

Zheng, K., Wang, Y., Na, Z., Jia, Q., and Wang, S. (2017). Involvement of paclobutrazol resistance6/KIDARI, an atypical bHLH transcription factor, in auxin responses in arabidopsis. *Front. Plant Sci.* 8:1813. doi: 10.3389/fpls.2017.01813

Zhou, X., Welsch, R., Yang, Y., Álvarez, D., Riediger, M., Yuan, H., et al. (2015). Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3558–3563. doi: 10.1073/pnas.1420831112

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yu, Wang, Zhang, Wu, Yang, Zhu, Ye, Leng, Yang and Feng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 June 2022
doi: 10.3389/fpls.2022.910249

Breeding of Penta Null Soybean [*Glycine max* (L.) Merr.] for Five Antinutritional and Allergenic Components of Lipoxygenase, KTI, Lectin, 7S α′ Subunit, and Stachyose

*Sang Woo Choi, Sarath Ly, Jeong Hwan Lee, Hyeon Su Oh, Se Yeong Kim, Na Hyeon Kim and Jong Il Chung**

Department of Agronomy, Gyeongsang National University, Jinju, South Korea

OPEN ACCESS

Edited by:
Jose C. Jimenez-Lopez, Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidin (CSIC), Spain

Reviewed by:
Carmen Cuadrado, Instituto Nacional de Investigación y Tecnología Agroalimentaria (INIA), Spain
Javaid Akhter Bhat, Nanjing Agricultural University, China

***Correspondence:**
Jong Il Chung
jongil@gnu.ac.kr

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *01 April 2022*
Accepted: *19 May 2022*
Published: *07 June 2022*

Citation:
Choi SW, Ly S, Lee JH, Oh HS, Kim SY, Kim NH and Chung JI (2022) Breeding of Penta Null Soybean [Glycine max (L.) Merr.] for Five Antinutritional and Allergenic Components of Lipoxygenase, KTI, Lectin, 7S α′ Subunit, and Stachyose. Front. Plant Sci. 13:910249. doi: 10.3389/fpls.2022.910249

Soybean [*Glycine max* (L.) Merr.] is an excellent source of protein, oil, carbohydrates and many other bioactive ingredients for humans. However, several antinutritional and allergenic components such as lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose exist in the raw mature seed. Genetic removal of these components would be the best method to improve soybean food quality. The objectives of this research were to breed a new soybean line with penta null recessive alleles (*lox1*/*lox1*/*lox2*/*lox2*/*lox3*/*lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*) for these five components and to evaluate agronomic traits for a breeding line with penta null alleles. Seven germplasms were used to breed the penta null strain. Analysis of lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose components in mature seeds was conducted by SDS-PAGE, western blot, and HPLC. One breeding line with penta null recessive alleles was developed. The breeding line has purple flowers, tawny pubescence, a determinate growth habit, and light yellow pods at maturity. The seed of the breeding line has a yellow hilum and yellow seed coat color. The stem height of the breeding line was 53.0 cm. The stachyose content of the breeding line was 2.9 g/kg. The 100-seed weight of the breeding line was 31.1 g and yield (t/ha) was 2.80. This is the first soybean strain with the penta null (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*) genotype (free of lipoxygenase, KTI, lectin, and 7S α′ subunit proteins, and with low stachyose content).

Keywords: penta null, lipoxygenase, KTI, lectin, 7S α′ subunit, stachyose

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is a major leguminous crop that has been cultivated for thousands of years. Soybean seeds are an important source of protein, oil, carbohydrates, isoflavones, and many other nutrients for human food and animal feed. Generally, soybean seeds contain about 40% protein, 20% oil, 30% carbohydrates, and various health functional ingredients. However, there are also antinutritional and harmful immunological components

such as lipoxygenase protein, Kunitz trypsin inhibitor (KTI) protein, 7S α′ subunit protein, and stachyose in mature soybean seeds.

Lipoxygenase protein causes undesirable grassy and beany flavors in foods containing soybean due to the oxidation of polyunsaturated fatty acids. There are three lipoxygenases (*Lox1*, *Lox2,* and *Lox3*) in mature seeds. Previous studies demonstrated that the absence of each enzyme is under the control of three null alleles, *lox1*, *lox2*, and *lox3*, which are inherited as simple recessive alleles (Hildebrand and Hymowitz, 1982; Kitamura et al., 1983; Davies and Nielsen, 1986).

Lipoxygenase-free genotypes are better accepted due to the production of very low levels of hexanal compounds (Kobayashi et al., 1995). The development of lipoxygenase-free genotypes through genetic elimination is the key to removing the beany flavor. Several lipoxygenase protein-free cultivars have been improved (Kim et al., 1997, 2015; Chung, 2014). Recently, lipoxygenase-free mutants were obtained using a pooled CRISPR-Cas9 system (Wang et al., 2020). Soybean Kunitz trypsin inhibitor (KTI) protein, which was isolated and crystallized by Kunitz (1945), is a small and non-glycosylated protein possessing 181 amino acid residues with 21.5 kDa. KTI protein severely inhibits trypsin, thus reducing food intake by diminishing digestion and absorption. Orf and Hymowitz (1979) have identified two soybean accessions (PI157440 and PI196168) lacking the KTI protein in the USDA Soybean Germplasm Collection. The presence of KTI protein was controlled at a single locus with a codominant multiple allelic series that identified a recessive allele designated *ti* that lacks the KTI protein (Orf and Hymowitz, 1979). Crude protein from the *titi* genotype showed a 30%–50% reduction in trypsin inhibitor activity compared with the *TiTi* genotype. The *Ti* gene was found to be on chromosome 8. Soybean with normal protein content was developed through a marker-assisted survey of 180 recombinant inbred lines (RILs) developed from the cross KTI + parent and KTI-free parent KTI-free (Rani et al., 2020). A breeder-friendly Kompetitive Allele-Specific PCR (KASP) marker linked to low Kunitz trypsin inhibitor (KTI) in soybean seeds has been developed (Rosso et al., 2021). A KTI-free vegetable soybean genotype was developed using *Ti*-linked simple sequence repeat (SSR) markers (Kumar et al., 2021). Recently, Lox-2 and KTI-free soybean lines were developed by introgressing null alleles of Lox-2 and KTI genes in the variety "JS97-52" through marker-assisted backcross breeding (Kumar et al., 2022). Soybean lectin protein with 120 kDa molecular weight is a major antinutritional element and can strongly inhibit degradation by proteases under *in vitro* and *in vivo* conditions (Pull et al., 1978). Consuming food including lectin protein can cause symptoms such as nausea, vomiting, and diarrhea (Katsuya et al., 2007). Soybean seed lectin is controlled by a single gene designated *Le* (*le*) and the *lele* genotype results in a lack of lectin in mature seed (Orf et al., 1978). The *Le* gene was found to be on chromosome 2. The two main seed storage proteins are glycinin (11S) and β-conglycinin (7S), which constitute about 70% of storage protein (Krishnan et al., 2009). Almost 90% of 7S protein is accumulated by β-conglycinin, which has three subunits (α, α′, β), and these three subunits are the major allergens of soybean protein (Medic et al., 2014). The three subunits of β-conglycinin (7S), α, α′, and β are dominated by genes *Cgy1*, *Cgy2*, and *Cgy3*, respectively (Davies et al., 1985). Among these three genes, a line with the homozygous recessive *Cgy1* gene cannot produce the 7S α′ subunit protein in mature seed (Kitamura et al., 1984). *Cgy1* gene was found to be on chromosome 10. A soybean experimental line (BSH-3) that is 7S α′ subunit protein free was developed by crossing the mutant donor line "HS99B" with the Chinese cultivar "Dongnong47" (Song et al., 2018).

Raffinose and stachyose are considered antinutritional factors because humans cannot digest them after absorption (Hata et al., 1991). Stachyose is the primary carbohydrate in soybean seed. Stachyose content ranges from 14 to 41 g/kg on a dry weight basis and is environmentally stable but genotypically dependent (Hymowitz et al., 1972). The amount of stachyose was controlled by a single gene or by a major quantitative trait locus (QTL; Skoneczka et al., 2009). The raffinose synthase 2 gene is a pathway for raffinose and stachyose biosynthesis. Soybean line PI200508 with a homozygous recessive genotype (*rs2rs2*) showed low raffinose and stachyose content (Dierking and Bilyeu, 2008). *RS2* locus was found to be located on chromosome 6. Two SSR markers, Sat_293 (LG-K/*chr*9) and Satt281 (LG-C2/*chr*6), were identified for stachyose in two F_2 populations (Jha et al., 2022).

Because of these antinutritional factors and allergens that exist in raw soybean seeds, heat treatment or other methods are needed to eliminate or reduce these components and secure the efficiency of nutrient absorption and food safety. But these treatments cause some changes and reduce soybean quality (Chen et al., 2019). Also, heat inactivation of the lipoxygenase at an industrial level not only incurs extra cost but also affects the solubility and functionality of proteins (MaCleod and Ames, 1988). Antigenic proteins remain in soybean food even after heat treatment and fermentation (Wilson et al., 2008). Genetic removal of lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose components that exist in mature soybean seed would be the best method for the soybean food industry. Only a few papers on soybean free of these antinutritional and allergenic components have been published. A soybean line with a triple null recessive genotype (*ti*/*ti-le*/*le-p34*/*p34*) for KTI, lectin, and P34 proteins was developed (Schmidt et al., 2015). A soybean line with a tetra null recessive genotype (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy*1/*cgy1*) for lipoxygenase, KTI, lectin, and 7S α′ subunit proteins was also developed (Choi et al., 2021). So far, a soybean line with a penta null recessive genotype for antinutritional and allergenic factors has not been developed. Therefore, the objectives of this research were to breed a new soybean line with a yellow seed coat color and penta null recessive alleles (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy*1/*cgy1-rs2*/*rs2*) for lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose components and to evaluate agronomic traits for a breeding line with penta null alleles.

MATERIALS AND METHODS

Breeding Materials

Seven germplasms were used to improve the new soybean strain with the penta null genotype for five components. The presence and absence of four proteins, stachyose content, seed coat, 100-seed weight (g), and origin of the seven germplasms used in this study are presented in **Table 1**. Three breeding lines (*lox1lox1*/*lox2lox2*/*lox3lox3* genotype—lipoxygenase protein free; *lox1lox1*/ *lox2lox2*/*lox3lox3-cgy1* genotype—lipoxygenase and 7S α′ subunit proteins free; *lox1lox1*/*lox2lox2*/*lox3lox3-ti-le-cgy1* genotype—lipoxygenase, KTI, lectin and 7S α′ subunit proteins free, Choi et al., 2021) and one germplasm (PI200508) were used to create a genetic population. The PI200508 parent has an *rs2rs2* genotype with low stachyose content (Dierking and Bilyeu, 2008). Three breeding lines and one germplasm have a yellow seed coat color in mature seeds.

Breeding Scheme

The parent with the *lox1lox2lox3*/*lox1lox2lox3* genotype was crossed with the PI200508 parent with the *rs2rs2* genotype to select a seed with a homozygous *lox1lox2lox3*/*lox1lox2lox3-rs2rs2* genotype. *lox1lox2lox3* alleles were determined by identifying the absence of lipoxygenase protein using SDS electrophoresis. The *rs2rs2* genotype, which determines the low stachyose content, was identified by DNA marker. A plant with a *lox1lox2lox3*/*lox1lox2lox3-rs2*/*rs2* genotype (lipoxygenase protein free and with low stachyose content) was developed. The plant was then crossed with the plant with the *lox1lox2lox3*/*lox1lox2lox3-cgy1*/*cgy1* genotype (lipoxygenase and 7S α′ subunit proteins free). The *cgy1cgy1* genotype was determined by identifying the absence of 7S α′ subunit protein using SDS electrophoresis. From this cross, a triple null genotype (*lox1lox2lox3*/*lox1lox2lox3-rs2*/*rs2-cgy1*/*cgy1*, free of lipoxygenase and 7S α′ subunit proteins, and with low stachyose content) was developed. During the summer of 2017, F_1 pollinations between *lox1lox2lox3*/*lox1lox2lox3-rs2*/*rs2-cgy1*/*cgy1* parent and *lox1lox2lox3* /*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1* parent were made in a greenhouse to produce seeds possessing penta null alleles (*lox1lox2lox3*/ *lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*). When mature, the hybrid seed borne on a female parent was collected and hand threshed. The resultant F_1 seeds were planted in the greenhouse on 25 February 2018. F_1 hybridity was confirmed based on pod color (i.e., tan female mated to brown male; the female self is tan, the F_1 hybrid is brown). All F_1 plants were individually harvested and were bulked after the confirmation of hybridity. A total of 172 F_2 seeds were obtained. *Titi* and *lele* genotypes were determined by identifying the absence of KTI and lectin proteins using the western blot technique. Each seed was analyzed to screen for genotypes with tetra recessive alleles [*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1* (absence of lipoxygenase, KTI, lectin, and 7S α′ subunit proteins)]. Among the 172 F_2 seeds, 10 F_2 seeds possessing tetra null alleles (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1*) were obtained and were planted in the greenhouse in July 2018. Young leaves from each of the 10 F_2 plants grown in the greenhouse were used to screen the plant of the *rs2rs2* genotype based on DNA marker (Yang et al., 2014; Choi et al., 2017). Only two plants were selected, and two F_2 plants were individually harvested. A random sample of 120 F_3 seeds for each strain was planted on 8 July 2019 in a field at the university. One F_3 strain was chosen based on plant type, maturation date, stem height, seed coat color, seed quality, seed weight, and yield. Then 360 random F_4 seeds of the strain were planted in the university field on 9 July 2020. After harvesting, random F_5 seeds were used to confirm recessive genotypes (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*) by observing the absence of lipoxygenase, KTI, lectin, and 7S α′ subunit proteins. Low stachyose content (*rs2rs2* genotype) for the strain developed was confirmed by HPLC (Sung et al., 2014). The scheme for the improvement of penta null alleles (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*) is presented in **Figure 1**.

Determination of Lipoxygenase and 7S α′ Subunit Proteins by SDS-PAGE

Crude protein from the random F_5 seeds of the breeding line and random seeds of the cultivar ("Daewon") was obtained to identify the presence ("+") or absence ("−" of lipoxygenase and 7S α′ subunit proteins. Lipoxygenase and 7S α′ subunit proteins were detected by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) performed according to Fling and Gregerson (1986). The cultivar ("Daewonkong") was used as a control for the presence of lipoxygenase and 7S α′ subunit proteins (*Lox1Lox2Lox3*/*Lox1Lox2Lox3-Cgy1*/*Cgy1* genotype). Fine powder samples of the two materials were incubated for

TABLE 1 | Seed coat, 100-seed weight, stachyose content, origin, presence or absence of lipoxygenase, Kunitz trypsin inhibitor (KTI), lectin and 7S α′ subunit for seven germplasms.

Germplasm name	Lipoxygenase			KTI	Lectin	7S α′ subunit	Stachyose	Seed coat	100-seed weight (g)	Origin
	1	2	3							
PI408251	Absent	Present	Present	Present	Present	Present	Normal	Black	6.1	Korea
PI86023	Present	Absent	Present	Present	Present	Present	Normal	Green	16.8	Japan
PI417458	Present	Present	Absent	Present	Present	Present	Normal	Yellow	11.8	Japan
PI200508	Present	Present	Present	Present	Present	Present	Low	Yellow	13.1	Japan
PI506876	Present	Present	Present	Present	Present	Absent	Normal	Yellow	15.1	Japan
PI157440	Present	Present	Present	Absent	Present	Present	Normal	Yellow	14.5	Korea
T102	Present	Present	Present	Present	Absent	Present	Normal	Black	6.7	United States

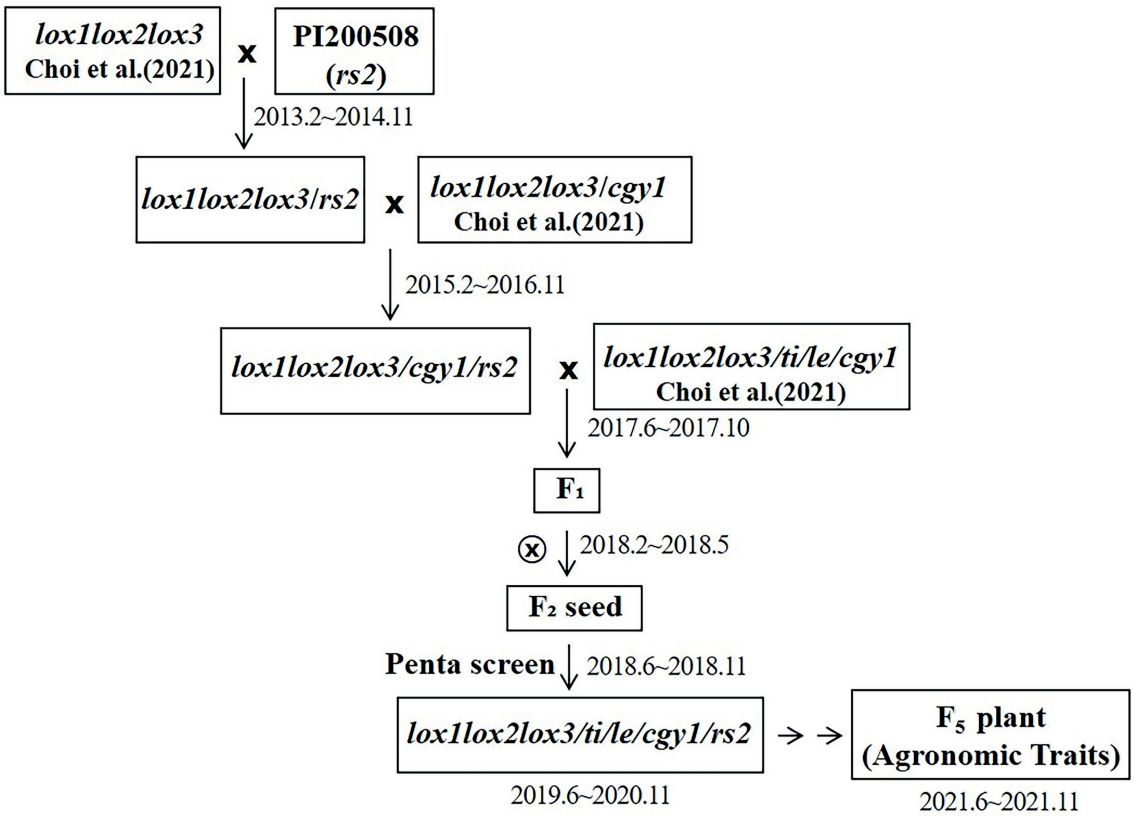


FIGURE 1 | Scheme for the development of penta null alleles (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2rs2*) for lipoxygenase, KTI, lectin, 7S α′ subunit and stachyose components.

30 min in 1 ml of Tris–HCl, pH 8.0, and 1.56% v/v β-mercaptoethanol. About 50 µl of the supernatant collected through centrifugation was added to an equivalent amount of 5× sample buffer containing 1 M Tris–HCl, pH 6.8, 50% v/v glycerol, 1.96% v/v β-mercaptoethanol, and 10% w/v sodium dodecyl sulfate (SDS). The sample obtained was boiled at 97°C for 5 min and centrifuged. About 2 µl of the supernatant was loaded on a 12% acrylamide SDS polyacrylamide gel electrophoresis medium from Owl Separation Systems Inc. (model: P9DS, Portsmouth, NH, United States). After electrophoresis at 120 V for 7 h, the gel was stained. The gel was then destained in destaining solution for several hours. A protein marker (Sigma Marker, Product Code: M4038) was used to identify the presence or absence of lipoxygenase protein (97 kDa) and 7S α' subunit protein (72 kDa).

Determination of KTI and Lectin Protein by Western Blot Analysis

Proteins obtained from the parental seed, each F_2 seed, and random F_5 seeds of the breeding line and cultivar ("Daewon") were separated by 10% or 12% SDS-PAGE and transferred onto an Immobilon-P membrane (PVDF, Millipore). Western blot analysis for the KTI protein was performed as previously described (Krishnan et al., 2000; Krishnan, 2001). Preparation of the antibody and western blot for lectin protein analysis was performed according to a previous method (Vodkin and Raikhel, 1986). The cultivar ("Daewonkong") was used as a control for the presence of KTI and lectin proteins (*TiTi-LeLe* genotype). After blocking for 2 h in TBS buffer containing 0.1% Tween 20, 20 mM Tris (pH 7.5), 150 mM NaCl, and 5% nonfat dried milk (Carnation, Glendale, CA), the membrane was incubated with the antibody of KTI and lectin protein for 1 h. The blot was incubated with a horseradish peroxidase conjugated secondary antibody after washing in TBS buffer. The complex was then visualized using an enhanced chemiluminescence kit (Amersham, Buckinghamshire, United Kingdom). The presence or absence of KTI and lectin proteins was determined visually. In F_2 seed generation, the ratio of segregation for the presence or absence of KTI and lectin proteins was determined by Chi-square analysis.

Determination of *rs2rs2* Genotype and Stachyose Component

Young leaves from each of the 10 F_2 plants possessing tetra null alleles (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*) growing in a greenhouse were used for DNA extraction to screen the plant of the *rs2rs2* genotype based on DNA marker. DNA from parental lines along with each F_2 individual were isolated according to the protocol described by Saghai Maroof et al. (1984). To analyze markers, primers (Forward:5′-

CGTGGAGCAGGTGTATGTGTGG-3′, Reverse:5′-GGCACCA GTCCAACT CCG TTAC-3′) were designed according to previous results (Dierking and Bilyeu, 2008). PCR for the genotype assay was carried out in a PTC-200 thermocycler (MJ Research/ Bio-Rad, Hercules, CA, United States) with the following conditions: 95°C for 5 min followed by 29 cycles of 95°C for 20 s, 65°C for 20 s, 72°C for 30 s, and a final extension at 72°C for 5 min. PCR products were electrophoresed in 2.5% 0.5× TBE agarose gels and were stained with EtBr. Gels were photographed under transmitted UV light. Stachyose content for the breeding line and cultivar ("Daewonkong") was analyzed using the high-performance liquid chromatography (HPLC) method (Sung et al., 2014). The cultivar ("Daewonkong") was used as a control for the normal content of stachyose (*Rs2Rs2* genotype).

Agronomic Traits of Penta Null Genotype

First, 360 random F_5 seeds of the breeding line and seeds of a cultivar ("Daewonkong") as a control were planted in the university field on 15 July 2021. The experimental field had a completely randomized design with three replications. The plots included four rows 3-m long spaced 0.65 m apart. The seeding rate was 30 seeds per row. The soil type was a silty clay loam. Soil K, Ca, Mg, and Na averaged 0.46, 8.84, 2.83, and 0.28 cmolc/kg, respectively. Soil pH was 6.8. Agronomic traits such as maturation date, stem height, number of pods per plant, number of seeds per plant, 100-seed weight, stachyose content, and yield were recorded for the F_5 plant generation of the breeding line (penta null genotype). The mean values of stem height, number of pods per plant, number of seeds per plant, 100-seed weight, stachyose content, and yield were compared by Duncan's multiple range test at the 5% level.

RESULTS

Selection of F_2 Seeds With KTI and Lectin Proteins Free

A total of 172 F_2 seeds were obtained from the cross of the *lox1lox2lox3/lox1lox2lox3-rs2/rs2-cgy1/cgy1* parent and *lox1lox2lox3 /lox1lox2lox3-ti/ti-le/le-cgy1/cgy1* parent. Each seed was analyzed for the segregation of KTI and lectin proteins. KTI protein of 21.5 kDa and lectin protein of 120 kDa were segregated in the F_2 seed generation (**Figure 2**). The segregation data for KTI and lectin proteins in the F_2 seed generation are presented in **Table 2**.

Among the 172 F_2 seeds, 138 F_2 seeds showed KTI protein and 34 F_2 seeds did not show KTI protein. Lectin protein existed in 126 F_2 seeds, and 46 F_2 seeds did not show lectin protein. The segregation ratio for the presence or absence of KTI and lectin proteins in the F_2 seed generation was fitted to an expected 3:1 ratio ($\chi^2 = 2.51$ for KTI and 0.28 for lectin proteins). Between KTI protein and lectin protein, the segregation ratios of 102 *Ti_Le_*: 36 *Ti_lele*: 24 *titiLe_*: 10 *titilele* were observed ($\chi^2 = 2.883$, $p = 0.5–0.1$). Ten F_2 seeds possessing tetra null alleles (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*) were selected and planted to select the plant in the greenhouse with the *rs2rs2* genotype based on the DNA marker.

Selection of F_2 Plants With *rs2rs2* Genotype Using DNA Marker

The *rs2* allele-specific DNA marker showed segregation according to the individual plant in the F_2 plant population consisting of 10 plants with the tetra null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*). Among 10 F_2 plants, eight showed a band (*Rs2*_genotype) and two showed no band (**Figure 3**). Thus, two F_2 plants were identified as having the *rs2rs2* genotype. At maturity, two F_2 plants identified

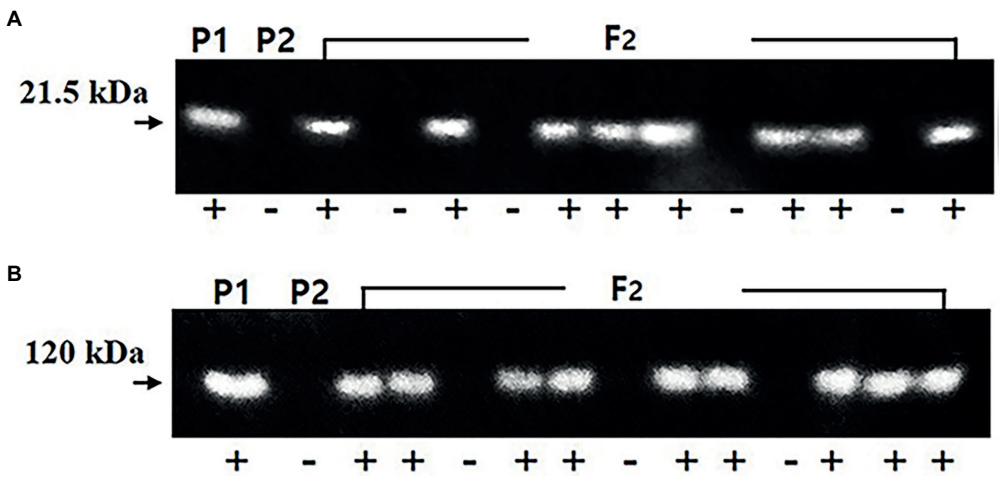


FIGURE 2 | Segregation of KTI **(A)** and lectin **(B)** proteins in the parents and F_2 seeds. Arrows indicate the KTI protein of 21.5 kDa and lectin protein of 120 kDa. P1: *lox1lox2lox3/lox1lox2lox3-Ti/Ti-Le/Le-cgy1/cgy1-rs2/rs2* genotype, P2: *lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-Rs2/Rs2* genotype. +, −: presence and absence of KTI and lectin proteins.

TABLE 2 | Segregation for the presence (+) or absence (–) of Kunitz trypsin inhibitor (KTI) and lectin proteins in F_2 seed generation derived from the cross of a *lox1lox2lox3/lox1lox2lox3-rs2/rs2-cgy1/cgy1* parent and *lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1* parent.

KTI	Lectin	Seed number		χ^2 value (9:3:3:1)	*p*-value
		Observed	Expected		
Present	Present	102	96.75	2.883	0.5–0.1
Present	Absent	36	32.25		
Absent	Present	24	32.25		
Absent	Absent	10	10.75		

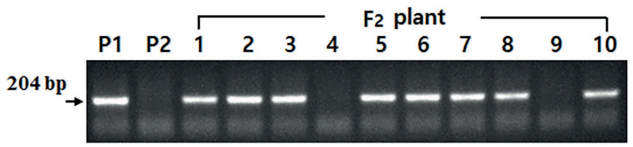


FIGURE 3 | Segregation of DNA marker based on *Rs2* allele in parents and 10 F_2 individual plants with the tetra null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*). P1: *lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-Rs2/Rs2* genotype. P2: *lox1lox2lox3/lox1lox2lox3-Ti/Ti-Le/Le-cgy1/cgy1-rs2/rs2* genotype. F_2 plant—4 and 9; *rs2rs2* genotype.

as having the penta null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*) were harvested separately. The F_3 seeds obtained were used to advance F_5 plant generation.

Confirmation of Penta Null Line

Random F_5 seeds for the breeding line were used to confirm the absence of lipoxygenase, KTI, lectin, and 7S α' subunit proteins (**Figure 4**). Proteins of lipoxygenase and 7S α' subunit were not observed in the mature F_5 seed of the breeding line by using SDS-PAGE analysis. Also, proteins of KTI and lectin were not observed by using western blot analysis. However, these four proteins were observed in the seed of the "Daewonkong" (*Lox1Lox2Lox3 /Lox1Lox2Lox3-Ti/Ti-Le/Le-Cgy1/Cgy1-Rs2/Rs2*) cultivar.

Agronomic Traits of Penta Null Line

The breeding line has purple flowers, tawny pubescence, a determinate growth habit, and light yellow pods at maturity. Some quantitative traits of the breeding line are shown in **Table 3**. The breeding line matured on 16 October, which was 3 days later than "Daewonkong." The stem height of the breeding line was 53.0 cm compared to the cultivar., "Daewonkong," at 48.0 cm. The number of pods per plant and seeds per plant for the breeding line was 52 and 84, respectively. The stachyose content of the breeding line was 2.9 g/kg, which was much lower than 12.7 g/kg of the cultivar., "Daewonkong." The 100-seed weight of the breeding line was 31.1 g, a little higher than that of "Daewonkong" (29.5 g). The yield of the breeding line was 2.80 t/ha, which was slightly higher than that of "Daewonkong" (2.78 t/ha). The plant type harvested and seeds of the breeding line with the penta null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*) are shown

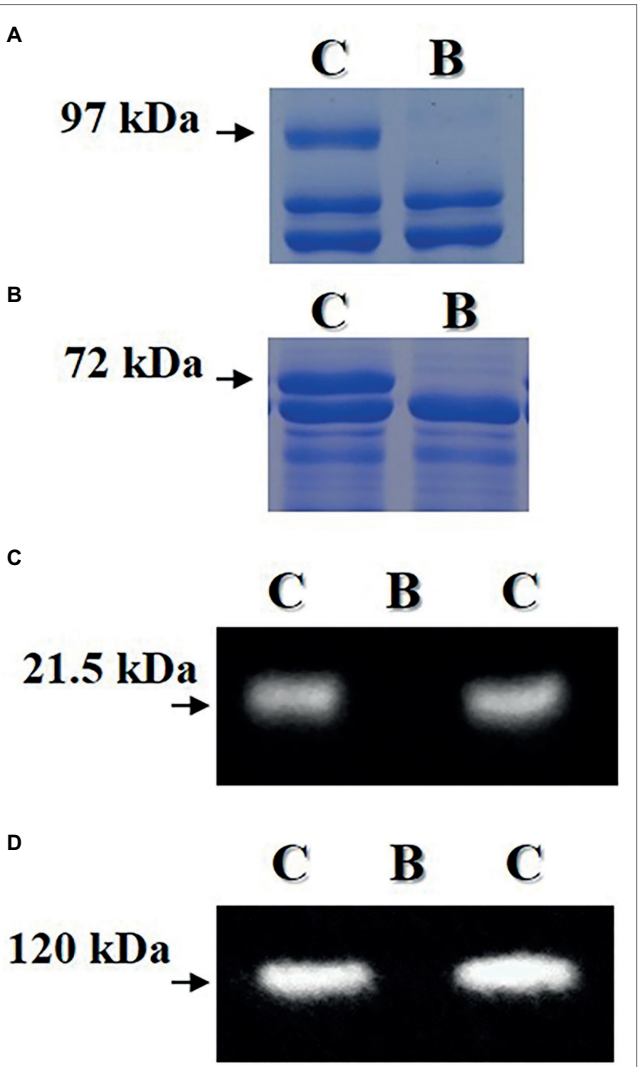


FIGURE 4 | Confirmation of absence of lipoxygenase **(A)**, 7S α′ subunit **(B)**, KTI **(C)**, and lectin **(D)** proteins in mature seed of the breeding line (B) with the penta null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*). Arrows indicate lipoxygenase protein of 97 kDa, 7S α′ subunit protein of 72 kDa, KTI protein of 21.5 kDa and lectin protein of 120 kDa. C: "Daewonkong" (*Lox1Lox2Lox3/Lox1Lox2Lox3-Ti/Ti-Le/Le-Cgy1/Cgy1-Rs2/Rs2* genotype.

in **Figure 5**. The seed of the breeding line has a yellow hilum and yellow seed coat color. The cotyledon color of the mature seed is yellow.

DISCUSSION

Demand for soybean and soybean products has increased in recent years because soybean is an excellent source of protein, oil, carbohydrates, and many other bioactive ingredients for humans. However, the lipoxygenase, KTI, lectin, 7S α' subunit, and stachyose components that exist in the raw mature seeds of soybean are considered antinutritional and allergenic factors

TABLE 3 | Quantitative characteristics of the cultivar ("Daewonkong") and breeding line strain in field conditions during 2021.

Cultivar/breeding line	Traits							
	Planting date	Maturing date	Stem height (cm)	NP/P	NS/P	Stachyose (g/kg)	SW (g)	Yield (Ton/ha)
"Daewonkong"	June 15	October 13	48.0[a]	55[a]	85[a]	12.7[a]	29.5[a]	2.78[a]
Breeding line	June 15	October 16	53.0[a]	52[a]	84[a]	2.9[b]	31.0[a]	2.80[a]

NP/P, number of pods per plant; NS/P, number of seeds per plant; SW, 100-seed weight. The same letters in the column are not significant at the 5% significance level by DMRT.

FIGURE 5 | Appearance of F_5 plants and F_6 seeds possessing penta null alleles (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*).

(Hata et al., 1991; Liener, 1995; Robinson et al., 1995; Katsuya et al., 2007; Krishnan et al., 2009). Heat treatment or other methods are needed to eliminate or reduce these components and secure the efficiency of nutrient absorption and food safety. But these treatments cause some changes and reduce soybean quality (Chen et al., 2019). Genetic removal or reduction of antinutritional and allergenic factors such as lipoxygenase, KTI, 7S α′ subunit, and stachyose components that exist in mature soybean seeds is needed to improve the nutritional values for the soybean food industry. Also, a cultivar with the penta null genotype (*lox1lox1lox2/lox2/ lox3/lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*) for lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose components enhances the utilization of soybean in food as well as feed. A total of 172 F_2 seeds were obtained from the cross of a *lox1lox2lox3/lox1lox2lox3-rs2/rs2-cgy1/cgy1* parent and *lox1lox2lox3 /lox1lox2lox3-ti/ti-le/le-cgy1/cgy1* parent to produce seeds possessing penta null alleles (*lox1lox2lox3/ lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*). KTI and lectin proteins were segregated in the F_2 seed generation (**Figure 2; Table 2**). The segregation ratio for the presence or absence of KTI and lectin proteins was fitted to an expected 3:1 ratio ($\chi^2 = 2.51$ for KTI and 0.28 for lectin protein). This result substantiates previous observations that the presence or absence of KTI and lectin proteins is controlled by a single gene (Orf et al., 1978; Orf and Hymowitz, 1979). Segregation ratios of independent inheritance between KTI protein and lectin protein were observed ($\chi^2 = 2.883$, $p = 0.5–0.1$). This result was consistent with previous studies reporting that both *Ti* and *Le* alleles were independently inherited (Orf and Hymowitz, 1979; Moraes et al., 2006; Lee et al., 2008; Choi et al., 2021). *Ti* allele was independently inherited with the *Le* allele in the F_2 population consisting of 24 plants (Moraes et al., 2006). Orf and Hymowitz (1979) observed that *Le* and *Ti* alleles were inherited independently by the F_2 population with 96 plants. Lee et al. (2008) observed that *Ti* and *Le* alleles were independently inherited in 173 F_2 seeds. Also, Choi et al. (2021) reported that *Le* and *Ti* alleles were independently inherited in F_2 seed generation consisting of 210 seeds. Ten F_2 seeds possessing tetra null alleles (*lox1lox2lox3/ lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*) were obtained (**Table 2**). In a previous study, two F_2 seeds possessing triple null alleles (*ti/ti-le/le-p34/p34*) were obtained from 150 F_2 seeds (Schmidt et al., 2015). Also, three F_2 seeds possessing tetra null alleles (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*) were obtained from 210 F_2 seeds (Choi et al., 2021). Among the 10 F_2 plants with the tetra null genotype (*lox1lox2lox3/lox1lox2lox3-ti/ti-le/le-cgy1/cgy1*), only two F_2 plants were selected by DNA marker (**Figure 3**). Its genotype was found to be *rs2rs2*, and it had low stachyose content compared to the normal cultivar with the *Rs2Rs2* genotype. The result that two individuals had the *rs2rs2* genotype in the F_2 population of 10 individuals was consistent with the results of a previous study (Skoneczka et al., 2009; Yang et al., 2014). The absence of lipoxygenase, KTI, lectin, and 7S *α′* subunit proteins was confirmed in random F_5 seeds for the breeding line possessing the penta null genotype (*lox1lox2lox3/ lox1lox2lox3-ti/ti-le/le-cgy1/cgy1-rs2/rs2*). However, in the "Daewonkong" (*Lox1Lox2Lox3 /Lox1Lox2Lox3-Ti/Ti-Le/Le-Cgy1/Cgy1-Rs2/Rs2*) cultivar., these four proteins were observed (**Figure 4**). For the *rs2* allele, the *rs2rs2* genotype of the breeding line was confirmed by analysis of stachyose content using random F_5 seeds (**Table 3**). The stachyose content of the breeding line was 2.9 g/kg, which was much lower than 12.7 g/kg of the check cultivar., "Daewon" (*Rs2Rs2* genotype). No negative effects on traits of field emergence, seed yield, maturity, height, and fatty acid content between lines derived from PI200508 containing the reduced stachyose content and wild types were reported (Neus et al., 2005). Song et al. (2018) reported that BSH-3 seeds with 7S α′ subunit protein-free accumulated high levels of free amino acids as compared with normal seeds, particularly arginine, and the amounts of several essential amino acids were significantly elevated in BSH-3 seeds. Rani et al. (2020) developed KTI-free soybean with normal protein content. Normal-protein KTI-free RILs were significantly higher in both acidic and basic subunit of glycinin and β-subunit of β-conglycinin fraction compared to low-protein KTI-free RILs. Kumar et al. (2021) developed 21 F_6 KTI-free lines exhibiting 100-fresh seed weight > 45 g, sucrose content > 7%, and morphologically similar to vegetable soybean. Moisture

content and pod yield of KTI-free lines at R6 stage were 64%–74.1% and 7.0–9.5 t•ha^{-1}, respectively. Genetic removal of lipoxygenase-2 improved the speed of emergence of sprouts and the length of the sprouts and sprouting, thereby enhancing the suitability of beans for sprouting. Genetic removal of the KTI gene did not have a significant effect on sprouting attributes, though it enhanced BBI concentration and improved protein digestibility (Kumar et al., 2022).

Quantitative traits of the breeding line with the penta null genotype are shown in **Table 3**. In spite of the absence of antinutritional and harmful immunological components such as lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose that exist in the mature soybean seed, the breeding line germinated, grew, flowered, and reproduced normally in the greenhouse and under field conditions when compared to the cultivar "Daewonkong" (**Figure 5**). Schmidt et al. (2015) reported that plants possessing triple null alleles (*ti*/*ti-le*/*le-p34*/*p34*) flowered and produced seeds without any overt differences in comparison with the standard "Williams 82″ cultivar. No significant differences were observed for stem height, number of pods per plant, number of seeds per plant, 100-seed weight, and yield between the breeding line with the penta null genotype and the "Daewonkong" cultivar. These results indicate that the penta null soybean line had no impact on these agronomic traits. These results suggest that stacking of recessive alleles for *Lox1*, *Lox2*, *Lox3*, *Ti*, *Le*, *Cgy1*, and *Rs2* genes results in a soybean cultivar with significantly reduced antinutritional and allergenic ingredients. This is the first soybean strain with the penta null (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1-rs2*/*rs2*) genotype (free of lipoxygenase, KTI, lectin, and 7S α′ subunit proteins, and low stachyose content). Since mature soybean seeds have more than 20 kinds of allergens, it is considered that some allergens are also present in the seeds of the breeding line with the penta null genotype obtained in this study. However, when compared to common soybean seeds, it seems that allergens are greatly reduced. The breeding line obtained through this study could be used as a valuable parent for the improvement of soybean cultivars that do not contain several antinutritional properties in mature seeds.

CONCLUSION

Lipoxygenase, KTI, lectin, 7S α′ subunit, and stachyose components that exist in mature soybean seeds are considered antinutritional and allergenic factors. The objective of this study was to breed a new soybean strain with a penta null genotype (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy*1/*cgy1-rs2*/*rs2*) for these five components. A total of 172 F_2 seeds were obtained from the cross of a *lox1lox2lox3*/*lox1lox2lox3-rs2*/*rs2-cgy1*/*cgy1* parent and *lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1* parent. Ten F_2 seeds possessing tetra null alleles (*lox1lox2lox3*/*lox1lox2lox3-ti*/*ti-le*/*le-cgy1*/*cgy1*) were selected and two F_2 plants with the *rs2rs2* genotype (low stachyose content) were developed. One F_3 strain with proper agronomical traits was selected. Proteins of lipoxygenase, KTI, lectin, and 7S α′ subunit were not observed in the random mature F_5 seeds of the breeding line with the penta null genotype. The penta null soybean line has purple flowers, a determinate growth habit, light yellow pod, yellow seed coat color, and yellow hilum. The breeding line matured on October 16. The stem height of the breeding line was 53.0 cm. The number of pods per plant and seeds per plant for the breeding line was 52 and 84, respectively. The stachyose content of the breeding line was 2.9 g/kg, which was much lower than 12.7 g/kg of the cultivar., "Daewonkong" (*Rs2Rs2* genotype). The 100-seed weight of the breeding line was 31.1 g. The yield of the breeding line was 2.80 t/ha.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SC, SL, JL, HO, SK, NK, and JC were involved in the experimental design, crossing, protein analysis, planting, harvesting, and data collection and interpretation as well as write-up of this research. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07045483) and Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (Research number: 119011-3).

REFERENCES

Chen, L., Teng, H., and Xiao, J. (2019). A value-added cooking process to improve the quality of soybean: protecting its isoflavones and antioxidant activity. *Food Sci. Hum. Wellness* 8, 195–201. doi: 10.1016/j.fshw.2019.05.001

Choi, S. W., Chae, W. G., Kang, G. Y., and Chung, J. I. (2021). Breeding of tetra null soybean (*Glycine max*) for lipoxygenase, kunitz trypsin inhibitor, lectin, and 7S α′ subunit proteins. *Plant Breed.* 140, 123–129. doi: 10.1111/pbr.12870

Choi, S. W., Han, S. J., Kang, D. H., Kim, J. A., Lee, S. J., and Chung, J. I. (2017). Selection of low Stachyose content of soybean line using *rs2* gene marker. *Korean J. Breed. Sci.* 49, 318–323. doi: 10.9787/KJBS.2017.49.4.318

Chung, J. I. (2014). A new soybean cultivar 'Jinyang': yellow soybean cultivar with Lipoxygenase 1,2,3 protein-free. *Korean J. Breed. Sci.* 46, 328–331. doi: 10.9787/KJBS.2014.46.3.328

Davies, C. S., Coates, J. B., and Nielsen, N. C. (1985). Inheritance and biochemical analysis of four electrophoretic variants of β-conglycinin from soybean. *Theor. Appl. Genet.* 71, 351–358. doi: 10.1007/BF00252079

Davies, C. S., and Nielsen, S. S. (1986). Genetic analysis of a null-allele for lipoxygenase −2 in soybean. *Crop. Sci.* 26, 460–463. doi: 10.2135/cropsci1986.0011183X002600030003x

Dierking, E. C., and Bilyeu, K. D. (2008). Association of a soybean raffinose synthase gene with low raffinose and stachyose seed phenotype. *Plant Genome.* 1, 135–145. doi: 10.3835/plantgenome2008.06.0321

Fling, S. P., and Gregerson, D. S. (1986). Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris-buffer system without urea. *Anal. Biochem.* 155, 83–88. doi: 10.1016/0003-2697(86)90228-9

Hata, Y., Yamamoto, M., and Nakajima, K. (1991). Effects of soybean oligosaccharides on human digestive organs: estimation of fifty percent effective dose and maximum non-effective dose based on diarrhea. *J. Clin. Biochem. Nutr.* 10, 135–144. doi: 10.3164/jcbn.10.135

Hildebrand, D. F., and Hymowitz, T. (1982). Inheritance of Lipoxygenase-1 activity in soybean seeds. *Crop. Sci.* 22, 851–853. doi: 10.2135/cropsci1982.0011183X002200040036x

Hymowitz, T., Collins, F. I., Panezner, J., and Walker, W. M. (1972). Relationship between the content of oil, protein, and sugar in soybean seed. *Agron. J.* 64, 613–616. doi: 10.2134/agronj1972.00021962006400050019x

Jha, P., Kumar, V., Rani, A., and Kumar, A. (2022). Genomic regions governing the biosynthesis of sucrose and raffinose family oligosaccharides in soybean. *J. Plant Biohem. Biotech.* doi: 10.1007/s13562-021-00756-z

Katsuya, M., Toru, T., and Paul, L. M. (2007). Lectin-based food poisoning: A new mechanism of protein toxicity. *PLoS One* 2:e687. doi: 10.1371/journal.pone.0000687

Kim, S. D., Kim, Y. H., Park, K. Y., Yun, H. T., Lee, S. H., and Lee, Y. H. (1997). A new beany tasteless soybean variety "Jinpumkong 2" with good quality. *Korean J. Breed.* 29:502.

Kim, H. T., Ko, J. M., Han, W. Y., Kang, B. K., Lee, Y. H., Lee, B. W., et al. (2015). "Large seed, shattering resistance, and non-beany flavor soybean cultivar Miso." in *2015 Cooperated Symposium of Korean Society of Breeding Science, Next Generation BioGreen 21 Project, and Golden Seed Project*; September 24-28, 1997; 48.

Kitamura, K., Davies, C. S., Kaizuma, N., and Nielsen, N. C. (1983). Genetic analysis of a null-allele for lipoxygenase-3 in soybean seeds. *Crop. Sci.* 23, 924–927. doi: 10.2135/cropsci1983.0011183X002300050026x

Kitamura, K., Davies, C. S., and Nielsen, N. C. (1984). Inheritance of alleles for *Cgy1* and *Cgy4* storage protein genes in soybean. *Theor. Appl. Genet.* 68, 253–257. doi: 10.1007/BF00266899

Kobayashi, A. T., Hirata, N., Kubota, K., and Kitamura, K. (1995). Aroma constituents of soybean milk lacking lipoxygenase isozymes. *J. Agric. Food Chem.* 43, 2449–2452. doi: 10.1021/jf00057a025

Krishnan, H. B. (2001). Characterization of a soybean [*Glycine max* (L.) Merr.] mutant with reduced levels of Kunitz trypsin inhibitor. *Plant Sci.* 160, 979–986. doi: 10.1016/S0168-9452(01)00346-6

Krishnan, H. B., Jiang, G., Krishnan, A. H., and Wiebold, W. J. (2000). Seed storage protein composition of non-nodulating soybean (*Glycine max* (L.) Merr.) and its influence on protein quality. *Plant Sci.* 157, 191–199. doi: 10.1016/S0168-9452(00)00283-1

Krishnan, H. B., Kim, W. S., Jang, S., and Kerley, M. S. (2009). All three subunits of soybean β-conglycinin are potential food allergens. *J. Agric. Food Chem.* 57, 938–943. doi: 10.1021/jf802451g

Kumar, V., Rani, A., Shukla, S., and Jha, P. (2021). Development of Kunitz trypsin inhibitor free vegetable soybean genotypes through marker-assisted selection. *Int. J. Veg. Sci.* 27, 364–377. doi: 10.1080/19315260.2020.1800886

Kumar, V., Rani, A., Tayalkar, T., Jha, P., Mittal, P., Anshu, A. K., et al. (2022). Genetic elimination of lipoxygenase-2 improves sprouting and tocopherols, and removal of Kunitz trypsin inhibitor enhances in vitro protein digestibility in soybean (*Glycine max*). *Plant Breed.* 141, 63–70. doi: 10.1111/pbr.12998

Kunitz, M. (1945). Crystallization of a soybean trypsin inhibitor from soybean. *Science* 101, 668–669. doi: 10.1126/science.101.2635.668

Lee, K. J., Park, M. S., Sung, M. K., Kim, M. S., and Chung, J. I. (2008). Inheritance between *Le* gene and *Ti* gene in soybean (*Glycine max* L.). *Korean J. Breed Sci.* 40, 97–100.

Liener, I. E. (1995). Possible adverse effects of soybean anticarcinogens. *J. Nutr.* 125, 744–750.

MaCleod, G., and Ames, J. (1988). Soya flavor and its improvement. *CRC Crit. Rev. Food Sci. Nutr.* 27, 219–400. doi: 10.1080/10408398809527487

Medic, J., Atkinson, C., and Hurburgh, C. R. (2014). Current knowledge in soybean composition. *J. Am. Oil Chem. Soc.* 91, 363–384. doi: 10.1007/s11746-013-2407-9

Moraes, R. M. A., Soares, T. C. B., Colombo, L. R., Salla, M. F. S., Barros, J. G. A., and Piovesan, N. D. (2006). Assisted selection by specific DNA markers for genetic elimination of the Kunitz trypsin inhibitor and lectin in soybean seeds. *Euphytica* 149, 221–226. doi: 10.1007/s10681-005-9069-0

Neus, J. D., Fehr, W. R., and Schnebly, S. R. (2005). Agronomic and seed characteristics of soybean with reduced raffinose and stachyose. *Crop. Sci.* 45, 589–592. doi: 10.2135/cropsci2005.0589

Orf, J. H., and Hymowitz, T. (1979). Inheritance of the absence of the Kunitz trypsin inhibitor in seed protein of soybeans. *Crop. Sci.* 19, 107–109. doi: 10.2135/cropsci1979.0011183X001900010026x

Orf, J. H., Hymowitz, T., Pull, S. P., and Pueppke, S. G. (1978). Inheritance of a soybean seed lectin. *Crop. Sci.* 18, 899–900. doi: 10.2135/cropsci1978.0011183X001800050058x

Pull, S. P., Pueppke, S. G., Hymowitz, T., and Orf, J. H. (1978). Soybean lines lacking the 120,000 Daltons seed lectin. *Science* 200, 1277–1279. doi: 10.1126/science.200.4347.1277

Rani, A., Kumar, V., Shukla, S., Jha, P., Tayalkar, T., and Mittal, P. (2020). Changes in storage protein composition on genetic removal of Kunitz trypsin inhibitor maintain protein content in soybean (*Glycine max*). *J. Agric. Food Res.* 2:100065. doi: 10.1016/j.jafr.2020.100065

Robinson, D. S., Wu, Z., Domoney, C., and Casey, R. (1995). Lipoxygenases and the quality of foods. *Food Chem.* 54, 33–43. doi: 10.1016/0308-8146(95)92659-8

Rosso, M., Shang, C., Song, Q., Escamilla, D., Gillenwater, J., and Zhang, B. (2021). Development of breeder-friendly KASP markers for low concentration of Kunitz trypsin inhibitor in soybean seeds. *Int. J. Mol. Sci.* 22:2675. doi: 10.3390/ijms22052675

Saghai Maroof, M. A., Soliman, K. M., Jorgensen, R. A., and Allard, R. W. A. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 81, 8014–8018. doi: 10.1073/pnas.81.24.8014

Schmidt, M. A., Hymowitz, T., and Herman, E. M. (2015). Breeding and characterization of soybean triple null; a stack of recessive alleles of Kunitz trypsin inhibitor, soybean agglutinin, and P34 allergen nulls. *Plant Breed.* 134, 310–315. doi: 10.1111/pbr.12265

Skoneczka, J. A., Saghai Maroof, M. A., Shang, C., and Buss, G. R. (2009). Identification of candidate gene mutation associated with low stachyose phenotype in soybean lines PI200508. *Crop. Sci.* 49, 247–255. doi: 10.2135/cropsci2008.07.0403

Song, B., Oehrle, N. W., Liu, S., and Krishnan, H. B. (2018). Development and characterization of a soybean experimental line lacking the α′ subunit of β-Conglycinin and G1, G2, and G4 Glycinin. *J. Agric. Food Chem.* 66, 432–439. doi: 10.1021/acs.jafc.7b05011

Sung, M. K., Han, S. J., Seo, H. J., Choi, S. W., Nam, S. H., and Chung, J. I. (2014). Genotype and environment influence on raffinose and stachyose content of soybean seed. *Korean J. Crop Sci.* 59, 319–324. doi: 10.7740/kjcs.2014.59.3.319

Vodkin, L. O., and Raikhel, N. V. (1986). Soybean Lectin and related proteins in seeds and roots of Le+ and Le- soybean varieties. *Plant Physiol.* 81, 558–565. doi: 10.1104/pp.81.2.558

Wang, J., Kuang, H., Zhang, Z., Yang, Y., Yan, L., Zhang, M., et al. (2020). Generation of seed lipoxygenase-free soybean using CRISPR-Cas9. *Crop J.* 8, 432–439. doi: 10.1016/j.cj.2019.08.008

Wilson, S. C., Martinez, V., and Eg, D. M. (2008). Purification, thermal stability, and antigenicity of the immunodominant soybean allergen P34 in soy cultivars, ingredients, and products. *J. Food Sci.* 73, T106–T114. doi: 10.1111/j.1750-3841.2008.00834.x

Yang, K. W., Ko, J. M., Ha, T. J., Lee, Y. H., Back, I. Y., Yang, T. J., et al. (2014). Development of molecular markers for low raffinose and stachyose in Korean soybean cultivars. *Plant Breed. Biotech.* 2, 151–157. doi: 10.9787/PBB.2014.2.2.151

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Choi, Ly, Lee, Oh, Kim, Kim and Chung. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Sustainable Food Systems

ORIGINAL RESEARCH
published: 08 July 2022
doi: 10.3389/fsufs.2022.908231

Prospects and Potential of *Bradyrhizobium diazoefficiens* Based Bio-Inoculants on Soybean Production in Different Agro-Ecologies of Mozambique

Canon Engoke Norris Savala[1]*, *Alexander Nimo Wiredu*[1], *David Chikoye*[2] and *Stephen Kyei-Boahen*[1]

[1] Plant Production and Health, International Institute of Tropical Agriculture, Nampula, Mozambique, [2] Plant Production and Health, International Institute of Tropical Agriculture, Lusaka, Zambia

Soybean production in sub-Saharan Africa is increasing as farmers open more land areas for cultivation and replace other crops, such as tobacco, in favor of this legume crop. Despite the production is increased in Mozambique, demand for animal feed and oil is not satisfied. As such, farmers explore ways to improve yield per unit area of soybean by using bio-inoculants from various sources and agroecological adaptability. These bio-inoculants are seldom available during planting time, and retail at almost similar prices although yield varied based on the product source, handling, and the rhizobia strain carrier. Mozambique does not produce bio-inoculants, so it obtains the product from neighboring countries or as far as the South American continent. In this study, we evaluated the performance, ecological adaptability, and soybean productivity of seven *Bradyrhizobium diazoefficiens* strain-based bio-inoculants from several countries with different carrier materials: Biofix, Masterfix, Nitrofix, NitroZam, N-Fixer, Soygro Peat, and Soygro Liquid against a control (non-inoculated) on two soybean varieties Storm and TGx 1904-6F. The trial was conducted in the 2016 and 2017 growing seasons in three agroecologies of Mozambique at Angonia, Nampula, and Ruace. Data on nodulation, plant growth, biomass nitrogen content at beginning of podding (R3) stage, yield, and yield components of soybean were evaluated. Analysis of variance and contrast comparisons were performed on the Statistical Analysis System® 9.4. Nodule weight per plant variedly increased from 7.7 to 167.6 mg with inoculation of both varieties across environments. Plant tissue nitrogen content at the R3 stage was higher in inoculated non-promiscuous variety at 3.9% than the promiscuous counterpart with 3.7%. Storm, a non-promiscuous short-maturity variety of soybean, responded to inoculation and accumulated more N than the medium-to-late maturity promiscuous TGx 1904-6F. Higher N tissue content is an indicator of better nutritive value, as well as high-quality recyclable biomass of inoculated soybean. Both Storm and TGx 1904-6F responded to all inoculants variedly with NitroZam yield of 2,750 kg ha^{-1} being highest, while Soygro Liquid was lowest with 2,051 kg ha^{-1} but more than the check with 1,690 kg ha^{-1} across sites.

OPEN ACCESS

Edited by:
Petr Smýkal,
Palacký University Olomouc, Czechia

Reviewed by:
Tomas Vymyslicky,
Agricultural Research, Ltd., Czechia
Arvind Kumar Rai,
Indian Council of Agricultural Research (ICAR), India

***Correspondence:**
Canon Engoke Norris Savala
cnsavala@gmail.com

Specialty section:
This article was submitted to Agroecology and Ecosystem Services, a section of the journal Frontiers in Sustainable Food Systems

Received: *30 March 2022*
Accepted: *03 June 2022*
Published: *08 July 2022*

Citation:
Savala CEN, Wiredu AN, Chikoye D and Kyei-Boahen S (2022) Prospects and Potential of Bradyrhizobium diazoefficiens Based Bio-Inoculants on Soybean Production in Different Agro-Ecologies of Mozambique. Front. Sustain. Food Syst. 6:908231. doi: 10.3389/fsufs.2022.908231

There were varietal differences in 100-seed weight after inoculation where Storm (15.4 g) had heavier seeds than TGx 1904-6F (13.1 g). The results show that inoculation improved plant growth and development, increased nodulation, and gave higher yields for better economic returns among farmers. Inoculation has the potential of increasing soybean yield, nutritive value, and biomass quality within Mozambique.

Keywords: biological nitrogen fixation, ecology, nodulation, non-promiscuous, promiscuous, rhizobia, yield

INTRODUCTION

Soybean production is becoming an important crop among the farming communities in Mozambique. Efforts by several agricultural development organizations, such as the International Institute of Tropical Agriculture (IITA), have increased the adoption of improved soybean varieties to 89% (AGRA, 2016). For instance, IITA and the *Instituto de Investigação Agrária de Moçambique* (IIAM) (Agricultural Research Institute of Mozambique) have released and promoted 14 improved soybean varieties. The soybean production system in Mozambique is dominated by smallholder farmers who face a myriad of limitations. These farmers seldom use the necessary inputs, such as quality seeds, fertilizers, and pesticides (Matteo et al., 2016). This situation is occasioned by a lack of access to and unaffordable prices at the time of planting and during the production period. However, the cost of using fertilizer in particular nitrogen could be averted through the use of bio-inoculants. Bio-inoculants have been shown to increase the yield of soybean (Chibeba et al., 2020; Savala et al., 2021) and other legumes, such as cowpea (Kyei-Boahen et al., 2017), in Mozambique. The use of bio-inoculants is also limited by the lack and inaccessibility of the input at the time of planting. Mozambique as a country does not have the capacity to produce bio-inoculants despite studies showing that locally isolated rhizobia perform comparably well to established strains *Bradyrhizobium diazoefficiens* (USDA110) (Chibeba et al., 2017). Instead, the country relies on the government or selected agricultural development organizations that import bio-inoculants from around the region or as far as the South American continent. The quality of the imported bio-inoculant is often compromised due to a lack of proper transportation and handling facilities. With all these prevailing challenges, soybean production in Mozambique has remained below its production potential, and therefore does not meet the growing soybean grain demand from the local livestock feed industry and for human consumption (Matteo et al., 2016). As such the country spends a lot of foreign exchange on the importation of soybean and its related products in addition to bio-inoculants to meet local demand. Fortunately, most of the available varieties are promiscuous that can freely form a symbiotic association and nodulate with existing native rhizobia strains in the soil (Gyogluu et al., 2016), although the application of bio-inoculants improves nodulation and efficiency of nitrogen fixation.

Not only does Mozambique lack the capacity of producing bio-inoculants but also accompanying policies and information on the handling, trading, and use are absent. This leaves the trade and use of bio-inoculants unregulated resulting in the importation of different bio-inoculant types/grades depending on the accessibility. These bio-inoculants grades (liquid or solid mostly peat based) are often transported like regular cargo from the importing agents to the farmers' fields. No standards exist for the bio-inoculant handlers to follow, hence jeopardizing their effectiveness to improve the biological nitrogen fixation (BNF) process that would have otherwise resulted in increased legume yields. Farmers in Mozambique use seed inoculation of both liquid- and peat-based bio-inoculants plus other good management practices to enhance the efficiency of the BNF process in soybean (Matteo et al., 2016; Savala and Kyei-Boahen, 2020). Although the use of peat-based bio-inoculants is higher than the liquid-based counterparts, their quality varies a lot depending on the source due in part to contamination as most of the bio-inoculants use peat as the carrier material (Woomer et al., 2014). Studies have established that for resource-poor farmers, the use of bio-inoculant alone could go a long way to improve legume yields due to the economic incentive on returns associated with the practice (Kyei-Boahen et al., 2017). Countries that supply Mozambique with bio-inoculants include Argentina, Brazil, Kenya, Malawi, South Africa, Zambia, and Zimbabwe. Past BNF trials conducted in Mozambique have also included bio-inoculants from the United States of America (Savala et al., 2022). The experiments demonstrated the importance of both liquid- and peat-based bio-inoculants applied either on seed or directly to the soil.

The use of bio-inoculants in soybean production tends to be more beneficial when an adequate supply of phosphorus fertilizer plus other necessary agronomic practices are maintained (Singleton et al., 1985, 2001; Tittabutr et al., 2007). Some of the soybean agronomic management practices used by farmers include adjusting the planting time to take advantage of the rainfall condition, variety selection, especially the promiscuous genotypes, and weeds and pest control. However, the use of inorganic fertilizer is limited among Mozambican farmers due to its obnoxiously high price (Zavale et al., 2020). As such non-monetary options like site selection or low-cost investment, such as bio-inoculant use, are often preferred. Soybean is mainly produced in the medium-to-high altitude agroecologies of Mozambique (Walker et al., 2006; AGRA, 2016; Matteo et al., 2016). These ecologies are in Manica, Nampula, Tete, and Zambezia provinces where farmers are increasingly growing promiscuous varieties. Promiscuous varieties can self-nodulate with native rhizobia strains that are present in the soils (Chibeba et al., 2017). Promiscuous soybean varieties can be grown without bio-inoculants, although studies indicate that

TABLE 1 | Soil characteristics of the three sites used for the inoculation trial.

Location	pH	P±	K	Ca	Mg	Na	EC	CEC	N	Sand	Silt	Clay
		ppm					dS/cm	cmolc/kg	%			
Angonia	6.4	25.0	122.8	772.8	165.5	29.4	0.059	6.6	0.09	64.0	6.6	29.4
Nampula	6.6	0.3	90.4	800.5	113.0	29.3	0.050	6.0	0.13	63.2	2.0	34.8
Ruace	6.2	44.8	421.0	1,755.0	301.8	17.9	0.057	15.0	0.15	56.8	12.1	31.1

±*Phosphorus determined using the Olsen method.*

nodulation and more yield is realized following inoculation. Inoculation of promiscuous soybean varieties yielded more than the non-inoculated ones in many sub-Saharan African countries (Majengo et al., 2011; van Heerwaarden et al., 2018; Chibeba et al., 2020). Although the information on differences in bio-inoculant carriers is lacking, the effectiveness of either peat- or liquid-based bio-inoculants is not documented in the country to guide farmers on selection. Such data could be important in drafting policies on the price, handling, transportation, and use of bio-inoculants in the country. Moreover, this would guide investors to the most suitable bio-inoculant carrier to use if a facility was to be established in Mozambique. Therefore, our study evaluated the performance, ecological adaptability, and soybean productivity of seven *Bradyrhizobium diazoefficiens* strain-based bio-inoculants acquired from several countries with different carrier materials against a control on nodulation and yield gain of two soybean varieties Storm and TGx 1904-6F. Secondly, we seek to determine the agroecological suitability of bio-inoculants for soybean production in Mozambique.

MATERIALS AND METHODS

Site Descriptions

Soybean field studies were conducted in the 2016 and 2017 growing seasons at three locations in Mozambique: Angonia 14.6273° S, 33.9812° E, 1,115 meters above sea level (m a.s.l.), Ruace 15.2343° S, 36.6883° E, 677 m a.s.l., and Nampula 15.2738° S, 39. 3143° E, 377 m a.s.l.. The sites have monomodal rainfall between December and April averaging at Angonia 900–1,400 mm, Ruace 800–1,400 mm, and Nampula 600–1,100 mm. As such the Angonia and Ruace are characterized by a maize legume cropping system, while Nampula is dominated by a cassava legume cropping system. Soybean is largely produced in these systems as a monoculture for commercial purposes as poultry feed although a little amount is utilized in households to enrich/fortify pastries. In both cropping systems, farmers rarely use fertilizers but mostly rely on the inherent soil fertility with limited residue recycling back to the fields. For each season, new fields previously under maize for two growing periods were used. According to the Soils Atlas of Africa, the predominant soil type at the sites in Angonia and Ruace are Chromic Luvisols, while that in Nampula is Haplic Lixisols (Jones et al., 2013). Ten soil samples were taken from 0 to 30 cm soil layer using a soil auger in a W pattern across the field for the trial before plowing or harrowing. Soils from each site were combined into a composite sample and four subsamples were drawn for chemical and particle-size analysis (**Table 1**). The pH was determined using a high impedance voltmeter on 1:2 soil–water suspension. Total organic carbon was determined by Walkley–Black method, total N by the Kjeldahl method, P by Olsen's method, and K using ICP-OES after extraction with Mehlich 3. The soil rhizobial population of the experimental field was determined from soil samples taken from all three locations (**Table 2**) using the most probable number method (Vincent, 1970; Woomer et al., 1990).

TABLE 2 | Soybean rhizobia most probable number (MPN) in soils collected from the experimental site with and without a history of soybean cultivation.

Soil sample	Estimated Cells/g	
	Soybean history	No soybean history (years)
Angonia	1.4×10^3	75
Nampula	1.1	0
Ruace	1.0×10^3	2

Bio-Inoculants

Several *Bradyrhizobium diazoefficiens* strains USDA110 formerly known as *Bradyrhizobium japonicum* (Delamuta et al., 2013) bio-inoculant products obtained from the regions across Africa and South America were used for the study (**Table 3**). The bio-inoculants namely Biofix, Masterfix, Nitrofix, NitroZam, N-Fixer, Soygro Peat, and Soygro Liquid. Two bio-inoculants Biofix and Masterfix were imported from outside the region. Biofix a peat-based bio-inoculant was sourced from a fertilizer-producing company in Kenya, while Masterfix was acquired from Brazil. Three peat-based bio-inoculants used in the study were obtained from the National Agricultural Research Institutes of Malawi (Nitrofix), Zambia (NitroZam), and Zimbabwe (N-Fixer). Also, peat- and liquid-based bio-inoculants by the trade name Soygro were sourced from South Africa. All the peat-based bio-inoculants were applied to the seed using inert gum Arabic solution as the adhesive agent. According to the manufacturer's recommendation, the bio-inoculants were applied to supply $\sim 10^6$ rhizobia cells seed^{-1}.

Experimental Design

Two soybean varieties, Zamboane (TGx 1904-6F), a promiscuous variety released by IITA and IIAM and Storm (non-promiscuous Seed Co. variety) were used in the experiment. Bio-inoculant products were evaluated against a control. The field experimental

TABLE 3 | Minimum manufacturer's specifications for colony-forming units per gram (CFU g^{-1}) or milliliter (CFU ml^{-1}) and carrier of the experimental inoculants.

Inoculant	Formulation/carrier	Colony forming units
Biofix	Sugar bagasse peat	1.0×10^9
MasterFix*	Peat	5.0×10^9
Nitrofix	Peat	4.0×10^9
NitroZam	Peat	1.0×10^7
N-Fixer	Peat	–
Soygro liquid§	Liquid	2.5×10^9
Soygro peat	Peat	1.0×10^9

**Bradyrhizobium elkanii and Bradyrhizobium japonicum.*
§*Colony-forming units per milliliter.*

design was a randomized complete block with four replications for each of the bio-inoculants. Plots consisted of five rows of 9 m length, 0.50 m row-spacing, and 0.1 m space between plants within rows. Land preparation was accomplished with disc plowing followed by two passes with a disc harrow. Seeds were inoculated in the field under a shade by weighing 1.0 kg of seeds of each cultivar into separate plastic bags for each bio-inoculant then a respective adhesive agent as per the manufacturer's guidelines was added. The seeds were mixed thoroughly to enable coating of each seed with the gum Arabic adhesive agent. Thereafter, respective bio-inoculants were added and further mixed thoroughly until all the seeds were completely covered with bio-inoculant (**Table 3**) before planting. In both the 2016 and 2017 seasons, the experiments were planted between 18 and 30 December, although scheduling was site-specific depending on the onset of rains in each location. The risk of contamination was avoided by planting the non-inoculated plots first before planting the inoculated ones. Furthermore, the same individuals handled the planting of each bio-inoculant for all four replicates. Planting and weeding (twice) were done manually using a hand-held hoe. To control pests, insecticides were applied using 100 ml of Cypermethrin (200 g active ingredient/l) and 50 ml of Lambda Cyhalothrin (50 g active ingredient/l) in 15 l knapsack sprayers at beginning of flowering.

Data Collection

Data on nodulation, plant growth, biomass nitrogen content at the R3 stage (beginning of podding), yield, and yield components were evaluated. At the R3 growth stage, when pods had reached 10–12 mm long at one of the four uppermost nodes on the main stem, it is regarded as the peak demand for nitrogen in soybean; five soybean plants were randomly selected from each plot and the roots were excavated using a hoe and spade. The soil was carefully washed out from the roots to ensure that all the nodules were recovered as much as possible. The nodules were plucked from the roots by hands, counted, subsequently placed in envelopes, and dried in an oven at 60°C for 48 h. The dried nodules were weighed to determine the dry weight. The remaining plant biomass was dried in an oven at 60°C until a stable weight was attained, and dry weight was determined. Later the biomass was ground to pass through a 2-mm mesh sieve for plant tissue N analysis using inductively coupled plasma optical emission spectrometry (ICP-OES), after extraction with nitric acid and hydrochloric acid. At maturity, 1 m^2 area (two middle rows, 1 m long) was randomly selected and harvested to determine above-ground biomass, pod density, and 100-seed weight. The rest of the plots were harvested manually and pods from each plot were threshed manually, and grain yield was determined. The moisture content of grain samples from each plot was measured using the Farmex MT-16 Grain Moisture Tester (AgraTronix LLC, Streetsboro, Ohio, USA) and grain yield in kg/ha was adjusted to 13% moisture content. After threshing the pods, the above-ground plant biomass from the 1 m^2 plot area was sun-dried for 10 days to 10% moisture content to determine dry matter yield. Except for land preparation, all field activities were done manually.

Data Analysis

Analyses of variance (ANOVAs) were performed on the data collected using PROC GLM in the Statistical Analysis System (SAS)® 9.4 (SAS/STAT®, 2018). A combined analysis across locations and cropping seasons was performed. First, an all locations combined ANOVA was performed to evaluate the effects of the environment, variety, bio-inoculants, and their interactions. As ecological conditions at the locations were different and dominant for all the variables analyzed, effects associated with the growing period were confounded with the weather within seasons (Moore and Dixon, 2015). Thus, the environment was considered a random effect. Environmental effects were significant for all the variables. Consequently, environment and replication were treated as random effects (Moore and Dixon, 2015). A second ANOVA for each environment was performed with a variety and bio-inoculant factors considered as fixed effects. Means were determined for treatments, and comparisons were done using the least significant difference (LSD) at a $P \leq 0.05$ significance level (SAS/STAT®, 2018). Correlation analysis of yield components was conducted to determine the effect of the variable on grain yield. Also, contrast analyses by environment and by locations for combined growing seasons were performed on mean grain yield. In this study, we evaluated the performance, ecological adaptability, and soybean productivity of seven *Bradyrhizobium diazoefficiens* strain-based bio-inoculants with different carrier materials on two soybean varieties.

RESULTS

Soil and Bio-Inoculant Properties

The soils at the three sites were sandy clay loam with pH ranging from 6.2 to 6.6 but the CEC for the Ruace soil was more than double that for the Angonia and Nampula sites (**Table 1**). The soil available P, K, and Mg were lower in Nampula than Angonia and Ruace. Total N was relatively lower in Angonia (0.09%) compared to that for Nampula (0.13%) and Ruace (0.15%). The indigenous rhizobial population was relatively high in Angonia and Ruace, whereas the population in Nampula was very limited (**Table 2**). The low populations of indigenous rhizobia found in Angonia, Nampula, and Ruace soils were not adequate for an

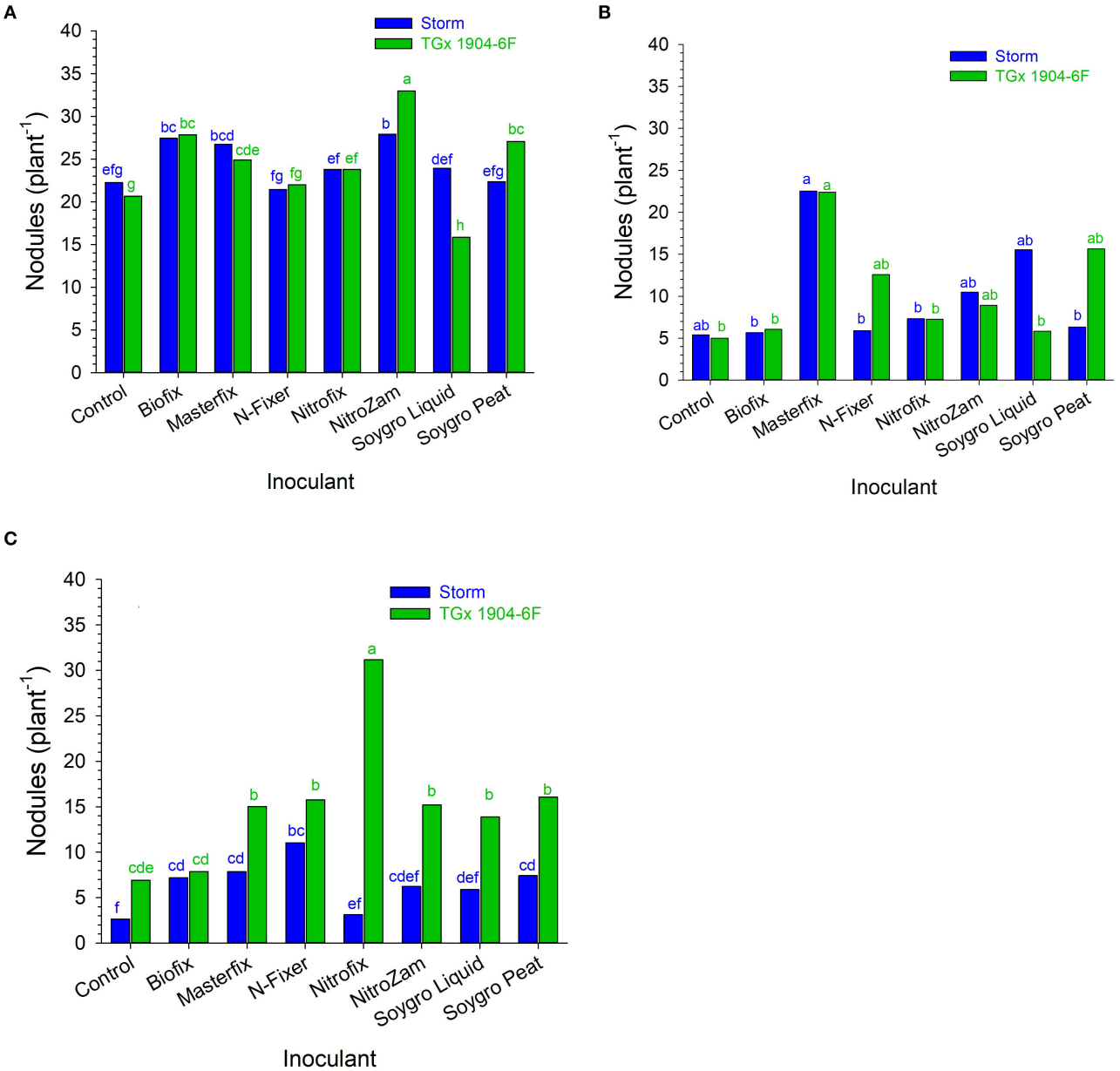


FIGURE 1 | Nodule population (plant^{-1}) following inoculation of soybean at **(A)** Nampula 2016, **(B)** Ruace 2016, and **(C)** Ruace 2017. Small letters on the bars indicate no significance while different letters show significant differences in means.

effective BNF process necessitating the use of bio-inoculants. **Table 3** shows the populations of the colony-forming units/cells per g or ml of the carrier for the inoculant products. NitroZam inoculant from Zambia had the lowest concentration of colony-forming units, whereas Masterfix had the highest concentration.

Influence of Bio-Inoculants on Nodulation

The main purpose of inoculation is to improve nodulation in legumes through the symbiotic association of the rhizobia strain and the host plant. It is presumed that improved nodulation leads to an increased BNF process that avails more N for plant growth and development. Overall, inoculation resulted in the formation of more nodules on soybean roots (**Figures 1**, **2**). Nodule population per plant was higher for the inoculated soybean than in the control fields. When segregated on a variety basis, TGx 1904-6F formed more nodules than Storm (**Figure 1**). The number of nodules was also distinct among bio-inoculants at different sites with Masterfix forming the highest nodules per plant in four out of the six environments. Nodule population on soybean roots also differed among sites. For instance, more

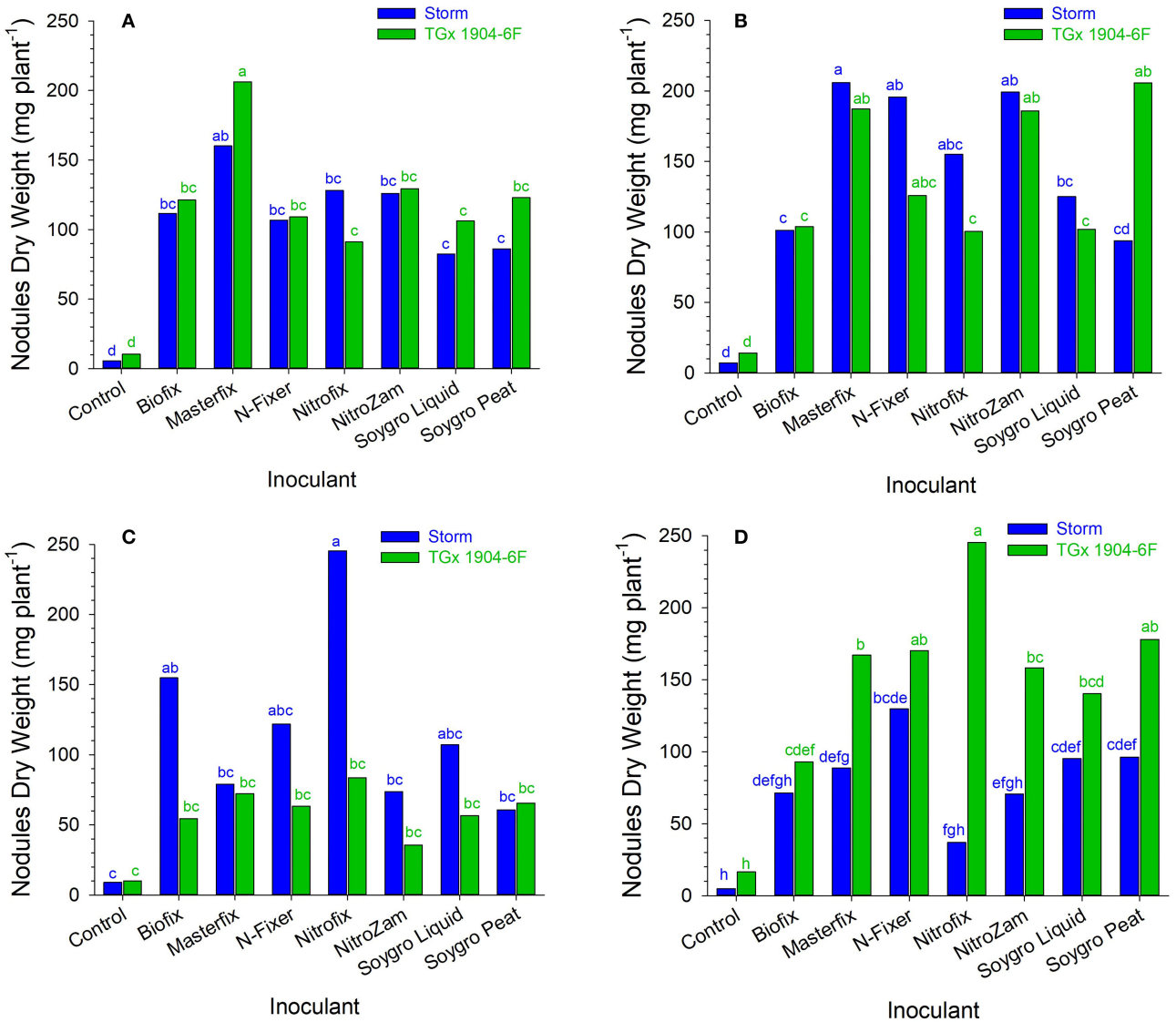


FIGURE 2 | Nodule dry weight (mg plant^{-1}) following inoculation of soybean at **(A)** Angonia 2016, **(B)** Ruace 2016, **(C)** Angonia 2017, and **(D)** Ruace 2017. Small letters on the bars indicate no significance while different letters show significant differences in means.

nodules were formed in Nampula than Ruace in the 2016 growing season (**Figure 1**). Upon referring to the most probable number of soybean rhizobia strains in the site soils, Nampula seldom had active cells to form the symbiotic relationship necessary for the BNF process (**Table 2**).

Like the nodule population, nodule dry weight per plant for all the bio-inoculants were significantly higher than that for the control (**Figure 2** and **Table 4**). Inoculation significantly increased nodule dry weight per plant from 11.4 mg (control) to an average of 135 mg (**Table 4**). The nodule dry weight of a particular treatment at a specific site generally was not related to those with nodule population due to the variation in the nodule sizes. Thus, treatment with the highest number of nodules may not necessarily have the highest nodule dry weight because of variability in nodule sizes. Although significantly more nodules were formed on soybean roots in Nampula (**Figure 1**), the nodule dry weight (data not presented) did not show any notable trend for the different bio-inoculants and seasons due to the high variability in the data. According to the manufacturer's specifications, Masterfix had the highest units forming colonies per gram (5.0×10^9 CFU g^{-1}) (**Table 3**). When bio-inoculants were averaged across environments, Masterfix formed heavier nodules per plant (163.0 mg) against the lowest value of 115.1 mg for Soygro liquid (**Table 4**). In addition, nodulation was distinct by season; nodules formed in 2016 (142.3 mg per plant) were significantly heavier than in the 2017

TABLE 4 | Site nodule weight (plant^{-1}) in the 2016 and 2017 growing seasons.

Inoculant	Angonia		Nampula		Ruace	
	2016	2017	2016	2017	2016	2017
Control	7.7	9.6	10.9	12.2	10.7	17.0
Biofix	116.5	104.7	102.5	82.2	151.7	163.3
Masterfix	183.2	75.7	196.6	127.8	202.8	192.0
N-Fixer	108.0	92.7	160.8	150.1	138.3	156.8
Nitrofix	109.7	153.7	127.7	126.2	158.0	156.9
NitroZam	127.7	54.7	192.5	114.5	177.3	184.8
Soygro liquid	94.3	81.9	113.5	117.7	122.3	160.8
Soygro peat	104.6	63.1	149.8	137.1	170.4	164.0
LSD	21.2	57.5	32.0	28.1	2.0	88.2
%CV	19.3	16.3	23.5	24.3	33.1	27.6

growing season with 126.7 mg when averaged across the bio-inoculant source although the reverse was observed in Ruace, where 2016 had lighter nodules than 2017.

Biomass Quality at the R3 Stage

The results indicated that there were no statistical differences in the plant tissue N content at podding stage across sites. However, nitrogen content in the plant tissue for the varieties in Angonia and Nampula in 2016 were different (**Table 5**) and there was a significant interaction between variety and bio-inoculant at Nampula in the same season. In the 2016 growing season, Storm plant tissue had a significantly higher percent N than that for TGx 1904-6F for both sites at Angonia and Nampula. The N tissue contents for Ruace 2016, Angonia, and Nampula 2017 were comparable at 4.4, 3.3, and 4.0%, respectively (**Table 5**). At Nampula in 2016, the plant tissue N content for the varieties varied with the bio-inoculant application. Plant tissue of Storm plants inoculated with Nitrofix had the highest N content (250.3 kg N ha^{-1}), while the control had the lowest (110.8 kg N ha^{-1}), which was not different from that of Soygro liquid bio-inoculant. On the other hand, TGx 1904-6F plants inoculated with N-fixer had the highest tissue N content (214.3 kg N ha^{-1}), whereas the control was the lowest (102.9 kg N ha^{-1}). The Storm had more N in biomass at 165.9 kg N ha^{-1} than TGx 1904-6F with 147.7 kg N ha^{-1} (**Figure 3**). Storm development is faster than the TGx 1904-6F, which is a medium maturity variety. There were more pods initiated on Storm than TGx 1904-6F by the time of sampling hence higher N demand for the variety. In general, the application of bio-inoculants resulted in more plant N uptake at the Nampula site, which had the lowest number of indigenous rhizobia.

Yield and Yield Components

The data showed that site effects were prominent for yield components. Therefore, the components were analyzed for each site and season averaged across varieties. Only Masterfix and NitroZam increased 100-seed weight at Angonia in 2016, whereas in 2017, four bio-inoculants, including Biofix, Masterfix, Nitrofix, and NitroZam, increased the 100-seed weight at this site (**Table 6**). No significant response to inoculation occurred for 100-seed weight at Nampula and Ruace except for the increase in 100-seed observed for N-Fixer at Ruace in 2017. Across all the sites, soybean seeds were bigger in the 2017 than in the 2016 season. Harvest biomass, on the other hand, was not affected by inoculation in Angonia and Ruace in both seasons. In Angonia, harvest biomass ranged from 5,077 kg ha^{-1} for the control to 8,899 kg ha^{-1} for the NitroZam treatment, whereas in Ruace it ranged from 4,533 kg ha^{-1} for the control to 6,343 kg ha^{-1} for the Nitrofix treatment (**Table 7**). However, the harvested biomass of all the bio-inoculant treatments in Nampula was significantly higher than that of the control in both seasons. NitroZam resulted in the highest harvest biomass in 2016, whereas plants inoculated with Nitrofix had the highest harvest biomass in 2017 in Nampula. As observed earlier for 100-seed weight, more harvest biomass was produced in 2017 than in 2016 averaging 7,360 and 5,126 kg ha^{-1}, respectively, across varieties and sites.

TABLE 5 | Varietal differences of mid-season biomass (R3) percent nitrogen content of inoculated soybean.

Location	Angonia		Nampula		Ruace
season	2016	2017	2016	2017	2016
Storm	3.8^{a}	3.4^{a}	4.3^{a}	3.9^{a}	4.5^{a}
TGx 1904-6F	3.0^{b}	3.2^{a}	3.9^{b}	4.0^{a}	4.3^{a}
LSD	0.18	0.2	0.2	0.28	0.29
%CV	10.3	12	9.9	12	12.7

Small letters denote no significance while different letters show significant differences in means.

Inoculation did not have any significant effect on the number of pods produced by individual soybean plants nor did it affect the number of seeds per pod. However, variety differences were observed for seed pod^{-1} and 100-seed weight for most of the environments except for Nampula in 2017 (**Table 8**). The number of seeds pod^{-1} was statistically higher for Storm than that for TGx 1904-6F in all environments. Storm had bigger seeds across sites and seasons with an average 100-seed weight of 15.4 g compared to that of TGx 1904-6F, which averaged 13.1 g (**Table 8**). Our data also showed that variety and bio-inoculant interaction had an inverse relationship between the seeds pod^{-1} and pods plant^{-1} at Ruace in the 2017 growing season (**Figure 4**). As the pods plant^{-1} increased, the seeds pod^{-1} declined linearly with an $r^2 = 0.84$. Yield components namely pods per plant, seeds per pod, and seed weight are important parameters that determine yield. These components have a complex relationship that could either be counteractive or complementary in determining yield within different agroecologies.

Bio-inoculant application increased grain yield in all environments compared to the control, although the yield differences were generally not significant (**Figure 5**). In Angonia, inoculation with N-fixer gave the highest yield increase of 806 kg ha^{-1} for Storm, whereas Masterfix resulted in the greatest yield increase of 768 kg ha^{-1} for TGx 1904-6F in 2016. Similarly, yield increases in Nampula due to inoculation were also higher

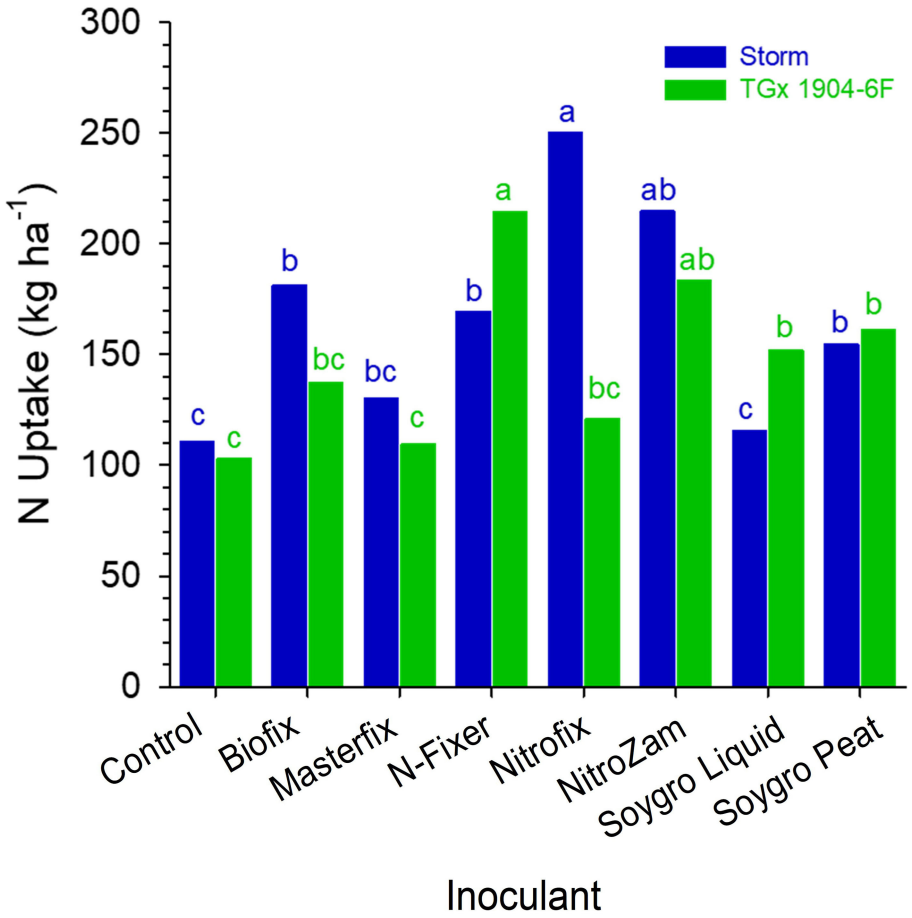


FIGURE 3 | N uptake by soybean at the beginning of podding (R3) stage in Nampula in 2016. Small letters on the bars indicate no significance while different letters show significant differences in means.

in 2016 than in 2017 for TGx 1904-6F. NitroZam had the best performance on both Storm and TGx 1904-6F in Nampula; almost tripling Storm yield from 701 kg ha^{-1} to 2,000 kg ha^{-1} and TGx 1904-6F by 2.3 times from 1,350 kg ha^{-1} to almost 3,100 kg ha^{-1}. At Ruace, the highest yield increase of 70% (1,000–1,700 kg ha^{-1}) occurred in 2016 for Storm when inoculated with either Biofix, Masterfix, Nitrofix, or NitroZam. TGx 1904-6F in combination with NitroZam produced the highest yield for that variety at Ruace in 2017. NitroZam doubled the yield of TGx 1904-6F from 1,350 to 2,900 kg ha^{-1}. Generally, grain yields for TGx 1904-6F were higher than that for Storm when average across bio-inoculant treatments, sites, and seasons (**Table 8**). All the bio-inoculants had statistically higher yields than the control for Storm in Angonia 2016 and Nampula 2016, 2017, and for TGx 1904-6F in Nampula 2016, 2017 and Ruace 2017. Inoculation yield was significantly higher than the control for Storm in Angonia 2016 at 1,550 vs. 966 kg ha^{-1}, Nampula 2016 at 1,583 vs. 701 kg ha^{-1}, and Nampula 2017 at 1,466 vs. 791 kg ha^{-1}. Inoculating TGx 1904-6F, on the other hand, yielded distinctively over the control in Nampula 2017 at 2,341 over 1,449 kg ha^{-1} and Ruace 2017 with 2,500 against 1,366 kg ha^{-1} (**Figure 5**). The interaction between variety and bio-inoculant on yield was diverse within each environment either numerically or statistically different from the control. For instance, in Angonia 2016, N-Fixer and Soygro liquid yields were numerically different from the control but not statistically distinct. In the same site in 2017, the application of bio-inoculants led to a numerically higher yield but not significantly above the control. Inoculation yielded 1,973 kg ha^{-1} against the control with 1,741 kg ha^{-1} for Storm, whereas only NitroZam bio-inoculant (2413 kg ha^{-1}) had a significantly higher yield than the control with 1,913 kg ha^{-1} (**Figure 5**). At Angonia, yield did not generally respond to the bio-inoculant application in both seasons except for NitroZam in 2017. Our results show that the production of soybean gave yield between 701 and 1,741 kg ha^{-1} for Storm and between 1,019 and 1,913 kg ha^{-1} for TGx 1904-6F, while the yield range for inoculating Storm was 1,086 kg ha^{-1} and for TGx 1904-6F was 1,357–2,941 kg ha^{-1} (**Figure 5**). These results show that inoculation leads to an increase in soybean yield.

Contrast analyses confirm the importance of inoculation and the differences in the variety productivity (**Figure 6**). However, the different yield gains due to the application of bio-inoculant

TABLE 6 | 100-Seed weight (g) of inoculants during the 2016 and 2017 growing seasons at three sites Angonia, Nampula, and Ruace.

Inoculant	Angonia		Nampula		Ruace	
	2016	2017	2016	2017	2016	2017
Control	12.9	13.2	12.4	15.5	14.0	15.7
Biofix	13.6	14.2	12.2	15.8	13.7	16.5
Masterfix	14.0	14.3	13.0	15.5	14.1	16.1
N-Fixer	12.5	13.5	11.1	16.3	14.0	16.9
Nitrofix	13.7	14.1	13.0	16.0	14.7	15.8
NitroZam	14.0	14.2	12.1	15.9	14.3	16.4
Soygro liquid	13.1	13.7	11.4	15.5	14.2	15.9
Soygro peat	13.3	13.6	11.5	15.7	13.2	16.2
LSD	0.8	0.8	1.4	1.2	1.5	0.8
%CV	5.8	5.5	11.3	7.4	10.3	4.5

and variety used were more pronounced in Nampula than Angonia and Ruace. At Angonia, bio-inoculant application regardless of grade/carrier led to a yield increase of 425 kg ha^{-1}. When segregated by the carrier, liquid-based bio-inoculant resulted in an average yield increase of 201 kg ha^{-1}, while the peat-based bio-inoculant gave a higher yield increase averaging at 462 kg ha^{-1} (**Figure 6**). At the Nampula site, the yield increase with the application of bio-inoculants was 1,103 kg ha^{-1}, which decreased for a liquid-based carrier to 610 kg ha^{-1}, but was higher for peat-based bio-inoculants (1,185 kg ha^{-1}). Similarly, yield increased by 608 kg ha^{-1} following the use of any bio-inoculant, but the average yield was 331 kg ha^{-1} for liquid-based bio-inoculant, whereas that for the peat-based bio-inoculants was 653 kg ha^{-1}. In essence, although inoculation led to an increase in yield, peat-based bio-inoculants gave a higher yield than liquid-based bio-inoculant. The variety selected also affected yield gains. At all the sites, changing from using TGx 1904-6F instead of Storm increased yield by 117 kg ha^{-1}, 1,019 kg ha^{-1}, and 790 kg ha^{-1} in Angonia, Nampula, and Ruace, respectively. More yield gain occurred in Nampula with the use of bio-inoculants, as well as changing from a non-promiscuous variety Storm to a promiscuous TGx 1904-6F.

DISCUSSIONS

Bio-Inoculants Attributes and Response to Inoculation

Soil analysis indicated differences in nutrient content between the sites. Apart from the bio-inoculant application, other soil nutrients are necessary for the effectiveness of the BNF process and fertilizer use. As such, soils in the Nampula site are likely to respond to P, K, and Mg application, while the effect of N application would be prominent in Angonia. Since P availability in the soil enhances the response of bio-inoculant applications, its use in Nampula would likely lead to improved biological N fixation. Soil characteristics indicate that response to the bio-inoculant application on soybean would be observed more in Angonia and Nampula than Ruace. The response to inoculation would also depend on the concentration and viability of the strain contained in the bio-inoculant. The concentration of viable rhizobia strains and their survival in the carrier is an indicator of the effectiveness of a bio-inoculant. In our study, Masterfix bio-inoculant had the highest CFU g^{-1} (**Table 3**) and was expected to form more nodules, especially in Nampula where the native rhizobial population was low, but this was demonstrated at Ruace in 2016 (**Figure 1**) and in terms of the average nodule dry weight (**Table 4**). Vandeir et al. (2020) reported that Masterfix bio-inoculant with 6.91×10^8 CFU mL^{-1}, which had a higher concentration of *Azospirillum* cells resulted in more agronomic activity leading to more root, shoot biomass, and yield than Nitro 100 with 4.79×10^8 CFU mL^{-1} in rice. These results agree with our finding that Masterfix with 5.0×10^9 CFU g^{-1} led to more nodulation in most of the environments. Perhaps due to the large number of CFU in the bio-inoculant, more cells were able to form associations with the host plant. Inoculation led to the formation of more nodules in Nampula than that at the other two sites (**Figure 1**), which could be due to the low population of indigenous rhizobia in the soil at Nampula. The presence of a large population of native rhizobia in the soil affects the effectiveness of bio-inoculant strains because of the competition for nodule formation and nodule occupancy on roots (Thies et al., 1991). Perhaps there was less competition from the resident strains in Nampula leading to more nodule formation compared to that in Angonia and Ruace, which had a relatively large indigenous population. Nodulation occurred in the non-inoculated plots, which indicates that the indigenous rhizobia also formed nodules but were not as effective as the bio-inoculant strains. Therefore, there is a need for bio-inoculant application to boost soybean production in Mozambique.

Bio-inoculant carrier influences the effectiveness of inoculation response (Vandeir et al., 2020). Peat-based bio-inoculants are the most popular in Mozambique, but there is a liquid-based bio-inoculant formulation produced in South Africa, which could be available to soybean producers in the country. Hence, the inclusion of the liquid-based bio-inoculant to test the efficacy in relation to the peat-based bio-inoculant. Peat has been shown to protect the rhizobia strains better than liquid carriers (Thao et al., 2002). Inoculating soybean with peat-based bio-inoculant resulted in more nodulation than the liquid-based counterparts in Vietnam (Thao et al., 2002). They reported that the peat-based bio-inoculants formed on average 16 nodules $plant^{-1}$ with a dry weight of 351 mg $plant^{-1}$, which was statistically higher than that for the liquid-based bio-inoculants with 12 $plant^{-1}$ and 280 mg dry $plant^{-1}$, respectively. These results agree with our findings where the peat-based bio-inoculants led to more nodulation and hence increased yield over the liquid-based bio-inoculant in Angonia 2016, Ruace 2016, and Ruace 2017. Therefore, it is evident that peat-based bio-inoculants hold a higher potential in increasing the nodulation of soybean in Mozambique than the liquid-based bio-inoculants. This could be attributed to the difference in the handling and management of the bio-inoculants before planting and the soil conditions in which the rhizobia strains are introduced. Peat seems to offer better protection to the bio-inoculant strains both during transportation, as well as when in the soil before attaching

TABLE 7 | Harvest biomass (kg ha^{-1}) inoculants during the 2016 and 2017 growing seasons at Angonia.

Inoculant	Angonia		Nampula		Ruace	
	2016	2017	2016	2017	2016	2017
Control	5,077	8,480	2,826	6,190	4,533	5,614
Biofix	5,247	8,554	4,172	7,896	5,618	6,071
Masterfix	6,198	8,573	4,016	7,386	5,619	5,881
N-Fixer	5,137	8,446	4,713	8,052	5,433	5,731
Nitrofix	5,537	8,616	4,573	7,360	6,343	6,126
NitroZam	5,831	8,899	6,449	7,508	5,489	6,306
Soygro liquid	5,104	8,850	3,685	7,628	5,891	5,833
Soygro peat	5,988	9,254	3,722	7,327	5,827	6,064
LSD	1,369.6	1,456.2	342.7	1,113.5	1,543.9	912.8
%CV	24.4	16.1	7.7	14	25.8	15.2

TABLE 8 | Yield components varietal differences of inoculated soybean.

Season	Angonia		Nampula		Ruace	
	2016	2017	2016	2017	2016	2017
Seeds Pod^{-1}						
Storm	2.03^{b}	2.20^{a}	2.27^{a}	2.25^{a}	2.27^{a}	2.47^{a}
TGx 1904-6F	2.29^{a}	2.04^{b}	2.16^{b}	2.17^{a}	2.08^{b}	2.35^{b}
LSD	0.14	0.13	0.08	0.11	0.1	0.07
%CV	12.4	11.3	6.7	9.4	9.1	5.5
100 Seed weight (g)						
Storm	14.5^{a}	15.1^{a}	13.0^{a}	16.3^{a}	15.5^{a}	17.9^{a}
TGx 1904-6F	12.3^{b}	12.6^{b}	11.2^{b}	15.2^{b}	12.6^{b}	14.5^{b}
LSD	0.39	0.38	0.72	0.6	0.74	0.38
%CV	5.8	5.5	11.3	7.4	10.3	4.5
Yield (kg ha^{-1})						
Storm	1,493^{a}	1,944^{b}	1,498^{b}	1,400^{b}	1,439^{b}	1,685^{b}
TGx 1904-6F	1,589^{a}	2,139^{a}	2,110^{a}	2,216^{a}	1,981^{a}	2,358^{a}
LSD	129	198	93	100	173	169
%CV	16.4	18.8	9.9	10.7	19.7	16.6

Small letters denote no significance while different letters show significant differences in means.

to the host plant (Vandeir et al., 2020). However, contrasting results were reported in Gauteng and Northwest province of South Africa where more nodules were formed on soybean with liquid than peat bio-inoculant (Gatabazi et al., 2021). Since rhizobia strains contained in the bio-inoculants are live cells, its packaging defines handling and transportation logistics for a country like Mozambique, where all the soybean bio-inoculants are imported from as far as the South American continent. Peat bio-inoculants have a better buffer capacity for the strains against environmental factors like temperature during transportation, storage, or soon after application. Peat-based bio-inoculants have also been shown to outperform liquid bio-inoculants in rice (Vandeir et al., 2020). In this study, peat-based bio-inoculant Nitro 1000 led to better plant development than its liquid counterparts. The low capacity of African agricultural systems to store bio-inoculants in good conditions for a long time has been reported (Woomer et al., 2014). Due to the lack of storage facilities among the smallholder farmers in Mozambique, the likelihood of liquid-based bio-inoculant quality deteriorating before the application is high because of the low buffering capacity offered to the rhizobia strains.

Nodulation and Biological Nitrogen Fixation

Plant tissue N is translocated to the reproductive parts during growth and development. However, some of the N remains in the nonreproductive plant parts and could be recycled back into the soil to improve soil fertility. Soybean acquires its N both from the soil directly and through the BNF process that takes place effectively when an appropriate rhizobia strain is present to form symbiotic associations with the host plant. The soils in Nampula had a very low population of soybean rhizobia strain (**Table 2**), because there was no competition from the resident rhizobia leading to more attachment by the introduced cells to

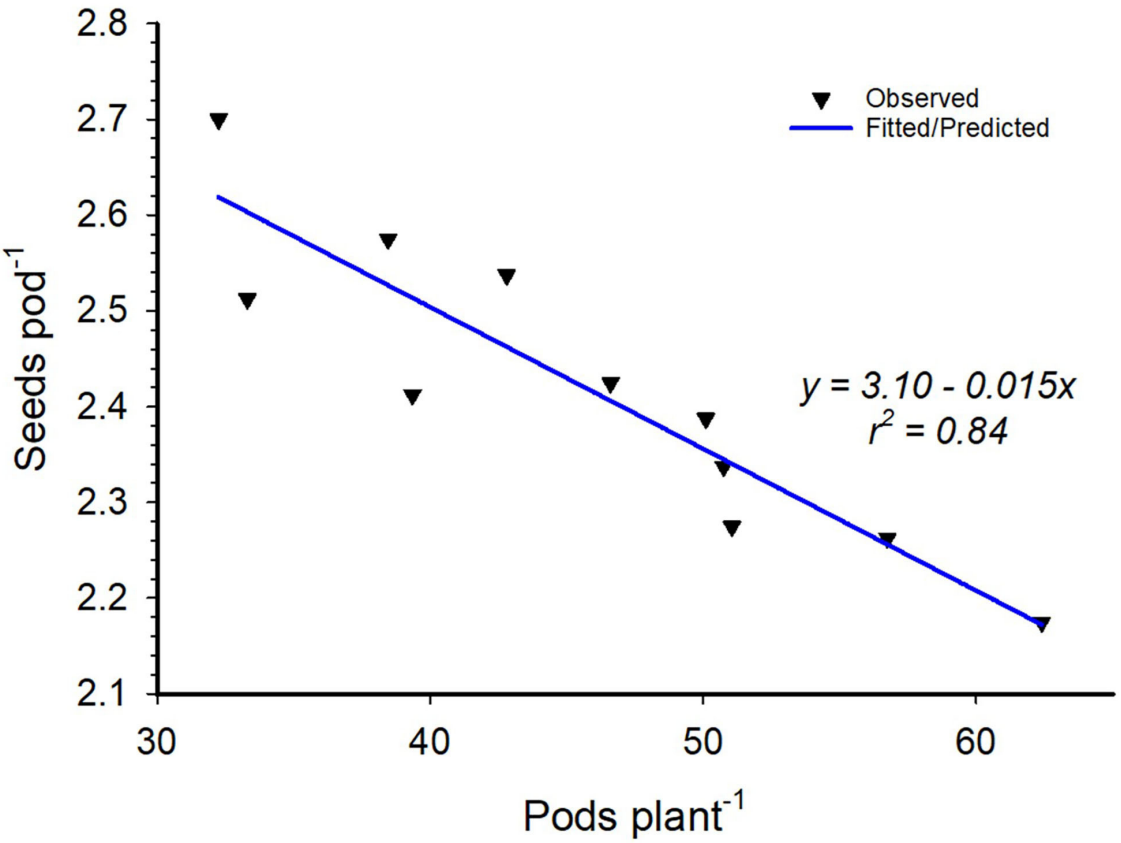


FIGURE 4 | Variety by inoculant interactions of pods per plant and seeds per pod at Ruace in the 2017 growing season.

the soybean roots. Legumes growing in soils with low populations of indigenous rhizobia increase nodulation when the right strain is applied (Sanginga et al., 2002; Majengo et al., 2011). Our finding corroborates with a study conducted in Eastern Ethiopia where nodulation increased in common beans planted in soils with different populations of indigenous rhizobia (Argaw and Angaw, 2015). In the same study, there were more nodules formed with an increase in the native rhizobia from 82.3 for soils with <100 common bean rhizobia g^{-1} soil to 186.7 for soils with more than 1,000 common bean rhizobia g^{-1} soil. As the indigenous rhizobia strains increased, the number of nodules formed on the common bean roots was also higher. Although promiscuous soybean varieties can form nodules freely with indigenous strains, the application of bio-inoculant improved nodulation (Sanginga et al., 2002; Osunde et al., 2003; Gitonga et al., 2011). Our results also showed that nodulation increased with the inoculation of the promiscuous variety TGx 1904-6F more than that for the non-promiscuous Storm.

Although our study did not determine nodule occupancy and activity of the nodules, plant N uptake at the R3 stage was measured. The amount of N accumulated in the plant tissue and yield obtained to serve as pseudo indicators of the BNF process following inoculation. It is presumed that improved nodulation leads to an increased BNF process that avails more N for plant growth and development. The soybean varieties had accumulated different amounts of N in their tissues at the R3 stage. Our results show that inoculated non-promiscuous variety Storm accumulated more N than promiscuous TGx 1904-6F at the R3 stage. Storm is an early maturing soybean variety, hence accumulated N at a faster rate during growth and development than the TGx 1904-6F. Inoculated soybean has been shown to accumulate more than 42% of the total N by the R3/R4 (beginning of podding to beginning of seed) than the inoculated counterparts (Sanginga et al., 1997). Early maturing promiscuous soybean varieties attain maximum N accumulation earlier than medium to late maturity groups. Sanginga et al. (1997) reported that IAC 100 reached maximum N accumulation at R1/R2 stages (beginning of flowering to full boom), while the late maturity promiscuous soybean variety TGx 1519-1D peak was at R3/R4 stage. Our findings where Storm had more tissue N content confirm this study's report (**Table 5**). Based on the biomass amount at the R3 stage, the total N accumulated in the plants per hectare was determined (**Figure 3**). Nitrogen accumulated in the plant biomass at the R3 stage increased with inoculation but was distinct with bio-inoculant interaction with the soybean variety. Other studies have reported an increase

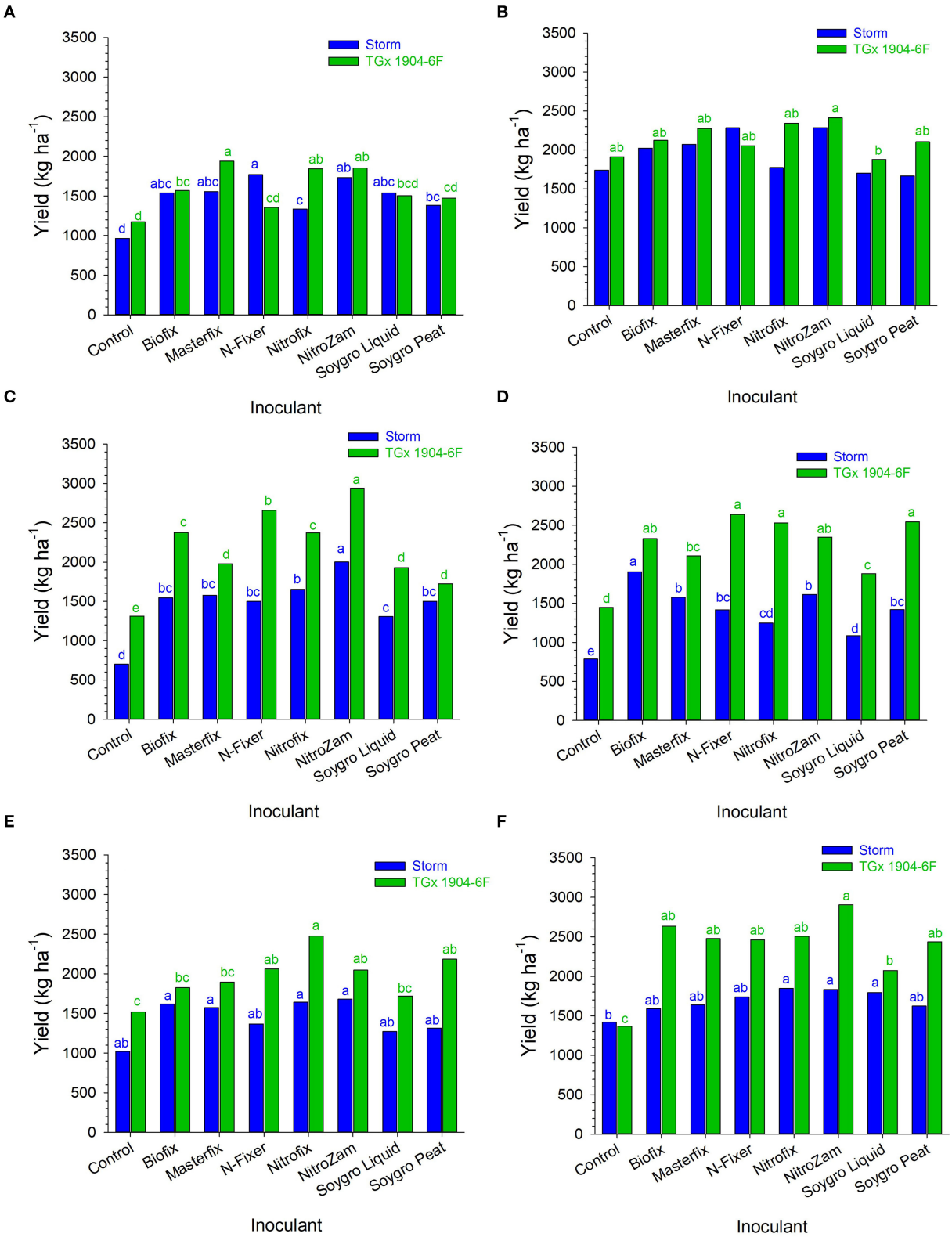


FIGURE 5 | Inoculated soybean yield in **(A)** Angonia 2016, **(B)** Angonia 2017, **(C)** Nampula 2016, **(D)** Nampula 2017, **(E)** Ruace 2016, and **(F)** Ruace 2017 growing seasons. Small letters on the bars indicate no significance while different letters show significant differences in means.

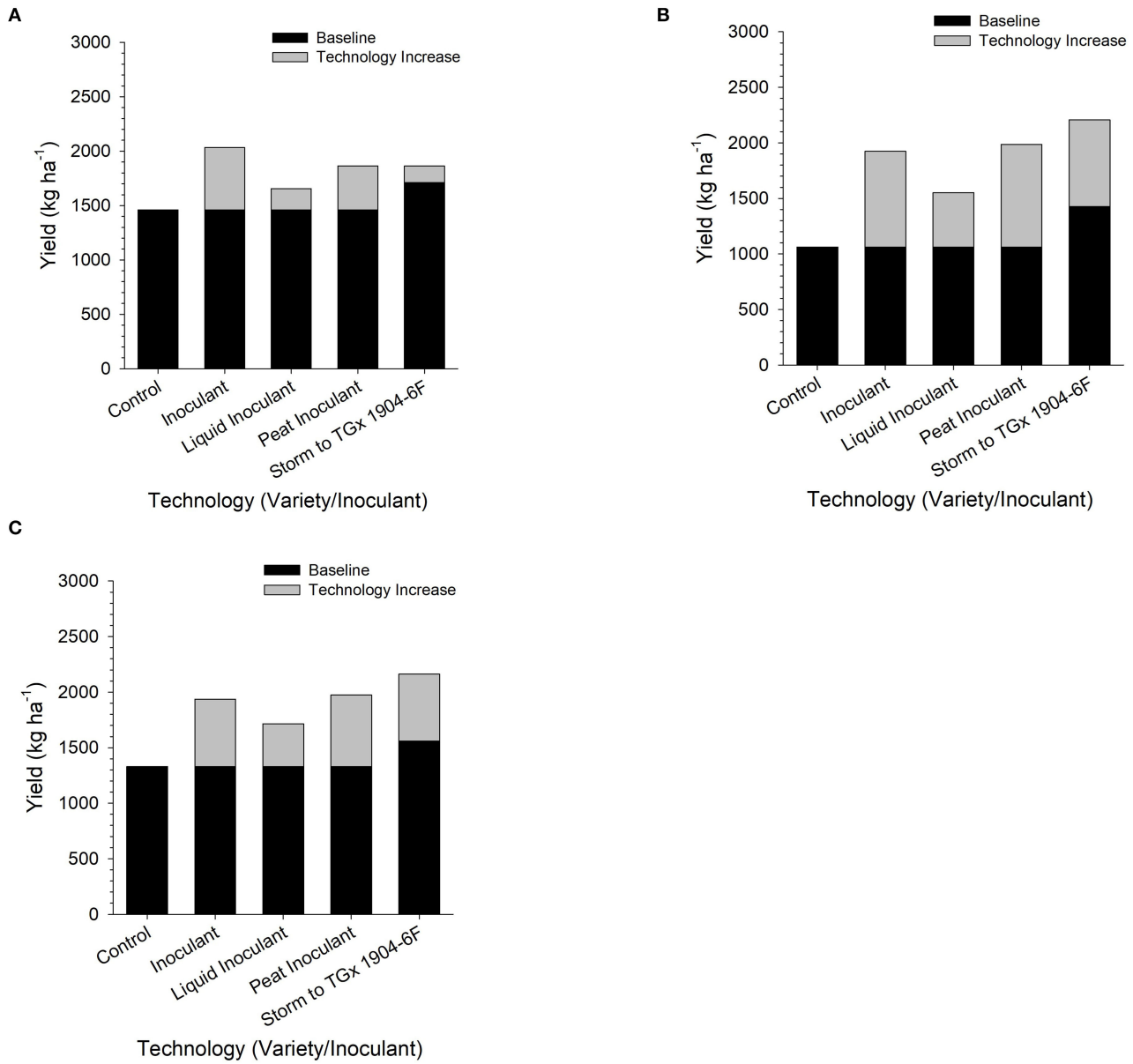


FIGURE 6 | Yield gain upon application of inoculants in Angonia **(A)**, Nampula **(B)**, and Ruace **(C)** across two growing seasons in 2016 and 2017.

in N amount accumulated in plant tissue due to inoculation (Gitonga et al., 2011; Savala et al., 2021, 2022; Omari et al., 2022). Storm accumulated more N per ha for the Biofix, Masterfix, Nitrofix, and NitroZam, while TGx 1904-6F had more N when inoculated with N-Fixer and Soygro. These findings contrasted with results from another study in the same locations of Angonia and Ruace where early maturity promiscuous soybean variety TGx 1740-2F accumulated more N by the R3 stage than Safari a non-promiscuous variety (Savala et al., 2021). This signifies that the amount of N tissue at the R3 stage could be variety-related rather than the differences in promiscuity. Not all the amount of biomass N accumulated is translocated to the reproductive parts. Part of the N remains in the biomass and could be recycled back into the soil and added to the residual N for use by the next crop. Inoculation leads to more biomass N accumulation that is important for improving soil N (Omari et al., 2022). Our results confirmed that inoculation leads to more biomass production in all the sites (**Table 7**). Therefore, we infer that inoculating soybean not only increases the yield but also leads to more biomass production that is important in improving the soil nitrogen through recycling. As such, recycling of inoculated soybean harvest biomass contributes to improving the overall soil health thereby reducing the need for external N sources.

Economic Importance of Bio-Inoculants for Soybean Production

The interaction between the environment, on one hand, and bio-inoculant or variety affected yield components, such as 100-seed weight, seed pod^{-1}, and pods plant^{-1}, differently. This interaction is significant in defining the suitability of bio-inoculants and soybean variety adaptation in specific locations. Although not all applied bio-inoculants resulted in an increase in 100-seed weight, Biofix, Masterfix, Nitrofix, and NitroZam gave higher values than the control in Angonia. The a pplication of bio-inoculants in soybean has been reported to result in increased seed quality and weight that positively increases yield (Gitonga et al., 2011; Tavanti et al., 2020). On the other hand, environment and variety interaction affected 100-seed weight, seeds pod^{-1}, and pods plant^{-1} (**Table 8**). Storm had a statistically higher 100-seed weight in all environments than TGx 1904-6F. The 100-seed weight was not affected by the bio-inoculant differences as it is a characteristic controlled more by the genotype's genetic capability. Although the seed weight slightly increased following inoculation, the difference was not statistically significant. Increased numerical values of 100-seed weight were also demonstrated in a study that assessed the productivity of two soybean cultivars upon application of N mineral fertilizers and seed inoculation with *Bradyrhizobium diazoefficiens* (Ksiezak and Jolanta, 2022). Other yield components that are affected with inoculation are seeds pod^{-1} and pods plant^{-1}. In our study, we confirmed that seeds pod^{-1} did not increase with the application of bio-inoculants but rather varied with variety. Storm gave statistically more seeds per pod in four environments, while TGx 1904-6F had many seeds pod^{-1} in Angonia 2016 only. Like 100-seed weight, seeds pod^{-1} plant trait is affected more by the genetics of the plant in soybean. However, past studies have reported a greater number of seeds pod^{-1} following inoculation. Lamptey et al. (2014) and Ksiezak and Jolanta (2022) demonstrated that soybean formed more seeds pod^{-1} due to inoculation. In these studies, bio-inoculants were combined with other nutrients, such as iron and nitrogen, making it complex to attribute the increase in seeds pod^{-1} to rhizobia strain.

Improvement in soybean yield due to inoculation has been widely documented. Our study did not fall short of ascertaining the same findings but also showed that variety selection affected soybean yield differently in the specific sites. It was more advantageous growing TGx 1904-6F in Nampula than Storm, whereas both varieties yielded comparable in Angonia. Gitonga et al. (2011) also reported a higher yield between 27.8 and 39.3% of Biofix inoculated TGx 1869 and TGx 1893 promiscuous soybean over the non-promiscuous Gazelle variety at Kaguru, Meru District in Kenya. Similar results have also been documented at Ruace in Mozambique where TGx 1740-2F out-yielded Safari (Savala et al., 2022). In the same study, inoculated Safari gave more yields than TGx 1740-2F at Angonia. Though our yield analysis indicates that promiscuous soybean TGx 1904-6F performed better than Storm, it is profound that there exists a difference in variety suitability for various specific agroecological sites. Therefore, variety selection should consider existing climatical and soil conditions for each site. For instance, Angonia and Ruace have suitable soil conditions and weather for soybean production, hence, both promiscuous and non-promiscuous varieties can be produced. However, the selection of the variety will depend on the weather forecast especially expected season rainfall distribution and amount. Yield also increased with inoculation, but the productivity varied with the site. Bio-inoculant effect on yield was distinct with carrier too. Peat-based bio-inoculants had higher yields than their liquid counterparts. As earlier reported, in the nodulation section, environmental factors could have affected the efficacy of the liquid-based bio-inoculants hence leading to low symbiotic associations for the BNF process. With reduced BNF process in liquid-based bio-inoculant across the sites, the low nitrogen necessary for seed development could have been fixed. This study confirms findings that peat bio-inoculant yielded more than liquid-based bio-inoculants. A study conducted in Vietnam also reported that peat-based bio-inoculants yielded 2.00 and 2.18 t ha^{-1} against the liquid-based counterparts 1.96 and 2.12 t ha^{-1} for two rhizobia strains designated as G5 and G6, respectively (Thao et al., 2001). Contrasting studies have also shown that liquid-based bio-inoculant yields more than peat-based bio-inoculants (Hungria et al., 2020; Gatabazi et al., 2021). This justifies that both liquid and peat-based bio-inoculants lead to increased yield in soybean although their effectiveness would depend on handling and the prevailing agroecological conditions. Although bio-inoculants lead to improved productivity in soybean, the key to their effectiveness could lie in the handling and prevailing environmental conditions. The bottom line is that inoculation is necessary for increasing soybean yield. The availability of bio-inoculants holds the potential for better soybean yields for smallholder farmers in Mozambique. This study showed that the use of bio-inoculants increased yield between 201 and 462 kg ha^{-1} (**Figure 6**). Considering the cost of bio-inoculants in Mozambique at USD 10 per hectare and the grain selling price of USD 0.90 kg^{-1}, the use of bio-inoculants would result in a monetary profit of USD 170.9–405.8 per hectare.

CONCLUSION

Application of bio-inoculants has the potential of increasing soybean production in Mozambique regardless of the strain carrier material. Both liquid- and peat-based bio-inoculants led to increased nodulation and yield of soybean over the control. Inoculation improved nodulation and yield at different proportions in soybean production regardless of the strain carrier material depending on the agroecological conditions, such as the population of indigenous rhizobia strain. There was more increase in nodulation in the Nampula site that had the lowest number of indigenous rhizobia strains. However, more nodules and yield were observed with the application of peat-based bio-inoculants than the liquid-based counterparts. The economic gain from yield improvement justifies the use of bio-inoculants for soybean production in Mozambique. This can be further increased if the country produces bio-inoculants locally. Locally producing bio-inoculants has the potential of reducing the cost

of the input and will make it accessible and affordable for smallholder farmers in Mozambique. As noted in the study, inoculation improved plant development and production of more biomass with high N content. The magnitude of plant development and biomass production due to the application of bio-inoculants depends on the prevailing agroecological conditions. There was more biomass production in Angonia and Ruace than Nampula site. Angonia and Ruace agroecologies receive relatively more rains and are more suitable for soybean production than Nampula. This biomass could be recycled back to the soil to improve the fertility status, especially N content that would benefit the subsequent crop. In addition to bio-inoculant use, more study is required by the breeders to produce better phosphorus scavenging promiscuous soybean that would effectively form symbiotic associations with indigenous rhizobia strains. Moreover, additional study is required to ascertain nodule occupancy and activity when bio-inoculants are applied in soils with native rhizobia, especially for the promiscuous soybean varieties.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CS, SK-B, and DC designed and supervised the experiment. CS participated in field evaluation, data compilation, and oversaw writing up. CS and AW conducted data analysis. All authors contributed to the preparation of the manuscript.

FUNDING

Funding was provided by the USAID through Grant #EEM-G-00-04-000136-00 Platform for Agricultural Research and Technology Innovation and Feed the Future Mozambique Improved Seeds for Better Agriculture (SEMEAR) Grant #AID-BFS-IO-17-00005 projects in Mozambique.

ACKNOWLEDGMENTS

We thank the IITA technical staff Rachad Ussene at Nampula, Irondino Saraiva at Angonia, Carlos Muananamuale, Francisco Domingos, and Mariano Aniceito at Ruace station in Mozambique for managing the trials and assisting in collecting the field related data.

REFERENCES

AGRA (2016). *Mozambique Early Generation Seed Study*. Available online at: https://agrilinks.org/sites/default/files/resource/files/mozambique_early_generation_seed_report.pdf (accessed March 16, 2022).

Argaw, A., and Angaw, T. (2015). Indigenous rhizobia population influences the effectiveness of rhizobium inoculation and need of inorganic N for common bean (*Phaseolus Vulgaris* L.) production in Eastern Ethiopia. *Chemical and Biological Technologies in Agriculture* 2, 1–13. doi: 10.1186/s40538-015-0047-z

Chibeba, A. M., Kyei-Boahen, S., Maria, F. G., Marco, A. N., and Mariangela, H. (2017). Isolation, characterization and selection of indigenous bradyrhizobium strains with outstanding symbiotic performance to increase soybean yields in mozambique. *Agricult. Ecosyst. Environ.* 246, 291–305. doi: 10.1016/j.agee.2017.06.017

Chibeba, A. M., Kyei-Boahen, S., Maria, F. G., Marco, A. N., and Mariangela, H. (2020). Towards sustainable yield improvement: field inoculation of soybean with bradyrhizobium and co-inoculation with azospirillum in mozambique. *Arch. Microbiol.* 202, 2579–90. doi: 10.1007/s00203-020-01976-y

Delamuta, J. R. M., Renan, A. R., Ernesto, O., Itamar, S. M., Esperanza, M., Mariangela, H., et al. (2013). Polyphasic evidence supporting the reclassification of bradyrhizobium japonicum group ia strains as *Bradyrhizobium diazoefficiens* Sp. *Nov. Int. J. Syst. Evol. Microbiol.* 63, 3342–51. doi: 10.1099/ijs.0.049130-0

Gatabazi, A., Barend, J. V., Mireille, A. M., Edgar, M., Robert, M., Ahmed, I. H., et al. (2021). Efficacy of peat and liquid inoculant formulations of bradyrhizobium japonicum strain wb74 on growth, yield and nitrogen concentration of soybean (*Glycine max* L.). *Nitrogen*. 2, 332–46. doi: 10.3390/nitrogen2030023

Gitonga, M. N., Gatheri, G. W., Cheruiyot, R., Gitonga, N. M., and Maingi, J. M. (2011). Nodulation and nitrogen fixation in promiscuous and non-promiscuous soybean (*Glycine max* (L.) Merrill) Varieties in Eastern Kenya. *J. Trop. Microbiol. Biotechnology* 6. doi: 10.4314/jtmb.v6i1.63744

Gyogluu, C., Kyei-Boahen, S., and Dakora, F. D. (2016). Response of Promiscuous-Nodulating Soybean (*Glycine max* L. *Merr.)* genotypes to bradyrhizobium inoculation at three field sites in mozambique. *Symbiosis* 69, 81–88. doi: 10.1007/s13199-015-0376-5

Hungria, M., Marco, A. N., Campos, L. J. M., Menna, P., Brandi, F., Ramos, Y. G., et al. (2020). Seed pre-inoculation with bradyrhizobium as time-optimizing option for large-scale soybean cropping systems. *Agron. J.* 112, 5222–36. doi: 10.1002/agj2.20392

Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., et al. (eds). (2013). *Soil Atlas of Africa*. Luxembourg: European Commission. ISBN 978-92-79-26715-4.

Ksiezak, J., and Jolanta, B. (2022). The effect of mineral n fertilization and bradyrhizobium japonicum seed inoculation on productivity of soybean (*Glycine max* (L.) Merrill). *Agriculture (Switzerland)*. 12, agriculture12010110. doi: 10.3390/agriculture12010110

Kyei-Boahen, S., Savala, C. E. N., Chikoye, D., and Abaidoo, R. (2017). Growth and yield responses of cowpea to inoculation and phosphorus fertilization in different. *Environments*. 8, 1–13. doi: 10.3389/fpls.2017.00646

Lamptey, S., Ahiabor, B., Yeboah, S., and Osei, D. (2014). Effect of rhizobium inoculants and reproductive growth stages on shoot biomass and yield of soybean (*Glycine max* (L.) Merril). *J. Agricult. Sci.* 6, v6n5p44. doi: 10.5539/jas.v6n5p44

Majengo, C. O., John Okalebo, R., Lesueur, D., Pypers, P., Ng'etich, W., Musyoki, M., et al. (2011). Interaction between nitrogen and phosphorus microbial inoculants on soybean production in Bungoma, Kenya. *Afr. Crop Sci. Conf. Proc.* 10, 117–19.

Matteo, F., Otsuki, K., and Schoneveld, G. (2016). "Soya bean expansion in mozambique: exploring the inclusiveness and viability of soya business models as an alternative to the land grab," in *The Public Sphere*. 60–86.

Moore, K. J., and Dixon, P. M. (2015). Analysis of Combined Experiments Revisited. *Agron. J.* 107, 763–71. doi: 10.2134/agronj13.0485

Omari, R. A., Yuan, K., Anh, K. T., Reckling, M., Halwani, M., Egamberdieva, D., et al. (2022). Enhanced soybean productivity by inoculation with indigenous bradyrhizobium strains in agroecological conditions of northeast germany. *Front. Plant Sci.* 12, 3276. doi: 10.3389/fpls.2021.707080

Osunde, A., Gwam, S., Bala, A., Sanginga, N., and Okogun, J. A. (2003). Responses to Rhizobial Inoculation by Two Promiscuous Soybean Cultivars in Soils of the Southern Guinea Savanna Zone of Nigeria. *Biolo. Fert. Soils*. 37, 274–79. doi: 10.1007/s00374-003-0609-2

Sanginga, N., Dashiell, K., Okogun, J. A., and Thottappilly, G. (1997). Nitrogen Fixation and N Contribution by Promiscuous Nodulating Soybeans in the Southern Guinea Savanna of Nigeria. *Plant and Soil* 195, 257–66. doi: 10.1023/A:1004207530131

Sanginga, N., Okogun, J. A., Vanlauwe, B., and Dashiell, K. (2002). The Contribution of Nitrogen by Promiscuous Soybeans to Maize Based Cropping the Moist Savanna of Nigeria. *Plant and Soil* 241, 223–31. doi: 10.1023/A:1016192514568

SAS/STAT® (2018). *15, 1*. User's Guide the GLM Procedure. http://support.sas.com/thirdpartylicenses.

Savala, C. E. N., Chikoye, D., and Kyei-Boahen, S. (2022). Inoculant Formulation and Application Determine Nitrogen Availability and Water Use Efficiency in Soybean Production. *In Soybean - Recent Advances in Research and Applications.* 102639. doi: 10.5772/intechopen.102639

Savala, C. E. N., and Kyei-Boahen, S. (2020). Potential of Inoculant and Phosphorus Application on Soybean Production in Mozambique. *Universal Journal of Agricultural Research 8, 46–57.* 2020.080204. doi: 10.13189/ujar.2020.080204

Savala, C. E. N., Wiredu, A. N., Okoth, J. O., and Kyei-Boahen, S. (2021). Inoculant, Nitrogen and Phosphorus Improves Photosynthesis and Water-Use Efficiency in Soybean Production. *Journal of Agricultural Science* 159, 349–62. doi: 10.1017/S0021859621000617

Singleton, P. W., Abdelmagid, H. M., and Tavares, J. W. (1985). Effect of Phosphorus on the Effectiveness of Strains of Rhizobium Japonicum. Soil Science Society of America Journal *49, 613–16.* 036159950049000300l6x. doi: 10.2136/sssaj1985.03615995004900030016x

Singleton, P. W., Keyser, H., and Sande, E. (2001). Development and Evaluation of Liquid Inoculants. Inoculants and Nitrogen Fixation of Legumes in Vietnam Edited by D. Herridge ACIAR Proceedings, 109e. https://www.researchgate.net/publication/242193371.

Tavanti, T. R., and Renan, F. R., Tavanti, F. S., Galindo, I. S., Larissa, S. D., and Sá, M. E. (2020). Yield and Quality of Soybean Seeds Inoculated with Bacillus Subtilis Strains. Revista Brasileira de Engenharia Agricola e Ambiental 24, 65–71. v24n1p65-71. doi: 10.1590/1807-1929/agriambi.v24n1p65-71

Thao, T. Y., Singleton, P. W., and Herridge, D. (2001). Inoculation Responses of Soybean and Liquid Inoculants as an Alternative to Peat-Based Inoculants. https://www.researchgate.~net/publication/240630719.

Thao, T. Y., Singleton, P. W., and Herridge, D. (2002). Inoculants and Nitrogen Fixation of Legumes in Vietnam Edited by D. Herridge ACIAR Proceedings. *Vol*. 109.

Thies, J. E., Singleton, P. W., and Bohlool, B. (1991). Influence of the Size of Indigenous Rhizobial Populations on Establishment and Symbiotic Performance of Introduced Rhizobia on Field-Grown Legumest. Applied and Environmental Microbiology. Vol. 57. https://journals.asm.org/journal/aem. doi: 10.1128/aem.57.1.19-28.1991

Tittabutr, P., Waraporn, P., Neung, T., Singleton, P. W., and Nantakorn, B. (2007). Growth, Survivaland Field Performance of Bradyrhizobial Liquid Inoculant Formulations with Polymeric Additives. ScienceAsia 33, 69–77. 33, 069. doi: 10.2306/scienceasia1513-1874.2007.33.069

van Heerwaarden, J., Baijukya, F., Kyei-Boahen, S., Adjei-Nsiah, S., Ebanyat, P., Kamai, N., et al. (2018). Soyabean response to rhizobium inoculation across sub-saharan africa: patterns of variation and the role of promiscuity. *Agricult. Ecosyst. Environ.* 261, 211–18. doi: 10.1016/j.agee.2017.08.016

Vandeir, F. G., Klein, J., Barbosa, F. M., and Kestring, K. D. (2020). Promotion of rice growth and productivity as a result of seed inoculation with *Azospirillum brasilense*. *Afr. J. Agricult. Res.* 16, 765–76. doi: 10.5897/AJAR2020.14723

Vincent, J. M. (1970). *A Manual for the Practical Study of the Root-Nodule Bacteria* (Oxford and Edinburgh: Blackwell Scientific Publ.),164.

Walker, T., Pitoro, R., Tomo, A., Sitoe, I., Salência, C., Mahanzule, R., et al. (2006). *Institute of Agricultural Research of Mozambique Directorate of Training, Documentation, and Technology Transfer Research Report Series Republic of Mozambique Priority Setting for Public-Sector Agricultural Research in Mozambique with the National Agricultural Survey Data*.

Woomer, P. L., Bennett, J., and Yost, R. (1990). Overcoming the inflexibility of most-probable-number procedures. *Agron. J.* 82, 349–53. doi: 10.2134/agronj1990.00021962008200020035x

Woomer, P. L., Mahamadi, D., Wafullah, T. N., Balume, I., Uwizerwa, M., Kyei-Boahen, S., et al. (2014). *N2Africa Putting Nitrogen Fixation to Work for Smallholder Farmers in Africa Improving Legume Inoculants and Developing Strategic Alliances for Their Advancement*. Available online at: http://www.N2Africa.org (accessed March 15, 2022).

Zavale, H., Matchaya, G., Vilissa, D., Nhemachena, C., Nhlengethwa, S., Duque, W., et al. (2020). Dynamics of the fertilizer value chain in Mozambique. *Sustainability (Switzerland)*. 12, su12114691. doi: 10.3390/su12114691

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Savala, Wiredu, Chikoye and Kyei-Boahen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 July 2022
doi: 10.3389/fpls.2022.934296

Grain Nutrients Variability in Pigeonpea Genebank Collection and Its Potential for Promoting Nutritional Security in Dryland Ecologies

*Dhanapal Susmitha[1,2], Thiyagarajan Kalaimagal[2], Ramachandran Senthil[1], Mani Vetriventhan[1], Swaminathan Manonmani[2], Prabhakaran Jeyakumar[3], Bellie Anita[4], Surender Reddymalla[1], Pushpajeet L. Choudhari[5], Chetna A. Nimje[5], Ovais H. Peerzada[1], Venkata Narayana Arveti[1], Vania C. R. Azevedo[6] and Kuldeep Singh[1]**

[1] Genebank, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, [2] Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, [3] Office of the Registrar, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, [4] Directorate of Open Distance Learning, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, [5] Charles Renard Analytical Laboratory, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, [6] International Potato Center (CIP), Lima, Peru

OPEN ACCESS

Edited by:
Jose C. Jimenez-Lopez,
Tenured Scientist at CSIC, Spain

Reviewed by:
Guanglong Hu,
Beijing Academy of Agricultural and Forestry Sciences, China
Showkat Ganie,
Royal Holloway, University of London, United Kingdom

***Correspondence:**
Kuldeep Singh
kuldeep.singh@cgiar.org

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 02 May 2022
Accepted: 02 June 2022
Published: 11 July 2022

Citation:
Susmitha D, Kalaimagal T, Senthil R, Vetriventhan M, Manonmani S, Jeyakumar P, Anita B, Reddymalla S, Choudhari PL, Nimje CA, Peerzada OH, Arveti VN, Azevedo VCR and Singh K (2022) Grain Nutrients Variability in Pigeonpea Genebank Collection and Its Potential for Promoting Nutritional Security in Dryland Ecologies. Front. Plant Sci. 13:934296. doi: 10.3389/fpls.2022.934296

Pigeonpea, a climate-resilient legume, is nutritionally rich and of great value in Asia, Africa, and Caribbean regions to alleviate malnutrition. Assessing the grain nutrient variability in genebank collections can identify potential sources for biofortification. This study aimed to assess the genetic variability for grain nutrients in a set of 600 pigeonpea germplasms conserved at the RS Paroda Genebank, ICRISAT, India. The field trials conducted during the 2019 and 2020 rainy seasons in augmented design with four checks revealed significant differences among genotypes for all the agronomic traits and grain nutrients studied. The germplasm had a wider variation for agronomic traits like days to 50% flowering (67–166 days), days to maturity (112–213 days), 100-seed weight (1.69–22.17 g), and grain yield per plant (16.54–57.93 g). A good variability was observed for grain nutrients, namely, protein (23.35–29.50%), P (0.36–0.50%), K (1.43–1.63%), Ca (1,042.36–2,099.76 mg/kg), Mg (1,311.01–1,865.65 mg/kg), Fe (29.23–40.98 mg/kg), Zn (24.14–35.68 mg/kg), Mn (8.56–14.01 mg/kg), and Cu (7.72–14.20 mg/kg). The germplasm from the Asian region varied widely for grain nutrients, and the ones from African region had high nutrient density. The significant genotype × environment interaction for most of the grain nutrients (except for P, K, and Ca) indicated the sensitivity of nutrient accumulation to the environment. Days to 50% flowering and days to maturity had significant negative correlation with most of the grain nutrients, while grain yield per plant had significant positive correlation with protein and magnesium, which can benefit simultaneous improvement of agronomic traits with grain nutrients. Clustering of germplasms based on Ward.D2 clustering algorithm revealed the co-clustering of germplasm from different regions. The identified top 10 nutrient-specific and 15

multi-nutrient dense landraces can serve as promising sources for the development of biofortified lines in a superior agronomic background with a broad genetic base to fit the drylands. Furthermore, the large phenotypic data generated in this study can serve as a raw material for conducting SNP/haplotype-based GWAS to identify genetic variants that can accelerate genetic gains in grain nutrient improvement.

Keywords: pigeonpea, protein, minerals, calcium, biofortification, landraces

INTRODUCTION

Malnutrition exists in most countries and across all socioeconomic classes. Undernutrition, micronutrient deficiency, and obesity are the implications of a nutritiously imbalanced diet (Food and Agriculture Organization [FAO], 2014). A healthy diet should meet the recommended dietary allowance of 54 g (men) and 46 g (women) protein, 1,000-mg phosphorus, 2,000-mg potassium, 1,000-mg calcium, 2-mg copper, 440 mg (men), and 370-mg (women) magnesium, 4-mg manganese, 19-mg (men) and 29-mg (women) iron, and 17-mg (men) and 13-mg (women) zinc per day (ICMR-NIN, 2020). Severe protein deficiency characterized by Kwashiorkor is widespread in developing countries. Similarly, micronutrient deficiencies of common occurrence are iron, vitamin A, and iodine (World Health Organization [WHO], 2021). Anemia outbreaks as a result of iron deficiency, and, globally, 1.8 billion people were anemic as of 2019, with South Asia, West Sub-Saharan Africa, and Central Sub-Saharan Africa regions having high prevalence (Safiri et al., 2021). Furthermore, poverty and malnutrition are interrelated to each other. Poverty compromises the dietary quality of food and results in the intake of inexpensive starchy food (Siddiqui et al., 2020). Imbalanced energy and protein intake result in protein-energy malnutrition (Bailey and West, 2015). The dietary protein intake can be of plant or animal origin. Furthermore, the source of protein origin has an impact on human health. Substituting foods rich in animal protein with plant protein can prolong longevity (Song et al., 2017; Naghshi et al., 2020). Animal protein production disturbs environmental sustainability (Aiking and De Boer, 2020), and its consumption also adds to the spread of zoonotic diseases (Andreoli et al., 2021). Comparatively, plant protein exerts less pressure on the environment. The only limitation associated with plant protein is the poor protein quality that is affected by the anti-nutritional factors contained in it, which in turn reduces the bio-availability of minerals (Pele et al., 2016; Ahnen et al., 2019).

Grain legumes were identified as the cheapest source of good quality protein (Dahiya et al., 1977). The nutritional profile states that legumes have two times the quantity of cereal protein, with no cholesterol and less fat (other than soybean and groundnut), and serves as a rich source of essential minerals, namely, Zn, Fe, Ca, Se, P, Cu, K, Mg, and Cr. The consumption of grain legumes dates to 5500 BC and is the second most consumed food crop across the globe next to cereals (Kouris-Blazos and Belski, 2016). Other than serving as high-quality food and feed, grain legumes defend the globe with reduced emission of greenhouse gases (5–7 times lesser than other crops). Carbon sequestration and atmospheric nitrogen fixation by grain legumes help to diversify crop cultivation, and reduced external inputs usage finds itself as a potential crop for sustainable agriculture (Stagnari et al., 2017).

Pigeonpea, also called as red gram, is a climate-resilient drylands legume and is widely cultivated in semiarid regions (Mula and Saxena, 2010). Globally, ∼5. million tonnes were produced from a planted area of 6.1 million hectares. Five countries, namely, India (82%), Myanmar (7%), Malawi (4%), Kenya (2%), and Tanzania (2%) account for 97% of the total cultivated area (FAOSTAT, 2021). Pigeonpea serves a variety of purposes, such as food, forage, feed, and meal for animals, piggery and fishery, fuel wood, green manure, barrier crop, rearing of lac insects, and roof thatches (Harinder et al., 1999; Upadhyaya et al., 2007; Mula and Saxena, 2010; Ohizua et al., 2017; Wangari et al., 2020). Nutritionally, pigeonpea grain is rich in protein, Ca, Mn, and crude fiber (Saxena et al., 2010). The variability for protein content, reported in previous studies, varied from 16.7 to 26.8% (Amarteifio et al., 2002; Sekhon et al., 2017; Obala et al., 2018; Cheboi et al., 2019; Choi et al., 2020; Jawalekar et al., 2020), whereas, in wild species, the range is from 16.3 to 33.8% (Upadhyaya et al., 2013). Very few studies enumerated mineral composition in pigeonpea (Singh et al., 1984; Oshodi et al., 1993; Amarteifio et al., 2002; Mishra and Acharya, 2018). Dahiya et al. (1977) reported the substantial influence of environment on protein, partial dominance of low protein over high protein, and negative correlation between seed yield and protein content. Furthermore, the nutrient accumulation varied with the seed developmental stage, Fe and Zn are rich at the green stage, whereas protein, starch, Mn, and Ca are high at the grain stage (Singh et al., 1984). Unlike cereals, the Fe and Zn are enriched in the cotyledons; thus the processing does not affect the availability of these minerals (Susmitha et al., 2022). Nutrient improvement in pigeonpea was done utilizing few wild species. Reddy et al. (1979) utilized wild species *Atylosia* from the tertiary gene pool to develop high protein lines, while Sharma et al. (2020) utilized *Cajanus platycarpus* to broaden the variability available for agronomic and grain nutritional traits. Saxena et al. (1987) identified high protein lines (HPL 2, HPL 7, HPL 40, and HPL 51) with 27–29% protein from five intergeneric crosses and mentioned the variable association between seed size and protein across crosses. Compared to normal lines (C 11 and ICPL 211), a hike of nearly 20% protein was observed in high protein lines (HPL 8 and HPL 40; Singh et al., 1990).

At ICRISAT, research on pigeonpea is focused primarily on the development of mid-early, early, and super-early varieties/hybrids with high yielding potential to attain self-sufficiency in the target areas. However, the identification of

genetic resources with superior grain nutrients can support pigeonpea biofortification and add nutritional security. The ICRISAT genebank conserves 13,787 pigeonpea germplasm.[1] This study was planned to characterize 600 diverse pigeonpea accessions for grain nutrients and important agronomic traits in 2 cropping years, with the objectives (i) to assess the variability for agronomic and grain nutritional traits, (ii) to understand the association between and among the agronomic traits and grain nutrients, and (iii) to identify trait-specific and multi-nutrient dense germplasm.

MATERIALS AND METHODS

Genetic Resources

The experimental material comprised 600 pigeonpea accessions conserved at Rajendra Singh Paroda Genebank, ICRISAT, India, along with four checks (**Supplementary Table 1**). The complete passport data of the germplasm are available in the ICRISAT Genebank database (see text footnote 1). These accessions represent 48 countries (**Figure 1**) across the globe, and five geographical regions, namely, Asia (357), Africa (148), America (80), Europe (11), and Oceania (4). The map depicting the country of collection was generated using R statistical software v.4.0.2 (R Core Team, 2021) using the "map" (South, 2011) and "ggplot2" (Wickham, 2016) packages. Out of 600 pigeonpea accessions, 577 were landraces, 19 improved cultivars/breeding lines, and four wild species accessions, *Cajanus acutifolius* (2), *C. cajanifolius* (1), and *C. sericeus* (1). ICP 11543 (an extra-early cultivar known as Pragati), ICP 6971 (an early cultivar known as UPAS-120), ICP 8863 (a medium duration cultivar, popularly known as Maruti), and ICP 7221 (a well-known long-duration cultivar named as Gwalior 3) were used as checks. Among these, ICP 11543 was developed by pedigree selection from the T 21 × JA 277 cross, while the remaining were developed through selection from the germplasm.

Field Experimental Design and Soil Properties

The experiment was laid in an augmented design with 20 blocks. Each block comprised of 30 test entries and four checks. Sowing was done in the last week of July in 2 cropping years, i.e., 2019 and 2020 at ICRISAT Patancheru, India (located at 17.51°N latitude, 78.27°E longitude, and 545 m above the mean sea level) in alfisols. Each accession was sown in a 4-meter row with an inter-row spacing of 75 cm and plant-plant spacing of 20–25 cm. As per the USDA soil taxonomy, the soil belongs to the fine loamy-mixed isohyperthermic family of Udic Rhodustalf. The first 30-cm soil of the experimental field in the 2019 rainy season had 7.22 pH, 0.07 dS/m EC, 0.42% organic matter, 7.5 mg/kg P, 67 mg/kg K, 1,116 mg/kg exchangeable Ca, 368 mg/kg exchangeable Mg, 6.1 mg/kg Fe, 1.39 mg/kg Zn, 1.34 mg/kg Cu, and 18.53 mg/kg Mn, and the 2020 rainy season had 6.97 pH, 0.08 dS/m EC, 0.45% organic matter, 18.67 mg/kg P, 79 mg/kg K, 1,057 mg/kg exchangeable Ca, 340 mg/kg of exchangeable Mg, 8.93 mg/kg Fe, 4.24 mg/kg Zn, 1.25 mg/kg Cu, and 17.54 mg/kg Mn.

[1]http://genebank.icrisat.org/

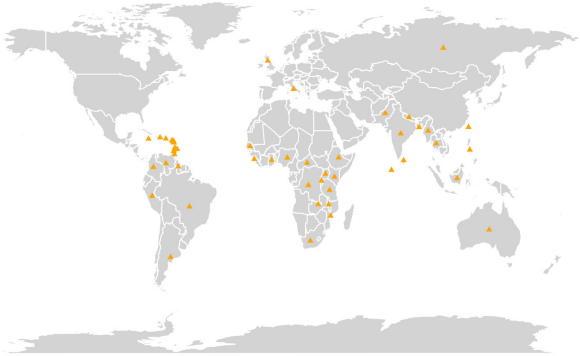


FIGURE 1 | A map depicting the country of collection of the 600 pigeonpea accessions and four checks conserved at Genebank, ICRISAT, India.

Agronomic Practices and Phenotyping

The agronomic practices started with the basal application of DAP (diammonium phosphate) at a rate of 100 kg/hectare. Thinning was practiced 21 days after sowing to maintain optimum plant density. Optimum field conditions were maintained following standard package of practises. Agronomic traits recorded were days to 50% flowering, days to maturity, 100-seed weight, and grain yield per plant. Days to 50% flowering was recorded on a plot basis. Days to maturity and grain yield per plant were recorded on a single-plant basis (5–21 plants) and averaged to represent the accession. The 100-seed weight was recorded from a random sample of 100 seeds drawn from the bulked single-plant yield of each accession. Grain nutrients analyses were performed on 598 accessions, while two checks (ICP 11543 and ICP 6971) and four accessions having poor germination/plant stand were excluded.

Grain Nutrients Estimation

The grain nutrients estimated in the study were protein, P, K, Ca, Cu, Mg, Mn, Fe, and Zn. Clean and dust-free grain samples weighing 15 g were taken from the bulked single-plant yield of each accession in each cropping year for grain nutritional analysis. The grain samples were submitted following the augmented design. The grain nutrients estimation was done at Charles Renard Analytical laboratory, ICRISAT, India. Protein estimation was done by digesting the grain sample by the sulfuric acid-selenium digestion method and analyzing the digests in a continuous flow autoanalyzer to obtain the total N value from which protein (%) is calculated by multiplying the total N with a 6.25 conversion factor (Sahrawat et al., 2002). Estimation of P, K, Ca, Cu, Mg, Mn, Fe, and Zn was done by digesting the plant samples with the nitric acid – hydrogen peroxide digestion method and analyzing the digests in Microwave Plasma Atomic Emission Spectrometry (MP-AES; Wheal et al., 2011).

Statistical Analysis

The components of variances for four agronomic and nine-grain nutritional traits for the individual years and pooled data over 2 years were analyzed by adopting the linear mixed model in residual maximum likelihood (REML) in GenStat 19 (VSN International, 2019). For the individual years, entry and block were assigned as random effects, whereas, in pooled data over years, the cropping year was kept fixed, and the factors, namely, entry, cropping year, and block were assigned random. Variance due to genotype (σ^2_g), genotype × environment ($\sigma^2_{g \times e}$), and error (σ^2_e) was estimated, while the significance of cropping years was tested by Wald's statistics (Wald, 1943). Heritability in broad sense for individual and pooled data over cropping years for each trait was estimated and categorized based on Johnson et al. (1955). Best linear unbiased predictors (BLUPs; Schönfeld and Werner, 1986) obtained for all the traits for each accession in each cropping year, and pooled analyses over cropping years were used for all downstream analyses. The accessions were broadly classified into three maturity groups as early (≤150 days to maturity), medium (151–180 days to maturity), and late (>180 days to maturity; Reddy, 1990). Newman–Keuls test (Newman, 1939; Keuls, 1952) and Levene's test (Levene, 1960) were used to compare the mean and test the homogeneity of variances in different groups formed based on the geographical region and maturity using R packages "agricolae" (de Mendiburu, 2021) and "car" (Fox and Weisberg, 2019). Histogram and a density graph depicting the distribution of agronomic and grain nutrients in each cropping year, geographical region, and maturity group were visualized using the package "ggplot2" (Wickham, 2016). The correlation coefficients among the agronomic and grain nutritional traits were performed using the native R function "cor ()" and visualized using the "corrplot" (Wei and Simko, 2021) package. The phenotypic distance matrix for four agronomic traits and nine-grain nutrients was constructed following the Gower's dissimilarity method using the R package "vegan" (Oksanen et al., 2020) and the dendrogram constructed based on the Ward.D2 method (Murtagh, 2014) using the R package "cluster" (Maechler et al., 2021), with a heatmap depicting the agronomic performance and grain nutrients content of each accession of the cluster using the package "heatmap3" (Zhao et al., 2021). The cluster means were compared using the Newman–Keuls test (Newman, 1939; Keuls, 1952). The circular stacked barplot depicting the contribution of each region to the sub-cluster was constructed using the "ggplot2" (Wickham, 2016) package. The nutrient-specific and multi-nutrient dense accessions were identified based on *per se* performance and superiority to the best check.

RESULTS

Components of Variance

The REML ANOVA indicated that the variance due to genotypes was highly significant ($p \leq 0.01$) for all agronomic and grain nutrients for individual cropping years and pooled analysis over cropping years except for P, K, and Mn of the 2019 rainy season (**Table 1**). The variance due to Genotype × Cropping year interaction ($\sigma^2_{g \times e}$) was significant for days to 50% flowering, 100-seed weight, grain yield per plant, and for grain nutrients, namely, protein, Fe, Zn, Mg, Cu, and Mn ($p \leq 0.05$) while insignificant for days to maturity, P, K, and Ca. However, the variance due to genotype (σ^2_g) was higher than the G × E variance ($\sigma^2_{g \times e}$) for days to 50% flowering, 100-seed weight, Mg, Cu, Mn, and Zn, but it was reverse for grain yield per plant, protein, and Fe. Wald's statistics for the environment (cropping years) revealed a significant difference between the cropping years (σ^2_e) for all agronomic and grain nutritional traits except for Ca.

Variability Parameters and Heritability (h^2)

The comparison of mean values of 598 pigeonpea accessions between two cropping years revealed a significant difference in the performance of the accessions for all the traits except for Ca (**Table 1**). Phenologically, the accessions flowered and matured earlier in the 2020 rainy season (116 ± 14.8 and 166 ± 14.5 days, respectively) compared to the 2019 rainy season (131 ± 16.5 and 182 ± 15.4 days, respectively; **Figures 2A,B**). Concerning agronomic performance, 100-seed weight was relatively higher in the 2019 rainy crop (10.67 ± 3.0 g) compared to the 2020 rainy crop (9.52 ± 2.6 g; **Figure 2C**), and grain yield per plant was higher in the 2020 rainy season crop (34.10 ± 7.7 g) compared with 2019 rainy crop (30.63 ± 7.6 g; **Figure 2D**). For grain nutrients, the 2019 rainy crop had high protein (27.30 ± 1.2%), P (0.45 ± 0.02%), Mn (10.75 ± 0.7 mg/kg), Fe (38.61 ± 1.9 mg/kg), and Zn (29.59 ± 1.8 mg/kg; **Figures 2E,F,K,L,M**), whereas the 2020 rainy crop had high K (1.55 ± 0.02%), Mg (1546.66 ± 77.6 mg/kg), and Cu (11.69 ± 0.8 mg/kg; **Figures 2G,I,J**). However, the Ca content (1537.58 ± 173.1 mg/kg and 1548.03 ± 170.3 mg/kg in 2019 and 2020 rainy seasons, respectively) observed no significant difference between the cropping years (**Figure 2H**). The heritability estimates in broad sense were found to be moderate to high (0.31–0.89 and 0.44–0.99 for the 2019 and 2020 rainy seasons, respectively) for all agronomic and grain nutritional traits in both the cropping years except for K (0.26) and Mn (0.19) in the 2019 rainy season.

The pooled analysis over cropping years presented that the average days to 50% flowering of 598 pigeonpea accession as 124 days encompassing 2-fold variation (67–166 days) and 275 accessions were found to be earlier than the trial mean (**Table 1**). Similarly, the days to maturity varied from 112 to 213 days with 261 accessions maturing earlier than the trial mean of 174 days. The accessions had wide variability for 100-seed weight, holding very small seeds (1.69 g) to large seeds (22.17 g), with an average 100-seed weight of 10.10 g and 236 accessions surpassed the trial mean. Grain yield per plant varied from 16.54 to 57.93 g, and 271 accessions yielded higher than the trial mean (32.37 g) and ICP 15241 recorded the highest grain yield per plant (57.93 g). For grain nutrients, the accessions varied from 23.35 - 29.50% for protein, 0.36–0.50% for P, 1.43–1.63% for K, 1,311.01–1,865.65 mg/kg for Mg, 8.56–14.01 mg/kg for

TABLE 1 | Variance, mean, range, co-efficient of variation (CV %), least significant difference ($LSD_{0.05}$), and heritability (broad sense) for agronomic and grain nutritional traits of pigeonpea accessions evaluated during 2019 and 2020 rainy seasons at ICRISAT, Hyderabad, India.

Trait	Environment (σ^2_e)		Genotype (σ^2_g)	G × E ($\sigma^2_{g \times e}$)	Mean ± SD	Range	CV (%)	$LSD_{0.05}$	h^2
Days to 50% flowering	Pooled	1730.62**	245.12**	17.87**	124 ± 15.6	67–166	3.84	13.21	0.92
	2019		284.60**		131 ± 16.5[a]	71–171	4.63	16.89	0.89
	2020		219.34**		116 ± 14.8[b]	62–160	1.38	4.46	0.99
Days to maturity	Pooled	2295.27**	230.11**	8.09	174 ± 14.9	112–213	2.86	13.84	0.93
	2019		259.97**		182 ± 15.4[a]	120–221	3.10	15.68	0.89
	2020		197.68**		166 ± 14.5[b]	105–206	1.31	6.03	0.98
100-seed weight (g)	Pooled	408.14**	7.26**	0.69**	10.10 ± 2.7	1.69–22.17	5.17	1.45	0.94
	2019		6.65**		10.67 ± 3.0[a]	1.96–23.28	10.66	3.16	0.84
	2020		6.32**		9.52 ± 2.6[b]	1.41–21.05	6.08	1.61	0.95
Grain yield per plant(g)	Pooled	28.22**	47.65**	51.67**	32.37 ± 6.9	16.54–57.93	22.22	19.98	0.48
	2019		112.34**		30.63 ± 7.6[b]	13.13–61.45	18.63	15.83	0.78
	2020		106.90**		34.10 ± 7.7[a]	13.73–61.81	19.28	18.31	0.71
Protein (%)	Pooled	21.87**	1.10**	1.47*	26.99 ± 1.0	23.35–29.50	6.90	5.17	0.31
	2019		2.67**		27.30 ± 1.2[a]	23.07–30.60	6.29	4.77	0.48
	2020		1.57**		26.69 ± 1.0[b]	23.54–29.11	5.21	3.86	0.45
Phosphorus (%)	Pooled	140.59**	0.0006**	0.0008	0.43 ± 0.02	0.36–0.50	12.91	0.15	0.25
	2019		0.0013		0.45 ± 0.02[a]	0.35–0.54	12.09	0.15	0.31
	2020		0.0006**		0.41 ± 0.02[b]	0.35–0.48	4.71	0.05	0.61
Potassium (%)	Pooled	408.10**	0.0016**	0.0002	1.50 ± 0.02	1.43–1.63	5.34	0.22	0.33
	2019		0.0019		1.46 ± 0.02[b]	1.39–1.58	5.00	0.20	0.26
	2020		0.0029**		1.55 ± 0.02[a]	1.48–1.67	3.78	0.16	0.46
Calcium (mg/kg)	Pooled	1.00	38700**	3357	1542.80 ± 171.2	1042.36–2099.76	11.13	476.83	0.70
	2019		42988**		1537.58 ± 173.1[a]	1025.98–2110.34	10.87	463.95	0.61
	2020		41289**		1548.03 ± 170.3[a]	1058.74–2089.18	9.53	410.15	0.65
Magnesium (mg/kg)	Pooled	37.93**	7314**	2287*	1530.20 ± 78.2	1311.01–1865.65	5.14	218.42	0.63
	2019		13443**		1513.73 ± 82.5[b]	1283.83–1835.19	3.62	152.18	0.82
	2020		7932**		1546.66 ± 77.6[a]	1338.19–1896.11	4.57	196.48	0.61
Copper (mg/kg)	Pooled	371.6**	0.75**	0.30**	11.17 ± 0.8	7.72–14.20	6.74	2.09	0.63
	2019		0.85**		10.65 ± 0.9[b]	6.58–13.45	8.58	2.54	0.51
	2020		0.84**		11.69 ± 0.8[a]	8.86–15.05	5.30	1.72	0.69
Manganese (mg/kg)	Pooled	139.73**	0.41**	0.38**	10.40 ± 0.6	8.56–14.01	7.55	2.18	0.45
	2019		0.29		10.75 ± 0.7[a]	8.51–14.80	10.22	3.05	0.19
	2020		0.43**		10.05 ± 0.6[b]	8.48–13.22	7.43	2.07	0.44
Iron (mg/kg)	Pooled	1631.43**	3.23**	3.41**	34.89 ± 1.7	29.23–40.98	6.65	7.13	0.39
	2019		8.22**		38.61 ± 1.9[a]	32.35–46.08	5.95	6.37	0.61
	2020		6.06**		31.18 ± 1.8[b]	25.14–37.91	7.51	6.50	0.52
Zinc (mg/kg)	Pooled	21.74**	2.78**	1.91**	29.27 ± 1.6	24.14–35.68	6.49	5.28	0.50
	2019		5.90**		29.59 ± 1.8[a]	22.44–36.50	6.60	5.43	0.61
	2020		4.14**		28.95 ± 1.6[b]	24.75–34.86	5.33	4.28	0.64

*SD, standard deviation; CV, co-efficient of variation; LSD, least significant difference; h^2, heritability in broad sense. *, **Significant at 0.05, 0.01 probability levels, respectively. The mean followed by the same letters is not significant at $p \leq 0.05$, and the mean followed by different letters is significant at $p \leq 0.05$.*

Mn, 29.23–40.98 mg/kg for Fe, and 24.14–35.68 mg/kg for Zn. With regard to the superiority of the germplasms to the trial mean, 294 accessions for protein (>26.99%), 223 accessions for P (>0.43%), 291 accessions for K (>1.50%), 290 accessions for Mg (>1,530.20 mg/kg), 288 accessions for Mn (>10.40 mg/kg), 291 accessions for Fe (>34.89 mg/kg), and 281 accessions for Zn (>29.27 mg/kg) were found to surpass the trial mean. However, a 2-fold variation was observed for Ca (1,042.36–2,099.76 mg/kg) and Cu (7.72–14.2 mg/kg), with 290 and 291 accessions exceeding the trial mean (1,542.80 mg/kg and 11.17 mg/kg, respectively). High heritability was observed for all agronomic traits except grain yield per plant (0.48). The heritability for grain nutrients was high for Ca (0.70), Mg (0.63), and Cu (0.63), whereas it was moderate for protein (0.31), K (0.33), Mn (0.45), Fe (0.39), and Zn (0.50). Grain phosphorus content had low heritability (0.25).

Mean Comparison Between Geographical Regions and Maturity Groups

Accessions from three regions, namely, Asia (358), Africa (148), and America (79) were considered for mean comparison, while

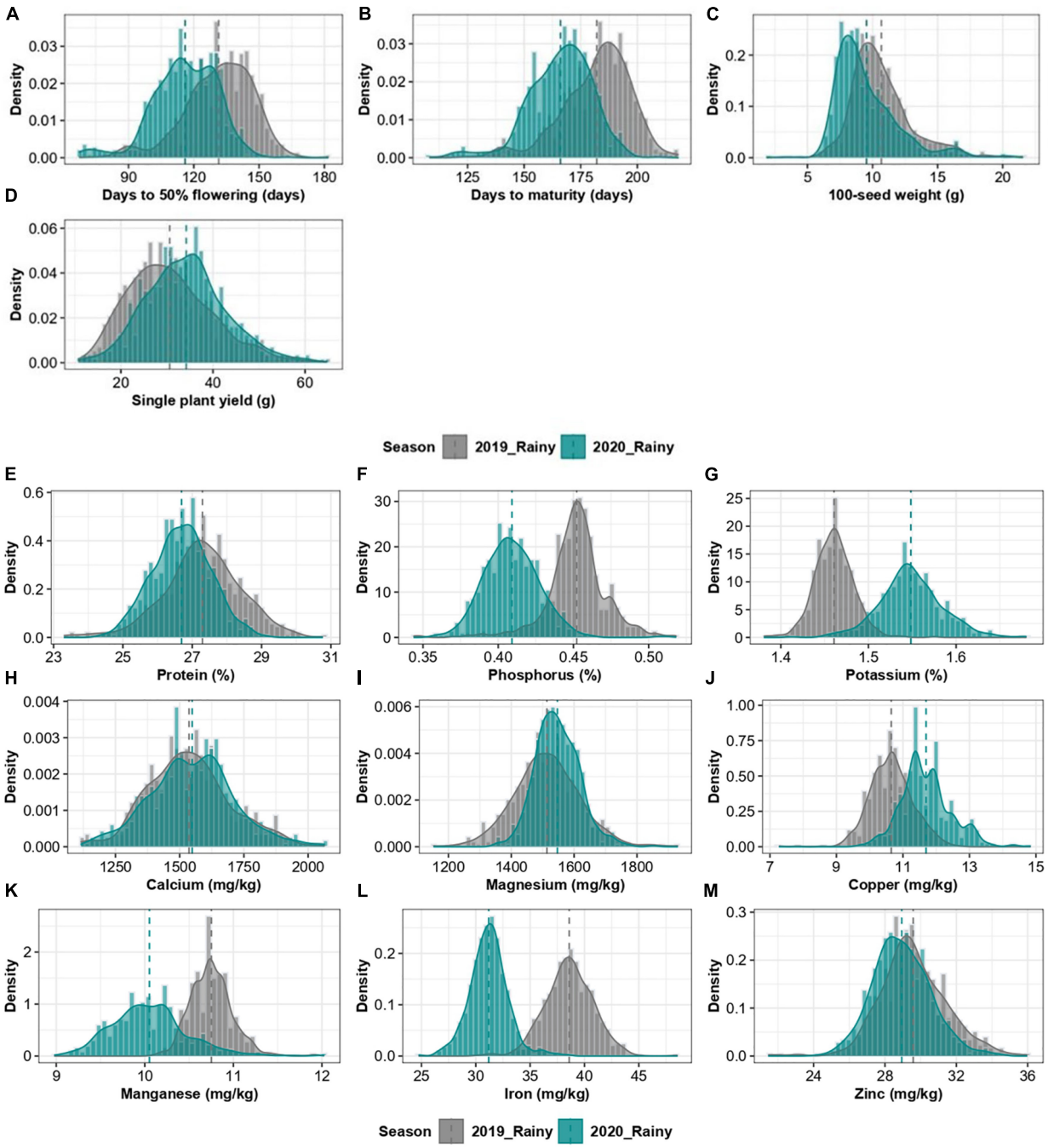


FIGURE 2 | Combined histogram and a density graph, depicting the density of agronomic traits **(A–D)** and grain nutrients **(E–M)** of 2019 and 2020 rainy season crops.

other regions with few accessions (Europe-11 and Oceania-2) were excluded. The region-wise mean comparison revealed that all agronomic and grain nutritional traits, except for Zn, varied significantly with the geographical region (**Table 2**). The traits – days to 50% flowering and days to maturity significantly differentiated the Asian (122 ± 16.8 and 172 ± 16.1 days, respectively) and American (123 ± 11.3 and 174 ± 10.8 days, respectively) regions from the African region (130 ± 11.6 and 180 ± 10.9 days, respectively; **Figures 3A,B**). The 100-seed weight varied significantly in all the three regions, with the American region having a higher 100-seed weight (12.47 ± 2.5 g), followed by the African (11.80 ± 2.4 g) and Asian regions (8.87 ± 2.2 g; **Figure 3C**). No significant difference was observed between the African and American

regions for grain yield per plant; however, the Asian region had a significantly higher yield (34.41 ± 7.1 g; **Figure 3D**). Protein was the only nutrient to differentiate all the three geographical regions, while other nutrients differentiated one of the three regions. Mean protein content was significantly higher in the Asian region (27.24 ± 1.%), which was followed by the African region (26.73 ± 0.7%) and the American region (26.44 ± 0.8%; **Figure 3E**). For other nutrients, one region stayed significantly distinct from the other two regions for the P (0.44 ± 0.02%) African region, the K (1.50 ± 0.02%), Mg (1,545.66 ± 77.7 mg/kg), and Cu (10.92 ± 0.8 mg/kg) Asian region, and the Ca (1,494.7 ± 181.2 mg/kg) and Mn (10.16 ± 0.7 mg/kg) American region (**Figures 3F–K**). For mean Fe content, a significant difference existed between the American (34.51 ± 1.6 mg/kg) and African regions (35.15 ± 1.5 mg/kg), while the Asian region (34.88 ± 1.8 mg/kg) was indifferentiable from the two regions (**Figure 3L**). For Zn, there was no significant difference between the geographical regions (**Figure 3M**). The variances remained heterogeneous for all agronomic traits and nutrients, namely, protein, P, and Fe (**Table 2**).

A comparison of agronomic traits and grain nutrients was made between maturity groups, early (32 accessions), medium (234 accessions), and late (332 accessions; **Table 2**). The mean days to 50% flowering and days to maturity significantly differentiated the maturity groups as the classification was based on the same (**Figures 4A, B**). Among the three maturity groups, the medium and late maturity groups showed no significant difference for 100-seed weight and grain yield per plant but were significantly higher than the early maturity group (**Figures 4C,D**). The nutrients, namely, protein, P, K, Ca, and Mn did not vary significantly between the maturity groups (**Figures 4E–H,K**). However, the mean Fe and Zn content in grain marked a significant difference between the maturity groups, with the early maturity group with high Fe (36.43 ± 1.7 mg/kg) and Zn (30.99 ± 1.7 mg/kg), followed by the medium duration group with intermediate Fe (35.05 ± 1.6 mg/kg) and Zn (29.40 ± 1.5 mg/kg) and the late maturity group with low Fe (34.47 ± 1.7 mg/kg) and Zn content (28.84 ± 1.5 mg/kg; **Figures 4L,M**). For Mg, the early (1,556.89 ± 54.9 mg/kg) and late maturity group (1,519.17 ± 86.6 mg/kg) varied significantly, while the medium duration group (1,535.40 ± 72.5 mg/kg) was indifferentiable between the two groups (**Figure 4I**). The Cu content in the early maturity group (11.80 ± 0.9 mg/kg) was high and varied significantly from the other two groups (**Figure 4J**). The variances were homogenous for all-grain nutrients except for Mg. Agronomic traits had heterogeneous variance except for grain yield per plant.

Correlation Between Agronomic Traits and Grain Nutrients

Among the agronomic traits, a highly significant and positive correlation was seen, except for a significant negative correlation between 100-seed weight and grain yield per plant ($r = -0.254$, $p \leq 0.01$; **Figure 5** and **Supplementary Table 2**). Protein, the nutrient of great significance in legumes, mostly had a significant positive correlation with all nutrients ($r = 0.136$–0.429, $p \leq 0.01$), except for a non-significant negative correlation with Ca ($r = -0.018$ and Cu ($r = -0.024$). The correlation between P and K was positive and highly significant ($r = 0.221$, $p \leq 0.01$), and similar correlations with nutrients, namely, protein, Cu, Fe, and Zn were seen. In addition, P, with a highly significant positive association with Mg ($r = 0.156$, $p \leq 0.01$), and K, with a highly significant negative association with Ca ($r = -0.235$, $p \leq 0.01$), were recorded. While the association of Ca was highly significant and positive with Mg, Mn, and Fe ($r = 0.115$–0.683, $p \leq 0.01$) and mostly non-significant with all other nutrients. Between Fe and Zn existed a highly significant and positive correlation ($r = 0.580$, $p \leq 0.01$). Other than the correlation for Fe with Ca, Fe ($r = 0.205$–0.340, $p \leq 0.01$) and Zn ($r = 0.148$–0.495, $p \leq 0.01$) had a highly significant positive correlation with all other nutrients. The association of days to 50% flowering and days to maturity with grain nutrients was mostly negative and was significant for protein, Mg, Cu, Fe, and Zn. However, Mn recorded a significant positive association with days to maturity ($r = 0.083$, $p \leq 0.05$). Although, 100-seed weight recorded a highly significant and positive correlation with K ($r = 0.158$, $p \leq 0.01$) and Cu ($r = 0.403$, $p \leq 0.01$), the association with most of the nutrients (protein, Ca, Mg, Mn, Fe, and Zn) was found to be negative and highly significant ($r = -0.140$ to -0.370, $p \leq 0.01$). Between grain yield per plant and nutrients namely, protein and Mg, a significant positive correlation was seen ($r = 0.104$ and 0.107, respectively, $p \leq 0.05$). On the other hand, the association of grain yield per plant with most of the nutrients, namely, P, K, Cu, Fe, and Zn was negative and highly significant ($r = -0.106$ to -0.377, $p \leq 0.01$). Withal, Ca recorded no significant correlation with most of the agronomic traits, except 100-seed weight.

A correlation study was conducted in three major geographical regions (Asia, Africa, and America) to identify significant and unique correlations existing between grain nutrients and agronomic traits in each region (**Supplementary Tables 3–5**). Across all the regions, the association among grain nutrients protein, Ca, Fe, and Zn remains unaltered from the general correlation, except for a non-significant association between Fe and Ca. In the Asian region, days to 50% flowering and days to maturity were significantly positively correlated with each other and with grain yield per plant ($r = 0.333$ and 0.332, $p \leq 0.01$), whereas, in the African region, was instead with 100-seed weight ($r = 0.458$ and 0.469, $p \leq 0.01$). Furthermore, 100-seed weight with a significant negative correlation with grain yield per plant ($r = -0.325$, $p \leq 0.01$) was observed only in the African region, and, in other regions, it was insignificant. In the American region, a significant positive correlation existed only between days to 50% flowering and days to maturity ($r = 0.985$, $p \leq 0.01$). Between agronomic traits and grain nutrients, namely, protein and Ca, the association was non-significant in all the three regions, except for a significant negative association with 100-seed weight in the Asian ($r = -0.292$ and $r = -0.173$, respectively, $p \leq 0.01$) and American regions ($r = -0.304$ and $r = -0.285$, respectively, $p \leq 0.01$). Concerning Fe and Zn, the association was significant and negative with all the agronomic traits in the Asian region. In the African and American regions,

TABLE 2 | Mean and range for agronomic and grain nutritional traits of 598 pigeonpea accessions belonging to different regions and maturity groups evaluated during the 2019 and 2020 rainy seasons at ICRISAT, Hyderabad, India.

Trait	Region							Maturity group						
	Mean ± SD			Range			Homogeneity of variance (*F* value)	Mean ± SD			Range			Homogeneity of variance (*F* value)
	Asia (358)	Africa (148)	America (79)	Asia (358)	Africa (148)	America (79)		Early (32)	Medium (234)	Late (332)	Early (32)	Medium (234)	Late (332)	
DFF	122 ± 16.8^{b}	130 ± 11.6^{a}	123 ± 11.3^{b}	67–166	88–152	85–148	22.75*	85 ± 7.3^{c}	117 ± 8.7^{b}	138 ± 5.6^{a}	66–100	96–150	126–166	35.77*
DM	172 ± 16.1^{b}	180 ± 10.9^{a}	174 ± 10.8^{b}	112–213	144–201	140–198	22.61*	137 ± 7.6^{c}	168 ± 8.2^{b}	188 ± 5.4^{a}	112–150	151–180	180–213	43.46*
SW (g)	8.87 ± 2.2^{c}	11.8 ± 2.4^{b}	12.47 ± 2.5^{a}	1.69–22.17	7.63–17.78	8.19–18.25	11.17*	8.57 ± 0.9^{b}	10.17 ± 2.8^{a}	10.20 ± 2.8^{a}	7.02–10.46	1.69–22.17	5.63–18.69	9.41*
GYP (g)	34.41 ± 7.1^{a}	29.79 ± 5.9^{b}	28.62 ± 5.8^{b}	16.54–57.93	17.42–45.74	16.66–43.71	4.15*	25.83 ± 5.6^{b}	32.40 ± 6.8^{a}	33.21 ± 7.1^{a}	16.54–41.12	16.66–57.93	17.42–53.61	1.70
Protein (%)	27.24 ± 1.0^{a}	26.73 ± 0.7^{b}	26.44 ± 0.8^{c}	23.35–29.5	25.03–28.68	24.68–28.04	12.92*	27.23 ± 0.9^{a}	26.97 ± 1.0^{a}	26.99 ± 1.0^{a}	24.68–29.12	23.35–29.5	24.77–29.02	0.62
P (%)	0.43 ± 0.02^{b}	0.44 ± 0.02^{a}	0.43 ± 0.02^{b}	0.36–0.48	0.4–0.5	0.39–0.48	5.71**	0.44 ± 0.02^{a}	0.43 ± 0.02^{a}	0.43 ± 0.02^{a}	0.40–0.46	0.37–0.49	0.36–0.50	0.61
K (%)	1.50 ± 0.02^{b}	1.51 ± 0.02^{a}	1.51 ± 0.02^{a}	1.43–1.63	1.45–1.58	1.47–1.58	0.36	1.51 ± 0.02^{a}	1.51 ± 0.02^{a}	1.50 ± 0.02^{a}	1.44–1.55	1.43–1.63	1.44–1.57	0.23
Ca (mg/kg)	1556.49 ± 171.2^{a}	1538.74 ± 160.7^{a}	1494.72 ± 181.2^{b}	1147.71–2099.76	1081.02–1913.49	1042.36–1881.99	1.07	1536.86 ± 148.9^{a}	1543.64 ± 171.7^{a}	1542.43 ± 171.5^{a}	1195.89–2006.41	1042.36–2099.76	1098.43–1965.28	0.28
Mg (mg/kg)	1545.66 ± 77.7^{a}	1510.35 ± 75.2^{b}	1498.04 ± 71.4^{b}	1327.51–1789.81	1335.08–1865.65	1311.01–1734.69	0.57	1556.89 ± 54.9^{a}	1535.4 ± 72.5^{ab}	1519.17 ± 86.6^{b}	1426.27–1684.17	1327.51–1750.32	1311.01–1865.65	4.49*
Cu (mg/kg)	10.92 ± 0.8^{b}	11.50 ± 0.7^{a}	11.68 ± 0.8^{a}	7.72–13.96	9.86–13.4	9.92–14.2	0.94	11.80 ± 0.9^{a}	11.24 ± 0.8^{b}	10.99 ± 0.8^{b}	10.5–13.96	9.43–14.2	7.72–13.4	0.70
Mn (mg/kg)	10.44 ± 10.4^{a}	10.43 ± 0.6^{a}	10.16 ± 0.7^{b}	8.89–14.01	9.19–12.05	8.56–12.51	0.40	10.24 ± 0.4^{a}	10.39 ± 0.6^{a}	10.43 ± 0.7^{a}	9.41–11.25	8.56–14.01	8.89–12.51	2.12
Fe (mg/kg)	34.88 ± 1.8^{ab}	35.15 ± 1.5^{a}	34.51 ± 1.6^{b}	29.23–40.98	31.96–39.69	31.00–39.08	3.40*	36.43 ± 1.7^{a}	35.05 ± 1.6^{b}	34.47 ± 1.7^{c}	33.62–39.81	30.00–39.69	29.23–40.98	0.41
Zn (mg/kg)	29.42 ± 1.6^{a}	29.04 ± 1.4^{a}	29.04 ± 1.6^{a}	24.14–35.68	25.33–34.18	25.51–33.78	1.80	30.99 ± 1.7^{a}	29.40 ± 1.5^{b}	28.84 ± 1.5^{c}	27.46–33.93	25.51–35.68	24.14–33.59	0.90

*The value inside the parenthesis represents the number of accessions in each category. SD, standard deviation; DFF, days to 50% flowering; DM, days to maturity; SW, 100-seed weight; GYP, grain yield per plant; P, phosphorus; K, potassium; Ca, calcium; Cu, copper; Mg, magnesium; Mn, manganese; Fe, iron; Zn, zinc. The mean followed by the same letters is not significant at $p \leq 0.05$, and the mean followed by different letters is significant at $p \leq 0.05$. *Homogeneity of variance tested by Levene's test is significant at $p \leq 0.05$.*

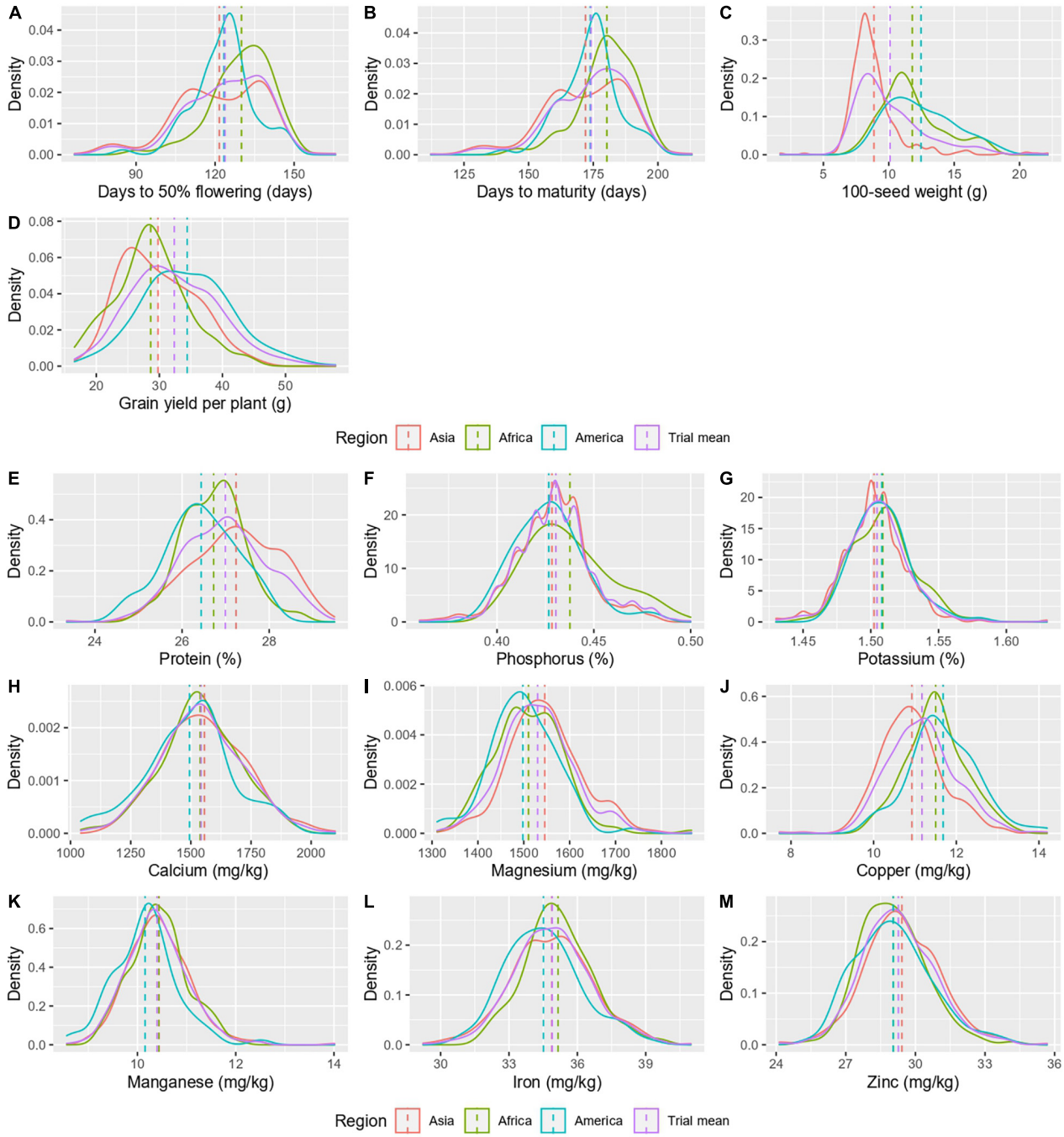


FIGURE 3 | A density graph depicting the distribution of agronomic traits **(A–D)** and grain nutrients **(E–M)** in different geographical regions.

the association of Fe and Zn with most of the agronomic traits was negative and significant, except for a non-significant negative association for Fe with grain yield per plant and Zn with 100-seed weight.

As the Fe and Zn content varied with the maturity group, the correlation study was conducted in each maturity group (**Supplementary Tables 6–8**). The association of the protein with days to 50% flowering, days to maturity, and 100-seed weight was non-significant in the early maturity group, significant and negative in the medium maturity group ($r = -0.221$ to -0.439, $p \leq 0.01$), and non-significant in the late maturity group, except for a significant negative association with 100-seed weight ($r = -0.288$, $p \leq 0.01$). Protein was significantly positively correlated with grain yield per plant in the early ($r = 0.392$, $p \leq 0.05$) and late maturity groups ($r = 0.125$, $p \leq 0.05$) and was non-significant in the medium maturity group. Across all

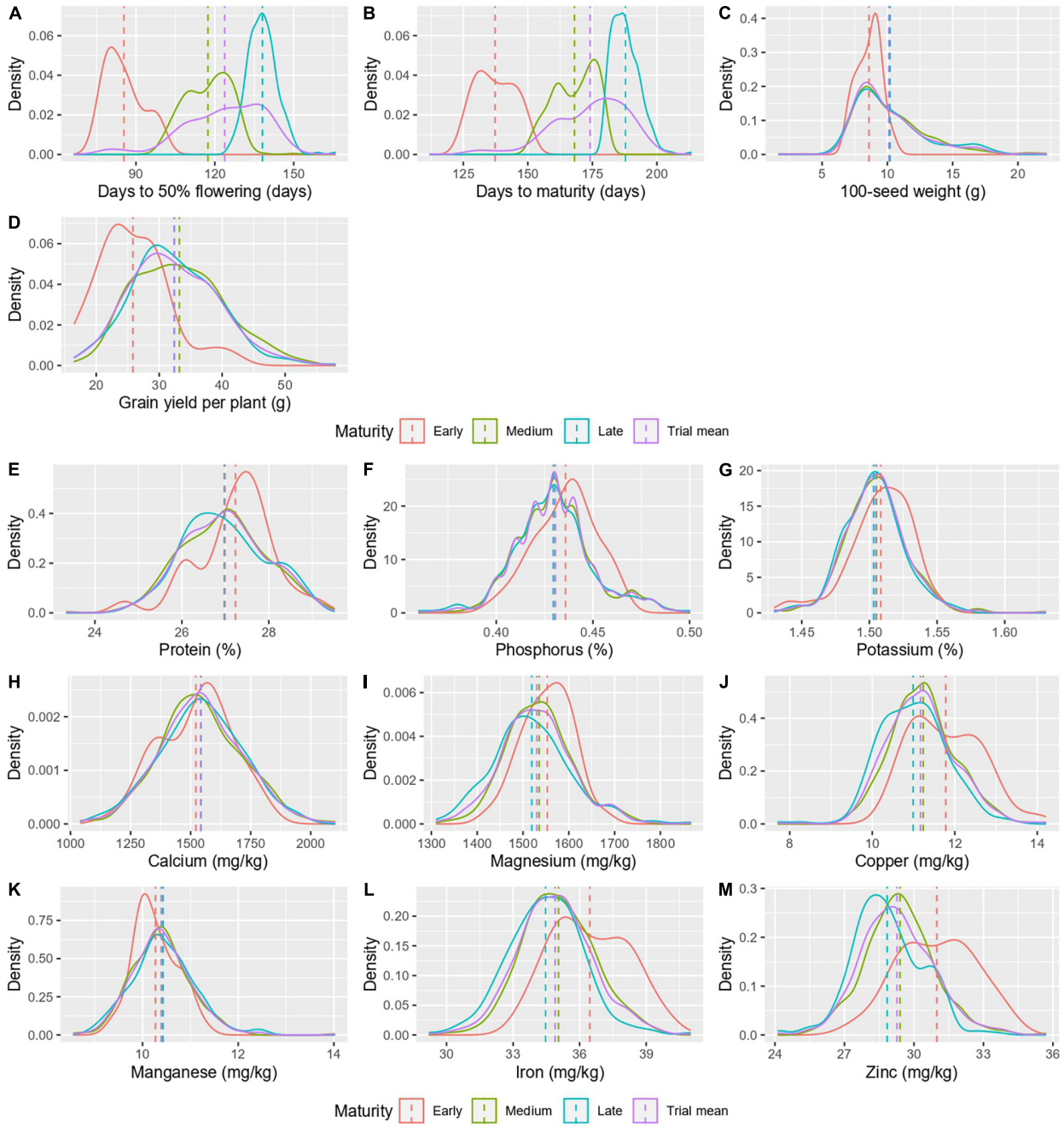


FIGURE 4 | A density graph depicting the distribution of agronomic traits **(A–D)** and grain nutrients **(E–M)** in different maturity groups.

the maturity groups, Ca was significantly negatively correlated with 100-seed weight ($r = -0.165$ to -0.452, $p \leq 0.01$) and non-significantly with all other agronomic traits, except for a significant positive association with days to maturity ($r = 0.143$, $p \leq 0.05$) in the late maturity group. In the early maturity group, Fe and Zn had a non-significant association with all agronomic traits, except for a significant negative association between Zn and grain yield per plant ($r = -0.493$, $p \leq 0.01$). In the medium maturity group, the association of Fe and Zn with all agronomic traits was negative and significant. In the late maturity group, Fe was significantly negatively correlated with days to maturity ($r = -0.143$, $p \leq 0.05$), whereas Zn had a significant positive association with days to 50% flowering ($r = 0.137$, $p \leq 0.05$) and a significant negative association with 100-seed weight ($r = -0.182$, $p \leq 0.01$).

The 100-seed weight marks the consumer preference, and it is noteworthy to study its association with grain nutrients (**Supplementary Tables 9–11**). Protein had a significant negative

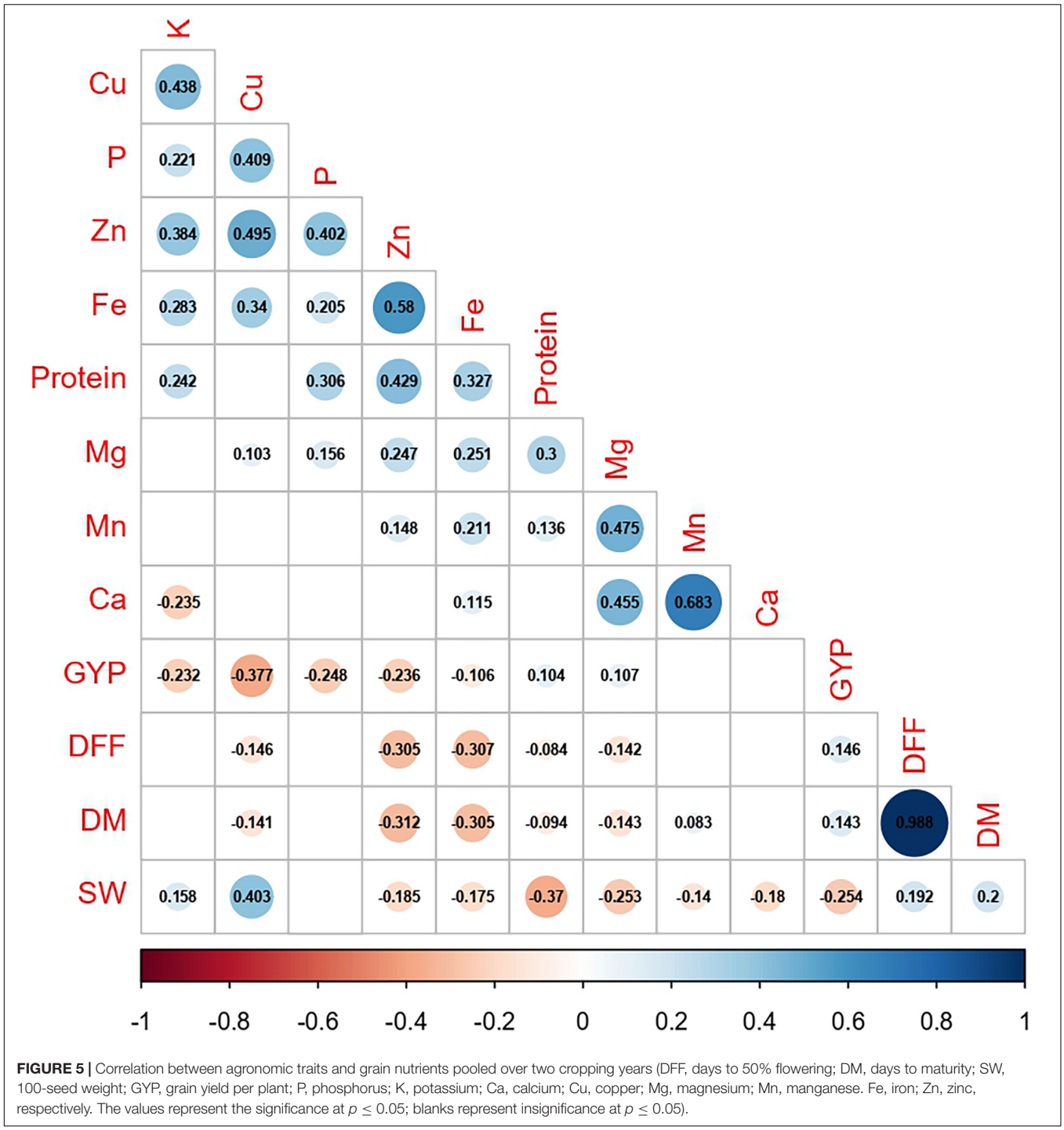


FIGURE 5 | Correlation between agronomic traits and grain nutrients pooled over two cropping years (DFF, days to 50% flowering; DM, days to maturity; SW, 100-seed weight; GYP, grain yield per plant; P, phosphorus; K, potassium; Ca, calcium; Cu, copper; Mg, magnesium; Mn, manganese. Fe, iron; Zn, zinc, respectively. The values represent the significance at $p \leq 0.05$; blanks represent insignificance at $p \leq 0.05$).

correlation with 100-seed weight of ≤ 10 g ($r = -0.156$, $p \leq 0.01$) and 10–15 g ($r = -0.222$, $p \leq 0.01$), and was non-significant beyond 15 g. With ≤ 10-g 100-seed weight, the grain Ca content was significantly negatively correlated ($r = -0.322$, $p \leq 0.01$) beyond which there was no significant correlation. Concerning Fe and Zn, mostly existed a non-significant association with 100-seed weight, except for a significant negative association with Fe ($r = -0.329$, $p \leq 0.05$) for >15-g 100-seed weight.

Cluster Analysis

The clustering based on Gower's distance matrix apportioned the 598 pigeonpea accessions into two major clusters with three sub-clusters each at h = 0.7 (**Figure 6** and **Supplementary Figure 1**). The minimum and maximum numbers of accessions were observed in sub-clusters 2 (52) and 4 (138), respectively. The other sub-clusters 1, 3, 5, and 6 had 59, 124, 92, and 133 accessions, respectively (**Supplementary Table 12**).

The region-wise contribution identified major Cluster I, with accessions predominantly from the Asian region, whereas the major Cluster II with the co-clustering of accessions from all regions (**Figure 7** and **Supplementary Figure 1**). Despite the domination of the Asian region in the major Cluster 1, the Sub-cluster 1 had accessions from all other regions (<10%). Within the major Cluster 2, Sub-cluster 6 had 62.41% accessions from the Asian region, along with 23.31, 12.78, and 1.50% accessions from African, American, and European regions, respectively. The co-clustering of accessions from different regions was predominantly found in Sub-clusters 4 and 5. The Sub-cluster 4 had 57.97, 26.81, and 15.22% accessions from the African, Asian, and American regions, respectively. In Subcluster 5, the Asian, African, American, and European regions contributed 36.96, 19.57, 35.87, and 7.61%, respectively. The genetic similarity/dissimilarity among accessions between and within sub-clusters was determined by inter and intra-cluster distances. The intra-cluster distance identified Sub-cluster 1 (d = 0.136) as the more diverse sub-cluster with maximum intra-cluster distance and Sub-cluster 2 with the least (d = 0.099; **Supplementary Table 13**). Similarly, the maximum inter-cluster distance was observed between Sub-clusters 1 and 6 (d = 0.227), followed by Sub-cluster 1 with Sub-clusters 2 and 4 (d = 0.187). Overall, Sub-cluster 1 had the maximum inter-cluster distance with all other sub-clusters. The least inter-cluster distance was observed between Sub-clusters 2 and 3 and, Subclusters 5 and 6 (d = 0.143).

The distribution of each grain nutrient and its corresponding agronomic performance in each sub-cluster is displayed in the heatmap, with a varying intensity of pink (low) to green (high) color, which characterizes the sub-cluster (**Figure 6**). The mean comparison between sub-clusters revealed that the sub-clusters varied significantly from each other for all agronomic traits and grain nutrients (**Supplementary Table 14**). Days to 50% flowering and days to maturity distinguished 4 out of 6 sub-clusters, with Sub-cluster 1 being the earliest in flowering, followed by Sub-clusters 3 and 5. The Sub-clusters 2, 4, and 6 were insignificantly different from each other for both the traits. Sub-clusters 4 and 5 had high 100-seed weight, whereas Sub-clusters 2, 3, and 6 had high grain yield per plant. Protein, Fe, and Zn distinguished five out of six sub-clusters. For specific nutrient sources, Sub-cluster 2 contained protein-dense accessions (28.15 ± 0.6%), Sub-cluster 3 for Ca (1583.83 ± 194.3 mg/kg), Sub-cluster 1 for Fe (36.62 ± 1.5 mg/kg) and Zn (31.21 ± 1.4 mg/kg). However, the Ca content in the Sub-clusters 1, 2, 4, and 5 was found to be indifferentiable from Sub-clusters 3 and 6. For other nutrients, the nutrient-dense accessions were found in Sub-clusters 1 and 4 for K and Cu, Subcluster 4 for P, and Sub-cluster 2 for Mg and Mn. Overall, high mean for four nutrients (Fe, Zn, K, and Cu) was observed in Sub-cluster 1 and for 3 nutrients in Sub-clusters 2 (Protein, Mg, and Mn) and 4 (P, K, and Cu).

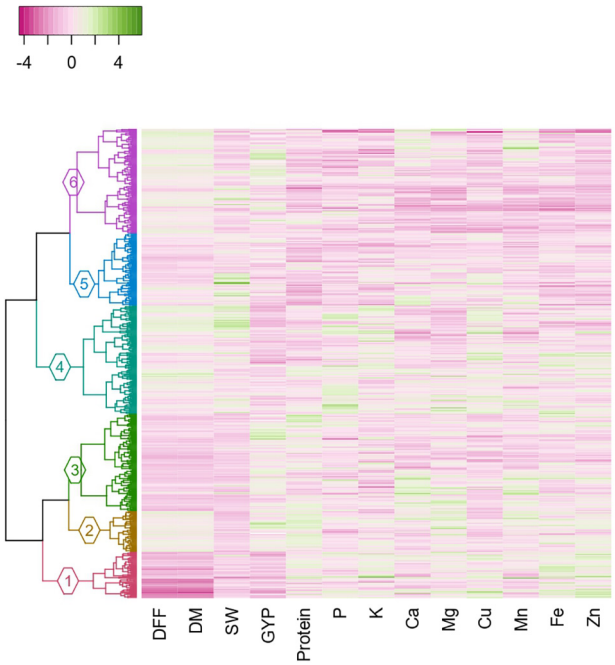


FIGURE 6 | Dendrogram constructed based on the Gower's distance matrix, adopting Ward. D2 clustering method with heatmap depicting the agronomic and grain nutrient content in each accession of the cluster (DFF, days to 50% flowering; DM, days to maturity; SW, 100-seed weight; GYP, grain yield per plant; P, phosphorus; K, potassium; Ca, calcium; Cu, copper; Mg, magnesium; Mn, manganese; Fe, iron; Zn, zinc).

Nutrient-Dense Accessions

Accessions with high nutrient density were identified based on the superiority to the trial mean and the superior check. Among the two checks, check ICP 8863 was found to have better nutrient content with 27.69% protein, 0.44% P, 1.51% K, 1,497.59 mg/kg Ca, 1,630.96 mg/kg Mg, 11.14 mg/kg Cu, 10.62 mg/kg Mn, 38.19 mg/kg Fe, and 32.58 mg/kg Zn. The number of superior accessions was 139 for protein, 107 for P, 171 for K, 290 for Ca, 53 for Mg, 291 for Cu, 197 for Mn, 21 for Fe, and 16 for Zn. The top 10 nutrient-specific accessions covered a range of 28.85–29.50% protein, 38.67–40.98 mg/kg Fe, 32.97–35.68 mg/kg Zn, and 1,923.79–2,099.76 mg/kg Ca. For other nutrients, the ranges were 0.48–0.49% for P, 1.55–1.58% for K, 1,710.14–1,865.65 mg/kg for Mg, 13.08–13.96 mg/kg for Cu, and 11.79–12.51 mg/kg for Mn. The multi-nutrient dense accessions were screened from the top 10 nutrient-specific accessions identified for each nutrient (**Table 3**). Fifteen accessions, representing eight countries and three geographical regions, were identified as superior sources for 3–7 nutrients (**Table 4**). Of these, eight out of 10 accessions in the Asian region are from India. These 15 accessions varied widely for days to 50% flowering and maturity (77–144 and 127–192 days, respectively). Among these, four accessions for 100-seed weight and six accessions for grain yield per plant were superior to the trial mean and check ICP 8863. However, the yield of these accessions (16.54–45.53 g) was not superior to the check ICP 7221 (48.93 g). These 15 accessions belonged to four sub-clusters (sub-clusters 1, 2, 3, and 4). Among the 15 accessions, ICP 7533 was identified as the best source for seven nutrients, followed by accessions ICP 8165, ICP 11485, ICP 12043, and ICP 13757 for six nutrients.

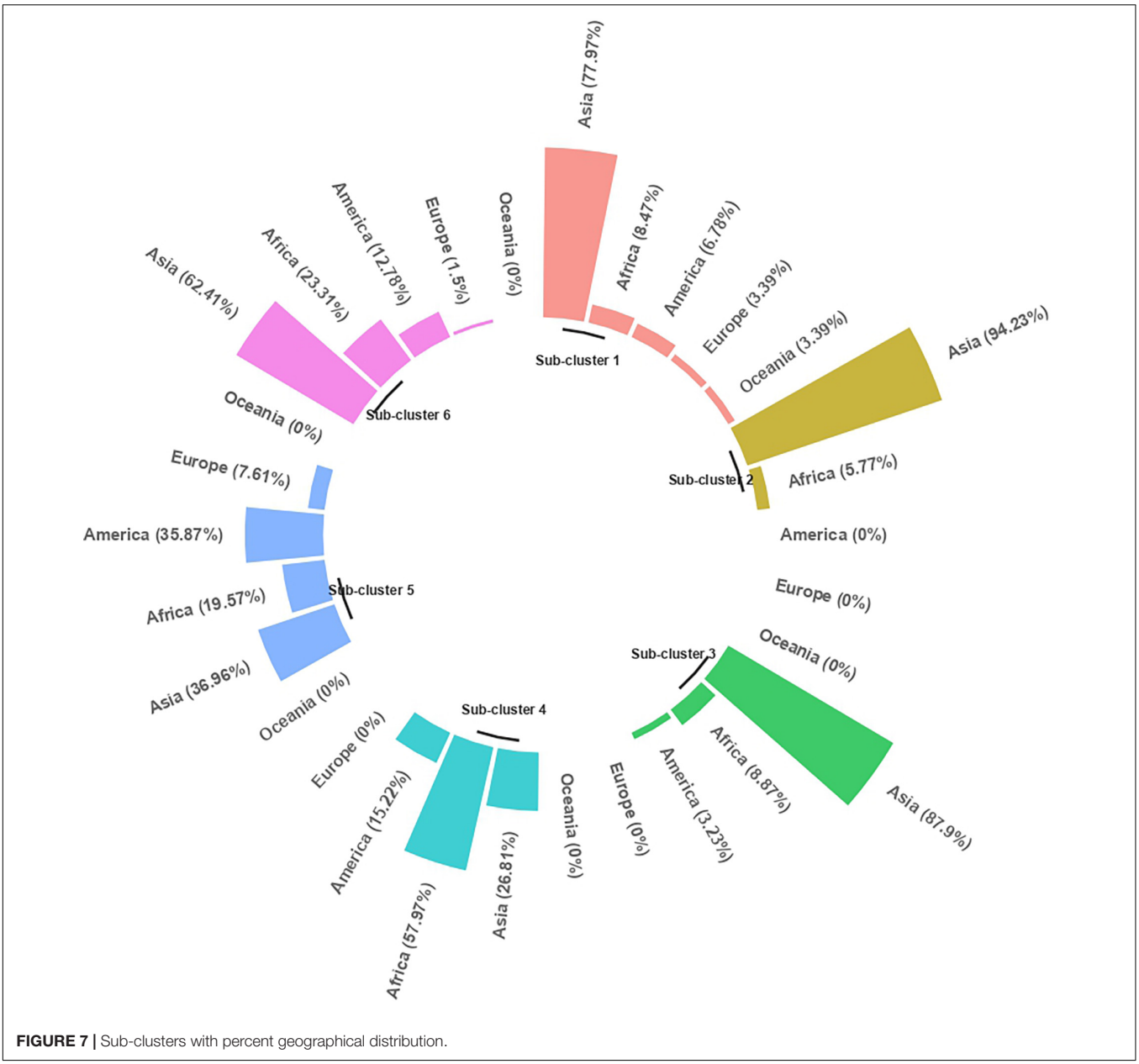


FIGURE 7 | Sub-clusters with percent geographical distribution.

DISCUSSION

Between germplasm availability and its subsequent utilization in crop improvement programs, there exists a huge gap. The attributable reasons are i) a lack of information about the genetic worth of the germplasm, ii) presence of undesirable linkages, difficulties, and expensiveness linked in screening for few elite lines from a vast ocean of germplasm, iii) risk of crossing program failure and the long time scale linked in the development of breeding lines, and iv) the possibility of toxins and allergens introduction into the elite cultivars during introgression (Upadhyaya et al., 2010; Mallikarjuna et al., 2014). Pigeonpea offers an affordable source of protein to the marginalized populations surviving in several developing countries of Asia and Africa. Other than protein, pigeonpea is rich in a few minerals too, and, more interestingly, the accumulation of Fe and Zn in the cotyledons benefits by overcoming the dehulling nutrient loss, which is common in cereals like wheat and rice (Susmitha et al., 2022). Identification of nutrient-rich germplasm can further enrich the breeders' crossing blocks for developing high-yielding and nutrient-rich varieties.

The REML analysis indicated the existence of adequate variability in the germplasm for all agronomic traits and grain nutrients. Other than Ca, the variance attributable to the environment was significant for all the traits, indicating that the extraneous factors contained in the cropping years were different and adequate in differentiating the accessions. The significant

TABLE 3 | Top ten pigeonpea accessions with high trait value identified from 598 pigeonpea accessions evaluated during the 2019 and 2020 rainy seasons at ICRISAT, India.

Trait	Trial mean	Mean+$LSD_{0.05}$	Top 10 accession (ICP#)	Nutrient range	DFF (days)	DM (days)	SW (g)	GYP (g)
Protein (%)	26.99	32.16	6027, 5369, 6165, 15249, 15247, 2860, 5960, 15237, 6219, 8165	28.85–29.50	110–140	159–188	7.85–9.06	29.3–52.70
P (%)	0.43	0.58	12788, 13046, 9185, 11350, 12064, 12043, 11863, 10086, 12828, 12832	0.48–0.49	105–141	155–187	6.76–14.69	21.62–41.64
K (%)	1.50	1.72	11485, 15109, 13315, 11350, 15106, 14109, 8035, 7399, 7028, 13757	1.55–1.58	81–144	135–192	7.18–14.36	19.35–33.79
Ca (mg/kg)	1542.8	2019.63	**7867, 8354**, 15597, 8392, 11472, 1458, 10298, 1514, 774, 10876	1923.79–2099.76	67–142	112–193	6.53–16.01	22.18–43.76
Mg (mg/kg)	1530.2	1748.62	**12043, 15269, 8588, 1514, 7439,** 13553, 7337, 13994, 11991, 12558	1710.14–1865.65	109–159	160–205	6.91–12.46	25.01–44.04
Cu (mg/kg)	11.17	13.26	**7533, 13857, 9146, 13807,** 12928, 13315, 14389, 7028, 15489, 13551	13.08–13.96	77–144	127–193	7.99–20.35	16.54–29.62
Mn (mg/kg)	10.40	12.58	13545, 14574, 12538, 8354, 12174, 9185, 7870, 7982, 11849, 7867	11.79–12.51	103–148	155–198	7.00–16.01	27.27–43.76
Fe (mg/kg)	34.89	42.02	10876, 9123, 13576, 7650, 8165, 1050, 13551, 8392, 8042, 7347	38.67–40.98	79–138	130–188	7.02–9.83	19.56–45.53
Zn (mg/kg)	29.27	34.55	**11485,** 16844, 14600, 7533, 7903, 11350, 8015, 8356, 6219, 13757	32.97–35.68	77–138	127–188	7.19–17.42	16.54–43.07

DFF, days to 50% flowering; DM, days to maturity; SW, 100-seed weight; GYP, grain yield per plant; P, phosphorus; K, potassium; Ca, calcium; Cu, copper; Mg, magnesium; Mn, manganese; Fe, iron; Zn, zinc. Bolded accession number represents superiority to trial mean + $LSD_{0.05}$.

G × E interaction for most of the traits indicated the sensitivity of nutrients accumulation to the environment. This suggests for further evaluation of the germplasm in multiple locations and multiple years to have a better insight into the G × E interaction existing for the traits (Murube et al., 2021) and selection thereafter. Low G × E interaction and moderate-to-high heritability for most of the traits studied suggest a better selection response. The heritability estimates of agronomic traits stay parallel with several studies (Kumara et al., 2013; Rangare et al., 2013; Obala et al., 2018; Shruthi et al., 2019; Sharma et al., 2020), while the estimates for protein content were variable across studies. The attributable reason may be due to the variable number of genotypes and the environment under evaluation (Obala et al., 2018). Wide variability, insensitivity to G × E interaction, and high heritability of Ca identify this nutrient to have stable trait-associated variants in genome-wide association studies (GWAS). The availability of reference genome sequence in pigeonpea (Varshney et al., 2012; Garg et al., 2021) facilitates the application of GWAS to understand the genetic basis of grain nutrient accumulation and to identify candidate genes or genomic regions associated with these nutrients in future studies to breed biofortified pigeonpea cultivars. However, earlier studies pertaining to the association of genomic regions with domestication and agronomic traits were reported (Varshney et al., 2017; Zhao et al., 2020).

Pulses are rich sources of protein, vitamins, and minerals. Combined with relatively low cost and wide access to the poor, pulses are characterized as "poor man's meat" (Malo and Hore, 2020). The variability observed for whole-grain protein in the present study (23.35–29.50%) was higher than the protein content (16.76–26.82%) reported in previous studies (Amarteifio et al., 2002; Sekhon et al., 2017; Obala, 2018; Cheboi et al., 2019; Choi et al., 2020; Jawalekar et al., 2020) and is comparable with the *dhal* protein content of high protein lines (27–29%; Saxena et al., 1987). The protein content in *dhal* is higher than that in the whole grain (Susmitha et al., 2022), signifying that *dhal* nutritional analysis of the superior accessions in this study may still have higher protein than the high protein lines reported by Saxena et al. (1987). This indicated the availability of superior parental sources for protein biofortification. In specific, the protein content of wild species *Cajanus cajanifolius and C. sericeus* (∼29%) was similar to the previous study (Upadhyaya et al., 2013). On par with wild species, few landraces *viz.* ICP 6027, ICP 5369, ICP 15249, ICP 15247, and ICP 6165 had similar protein content (∼29%) and belonged to medium and late maturity groups. These sources from the primary gene pool can make crossing or gene transfer easy compared to those involving the secondary gene pool (Harlan and de Wet, 1971).

The pigeonpea is found to be rich in calcium (Saxena et al., 2010; Susmitha et al., 2022), and the results of this study inferred that the Ca content in pigeonpea (154. 28 mg/100 g) was found to be higher than many staple cereals (7.49–39.36 mg/100 g), such as rice, wheat, maize, pearl millet, sorghum, and barley but lesser to Ca-dense finger millet (364 mg/100 g). Among grain legumes, pigeonpea stands next to soybean (239 mg/100 g) in whole-grain-Ca content (Longvah et al., 2017). Furthermore, a good amount of K (15,000 mg/kg) and Mg (1,530.20 mg/kg) is accumulated in the pigeonpea whole grain, which can reduce the risk of cardiovascular diseases and diabetes when included in the diet (Schulze et al., 2007; Cherbuin, 2016; Stone et al., 2016). In pigeonpea, the Fe and Zn content in cotyledon is indifferentiable from the whole-grain Fe and Zn (Susmitha et al., 2022). This indicates that the Fe and Zn content reported in this study not only represents the whole grain but also the cotyledon. The Fe content in pigeonpea (3.49 mg/100 g) is low when compared to other pulses like chickpea, black gram, horse gram (5.97–8.76 mg/100 g), while the Zn content (2.93 mg/100 g)

TABLE 4 | Promising multi-nutrient dense pigeonpea landraces identified from 598 pigeonpea accessions evaluated during the 2019 and 2020 rainy seasons at ICRISAT, India.

S.No	Accession	Country	Maturity group	Sub-cluster	DFF (days)	DM (days)	SW (g)	GYP (g)	Protein (%)	P (mg/kg)	K (mg/kg)	Ca (mg/kg)	Mg (mg/kg)	Cu (mg/kg)	Mn (mg/kg)	Fe (mg/kg)	Zn (mg/kg)
Asia																	
1	ICP 6219	India	Medium	3	122	171	8.12	**43.07**	**28.95**	**0.45**	1.50	1440.92	1521.33	**11.97**	10.38	36.62	**32.99**
2	ICP 8165	India	Late	2	138	188	9.06	**45.53**	**28.85**	0.43	**1.52**	**1657.89**	1565.65	**11.60**	**11.66**	**39.18**	31.00
3	ICP 7867	India	Medium	3	127	179	**16.01**	**43.76**	26.37	0.40	1.49	**2099.76**	**1654.93**	10.81	**11.79**	33.85	27.84
4	ICP 1514	India	Medium	3	111	160	7.4	**41.25**	26.83	0.42	1.47	**1955.06**	**1750.32**	10.84	**11.68**	34.34	27.80
5	ICP 7028	India	Early	1	81	135	9.5	19.35	27.13	**0.46**	**1.55**	1495.66	1552.74	**13.09**	10.28	35.36	31.14
6	ICP 7533	India	Early	1	77	127	7.99	16.54	27.26	**0.46**	**1.53**	**1672.91**	**1614.33**	**13.96**	**10.90**	38.17	**33.76**
7	ICP 8354	India	Medium	3	104	155	8.32	**35.13**	27.00	0.41	1.49	**2049.67**	1567.97	**11.23**	**12.08**	37.08	31.45
8	ICP 8392	India	Medium	3	111	163	7.74	**41.57**	**28.03**	0.42	1.51	**1968.65**	1523.76	**11.59**	**11.61**	**39.00**	32.54
9	ICP 11350	Nepal	Late	4	138	188	9.2	28.38	**28.65**	**0.48**	**1.56**	1293.94	1558.29	**11.98**	9.89	35.79	**33.59**
10	ICP 11485	Thailand	Medium	1	103	156	7.37	20.34	**28.63**	**0.47**	**1.58**	1188.95	1505.61	**12.35**	9.35	**38.57**	**35.68**
Africa																	
11	ICP 9185	Kenya	Medium	4	129	179	**12.91**	27.27	26.89	**0.49**	1.51	**1716.81**	1587.93	**11.69**	**12.05**	35.21	31.23
12	ICP 12043	Tanzania	Late	4	134	187	**10.89**	28.58	**28.13**	**0.48**	1.51	**1751.98**	**1865.65**	**11.37**	**11.74**	33.58	29.89
13	ICP 13315	Rwanda	Late	4	144	192	**11.48**	23.23	**28.66**	**0.46**	**1.57**	1205.61	1439.64	**13.14**	9.19	34.95	30.12
America																	
14	ICP 13757	Trinidad and Tobago	Medium	4	129	179	10.07	30.64	**28.04**	**0.45**	**1.55**	**1594.69**	1504.33	**12.39**	10.24	35.90	**32.97**
15	ICP 13551	Antigua and Barbuda	Medium	4	126	177	8.71	19.56	27.08	**0.44**	1.51	1448.73	1599.75	**13.08**	9.58	**39.08**	**32.87**
Minimum					77	127	7.37	16.54	26.37	0.40	1.47	1188.95	1439.64	10.81	9.19	33.58	27.8
Maximum					144	192	16.01	45.53	28.95	0.49	1.58	2099.76	1865.65	13.96	12.08	39.18	35.68

DFF, days to 50% flowering; DM, days to maturity; SW, 100-seed weight; GYP, grain yield per plant; P, phosphorus; K, potassium; Ca, calcium; Cu, copper; Mg, magnesium; Mn, manganese; Fe, iron; Zn, zinc. Bold values indicate the superiority over the trial mean and superior check.

is comparable with these pulses (2.71–3.37 mg/100 g; Longvah et al., 2017). This necessitates their subsequent improvement through intra or inter-specific hybridization. To enhance the variability for Fe and Zn in the primary gene pool, Sharma et al. (2020) attempted interspecific crosses with *Cajanus platycarpus*. Despite this, a good response to agronomic biofortification for Fe and Zn was reported in pigeonpea (Gopalakrishnan et al., 2016; Hanumanthappa et al., 2018; Behera et al., 2020). However, Upadhyaya et al. (2010) identified 14 high Zn accessions from core and mini-core collections of pigeonpea available in Genebank at ICRISAT, India. Furthermore, two accessions for Ca (2,049.67–2,099.76 mg/kg), four accessions for Mg (1,750.32–1,865.65 mg/kg), five accessions for Cu (13.34–14.20 mg/kg), and one accession for Zn (35.68 mg/kg), with significantly higher nutrient content than the trial mean identified in this study, enlighten the presence of potential germplasm for mineral biofortification in the ICRISAT Genebank.

The nutrients among themselves were positively correlated with one another, thus facilitating combined multi-nutrient biofortification. The protein improvement in pigeonpea is favored by selection for nutrients, namely, P, K, Mg, Mn, Fe, and Zn. The nutrients, Fe and Zn, are highly positively correlated with each other, and, hence, their improvement together stays significant. This correlation existed across several legumes, such as pigeonpea (Mishra and Acharya, 2018; Sharma et al., 2020), common bean (Celmeli and Sari, 2018), cowpea (Dakora and Belane, 2019), and green gram (Singh et al., 2018). Furthermore, the positive correlation of Fe with all other nutrients offers opportunities for reciprocal nutrient improvement. For Ca improvement, the selection can be done for Mg, Mn, and Fe or against K. This relation stays analogous to the results of Sharma et al. (2020) for the association of Ca with Fe and Mg and Gerrano et al. (2019) for Ca with K.

In recent years, extensive research has been carried out to develop more super-early, extra-early, and early types as photo insensitivity is directly related to earliness, which can break an adaptation barrier and help in the introduction of the crop in new niches and can diversify traditional cereal-based cropping systems (Saxena et al., 2018, 2019). The variability for days to maturity identified the presence of extra-early accession (ICP 15597), a released cultivar (MN1), which has been exploited for breeding high-yielding super-early varieties (Srivastava et al., 2012). Interestingly, the days to 50% flowering and the days to maturity were found to be negatively associated with protein, Mg, Cu, Fe, and Zn, which complements the development of early lines with high nutrient content. However, the pigeonpea cultivated worldwide belongs to medium and late-maturity groups. The presence of indifferentiable Ca and protein content across different maturity groups stands as an advantage for improving Ca and protein in different maturity groups, which can fit into different cropping systems across the globe. Furthermore, Zn exhibited a non-significant association with grain yield per plant among the early- and late-maturity groups, which is of great significance in promoting food security and overcoming Zn deficiency worldwide.

The choice of pigeonpea varieties cultivated across different geographical regions is decided by the market value and/or consumer preference. The seed size defines the consumer preference, and the most preferred seed size in Indian market is 10–14 g/100 seeds (Varshney et al., 2017), whereas, in African and the Caribbean regions, it is about 18 g/100 seeds (Saxena et al., 1987). The mean 100-seed weight (10.10 g) indicated that most of the accessions were distributed around the mean, which is preferable in the Indian market. Forty-two accessions recorded more than 15 g per 100 seeds, of which African and American regions alone contributed 19 and 14 accessions, respectively, reflecting their seed size preferences. The correlation analysis revealed that the nutrient improvement (protein, Ca, Fe, and Zn) in pigeonpea is favored by selection for a small seed size (less 100-seed weight). This can be related to most of the wild species, with small seed size having high nutrient content in pigeonpea. The region-based correlation analysis revealed that protein and Ca improvement in the African region is unaffected by 100-seed weight. Furthermore, the100-seed weight does not affect the improvement of Ca and Zn beyond 10 g, Fe up to 15 g, and protein beyond 15 g. Similar to this, an earlier report on variable association of protein with 100-seed weight in different intergeneric crosses was reported by Saxena et al. (1987).

The yield of majority of the staple crops was stagnated and/or unable to meet global demand. For further genetic improvement, variability for the trait is essential. The grain yield per plant recorded good variability and inhibited a positive correlation with protein and Mg, and a non-significant association with Ca. These useful correlations can be utilized in enhancing the nutrient content along with yield, which can promote combined food and nutritional security. However, the high coefficient of variation observed for the trait is attributed toward the variable number of plants across accessions.

Trading played a key role in the introduction of landraces from India to Africa (Hillocks et al., 2000) and from Africa to America (Van Der Maesen, 1980), which created the possibility for the existence of allochthonous landraces in these regions, which, over time, might have crossed with autochthonous landraces of the region and evolved as autochthonous landraces, sharing some common features between regions (Zeven, 1998), leading to the co-clustering of accessions from different regions within a cluster. Geographical diversity combined with high nutrient density in the Sub-clusters 1, 2, and 4 can provide a valuable parental source for introducing new variability in the primary gene pool of pigeonpea for grain nutrient improvement in different regions. Furthermore, the 10-trait specific and 15 multi-nutrient dense accessions identified based on the *per se* performance and superiority to the nutrient dense check belong to different geographical regions and exhibited wide variation for agronomic traits. These germplasms can be utilized to improve the grain nutrient content under different seed sizes and maturity categories. Furthermore, the pigeonpea breeding community across the globe can get access to the limited quantity of the seed through the Standard Material Transfer Agreement.

CONCLUSION

The study revealed the presence of considerable variability and moderate-to-high heritability for the agronomic traits and grain nutrients in the primary gene pool of pigeonpea germplasm.

The distribution and the association of grain nutrients among themselves and with agronomic traits were variable across the geographical region and maturity groups, which could benefit the breeders in identification of region- and maturity-group-specific sources and associations, respectively, which can eliminate the risk of acclimatization in the newly breed cultivars. The trait-specific sources identified for grain nutrients content can provide a new parental base in the biofortification program for the development of nutrient-dense cultivars in a desirable agronomic background that can promote food and nutritional security. However, with the available low-cost sequencing technology, genotyping of the 600 accessions in the future, combined with the large phenotypic data generated in this study, can serve as a valuable raw material for conducting SNP/haplotype-based GWAS to identify genetic variants associated with the nutrients that can accelerate genetic gains in pigeonpea biofortification.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

KS, RS, and MV to the conception and design of the study (this work is part of DS's Ph. D. thesis research). TK, VA, and KS supported student research as supervisors. OP provided resources (seed material) for the study. PC and CN performed laboratory analysis. SR supported in data collection and VNA in data documentation. DS, RS, MV, TK, and KS curated the data and performed the formal data analysis. SM, PJ, DS, MV, and BA did data validation and visualization. DS, MV, and KS were involved in writing the original draft, reviewing, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was undertaken as a part of the CGIAR Genebank Platform coordinated by Crop Trust, and the CGIAR Research Program on Grain Legumes and Dryland Cereals.

ACKNOWLEDGMENTS

We are thankful to the scientific officer, Naresh Reddy, and the research technicians, Judson Babu and Ravi, working at Genebank, ICRISAT, for their technical support; also, they duly acknowledge Farm & Engineering Services, ICRISAT, for sharing the soil properties of the experimental fields. The research scholar is grateful to ICRISAT for providing a stipend for the entire study program at ICRISAT and duly acknowledges Learning Systems Unit, ICRISAT for facilitation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.934296/full#supplementary-material

Supplementary Figure 1 | Dendrogram constructed based on Gower's distance using Ward. D2 clustering algorithm for 598 pigeonpea accessions.

Supplementary Table 1 | Metadata of the 600 pigeonpea accessions and four checks evaluated for agronomic traits and grain nutrients during the 2019 and 2020 rainy seasons at ICRISAT, India.

Supplementary Table 2 | Correlation among agronomic traits and grain nutrients of 598 pigeonpea accessions evaluated during the 2019 and 2020 rainy seasons at ICRISAT, India.

Supplementary Table 3 | Correlation between agronomic traits and grain nutrients for accessions from Asian region.

Supplementary Table 4 | Correlation between agronomic traits and grain nutrients for accessions from African region.

Supplementary Table 5 | Correlation between agronomic traits and grain nutrients for accessions from American region.

Supplementary Table 6 | Correlation between agronomic traits and grain nutrients for accessions from the early maturity group.

Supplementary Table 7 | Correlation between agronomic traits and grain nutrients for accessions from the medium maturity group.

Supplementary Table 8 | Correlation between agronomic traits and grain nutrients for accessions from the late maturity group.

Supplementary Table 9 | Correlation between agronomic traits and grain nutrients for accessions with $\leq$ 10 g 100-seed weight.

Supplementary Table 10 | Correlation between agronomic traits and grain nutrients for accessions with 10-15 g 100-seed weight.

Supplementary Table 11 | Correlation between agronomic traits and grain nutrients for accessions with $>$ 15 g 100-seed weight.

Supplementary Table 12 | Contribution of accessions from different regions to each sub-cluster.

Supplementary Table 13 | Within and between cluster distances based on Gower's phenotypic distance matrix for 598 pigeonpea accessions evaluated during the 2019 and 2020 rainy seasons at ICRISAT, India.

Supplementary Table 14 | Mean and range comparison for agronomic traits and grain nutrients between six sub-clusters of 598 pigeonpea accessions evaluated during the 2019 and 2020 rainy seasons at ICRISAT, India.

REFERENCES

Ahnen, R. T., Jonnalagadda, S. S., and Slavin, J. L. (2019). Role of plant protein in nutrition, wellness, and health. *Nutr. Rev.* 17, 735–747. doi: 10.1093/nutrit/nuz028

Aiking, H., and De Boer, J. (2020). The next protein transition. *Trends Food Sci. Technol.* 105, 515–522. doi: 10.1016/j.tifs.2018.07.008

Amarteifio, J. O., Munthali, D. C., Karikari, S. K., and Morake, T. K. (2002). The composition of pigeon peas (*Cajanus cajan* (L .) Millsp.) grown in Botswana. *Plant Foods Hum. Nutr.* 57, 173–177. doi: 10.1023/a:1015248326920

Andreoli, V., Bagliani, M., Corsi, A., and Frontuto, V. (2021). Drivers of protein consumption : a cross-country analysis. *Sustainability* 13:7399. doi: 10.3390/su13137399

Bailey, L., and West, K. P. Jr. (2015). deficiencies often occur as part of an intergenerational cycle The Epidemiology of Global Micronutrient Deficiencies The Epidemiology of Global Micronutrient. *Ann. Nutr. Metab.* 66, 22–33. doi: 10.1159/000371618

Behera, S. K., Shukla, A. K., Tiwari, P. K., Tripathi, A., Singh, P., Trivedi, V., et al. (2020). Classification of pigeonpea (*Cajanus cajan* (L.) Millsp.) genotypes for zinc efficiency. *Plants* 9:952. doi: 10.3390/plants9080952

Celmeli, T., and Sari, H. (2018). The nutritional content of common bean (*Phaseolus vulgaris* L .) landraces in comparison to modern varieties. *Agronomy* 8:166. doi: 10.3390/agronomy8090166

Cheboi, J. J., Kinyua, M. G., Kimurto, P. K., Kiplagat, O. K., Ng, F., and Sita, R. (2019). Biochemical composition of pigeonpea genotypes in Kenya. *Aust. J. Crop Sci.* 13, 1848–1855. doi: 10.21475/ajcs.19.13.11.p1886

Cherbuin, N. (2016). *Dietary Mineral Intake (Magnesium, Calcium, and Potassium) and the Biological Processes of Aging*. Amsterdam: Elsevier Inc. doi: 10.1016/B978-0-12-801816-3.00038-8

Choi, Y., Hyun, D. Y., Lee, S., Yoon, H., Lee, M., Oh, S., et al. (2020). Agricultural characters, phenolic and nutritional contents, and antioxidant activities of Pigeon Pea (*Cajanus cajan*) germplasms cultivated in the Republic of Korea. *Korean J. Plant Reour.* 33, 50–61. doi: 10.7732/kjpr.2020.33.1.50

Dahiya, B. S., Brar, J. S., and Bhullar, B. S. (1977). Inheritance of protein content and its correlation with grain yield in pigeonpea (*Cajanus cajan* (L.) Millsp.). *Plant Food Hum. Nutr.* 27, 327–334. doi: 10.1007/BF01092325

Dakora, F. D., and Belane, A. K. (2019). Evaluation of protein and micronutrient levels in edible cowpea (*Vigna unguiculata* L. Walp.). *Front. Sustain. Food Syst.* 3:70. doi: 10.3389/fsufs.2019.00070

de Mendiburu, F. (2021). *Agricolae: Statistical Procedures for Agricultural Research*. Available online at: https://cran.r-project.org/package=agricolae (accessed 9 Feburary 9, 2022).

FAOSTAT (2021). *Crops and Livestock Products*. Available online at: https://www.fao.org/faostat/en/#data/QCL (accessed January 25, 2022).

Food and Agriculture Organization [FAO] (2014). *Understanding the True Cost of Malnutrition*. Available online at: https://www.fao.org/zhc/detail-events/en/c/238389/ (accessed January 25, 2022).

Fox, J., and Weisberg, S. (2019). *An {R} Companion to Applied Regression*, 3rd Edn. Available online at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/ (accessed February 9, 2022).

Garg, V., Dudchenko, O., Wang, J., Khan, A. W., Gupta, S., Kaur, P., et al. (2021). Chromosome-length genome assemblies of six legume species provideinsights into genome organization, evolution, and agronomic traits forcrop improvement. *J. Adv. Res.* (in press). doi: 10.1016/j.jare.2021.10.009

Gerrano, A. S., Jansen van Rensburg, W. S., Venter, S. L., Shargie, N. G., Amelework, B. A., Shimelis, H. A., et al. (2019). Selection of cowpea genotypes based on grain mineral and total protein content. *Acta Agric. Scand. Sect. B Soil Plant Sci.* 69, 155–166. doi: 10.1080/09064710.2018.1520290

Gopalakrishnan, S., Vadlamudi, S., Samineni, S., and Kumar, C. V. S. (2016). Plant growth - promotion and biofortification of chickpea and pigeonpea through inoculation of biocontrol potential bacteria, isolated from organic soils. *Springerplus* 5:1882. doi: 10.1186/s40064-016-3590-6

Hanumanthappa, D., Maruthi, S. N., and Shakuntala, J. B. (2018). Enrichment of iron and zinc content in pigeonpea genotypes through agronomic biofortification to mitigate malnutrition. *Int. J. Curr. Microbiol. Appl. Sci.* 7, 4334–4342.

Harinder, K., Kaur, B., and Sharma, S. (1999). Studies on the baking properties of wheat: pigeonpea flour blends. *Plant Foods Hum. Nutr.* 54, 217–226. doi: 10.1023/A:1008158208549

Harlan, J. R., and de Wet, J. M. J. (1971). Toward a rational classification of cultivated plants. *Int. Assoc. Plant Taxon.* 20, 509–517. doi: 10.2307/1218252

Hillocks, R. J., Minja, E., Mwaga, A., Nahdy, M. S., and Subrahmanyam, P. (2000). Diseases and pests of pigeonpea in eas tern Africa : a review. *Int. J. Pest Manag.* 46, 7–18.

ICMR-NIN (2020). *Nutrient Requirements for Indians*. ICMR-NIN. Available online at: https://www.metabolichealthdigest.com/nutrient-requirements-for-indians-icmr-nin-2020/ (accessed February 9, 2022).

Jawalekar, K. V., Baishya, S., Rathi, S., and Gade, K. (2020). Assessment of nutritional composition in prominent *Cajanus cajan* germplasm in India assessment of nutritional composition in prominent *Cajanus cajan* Germplasm in India B-complex. *Int. J. Curr. Microbiol. Appl. Sci.* 9, 387–397. doi: 10.20546/ijcmas.2020.912.xx

Johnson, H. W., Robinson, H. F., and Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybean. *Agron. J.* 47, 314–318.

Keuls, M. (1952). The use of the "studentized range" in connection with an analysis of variance. *Euphytica* 1, 112–122. doi: 10.1007/BF01908269

Kouris-Blazos, A., and Belski, R. (2016). Health benefits of legumes and pulses with a focus on Australian sweet lupins. *Asia Pac. J. Clin. Nutr.* 25, 1–17. doi: 10.6133/apjcn.2016.25.1.23

Kumara, N., Santoshagowda, G. B., Nishanth, G. K., and Dharmaraj, P. S. (2013). Genetic diversity, variability and correlation studies in advanced genotypes of Genetic diversity, variability and correlation studies in advanced genotypes of Pigeonpea [*Cajanus cajan* (L.) Millsp .]. *Acta Biol. Indica* 2, 406–411.

Levene, H. (1960). "Robust tests for equality of variances," in *Contributions to Probability and Statictics:Essays in Honorof Harold Hotelling*, eds I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, and H. Mann (Palo Alto, CA: Stanford University Press), 278–292.

Longvah, T., Ananthan, R., Bhaskarachary, K., and Venkaiah, K. (2017). *Indian Food Compostion Tables*. Hyderabad: National Institute of Nutrition.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2021). *Cluster: Cluster Analysis Basics and Extensions*. Available online at: https://cran.r-project.org/package=cluster (accessed February 15, 2022).

Mallikarjuna, N., Srikanth, S., Kumar, C. V. S., Srivastava, R. K., Saxena, R. K., and Varshney, R. K. (2014). "Pigeonpea," in *Broadening the Genetic Base of Grain Legumes*, eds M. Singh, M. Singh, I. Bisht, and M. Dutta (New Delhi: Springer), 149–159. doi: 10.1007/978-81-322-2023-7

Malo, M., and Hore, J. (2020). "Pulse Production in India: major Constraints and Way Forward," in *Research Trends in Multidisciplinary Research and Development*, ed. S. Fedorov (Zittau: Weser Books), 35–63. doi: 10.3389/fpls.2022.861191

Mishra, S., and Acharya, S. (2018). Targeted Traits for Enhancement of Seed Iron and Zinc Concentrations in Pigeonpea. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* 88, 1199–1205. doi: 10.1007/s40011-017-0857-9

Mula, M., and Saxena, K. (2010). *Lifting the Level of Awareness on Pigeonpea – A Global Perspective*. Patancheru: ICRISAT.

Murtagh, F. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Fionn. J. Classif.* 31, 274–295. doi: 10.1007/s00357-014-9161-z

Murube, E., Beleggia, R., Pacetti, D., Nartea, A., Frascarelli, G., Lanzavecchia, G., et al. (2021). Characterization of nutritional quality traits of a common bean germplasm collection. *Foods* 10:1572. doi: 10.3390/foods10071572

Naghshi, S., Sadeghi, O., Willett, W. C., and Esmaillzadeh, A. (2020). Dietary intake of total, animal, and plant proteins and risk of all cause, cardiovascular, and cancer mortality : systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* 370:m2412. doi: 10.1136/bmj.m2412

Newman, D. (1939). The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika* 31:20. doi: 10.2307/2334973

Obala, J. (2018). Development of sequence-based markers for seed protein content in pigeonpea. *Mol. Genet. Genomics* 294, 57–68. doi: 10.1007/s00438-018-1484-8

Obala, J., Saxena, R. K., Singh, V. K., Vechalapu, S., Das, R., Sameer-kumar, C. V., et al. (2018). Genetic variation and relationships of total seed protein content with some agronomic traits in pigeonpea (*Cajanus cajan* (L .) Millsp .). *Aust. J. Crop Sci.* 12, 1859–1865. doi: 10.21475/ajcs.18.12.12.p1138

Ohizua, E. R., Adeola, A. A., Idowu, M. A., Sobukola, O. P., Afolabi, T. A., Ishola, R. O., et al. (2017). Nutrient composition, functional, and pasting properties of unripe cooking banana, pigeon pea, and sweetpotato flour blends. *Food Sci. Nutr.* 5, 750–762. doi: 10.1002/fsn3.455

Oksanen, F. J., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, R. P., et al. (2020). *vegan: Community Ecology Package*. Available online at: https://cranr-project.org/package=vegan%0A (accessed February 15, 2022).

Oshodi, A. A., Plaofe, O., and Hall, G. (1993). Amino acid, fatty acid and mineral composition of pigeon pea. *Int. J. Food Sci. Nutr.* 43, 187–191.

Pele, G. I., Oladiti, E. O., Bamidele, P. O., and Fadipe, E. A. (2016). Influence of processing techniques on the nutritional and anti-nutritional properties of pigeon pea (*Cajanus cajan*). *Int. J. Eng. Appl. Sci.* 3, 92–94.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rangare, N. R., Reddy, E., and Ramesh Kumar, S. (2013). Study of heritability, genetic advance and variability for yield contributing characters in pigeonpea (*Cajanus cajan* L. Millspaugh). *Trends Biosci.* 6, 660–662.

Reddy, L. J. (1990). "Pigeonpea: morphology," in *The Pigeonpea*, eds Y. L. Nene, S. D. Hall, and V. K. Sheila (Wallingford: C.A.B International), 47–87.

Reddy, L. J., Singh, J. M., Bisen, S. S., and Jambunathan, R. (1979). "Seed protein studies on *Cajanus cajan*, *Atylosia* spp. and so hybrid derivatives," in *Proceedings of the International Symposium on Seed Protein improvement in Cereals and Grain Legumes* (Vienna: IAEA), 105.

Safiri, S., Kolahi, A. A., Noori, M., Nejadghaderi, S. A., Karamzad, N., Bragazzi, N. L., et al. (2021). Burden of anemia and its underlying causes in 204 countries and territories, 1990 – 2019 : results from the Global Burden of Disease Study 2019. *J. Hematol. Oncol.* 14:185. doi: 10.1186/s13045-021-01202-2

Sahrawat, K. L., Kumar, G. R., and Murthy, K. V. S. (2002). Sulfuric acid-Selenium digestion for multi-element analysis in a single plant digest. *Commun. Soil Sci. Plant Anal.* 33, 3757–3765. doi: 10.1081/CSS-120015920

Saxena, K. B., Choudhary, A. K., Saxena, R. K., and Varshney, R. K. (2018). Breeding pigeonpea cultivars for intercropping: synthesis and strategies. *Breed. Sci.* 68, 159–167. doi: 10.1270/jsbbs.17105

Saxena, K. B., Faris, D. G., Singh, U., and Kumar, R. V. (1987). Relationship between seed size and protein content in newly developed high protein lines of pigeonpea. *Qual. Plant. Plant Foods Hum. Nutr.* 36, 335–340. doi: 10.1007/BF01892354

Saxena, K. B., Kumar, R. V., and Sultana, R. (2010). Quality nutrition through pigeonpea — a review. *Health (Irvine Calif.)* 2, 1335–1344. doi: 10.4236/health.2010.211199

Saxena, K., Choudhary, A. K., Srivastava, R. K., Bohra, A., and Saxena, R. K. (2019). Origin of early maturing pigeonpea germplasm and its impact on adaptation and cropping systems. *Plant Breed.* 138, 243–251. doi: 10.1111/pbr.12696

Schönfeld, P., and Werner, H. J. (1986). "Ökonomische progress," in *Entscheidungsund Gleichgewichts-Modelle*, ed. W. Krelle (Weinheim: VCH Verlagsgesellschaft), 251–262.

Schulze, M. B., Schulz, M., Heidemann, C., Schienkiewitz, A., Hoffmann, K., and Boeing, H. (2007). Fiber and magnesium intake and incidence of type 2 diabetes: a prospective study and meta-analysis. *Arch. Intern. Med.* 167, 956–965. doi: 10.1001/archinte.167.9.956

Sekhon, J., Grewal, S. K., Singh, I., and Kaur, J. (2017). Evaluation of nutritional quality and antioxidant potential of pigeonpea genotypes. *J. Food Sci. Technol.* 54, 3598–3611. doi: 10.1007/s13197-017-2818-y

Sharma, S., Paul, P. J., Kumar, C. V. S., and Nimje, C. (2020). Utilizing wild *Cajanus platycarpus*, a tertiary genepool species for enriching variability in the primary genepool for pigeonpea improvement. *Front. Plant Sci.* 11:1055. doi: 10.3389/fpls.2020.01055

Shruthi, H. B., Hingane, A. J., Sekhar, M. R., Kumar, C. V. S., and Srivarsha, J. (2019). Genetic variability for yield, physiological and quality traits in novel super-early pigeonpea (*Cajanus cajan* (L.) Millsp.). *India J. Pure Appl. Biosci.* 7, 378–385. doi: 10.18782/2582-2845.7941

Siddiqui, F., Salam, R. A., Lassi, Z. S., and Das, J. K. (2020). The intertwined relationship between malnutrition and poverty. *Front. Public Heal.* 8:453. doi: 10.3389/fpubh.2020.00453

Singh, R., Van Heusden, A. W., Kumar, R., and Visser, R. G. F. (2018). Genetic variation and correlation studies between micronutrient (Fe and Zn), protein content and yield attributing traits in mungbean (*Vigna radiata* L .). *Legum. Res.* 41, 167–174. doi: 10.18805/lr.v0i0.7843

Singh, U., Jain, K. C., Jambunathan, R., and Faris, D. G. (1984). Nutritional quality of vegetable pigeonpeas [*Cajanus cajan* (L.) Millsp.]: dry matter accumulation, carbohydrates and proteins. *J. Food Sci.* 49, 799–802.

Singh, U., Jambunathan, R., Saxena, K., and Subrahmanyam, N. (1990). Nutritional quality evaluation of newly developed high-protein genotypes of pigeonpea (*Cajanus cajan*). *J Sci Food Agric* 50, 201–209.

Song, M., Fung, T. T., Hu, F. B., Walter, C., Longo, V., Chan, A. T., et al. (2017). Animal and plant protein intake and all-cause and cause-specific mortality: results from two prospective US cohort studies. *AMA Intern. Med.* 176, 1453–1463. doi: 10.1001/jamainternmed.2016.4182. Animal

South, A. (2011). *rworldmap: A New R Package for Mapping Global Data. R J*, Vol. 3. 35–43. Available online at: http://journal.r-project.org/archive/2011-1/RJournal_2011-1_South.pdf (accessed April 13, 2022).

Srivastava, R. K., Vales, M. I., Sultana, R., and Saxena, K. (2012). Development of 'super-early' pigeonpeas with good yield potential from early development of 'super-early' pigeonpeas with good yield potential from early × early crosses. *SAT Agric. Res.* 10:6.

Stagnari, F., Maggio, A., Galieni, A., and Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: an overview. *Chem. Biol. Technol. Agric.* 4:2. doi: 10.1186/s40538-016-0085-1

Stone, M. S., Martyn, L., and Weaver, C. M. (2016). Potassium intake, bioavailability, hypertension, and glucose control. *Nutrients* 8:444. doi: 10.3390/nu8070444

Susmitha, D., Kalaimagal, T., Senthil, R., Vetriventhan, M., Anitha, S., Manonmani, S., et al. (2022). Calcium-Rich pigeonpea seed coat : a potential byproduct for food and pharmaceutical industries. *Sustainability* 14:4918. doi: 10.3390/su14094918

Upadhyaya, H. D., Reddy, K. N., Gowda, C. L., and Singh, S. (2007). Phenotypic diversity in the pigeonpea (*Cajanus cajan*) core collection phenotypic diversity in the pigeonpea (*Cajanus cajan*) core collection. *Genet. Resour. Crop Evol.* 54, 1167–1184. doi: 10.1007/s10722-006-9008-5

Upadhyaya, H. D., Reddy, K. N., Singh, S., and Gowda, C. L. L. (2013). Phenotypic diversity in *Cajanus* species and identification of promising sources for agronomic traits and seed protein content. *Genet. Resour. Crop Evol.* 60, 639–659. doi: 10.1007/s10722-012-9864-0

Upadhyaya, H. D., Yadav, D., Dronavalli, N., Gowda, C. L. L., and Singh, S. (2010). Mini core germplasm collections for infusing genetic diversity in plant breeding programs. *Electron. J. Plant Breed.* 1, 1294–1309.

Van Der Maesen, L. J. G. (1980). "India is the native of the pigeonpea," in *Liber Gratulatorius in nonerem H.C.D. de Wit*, eds J. C. Arends, G. Boelema, C. T. Groot, and A. J. M. Leeuwenberg (Wageningen: Wageningen Agricultural University), 257–262.

Varshney, R. K., Chen, W., Li, Y., Bharti, A. K., Saxena, R. K., Schlueter, J. A., et al. (2012). Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30, 83–89. doi: 10.1038/nbt.2022

Varshney, R. K., Saxena, R. K., Upadhyaya, H. D., Khan, A. W., Yu, Y., Kim, C., et al. (2017). Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nat. Genet.* 49, 1082–1088. doi: 10.1038/ng.3872

VSN International (2019). *Genstat for Windows*, 19th Edn. Available online at: Genstat.co.uk (accessed January 20, 2022).

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Am. Math. Soc.* 54, 426–482.

Wangari, C., Mwema, C., Siambi, M., Silim, S., Ubwe, R., Malesi, K., et al. (2020). Changing perception through a participatory approach by involving adolescent school children in evaluating smart food dishes in school feeding programs–real-time experience from Central and Northern Tanzania. *Ecol. Food Nutr.* 59, 472–485. doi: 10.1080/03670244.2020.1745788

Wei, T., and Simko, V. (2021). *R package "corrplot": Visualization of a Correlation Matrix*. Available online at: https://github.com/taiyun/corrplot (accessed February 9, 2022).

Wheal, M. S., Fowles, T. O., and Palmer, L. T. (2011). A cost-effective acid digestion method using closed polypropylene tubes for inductively coupled plasma optical emission spectrometry (ICP-OES) analysis of plant essential elements. *Anal. Methods* 3, 2854–2863. doi: 10.1039/c1ay05430a

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Available online at: https://ggplot2.tidyverse.org (accessed February 25, 2022).

World Health Organization [WHO] (2021). *Malnutrition*. Available online at: https://www.who.int/news-room/fact-sheets/detail/malnutrition (accessed January 25, 2022).

Zeven, A. C. (1998). Landraces : a review of definitions and classifications. *Euphytica* 104, 127–139.

Zhao, J., Bayer, P. E., Ruperao, P., Saxena, R. K., Khan, A. W., Golicz, A. A., et al. (2020). Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnol. J.* 18, 1946–1954. doi: 10.1111/pbi.13354

Zhao, S., Yin, L., Guo, Y., and Shyr, Y. (2021). *heatmap3: An Improved Heatmap Package*. Available online at: https://cran.r-project.org/package=heatmap3 (accessed April 29, 2022).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Susmitha, Kalaimagal, Senthil, Vetriventhan, Manonmani, Jeyakumar, Anita, Reddymalla, Choudhari, Nimje, Peerzada, Arveti, Azevedo and Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 August 2022
DOI 10.3389/fpls.2022.933277

Check for updates

OPEN ACCESS

EDITED BY
Jaroslaw Czubinski,
Poznan University of Life Sciences, Poland

REVIEWED BY
Aamir Raina,
Aligarh Muslim University, India
Karansher Singh Sandhu,
Bayer Crop Science, United States

*CORRESPONDENCE
Shweta Kalve
shweta.kalve@usask.ca

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 30 April 2022
ACCEPTED 17 June 2022
PUBLISHED 19 August 2022

CITATION
Kalve S, Gali KK and Tar'an B (2022) Genome-wide association analysis of stress tolerance indices in an interspecific population of chickpea.
Front. Plant Sci. 13:933277.
doi: 10.3389/fpls.2022.933277

COPYRIGHT
© 2022 Kalve, Gali and Tar'an. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genome-wide association analysis of stress tolerance indices in an interspecific population of chickpea

Shweta Kalve*, Krishna Kishore Gali and Bunyamin Tar'an

Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada

Chickpea is a cool season crop that is highly vulnerable to abiotic stresses such as heat and drought. High temperature during early flowering and pod development stages significantly reduces the crop yield. The wild relatives of chickpeas can be potential donors for the introgression of heat and drought tolerance into cultivated chickpeas for crop improvement. Initially, 600 interspecific lines were derived from crosses between two elite cultivars, CDC Leader (kabuli chickpea) and CDC Consul (desi chickpea), and 20 accessions of *Cicer reticulatum*. The F_5 interspecific lines were tested for agronomic and seed quality traits including reaction to ascochyta blight disease under field conditions at two locations in 2018. A subset of 195 lines were selected based on resistance to ascochyta blight and acceptable seed quality. These lines were evaluated for their performance under suboptimal conditions at Lucky Lake (2019 and 2020) and Moose Jaw (2019), Saskatchewan, Canada, and Yuma, Arizona, United States (2019–2020). The lines were grown and evaluated at two seeding dates, normal (SD1) and late (SD2) seeding dates, at each location and year. The same lines were genotyped using Cicer60K Axiom® SNP chip. The population structure was determined based on 35,431 informative SNPs using fastStructure, and the interspecific lines were clustered at a *k*-value of 15. Significant marker-trait associations were identified for seed yield from SD1 and SD2 seeding dates, and stress tolerance indices (ATI, K_1STI, MP, SSPI, and TOL) using phenotypic values both from individual locations and combined analyses based on BLUP values. SNP marker Ca2_34600347 was significantly associated with yield from both the seeding dates. This and other SNP markers identified in this study may be useful for marker-assisted introgression of abiotic stress tolerance in chickpea.

KEYWORDS

suboptimal conditions, interspecific crosses, marker-assisted introgression, wild chickpea, cultivars, stress tolerance indices

Introduction

Chickpea (*Cicer arietinum* L.) is the world's second most important pulse crop after the common bean (Varshney et al., 2013; Rani et al., 2020). It was one of the earliest domesticated legume crops and is currently grown in 59 countries. In 2019, the world production of chickpeas was around 14.2 million tons (FAO, 2019).

Among abiotic stresses, drought and heat are the major environmental constraints limiting chickpea production worldwide in recent years (Devasirvatham and Tan, 2018; Rani et al., 2020). It has been reported that heat and drought can cause more than 70% yield loss in chickpea (Varshney et al., 2019). Chickpeas are mostly grown under rainfed conditions without irrigation. Therefore, soil moisture deficit toward the end of the crop season (terminal drought) affects about two-thirds of the global chickpea area (Gaur et al., 2008, 2019). Moreover, being a cool season food legume, chickpea yield is sensitive to heat stress exposure during the reproductive stage (Devasirvatham et al., 2015; Gaur et al., 2015).

Extreme heat and dry conditions are among the main abiotic stresses that affect crop yield across Canada. The prolonged heat wave and lack of precipitation in recent years in Western Canada[1] has had an adverse impact on chickpea yield. Tolerance to abiotic stresses such as heat and drought is a complex trait that is the result of various morphological, physiological, and biochemical changes in plants (Kaloki et al., 2019). Moreover, these abiotic stresses are quantitatively inherited with a large effect of genotype x environment interaction (Jha et al., 2014). The development of cultivars with abiotic stress tolerance and yield stability is critical in chickpea breeding programs. However, very slow progress has been made in developing tolerant cultivars due to the physiological and genetic complexity of the trait. The variability and unpredictability of stress conditions during trials limit the selection efficiency. Therefore, knowledge of the traits responsible for the adaptation of chickpea to suboptimal environments is important for the development of cultivars with improved abiotic stress tolerance.

Crop wild relatives (CWR) preserve higher levels of genetic diversity as they have been challenged in natural environments for many years in comparison to domesticated cultivars. Hence, these wild relatives are crucial genetic resources used by plant breeders for crop improvement (Hajjar and Hodgkin, 2007). CWR has been used as a source of abiotic stress tolerance in many cultivated species (Hajjar and Hodgkin, 2007; Zhang et al., 2019). Recent advances in genotyping, breeding, and genomics have accelerated the use of CWR for crop improvement by marker-assisted introgression of wild alleles into cultivated germplasm (Bohra et al., 2021). Few examples included enhanced drought tolerance in cultivated germplasm of sunflower (Hussain et al., 2019), and improved drought related traits such as water use efficiency, earliness, and yield of cultivated groundnut by introgression of alleles from the wild groundnut species *Arachis duranensis* and *Arachis batizocoi* (Dutra et al., 2018). Another example is enhanced drought resistance and productivity of elite durum (*Triticum turgidum* ssp. *durum*) and bread wheat (*Triticum aestivum* L.) cultivars by the reintroduction of alleles from wild emmer wheat (Merchuk-Ovnat et al., 2017). More studies that have demonstrated wild progenitors as a valuable source for the enrichment of the domesticated gene pool for abiotic stress tolerance included reintroducing wild alleles in lentils (Gorim and Vandenberg, 2017) and wheat (Placido et al., 2013).

Genome-Wide Association Study (GWAS) provides a higher resolution of marker-trait association than classical QTL analysis using bi-parental populations (Korte and Farlow, 2013). Both GWAS and QTL mapping studies were used to identify the genetic loci associated with various abiotic stresses including drought and heat stress (Jha et al., 2021). Nested association mapping (NAM) has been a valuable approach to dissecting the genetic architecture of complex quantitative traits (Gangurde et al., 2020; Lakmes et al., 2022). NAM population consists of multiple families of recombinant inbred lines (RILs) derived from multiple inbred lines crossed to a single reference inbred line (Yu et al., 2008). It utilizes the combined power of QTL mapping and association mapping to identify the trait-associated markers (Buckler et al., 2009). A publicly available collection of wild chickpeas, especially *Cicer reticulatum*, which survives under suboptimal environment, is an important resource to improve stress tolerance in current chickpea cultivars (Singh et al., 2008; von Wettberg et al., 2018). Linkage drag is a known bottleneck for the introgression of QTLs from wild accessions to cultivated. To overcome the linkage drag and to introgress and expand the genetic basis of cultivated chickpea, introgression of wild alleles from multiple accessions of *C. reticulatum* into cultivated chickpea germplasm has been initiated at the University of Saskatchewan chickpea breeding program. The main objective of the current study was to examine the performance of the interspecific population derived from *C. arietinum* x *C. reticulatum* crosses under suboptimal conditions and to identify the genetic loci associated with the traits crucial for plant performance under suboptimal environments using genome-wide association analysis.

Materials and methods

Plant materials

A chickpea interspecific population consisting of 600 lines derived from crosses between elite cultivars (*C. arietinum*) and 20 accessions of *Cicer reticulatum* were developed. The elite cultivars were CDC Leader (kabuli chickpea) and CDC Consul (desi chickpea). The design of this population is like a nested association mapping (NAM) design. Each of the 20 *C. reticulatum* accessions (tester lines) was crossed with CDC Leader and CDC Consul (founder lines). A single F_1 seed from each cross was grown in the greenhouse in 2014 and selfed to produce the F_2 seed. The F_2 plants of each cross were advanced to F_5 as single seed

1 https://www.canada.ca/en/environment-climate-change/news/2021/06/environment-and-climate-change-canada-warns-western-canadians-about-dangerous-record-high-temperatures.html

descents (SSDs). The seeds of each F_5 line were bulked and used for the evaluation of their agronomic performance including their reaction to ascochyta blight disease and seed visual quality at two locations, Limerick and Lucky Lake, SK during the 2018 growing season. One hundred and ninety-five F_7 interspecific lines with improved resistance to ascochyta blight and acceptable visual seed quality were selected and used in the current study. All the selected lines were derived from crosses with CDC Leader as a common parent.

Growing conditions and phenotypic data analysis

The 195 F_7 interspecific lines were grown at Lucky Lake (2019 and 2020) and Moose Jaw (2019), Saskatchewan, Canada, and Yuma, Arizona, United States (2019–2020). The population was planted at two different seeding dates [normal (SD1) and late seeding (SD2)] at each location and year. The purpose of late seeding was to expose the plants to higher temperatures during flowering. The normal seeding date in Saskatchewan was in the second week of May (14 May in Moose Jaw, 2019, 17 May in Lucky Lake, 2019, and 12 May in Lucky Lake, 2020) while the late seeding was planted around 2 weeks after the normal seeding (28 May in all the locations; Supplementary Table 1). At Yuma, Arizona, the first seeding was on 6 November while the second seeding was on 14 January (Supplementary Table 1) to maximize the chance that the population was exposed to a temperature above 27°C during flowering. At each location, the lines were planted in a one square meter plot. In each plot, the seeds were planted in three rows with a density of 60 seeds m^{-2}. A randomized complete block design (RCBD) with three replications was used at each location and seeding date. Daily maximum and minimum temperatures and other meteorological data were recorded at each location. There were multiple occurrences of temperatures above 26°C–27°C in all the locations during flowering on late seeding dates (Supplementary Table 1).

Days to flowering (DTF) were recorded for each line on the plot basis when 50% of the plants had flowered. Days to maturity (DTM) was also documented for each line when 50% of the plants in each plot were matured. Plant height was measured from the ground level to the tip of the plants when the pods reached physiological maturity. Flower color was recorded based on visual observation during flowering. Seed weight was measured by weighing 1,000 seeds per line after harvest. Reaction to ascochyta blight of each line was recorded on a plot basis, using a mixed quantitative and qualitative 0–9 score scale as described by Chongo et al. (2004). Seed yield was measured on a plot basis. Seed yield under stressed (late seeding) and non-stressed (normal seeding) conditions were used to predict the stress tolerance index of each line. The different stress indices used to assess stress tolerance included tolerance index (TOL), mean productivity (MP), abiotic tolerance index (ATI), stress susceptibility percentage index (SSPI), and modified stress tolerance index (K_1STI). These have been reported as reproducible indices under severe stress conditions in chickpeas (Farshadfar and Geravandi, 2013). The following formulas were used to calculate the stress indices.

TOL = Yp−Ys (Rosielle and Hamblin, 1981).
MP = (Yp + Ys)/2 (Rosielle and Hamblin, 1981).
ATI = [(Yp−Ys)/($\bar{Y}$p/$\bar{Y}$s)] × 100 (Moosavi et al., 2008).
SSPI = [(Yp−Ys)/(2$\bar{Y}$p)] × 100 (Moosavi et al., 2008).
K_1STI = (Yp2/$\bar{Y}$p2) × [(Yp + Ys)/$\bar{Y}$p2)] (Farshadfar and Sutka, 2003).

In the above formulas, Ys, Yp, $\bar{Y}$s, and $\bar{Y}$p represent seed yield in stress and non-stress conditions for each genotype, and mean seed yield in stress and non-stress conditions for all genotypes, respectively.

For field evaluation, both combined analyses across locations and years for each seeding date, and separate analyses on each year (2019 and 2020) for each seeding date were conducted. ANOVA was done using PROC MIXED in which genotypes were considered as a fixed factor and years as a random factor. The LSMEANS statement was used to compute the average phenotypic score for each line. For a separate analysis of each year, ANOVA was done using the PROC MIXED procedure, in which the lines were considered as a fixed factor and replication was considered as a random factor. To estimate the broad sense heritability (H^2), variance components were calculated using the SAS PROC VARCOMP procedure (SAS Institute Inc, 1999). The H^2 of various phenotypic responses at plot level based on individual experiments and over the years were estimated using the following two equations, respectively

$$H^2 = \frac{\sigma^2 G}{\sigma^2 G + \sigma^2 \mathrm{er}} \text{ and } H^2 = \frac{\sigma^2 G}{\sigma^2 G + \sigma^2 GY + \sigma \mathrm{er}}$$

where σ^2G, σ^2Y, σ^2GY, and σ^2er are estimates of genotype, site-year, genotype by site-year interaction, and error variance, respectively (Singh et al., 1993). Spearmen's rank correlation coefficients between the seed yields obtained under stress and non-stress conditions and the tolerance indices for each site year were calculated.

Genotyping of mapping population

The interspecific population was genotyped using a 60 K Axiom® SNP array (61,335 SNPs) at Eurofins, WI, United States. Individual plants from each line were grown under controlled conditions in phytotron chambers. Young leaf tissue was harvested from 2 to 3 weeks old plants. These leaf tissues were freeze dried using Labconco FreeZone 6-L Console Freeze Dry System and

sent to Eurofins BioDiagnostics, WI, United States, for DNA extraction and genotyping.

Population structure and linkage disequilibrium analysis

The program fastSTRUCTURE (Raj et al., 2014) was used to estimate the most likely number of clusters (*K*) into which the interspecific population can be grouped, and their degree of admixtures, based on genotypic data of 35,431 SNPs. The value of *K* was determined based on the lowest prediction error, and the smallest number of iterations for convergence. For each line, the value of *Q,* which is the probability of belonging to one of the clusters was derived from the matrix of contributions. A shared allele index derived from the dissimilarity matrix estimated from the SNP genotypic data were used to construct an unweighted neighbor-joining tree (Perrier et al., 2003).

Linkage disequilibrium (LD) of each chromosome was calculated as the correlation between marker-pairs among the SNP markers of each chromosome. The calculated Pearson correlation coefficient [r] was used to calculate LD decay by Quantile regression (R package "quantreg"; Koenker, 2017). The LD decay was calculated by plotting r^2 values as a function of genetic distance.

Association mapping

The software program GAPIT (Genome Association and Prediction Integrated Tool—R package: Lipka et al., 2012) was used to determine the association between SNP genotypes and the tolerance indices of the population. The phenotypes measured in individual environments were used for association analysis. Simultaneously, the best linear unbiased predictors (BLUP) of each trait of four environments tested were calculated using the "lme4" package of the R3.6.1 software,[2] considering environmental effects as fixed and genotype as random: y ~ (1 |rep% in % env) + (1|env) + (1 |lines) + (1 |env: lines), where rep% in % env represents replications were nested within the environments. The formula for the best linear unbiased estimate (BLUP) of phenotype is: y = Xb + Zu + e, where y, b, u, and e represent the observed phenotype, fixed effect vector, random effect vector, and residual, respectively, and X and Z represent incidence matrices. For the association analysis, the kinship coefficient matrix (K) values calculated by GAPIT by identity-by-state (IBS), were used. Multiple models of the GAPIT program including mixed linear model (MLM), multi-linear mixed model (MLMM), compressed mixed linear model (CMLM), general linear model (GLM), settlement of mixed linear models under progressively exclusive relationship (SUPER), fixed and random model circulating probability unification (FarmCPU) and factored spectrally transformed linear mixed models (FaST-LMM'; Tang et al., 2016; Kaler et al., 2020) were tested for comparison of the association analysis (Supplementary Figure 1). By comparison of the Q–Q plots of each model, which are drawn by plotting the observed and expected $\log_{10}$ *p*-values, the MLMM model was chosen to report marker-trait associations.

Results

Phenotypic evaluation

Various phenotypic traits were recorded across the locations and seeding dates in 2019 and 2020 (Tables 1A,B). ANOVA showed significant differences among the interspecific lines for days to flowering, days to maturity, seed yield, flower color, plant height, seed size, and ascochyta blight severity across all locations for normal (SD1) and late seeding (SD2) treatments in 2019 and 2020. There were significant effects of the environment (year) and genotype by environment (year) interaction in all the phenotypic traits recorded for both seeding dates. The interaction of genotype and site year was not significant for plant height in SD2 (Table 1B). Some of the phenotypic traits were not recorded in all four field trials because of a lack of resources and the restrictions due to COVID-19, especially during the 2019–2020 trials at Yuma, AZ. Plot yield was among the phenotypic traits measured in all years, locations, and seeding dates. Moreover, in previous studies seed yield was found to be an important trait to measure tolerance under stress and non-stress environments. The seed yield data were used for the calculation of stress tolerance indices and association analysis. The mean value and range of population for each trait are presented in Table 2.

Correlation of tolerance indices under field conditions

Late seeding (stress condition) lowered the yield in comparison to normal seeding (non-stress condition) in all site years, except Lucky Lake in 2020 where the difference in seed yield was marginal but significant ($p < 0.02$; Table 2). Spearman's rank correlation coefficients between the seed yields obtained under normal and late seeding and the stress tolerance indices for each year were calculated (Table 3). The correlation between the seed yield ranks under normal (non-stress) and late seeding (stress) conditions were positive and significant ($p \leq 0.001$) for all site years, indicating a change in the ranking of the genotypes for seed yield production due to stress.

Positive significant relationships were observed between the seed yield under non-stress conditions (Yp) and TOL, MP, ATI, SSPI, and K_1STI in all site years. On the other hand, the seed yield under stress conditions (Ys) was positively correlated with MP and K_1STI, but negatively correlated with TOL, ATI, and SSPI in all site years.

2 www.r-project.org

TABLE 1 ANOVA for days to flowering (DTF), days to maturity (DTM), seed yield, flower color, plant height, seed size (1,000 seed weight), and ascochyta blight rating under (A) normal seeding (SD1) and (B) late seeding (SD2) for the interspecific population was evaluated under field conditions at Lucky Lake and Moose Jaw, Saskatchewan in 2019, Yuma, Arizona in 2019–2020, and Lucky Lake, Saskatchewan in 2020.

Year/locations	Effect	*F* value						
		DTF	DTM	Seed yield (g)	Flower color	Plant height (cm)	Seed size (g)	Ascochyta blight rating
(A)								
Field combined	G	7.36***	3.38***	5.49***	12.08***	2.82***	25.86***	Only 1 year data
years (2019 and	Y	404.08***	252.36***	237.41***	1.15***	179.43***	1115.76***	available
2020)	G*Y	1.27***	3.41***	2.01***	3.24***	1.51***	3.29***	
	σ^2G	3.79	−0.32	3887.6	0.19	10.06	897.73	
	σ^2Y	3.9	6.79	5470.2	0.02	13.9	910.83	
	σ^2GY	0.59	12.08	4714.3	0.1	8.66	348.55	
	σ^2er	5.5	14.51	12864.9	0.12	46.81	465.58	
	H^2	0.38	−0.01	0.18	0.46	0.15	0.52	
Individual year/location								
Lucky Lake, 2019	G	7.32***	3.12***	3.27***	7.6***	2.07***	11.93***	No symptoms
	σ^2G	2.46	5.59	3860.8	0.23	15.31	1685.3	
	σ^2er	1.23	7.62	5095.7	0.1	41.87	452.63	
	H^2	0.67	0.42	0.43	0.7	0.27	0.79	
Moose Jaw, 2019	G	2.85***	NA	2.31***	NA	NA	13.56***	NA
	σ^2G	5.35	NA	2588.3	NA	NA	1663.8	NA
	σ^2er	8.26	NA	5994.7	NA	NA	376.24	NA
	H^2	0.39	NA	0.3	NA	NA	0.82	NA
Yuma, 2019–-2020	G	NA	NA	2.69***	NA	NA	6.39***	NA
	σ^2G	NA	NA	23901.2	NA	NA	672.38	NA
	σ^2er	NA	NA	36115.4	NA	NA	423.19	NA
	H^2	NA	NA	0.4	NA	NA	0.61	NA
Lucky Lake, 2020	G	3.24***	3.53***	4.4***	7.9***	2.3***	5.57***	1.81***
	σ^2G	5.24	17.31	5346.4	0.33	22.23	932.14	0.36
	σ^2er	7.01	21.87	4060.6	0.15	51.42	644.15	1.33
	H^2	0.43	0.44	0.57	0.7	0.3	0.59	0.21
(B)								
Field combined	G	1.92***	2.15***	6.5***	15.05***	3.32***	16.3***	1.84***
years (2019 and	Y	66.36***	2540.84***	355.52***	6.83**	6.96**	131.91***	4574.95***
2020)	G*Y	1.21*	1.35**	1.77***	2.04***	0.93^{ns}	4.15***	1.42**
	σ^2G	1.98	3.27	1929.5	0.31	13.93	825.95	0.001
	σ^2Y	1.95	91.04	2,796	0.0009	0.38	91.82	3.57
	σ^2GY	0.98	2.06	1242.6	0.05	−0.44	421.05	0.09
	σ^2er	17.28	20.18	3970.8	0.14	34.77	388.51	0.47
	H^2	0.1	0.13	0.27	0.61	0.29	0.51	0.002
Individual year/location								
Lucky Lake, 2019	G	1.29*	1.23*	2.32***	7.19***	2.17***	NA	1.35**
	σ^2G	3.02	1.37	2341.8	0.33	13.27	NA	0.03
	σ^2er	32.57	27.7	5111.5	0.16	33.62	NA	0.36
	H^2	0.08	0.05	0.32	0.67	0.28	NA	0.07
Moose Jaw, 2019	G	NA	NA	2.81***	NA	NA	NA	NA
	σ^2G	NA	NA	3026.2	NA	NA	NA	NA
	σ^2er	NA	NA	4649.5	NA	NA	NA	NA
	H^2	NA	NA	0.39	NA	NA	NA	NA

(Continued)

TABLE 1 Continued

Year/locations	Effect	*F* value						
		DTF	DTM	Seed yield (g)	Flower color	Plant height (cm)	Seed size (g)	Ascochyta blight rating
Yuma, 2019–2020	G	NA	NA	2.39***	NA	NA	8.2***	NA
	σ^2G	NA	NA	1335.4	NA	NA	1340.5	NA
	σ^2er	NA	NA	2411.9	NA	NA	513.19	NA
	H^2	NA	NA	0.36	NA	NA	0.72	NA
Lucky Lake, 2020	G	4.55***	2.74***	4.97***	10.39***	2.14***	14.04***	1.87***
	σ^2G	2.9	8.1	5139.7	0.38	13.67	1158.2	0.16
	σ^2er	2.45	13.94	3798.7	0.12	35.89	269.43	0.56
	H^2	0.54	0.37	0.58	0.76	0.27	0.81	0.22

G, genotype interaction; Y, site year interaction; G*Y, genotype by site year interaction; σ^2G, σ^2Y, σ^2GY, and σ^2er are estimates of genotype, site year, genotype by site year interaction, and error variance, respectively; H^2 is broad sense heritability; ns, not significant; NA, data not available.
***Indicates a significant difference at $P \leq 0.001$.
**Indicates a significant difference at $P \leq 0.01$.
*Indicates a significant difference at $P \leq 0.05$.

TABLE 2 Mean and range values of the interspecific population for characters assessed under field conditions.

Character	Location	Population			
		SD1		SD2	
		Mean	Range	Mean	Range
Days to flowering	Lucky Lake, 2019	50.75	45–56	49.57	42–60
	Moose Jaw, 2019	NA	NA	NA	NA
	Yuma 2019–2020	NA	NA	NA	NA
	Lucky Lake 2020	54.28	48–64	47.59	42–56
Days to maturity	Lucky Lake, 2019	98.88	94–108	102.61	89–110
	Moose Jaw, 2019	NA	NA	NA	NA
	Yuma 2019–2020	NA	NA	NA	NA
	Lucky Lake 2020	102.37	90–118	89.23	79–97
Seed yield (g)	Lucky Lake, 2019	227.97	28–443	192.62	22–395
	Moose Jaw, 2019	306.6	90–486	210.14	43–385
	Yuma 2019–2020	259.24	5–1,638	89.54	3–282
	Lucky Lake 2020	128.85	6–466	146.67	9–467
Plant height (cm)	Lucky Lake, 2019	42.67	29–61	44.99	35–67
	Moose Jaw, 2019	NA	NA	NA	NA
	Yuma 2019–-2020	NA	NA	NA	NA
	Lucky Lake 2020	47.81	27–66	45.7	31–63
Seed size (1,000 seed weight in g)	Lucky Lake, 2019	172.88	77–343	NA	NA
	Moose Jaw, 2019	202.54	111–330	NA	NA
	Yuma 2019–2020	131.09	48–261	165.67	55–316
	Lucky Lake 2020	165.86	61–267	177.67	87–303
Ascochyta blight rating	Lucky Lake, 2019	No symptoms	No symptoms	5.27	4.5–7.5
	Moose Jaw, 2019	NA	NA	NA	NA
	Yuma 2019–2020	NA	NA	NA	NA
	Lucky Lake 2020	0.43	0–4	2.59	2–5

Genotyping

A total of 35,429 SNPs were used to determine the population structure. Of these SNPs, 32,228 were located on the eight chromosomes of chickpea and 3,101 SNPs were on scaffolds and are not located on any of the chromosomes. After filtering for a minimum allele frequency (MAF) of 0.05, 20,679 SNPs were used for association analysis.

10.3389/fpls.2022.933277

TABLE 3 Spearmen's rank correlation between different stress tolerance indices for each site years.

Year/location	Parameters	Yp	Ys	TOL	MP	ATI	SSPI
Lucky Lake, 2019	Ys	0.39***					
	TOL	0.58***	−0.46***				
	MP	0.83***	0.81***	0.09*			
	ATI	0.58***	−0.46***	1***	0.09*		
	SSPI	0.58***	−0.46***	1***	0.09*	1***	
	K_1STI	0.98***	0.52***	0.45***	0.91***	0.45***	0.45***
Moose Jaw, 2019	Ys	0.36***					
	TOL	0.58***	−0.51***				
	MP	0.82***	0.80***	0.06^{ns}			
	ATI	0.58***	−0.51***	1***	0.06^{ns}		
	SSPI	0.58***	−0.51***	1***	0.06^{ns}	1***	
	K_1STI	0.98***	0.50***	0.43***	0.91***	0.43***	0.43***
Yuma 2019–2020	Ys	0.31***					
	TOL	0.91***	-0.04^{ns}				
	MP	0.95***	0.55***	0.76***			
	ATI	0.91***	-0.04^{ns}	1***	0.76***		
	SSPI	0.91***	-0.04^{ns}	1***	0.76***	1***	
	K_1STI	0.97***	0.33***	0.89***	0.94***	0.89***	0.89***
Lucky Lake 2020	Ys	0.55***					
	TOL	0.41***	−0.48***				
	MP	0.85***	0.89***	-0.06^{ns}			
	ATI	0.41***	−0.48***	1***	-0.06^{ns}		
	SSPI	0.41***	−0.48***	1***	-0.06^{ns}	1***	
	K_1STI	0.99***	0.65***	0.29***	0.91***	0.29***	0.29***

Yp, seed yield under non-stress; Ys, seed yield under stress; TOL, tolerance index; MP, mean productivity; ATI, abiotic tolerance index; SSPI, stress susceptibility percentage index; K_1STI, modified stress tolerance index.
*Indicates significance at $p \leq 0.05$.
***Indicates significance at $p \leq 0.001$.
^{ns}not significant.

Linkage disequilibrium analysis

LD decay was calculated as the Pearson correlation coefficient (r^2) between marker pairs of each chromosome. The LD decay differed among the eight chromosomes. The $r^2_{max,90,}$ which is calculated as the maximum r^2 in the 90th percentile of each chromosome, for chromosomes 1 to 8 is between 0.24 and 0.26. The physical distance in Mb at which LD of each chromosome has decayed to half of $r^2_{max,90}$ is 0.51, 0.15, 0.25, 0.21, 0.38, 0.23, 0.24, and 0.43 Mb, for chromosomes 1 to 8, respectively (Figure 1).

Population genetic structure

To determine the population structure, the most likely number of clusters (*k*) was tested from 2 to 20. A *k*-value of 15 is best suited to describe the genetic structure of 195 chickpea interspecific lines along with the two parents CDC Consul and CDC Leader, and six CDC cultivars used as checks, CDC Orion, CDC Cory, ILC 533, ICC 4958, and ILC 3279. The probability of membership of individual lines in each cluster was estimated by admixture analysis (Figure 2; Supplementary Table 2). The grouping of the interspecific population into phylogenetic clusters based on the neighbor-joining (NJ) tree differed from the population structure analysis that the population is divided into eight clades (Figure 3).

Association mapping

The seed yield of interspecific lines measured both in normal and heat stress treatments, and the heat stress indices calculated based on the seed yield in both the treatments, were used to identify the associated SNP markers. The BLUP values calculated based on the measurements of each trait/index in four station-years were used for association analysis. SNP markers were identified for association with seed yield both in normal and heat stress environments, and the indices calculated based on the seed yield in these two treatments are listed in Table 4. SNP marker Ca2_34600347 was identified for association with grain yield both in normal and stress (late seeding) treatments. The markers for ATI and TOL are located on chromosome 1, while the markers for other indices K_1STI and SSPI are located on chromosome 4. The Manhattan plots representing the genome-wide marker-trait associations, concerning the seven traits/indices measured, and their corresponding Q–Q plots are presented in Figure 4.

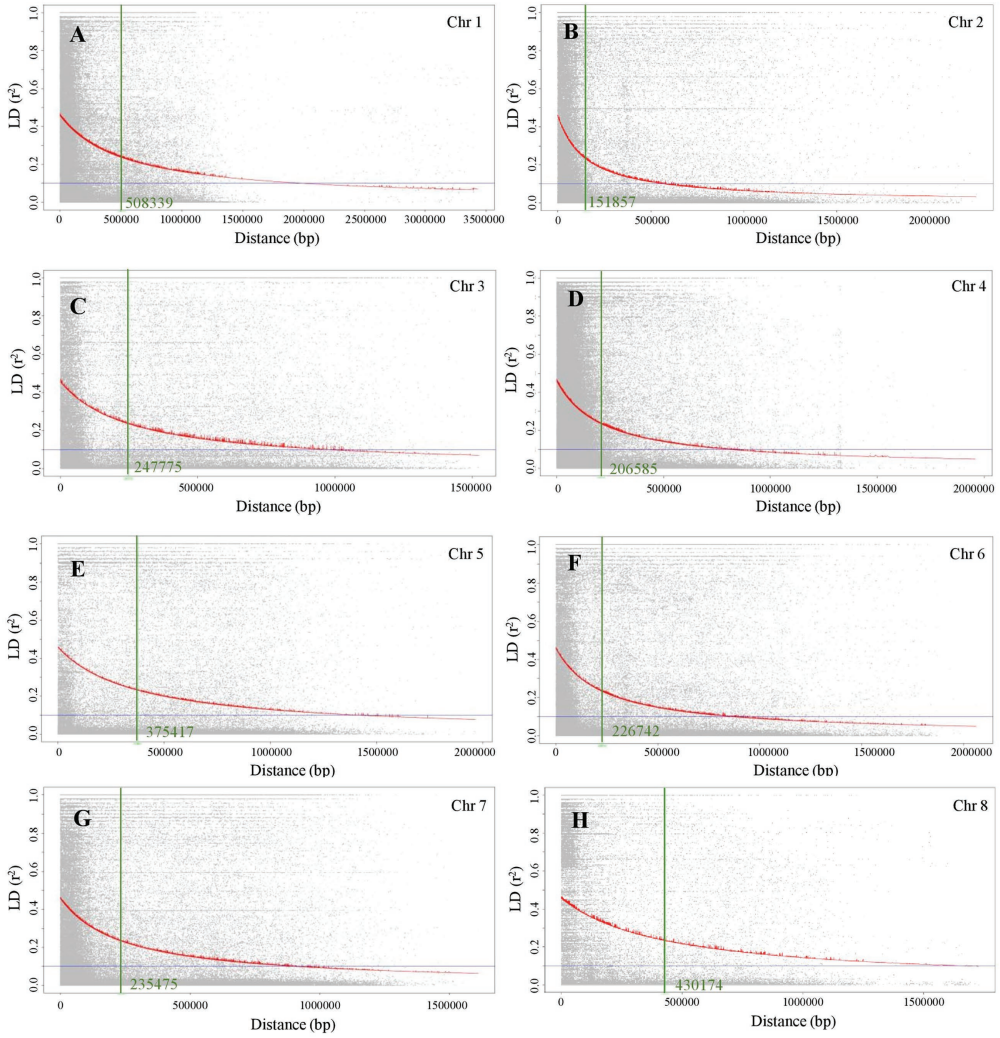


FIGURE 1
Linkage disequilibrium (LD) decay plots of eight chromosomes of chickpea. **(A)** Chromosome 1, **(B)** Chromosome 2, **(C)** Chromosome 3, **(D)** Chromosome 4, **(E)** Chromosome 5, **(F)** Chromosome 6, **(G)** Chromosome 7, and **(H)** Chromosome 8.

Discussion

The wild relatives of chickpea have been an invaluable source for improving elite chickpea germplasm through resistance to biotic stress (Sharma et al., 2006; Singh et al., 2008). Moreover, crop wild relatives were reported to be an important resource for genetic improvement for various abiotic stresses (Singh et al., 2008; Coyne et al., 2020; Bakir et al., 2021). Collection of various wild chickpeas, mainly *C. reticulatum* was reported to have biotic and abiotic stress tolerance (Singh et al., 2008; von Wettberg et al., 2018; Bakir et al., 2021). *C. reticulatum* is in the primary gene pool of chickpea (Bakir et al., 2021). It is known for its crossing-compatibility with domesticated chickpea (*C. arietinum*) and generally produces fertile progeny because of good chromosome pairing (Bakir et al., 2021). Therefore, in the current study, we have used *C. reticulatum* to introgress stress tolerance in cultivated kabuli and desi chickpea (*C. arietinum*) germplasm. In this study, we developed progeny from interspecific crosses between adapted elite cultivars, CDC Leader and CDC Consul, and 20 *C. reticulatum* accessions. Further lines with improved response to ascochyta blight and acceptable seed quality were selected and used in this study. Interestingly, all these selected lines were derived from CDC Leader (kabuli chickpea) and 17 wild accessions. The design of our population is the same as the nested association mapping (NAM) design. Generally, the bi-parental population has a lack of mapping precision and low genetic diversity. Therefore, using NAM helps to capture additional genetic diversity and increase genetic recombination.

Chickpea is a dry season food legume that is mostly grown on residual moisture after the rainy season. Toward the end of the growing season chickpeas often experience terminal drought stress. Moreover, if the sowing is delayed, the crop may deal with heat stress during the reproductive phase (Gaur et al., 2019; Maphosa et al., 2020; Rani et al., 2020). Low moisture and high temperatures during the flowering and early pod filling stage can substantially influence chickpea yield by forcing early maturity resulting in low biomass and a low number of pods and seeds per

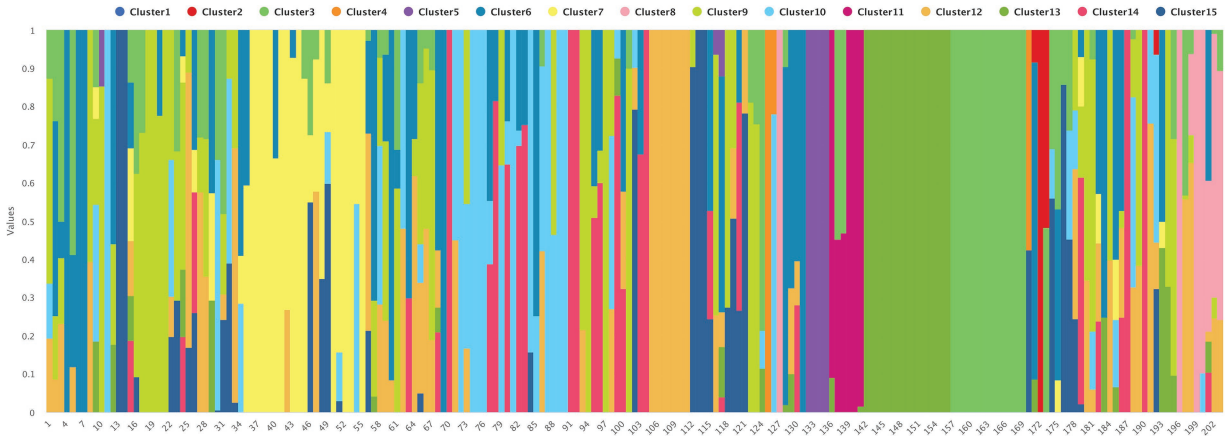


FIGURE 2
The population structure of 203 chickpea interspecific lines and accessions based on k = 15.

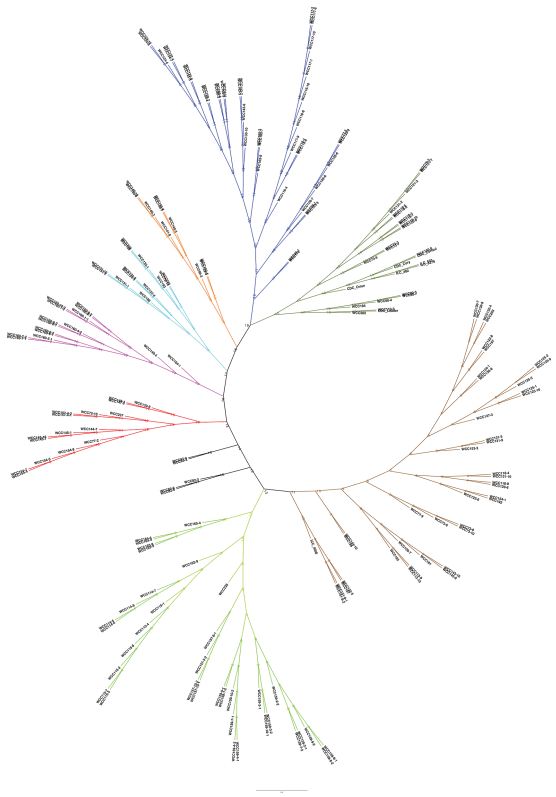


FIGURE 3
Genetic relatedness among the 203 chickpea accessions, estimated by neighbor-joining method.

plant (Maphosa et al., 2020; Rani et al., 2020). Therefore, to understand the performance of interspecific populations under a suboptimal environment, the population in this study was planted at two different seeding dates (normal and late seeding). The main objective of late seeding was to expose the plants to stress conditions such as higher temperatures and low moisture during flowering. Previous studies have also used the same approach to understand the effect of stress on crop development and yield in canary seeds (Miller, 1999). Furthermore, it was observed that late seeding has a negative effect on chickpea yield (Machado et al., 2006; McKenzie et al., 2006).

The current study evaluated various phenotypic traits such as days to flowering, days to maturity, seed yield, flower color, plant height, seed size, and ascochyta blight disease resistance on both the seeding dates. There were significant effects of environment and genotype by environment interaction in all the phenotypic traits documented for both seeding dates which clearly shows that a suboptimal environment negatively impacts crop development in chickpea. Similarly, other studies have shown the negative effect of stress on various phenotypic traits grown under stress and non-stress environments (Arif et al., 2021; Jeffrey et al., 2021; Jha et al., 2021). Yield is a crucial trait and an important indicator to define tolerance between stress and non-stress conditions and has been used to describe the performance of any genotype while screening in various environments (Kaloki et al., 2019; Pour-Aboughadareh et al., 2019). Moreover, in the current study seed yield was among the phenotypic traits measured in all years, locations, and seeding dates. Therefore, we have used seed yield to calculate various stress tolerance indices. Furthermore, crop yield under stress and non-stress conditions and stress tolerance indices were used for association analysis. Seed yield was significantly reduced in the late seeded population in all the locations as compared to the normal seeded population except Lucky Lake, 2020 which could be because of less stressed conditions while the late seeded population in Lucky Lake. Similar to our findings, previous studies have also shown reduced yield under late seeded populations (Machado et al., 2006; McKenzie et al., 2006). The calculation of spearman's rank correlation coefficients between seed yields acquired under normal and late seeding and the stress tolerance indices for each year was done. We found that the correlation between seed yield under stress and non-stress conditions was positive and significant which shows that the yield

TABLE 4 Selected marker-trait associations, identified based on BLUP values of phenotypes measured in four station-years during 2019 and 2020.

Phenotypic trait/index	SNP marker	*P*-value	maf
Ys (seed yield under stress conditions)	NW_9270594	$2.31E^{-04}$	0.31
	Ca3_15304269	$2.61E^{-04}$	0.09
	Ca6_3396299	$2.89E^{-04}$	0.31
	Ca7_43614232	$3.14E^{-04}$	0.41
	Ca4_37419513	$3.47E^{-05}$	0.40
	Ca2_34600347	$3.92E^{-04}$	0.07
Yp (seed yield under non-stress conditions)	Ca2_34600347	$3.25E^{-06}$	0.07
	Ca4_8694304	$1.35E^{-05}$	0.05
	Ca4_8737135	$5.55E^{-05}$	0.14
ATI (abiotic tolerance index)	Ca1_47259	$3.42E^{-06}$	0.46
	Ca1_56428	$6.91E^{-06}$	0.46
K_1STI (modified stress tolerance index)	Ca4_36637574	$6.05E^{-05}$	0.10
	Ca4_8646741	$7.12E^{-05}$	0.16
	Ca4_11276937	$9.00E^{-05}$	0.07
	Ca4_11277513	$9.00E^{-05}$	0.07
MP (mean productivity)	Ca2_34600347	$1.34E^{-05}$	0.07
SSPI (stress susceptibility percentage index)	Ca4_8694304	$2.37E^{-06}$	0.05
	Ca4_8313845	$4.17E^{-06}$	0.06
TOL (tolerance index)	Ca4_8694304	$2.87E^{-06}$	0.05
	Ca4_8670257	$1.04E^{-05}$	0.06
	Ca4_8313845	$1.19E^{-05}$	0.06
	Ca1_47259	$1.66E^{-05}$	0.46

maf, minimum allele frequency.

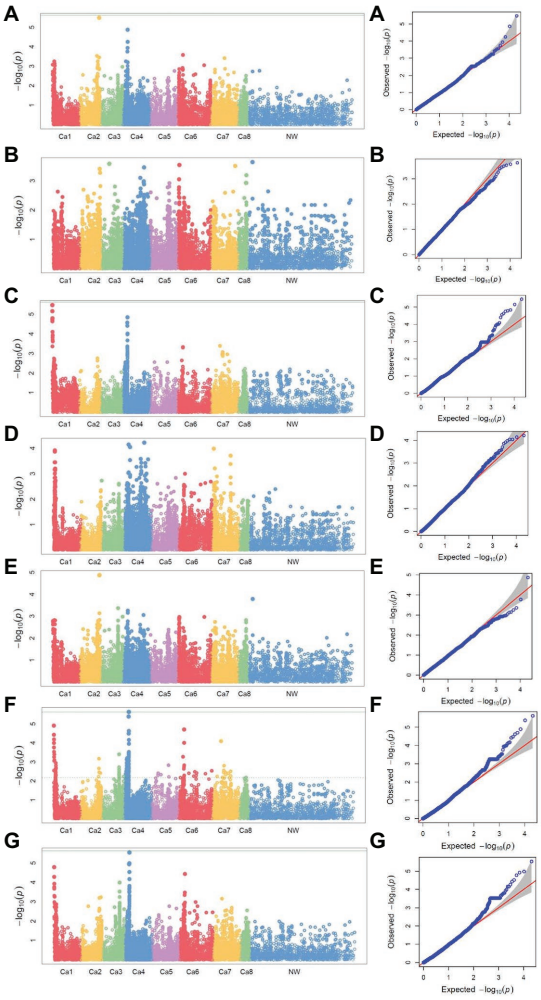


FIGURE 4
Manhattan plots of –log10 *p*-values and the corresponding quantile–quantile (Q–Q) plots of the association analysis for **(A)** seed yield under non-stress and **(B)** seed yield under stress conditions and yield indices **(C)** ATI, **(D)** K_1STI, **(E)** MP, **(F)** SSPI, and **(G)** TOL using a mixed linear model for four locations: Lucky Lake and Moose Jaw, Saskatchewan, 2019, Yuma, Arizona, 2019–2020, and Lucky Lake, Saskatchewan, 2020. For Manhattan plots: *y*-axis, –log10 *p*-values; *x*-axis, chromosome numbers. For Q–Q plots: *y*-axis, observed –log10 *p*-values; *x*-axis, expected –log10 *p*-values.

production was different in all the genotypes under non-stress conditions. Similar results were also found in previous studies (Moosavi et al., 2008; Farshadfar and Geravandi, 2013). Moreover, seed yield under non-stress conditions was positively correlated with TOL, MP, ATI, SSPI, and K_1STI which shows that selection of the genotypes based on these indices will improve the yield under non-stress conditions. Some studies have also found a positive correlation between non-stress conditions and MP, K_1STI (Farshadfar and Geravandi, 2013) and TOL, MP, SSPI, ATI (Moosavi et al., 2008). On the other hand, the seed yield under stress conditions was positively correlated with MP and K_1STI but negatively correlated with TOL, ATI, and SSPI indicating that the selection based on a higher value of MP and K_1STI will improve seed yield while selection based on TOL, ATI, and SSPI will reduce seed yield under stress conditions. Therefore, correlation analysis between seed yield and stress tolerance indices can be a good criterion for screening the best genotypes and indices used.

Linkage drag is one of the major bottlenecks to using wild species in the genetic improvement of crop plants including chickpea. Often, a large portion of entire chromosomes is affected by the linkage drag, which is a hindrance to the introgression of desirable alleles. In the current study, the LD decay observed in each chromosome is from 0.15 to 0.51 Mb, indicating that the interspecific population used in the current study is suitable for the objectives of this study. Genome-wide association study (GWAS) was used to understand the genetic basis of complex traits in chickpea (Kohli et al., 2020) and heat stress in other legumes (Tafesse et al., 2020, 2021). Since we intend to identify the heat stress loci in wild chickpea, and as well to introgess the same in an elite chickpea genetic background, we developed an inbred line population derived from crosses of wild accessions and elite chickpea cultivars. The 195 inbred line population used for association mapping of heat stress tolerance was an interspecific population derived from crosses of CDC Leader with 17 *C. reticulatum* accessions.

In the current study, we used the measurement of seed yield in heat stress conditions as an important criterion to identify heat stress tolerance in chickpea interspecific lines. Understanding the genetic basis of yield in heat stress conditions is important to develop heat stress tolerant high yielding chickpea varieties.

Previous studies have identified quantitative trait loci (QTLs) associated with yield (Barmukh et al., 2021), outlining component traits and heat tolerance (Thudi et al., 2014) in chickpea. Barmukh et al. (2021) identified yield-associated QTLs on LG1 and LG4 using a bi-parental mapping population of ICC 4958 × DCP 92–3. In the current study, we identified markers associated with seed yield and heat tolerance indices mainly on LG1 and LG4, while markers associated with seed yield in heat stress were distributed on multiple linkage groups including LG4. Yield is a complex trait and is determined by several components such as pod weight, pod number, and the number of reproductive nodes, etc. Tafesse et al. (2020, 2021) identified the correlation between multiple components of heat stress and SNP markers associated with these components in a GWAS study on peas. We have used the measurement of seed yield in heat stress conditions as a key indicator of heat tolerance, and the indices calculated based on seed yield in normal and heat stress conditions were among the other indicators used. The SNP marker Ca2_34600347, which was identified for its association with seed yield both in normal and heat stress environments, could be a valuable marker for the marker-assisted selection of heat tolerance in chickpea.

Conclusion

Wild relatives are a source of novel alleles to improve chickpea adaptation to suboptimal environments. In the current study, we used an interspecific population derived from *C. reticulatum* accessions. The late seeding and exposure of this population to suboptimal growth conditions has significantly reduced the seed yield. The study characterized the variation of the interspecific population for its yield performance and yield-related indices under suboptimal conditions. Individual lines that were identified for superior performance in suboptimal conditions can be used as trait donors in breeding for stress tolerance. The trait-associated SNP markers identified can be used for marker-assisted selection in the breeding pipeline. Future detailed analysis of these SNPs would allow us to identify the genes underlying tolerance to abiotic stress, providing new alleles and molecular markers to use in chickpea crop improvement.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SK conducted all the experiments, analyzed phenotypic data, and wrote the manuscript. KG analyzed genotypic data, conducted association analysis, and wrote the manuscript. BT conceived, designed, and supervised the research and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

We acknowledge the financial support from the Saskatchewan Ministry of Agriculture through the Agricultural Development Fund.

Acknowledgments

We acknowledge Donna Lindsay for technical assistance in developing the interspecific population, Carmen Breitkreutz for technical assistance with field work, and Brent Barlow and Alison Sackville for technical assistance in the preparation of field experiments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.933277/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Q–Q plots of the association analysis for seed yield under non-stress (Yp), seed yield under stress conditions (Ys) and yield indices ATI, K1STI, MP, SSPI, and TOL from various GWAS models (CMLM, FarmCPU, FaST-LMM, GLM, MLM, MLMM, SUPER).

SUPPLEMENTARY TABLE 1
Daily temperature and precipitation recorded at Lucky Lake and Moose Jaw, Saskatchewan, Canada in 2019, Yuma, Arizona, USA in 2019–2020 and Lucky Lake, Saskatchewan, Canada in 2020 during growing season.

References

Arif, A., Parveen, N., Waheed, M. Q., Atif, R. M., Waqar, I., and Shah, T. M. (2021). A comparative study for assessing the drought-tolerance of chickpea under varying natural growth environments. *Front. Plant Sci.* 11:607869. doi: 10.3389/fpls.2020.607869

Bakir, M., Sari, D., Sari, H., Waqas, M., and Atif, R. M. (2021). "Chickpea wild relatives: potential hidden source for the development of climate resilient chickpea varieties," in *Wild Germplasm for Genetic Improvement in Crop Plants*. eds. M. T. Azhar and S. H. Wani (Cambridge, MA, USA: Academic Press), 269–297.

Barmukh, R., Soren, K. R., Madugula, P., Gangwar, P., Shanmugavadivel, P. S., Bharadwaj, C., et al. (2021). Construction of a high-density genetic map and QTL analysis for yield, yield components and agronomic traits in chickpea (*Cicer arietinum* L.). *PLoS One* 16:e0251669. doi: 10.1371/journal.pone.0251669

Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., et al. (2021). Reap the crop wild relatives for breeding future crops. *Trends Biotechnol.* 40, 412–431. doi: 10.1016/j.tibtech.2021.08.009

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., et al. (2009). The genetic architecture of maize flowering time. *Science* 325, 714–718. doi: 10.1126/science.1174276

Chongo, G., Gossen, B., Buchwaldt, L., Adhikari, T., and Rimmer, S. (2004). Genetic diversity of *Ascochyta rabiei* in Canada. *Plant Dis.* 88, 4–10. doi: 10.1094/PDIS.2004.88.1.4

Coyne, C. J., Agrawal, S. K., von Wettberg, E. J. B., Marques, E., Berger, J., Redden, R. J., et al. (2020). Potential and limits of exploitation of crop wild relatives for pea, lentil, and chickpea improvement. *Legum. Sci.* 2:e36. doi: 10.1002/leg3.36

Devasirvatham, V., Gaur, P. M., Raju, T. N., Trethowan, R. M., and Tan, D. K. Y. (2015). Field response of chickpea (*Cicer arietinum* L.) to high temperature. *Field Crop Res.* 172, 59–71. doi: 10.1016/j.fcr.2014.11.017

Devasirvatham, V., and Tan, D. K. Y. (2018). Impact of high temperature and drought stresses on chickpea production. *Agronomy* 8:145. doi: 10.3390/agronomy8080145

Dutra, W. F., Guerra, Y. L., Ramos, J. P. C., Fernandes, P. D., Silva, C. R. C., Bertioli, D. J., et al. (2018). Introgression of wild alleles into the tetraploid peanut crop to improve water use efficiency, earliness and yield. *PLoS One* 13:e0198776. doi: 10.1371/journal.pone.0198776

FAO (2019) Chickpea production data. Available at: http://www.fao.org/faostat/en/#data/QCL/visualize

Farshadfar, E., and Geravandi, M. (2013). Repeatability of drought tolerance indices in chickpea genotypes over stress and non-stress environments. *Acta Agron. Hung.* 61, 123–137. doi: 10.1556/AAgr.61.2013.2.4

Farshadfar, E., and Sutka, J. (2003). Multivariate analysis of drought tolerance in wheat substitution lines. *Cereal Res. Commun.* 31, 33–40. doi: 10.1007/BF03543247

Gangurde, S. S., Wang, H., Yaduru, S., Pandey, M. K., Fountain, J. C., Chu, Y., et al. (2020). Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnol. J.* 18, 1457–1471. doi: 10.1111/pbi.13311

Gaur, P. M., Krishnamurthy, L., and Kashiwagi, J. (2008). Improving drought-avoidance root traits in chickpea (*Cicer arietinum* L.), current status of research at ICRISAT. *Plant Prod. Sci.* 11, 3–11. doi: 10.1626/pps.11.3

Gaur, P. M., Samineni, S., Krishnamurthy, L., Kumar, S., Ghanem, M. E., Beebe, S., et al. (2015). High temperature tolerance in grain legumes. *Legume Perspect.* 7, 23–24.

Gaur, P. M., Samineni, S., Thudi, M., Tripathi, S., Sajja, S. B., Jayalakshmi, V., et al. (2019). Integrated breeding approaches for improving drought and heat adaptation in chickpea (*Cicer arietinum* L.). *Plant Breed.* 138, 389–400. doi: 10.1111/pbr.12641

Gorim, L. Y., and Vandenberg, A. (2017). Root traits, nodulation and root distribution in soil for five wild lentil species and Lens culinaris (medik.) grown under well-watered conditions. *Front. Plant Sci.* 8:1632. doi: 10.3389/fpls.2017.01632

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156, 1–13. doi: 10.1007/s10681-007-9363-0

Hussain, M. M., Rauf, S., and Warburton, M. L. (2019). Development of drought-tolerant breeding lines derived from Helianthus annuus × H. argophyllus interspecific crosses. *Plant Breed.* 138, 862–870. doi: 10.1111/pbr.12731

Jeffrey, C., Trethowan, R., and Kaiser, B. (2021). Chickpea tolerance to temperature stress: status and opportunity for improvement. *J. Plant Physiol.* 267:153555. doi: 10.1016/j.jplph.2021.153555

Jha, U. C., Chaturvedi, S. K., Bohra, A., Basu, P. S., Khan, M. S., and Barh, D. (2014). Abiotic stresses, constraints and improvement strategies in chickpea. *Plant Breed.* 133, 163–178. doi: 10.1111/pbr.12150

Jha, U. C., Nayyar, H., Palakurthi, R., Jha, R., Valluri, V., Bajaj, P., et al. (2021). Major QTLs and potential candidate genes for heat stress tolerance identified in chickpea (*Cicer arietinum* L.). *Front. Plant Sci.* 12:655103. doi: 10.3389/fpls.2021.655103

Kaler, A. S., Gillman, J. D., Beissinger, T., and Purcell, L. C. (2020). Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Front. Plant Sci.* 10:1794. doi: 10.3389/fpls.2019.01794

Kaloki, P., Devasirvatham, V., and Tan, D. K. (2019). "Chickpea abiotic stresses: combating drought, heat and cold," in: *Abiotic and Biotic Stress in Plants* ed. OliveiraA. B. de (London, UK: IntechOpen).

Koenker, R. (2017). Quantile regression: 40 years on. *Annu. Rev. Econ.* 9, 155–176. doi: 10.1146/annurev-economics-063016-103651

Kohli, P. S., Verma, P. K., Verma, R., Parida, S. K., Thakur, J. K., and Giri, J. (2020). Genome-wide association study for phosphate deficiency responsive root hair elongation in chickpea. *Funct. Integr. Genomics* 20, 775–786. doi: 10.1007/s10142-020-00749-6

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29. doi: 10.1186/1746-4811-9-29

Lakmes, A., Jhar, A., Penmetsa, R. V., Wei, W., Brennan, A. C., and Kahriman, A. (2022). The quantitative genetics of flowering traits in wide crosses of chickpea. *Agriculture* 12:486. doi: 10.3390/agriculture12040486

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Machado, S., Humphreys, C., Tuck, B., and Corp, M. (2006). Seeding date, plant density, and cultivar effects on chickpea yield and seed size in eastern Oregon. *Crop Manag.* 5, 1–7. doi: 10.1094/CM-2006-0621-01-RS

Maphosa, L., Richards, M. F., Norton, S. L., and Nguyen, G. N. (2020). Breeding for abiotic stress adaptation in chickpea (*Cicer arietinum* L.): A comprehensive review. *Crop Breed. Genet. Genom.* 2:e200015. doi: 10.20900/cbgg20200015

McKenzie, R. H., Middleton, A. B., and Bremer, E. (2006). Effect of seeding date and rate on desi chickpea in southern Alberta. *Can. J. Plant Sci.* 86, 717–721. doi: 10.4141/P05-123

Merchuk-Ovnat, L., Fahima, T., Ephrath, J. E., Krugman, T., and Saranga, Y. (2017). Ancestral QTL alleles from wild emmer wheat enhance root development under drought in modern wheat. *Front. Plant Sci.* 8:703. doi: 10.3389/fpls.2017.00703

Miller, P. R. (1999). Effect of varying seeding date on crop development, yield and yield components in canaryseed. *Can. J. Plant Sci.* 80, 83–86. doi: 10.4141/P99-047

Moosavi, S. S., Yazdi-Samadi, B., Naghavi, M. R., Zali, A. A., Dashti, H., and Pourshahbazi, A. (2008). Introduction of new indices to identify relative drought tolerance and resistance in wheat genotypes. *Desert* 12, 165–178. doi: 10.22059/JDESERT.2008.27115

Perrier, X., Flori, A., and Bonnot, F. (2003). "Data analysis methods," in *Genetic Diversity of Cultivated Tropical Plants*. eds. P. Hamon, M. Seguin, X. Perrier and J. C. Glaszmann (Enfield, USA: Science publishers), 43–76.

Placido, D. F., Campbell, M. T., Folsom, J. J., Cui, X., Kruger, G. R., Baenziger, P. S., et al. (2013). Introgression of novel traits from a wild wheat relative improves drought adaptation in wheat. *Plant Physiol.* 161, 1806–1819. doi: 10.1104/pp.113.214262

Pour-Aboughadareh, A., Yousefian, M., Moradkhani, H., Vahed, M. M., Poczai, P., and Siddique, K. H. M. (2019). iPASTIC: an online toolkit to estimate plant abiotic stress indices. *Appl. Plant Sci.* 7:e11278. doi: 10.1002/aps3.11278

Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi: 10.1534/genetics.114.164350

Rani, A., Devi, P., Jha, U. C., Sharma, K. D., Siddique, K. H., and Nayyar, H. (2020). Developing climate-resilient chickpea involving physiological and molecular approaches with a focus on temperature and drought stresses. *Front. Plant Sci.* 10:1759. doi: 10.3389/fpls.2019.01759

Rosielle, A. A., and Hamblin, J. (1981). Theoretical aspects of selection for yield in stress and nonstress environments. *Crop. Sci.* 21, 943–946. doi: 10.2135/cropsci1981.0011183X002100060033x

SAS Institute Inc. (1999). *SAS/STAT® User's Guide, Version 8*, SAS Institute Inc., Cary, NC.

Sharma, H. C., Bhagwat, M. P., Pampapathy, G., Sharma, J. P., and Ridsdill-Smith, T. J. (2006). Perennial wild relatives of chickpea as potential sources of resistance to *Helicoverpa armigera*. *Genet. Resour. Crop. Evol.* 53, 131–138. doi: 10.1007/s10722-004-1951-4

Singh, M., Ceccarelli, S., and Hamblin, J. (1993). Estimation of heritability from varietal trials data. *Theor. Appl. Genet.* 86, 437–441. doi: 10.1007/BF00838558

Singh, R., Sharma, P., Varshney, R. K., Sharma, S. K., and Singh, N. K. (2008). Chickpea improvement: role of wild species and genetic markers. *Biotechnol. Genet. Eng. Rev.* 25, 267–314. doi: 10.5661/bger-25-267

Tafesse, E. G., Gali, K. K., Lachagari, V. B. R., Bueckert, R., and Warkentin, T. D. (2020). Genome-wide association mapping for heat stress responsive traits in field pea. *Int. J. Mol. Sci.* 21:2043. doi: 10.3390/ijms21062043

Tafesse, E. G., Gali, K. K., Lachagari, V. B. R., Bueckert, R., and Warkentin, T. D. (2021). Genome-wide association mapping for heat and drought adaptive traits in pea. *Genes* 12:1897. doi: 10.3390/genes12121897

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9. doi: 10.3835/plantgenome2015.11.0120

Thudi, M., Upadhyaya, H. D., Rathore, A., Gaur, P. M., Krishnamurthy, L., Roorkiwal, M., et al. (2014). Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *PLoS One* 9:e96758. doi: 10.1371/journal.pone.0096758

Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., et al. (2013). Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat. Biotechnol.* 31, 240–246. doi: 10.1038/nbt.2491

Varshney, R. K., Thudi, M., Roorkiwal, M., He, W., Upadhyaya, H. D., Yang, W., et al. (2019). Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nat. Genet.* 51, 857–864. doi: 10.1038/s41588-019-0401-3

von Wettberg, E. J. B., Chang, P. L., Başdemir, F., Carrasquila-Garcia, N., Korbu, L. B., Moenga, S. M., et al. (2018). Ecology and genomics of an important crop wild relative as a prelude to agricultural innovation. *Nat. Commun.* 9:649. doi: 10.1038/s41467-018-02867-z

Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178, 539–551. doi: 10.1534/genetics.107.074245

Zhang, H., Yasmin, F., and Song, B.-H. (2019). Neglected treasures in the wild — legume wild relatives in food security and human health. *Curr. Opin. Plant Biol.* 49, 17–26. doi: 10.1016/j.pbi.2019.04.004

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fpls.2022.987985

Check for updates

OPEN ACCESS

EDITED BY
Jaroslaw Czubinski,
Poznan University of Life Sciences, Poland

REVIEWED BY
Patrick Olusanmi Adebola,
International Institute of Tropical Agriculture (IITA), Nigeria
Rafiul Amin Laskar,
Bahona College, India

*CORRESPONDENCE
Haitham E. M. Zaki
haitham.zaki@mu.edu.eg;
haithamz.sur@cas.edu.om

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 06 July 2022
ACCEPTED 05 September 2022
PUBLISHED 27 September 2022

CITATION
Zaki HEM and Radwan KSA (2022) Estimates of genotypic and phenotypic variance, heritability, and genetic advance of horticultural traits in developed crosses of cowpea (*Vigna unguiculata* [L.] Walp). *Front. Plant Sci.* 13:987985. doi: 10.3389/fpls.2022.987985

COPYRIGHT
© 2022 Zaki and Radwan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Estimates of genotypic and phenotypic variance, heritability, and genetic advance of horticultural traits in developed crosses of cowpea (*Vigna unguiculata* [L.] Walp)

Haitham E. M. Zaki[1,2]* and Khlode S. A. Radwan[3]

[1]Horticulture Department, Faculty of Agriculture, Minia University, El-Minia, Egypt, [2]Applied Biotechnology Department, University of Technology and Applied Sciences-Sur, Sur, Oman, [3]Plant Pathology Department, Faculty of Agriculture, Minia University, El-Minia, Egypt

Cowpea, in addition to being a food and feed crop, plays a key role in sustainable farming. The present study's goal is to develop new high-yielding cowpea varieties. A Field experiment was carried out across 3 summer seasons and the breeding program included 28 distinct cowpea varieties, out of which five potential parents were selected for this investigation. Local cultivars, i.e., Cream 7 'Cr7', Dokki 331 'D331', Commercial 1 'Com1', and introduced cultivars, i.e., Colossus 'Col' and Asian Introduction 'AI' were utilized to produce six crosses in two generations apart; F_1 and F_2: Col x AI, Col x Com1, Cr7 x AI, Cr7 x Com1, D331 x AI, and D331 x Com1. 'AI' and 'Com1' were superior in pod length, pod diameter, number of seeds/pod and seeds weight/pod, whereas 'Col', 'Cr7' and 'D331' were superior in seeds yield/plant, number of pods/plant and the least number of aborted ovules/pod. The genotypes/crosses showed greater genotypic variance (GV) than phenotypic variance (PV) for number of pods/plant, pod length, number of seeds/pod, number of aborted ovules/pod, fresh pod weight, seeds weight/pod, and seeds yield/plant. All studied variables showed high heritability (H%) in genotypes/crosses, despite the exception of seeds weight/pod, which ranged from 29.14 in 'D331' to 83.7 in F_2 of Col x Com1. F_2 plants and their parents' genotypes showed greater H%. Cr7 x AI developed the most H%, 99.04% for number of pods/plant. D331 x Com1 and Cr7 x AI exhibited moderate H% for fresh pod weight in F_1, but all other crosses

had high H%. F_1 and F_2 crosses yielded moderate to high GCV and PCV for number of seeds/pod. Variations in parental genotypes and crossings reflect genetic diversity and the possibility of selection. Crossing with 'AI,' and 'Com1' genotypes enhanced the performance of the other varieties, 'Col', 'D331' and 'Cr7'. Cr7 x Com1 and D331 x AI were selected as the most promising crosses for cowpea breeding programs.

KEYWORDS

cowpea (*Vigna unguiculata* L. Walp.), cowpea crosses, pod traits, yield and yield components, phenotypic variance, genotypic variance, heritability

1 Introduction

Cowpea (*Vigna unguiculata* L. Walp.) is a diploid species (2n = 22), and it is classified as one of the most pertinent food sources in Africa's arid and semi-arid regions (Damarany, 1994a; Singh et al., 2002; Singh, 2012). Cowpea is a leguminous crop with a massive tendency to increase legume production; it is a member of the Phaseoleae (L.) tribe, a family of Leguminosae, and a self-pollinated dicotyledonous crop plant (Nameirakpam and Khanna, 2018). Cowpea cultivars are classified into five species: *unguiculata, sesquipedalis, textles, melanophtalmus, and biflora*. Crop seeds have a high calcium and iron content, as well as carbohydrates and protein, yet they are low in fat (Pavadai et al., 2009; Nwosu and Awa, 2013). Cowpea proteins are high in tryptophan and lysine when compared to other crop plants. As a result, cowpea constitutes an enormous part of the dietary protein, particularly for people living in tropical areas; in Africa, cowpea is alluded to as "poor man's meat" (Baudoin and Marechal, 2001; Tarawali et al., 2002; Reda et al., 2016). As per FAOSTAT (2019), the world's cowpea production was approximated at 6163 hg/ha, with Africa accounting for 6066 hg/ha, and Egypt accounting for 38748 hg/ha. The area under cowpea cultivation in Egypt is estimated to be 1853 ha. Cowpea is predominantly produced in more than 16 African countries (FAOSTAT, 2020). Because of its high nitrogen fixation ability, it is well adapted to growing in poor soils (Timko and Singh, 2008; EL-Ameen, 2018). It has a degree of tolerance to stress factors. However, salinity is one of the major abiotic stresses that severely affects cowpea crop production and quality (Hall, 2004; Eric et al., 2018). Yield is highly correlated with horticultural traits such as pod number per plant, pod length, and seed number per pod. As a result, any improved performance in these characteristics leads to an increase in yield (Ogunkanmi et al., 2013).

Cowpea genetic diversity has traditionally been assessed by measuring variability in phenotypic traits, which does not always have exceptionally sharp genetic relatedness (Patil et al., 2013; Eric et al., 2018). Furthermore, environmental factors have a strong impact on the interpretation of quantitative traits, limiting knowledge of the germplasm structure for the development of hybrids with specific environmental adaptive responses (Kumar, 1999; Animasaun et al., 2015; John and Clabe, 2016). Meanwhile, cross-breeding amongst cowpea genotypes is a very effective crop enhancement breeding approach. In the 1890s, the first crossings between crop cultivars and wild relatives to obtain disease-resistant varieties were made. Self-pollinating, hybridization to introduce inherited desired traits, and the use of lines and varieties as parents in crossover programs all contribute to cowpea's inherent narrow genetic diversity (Nwosu and Awa, 2013; Patel et al., 2016). Prior research did not include wild relatives because breeders were concerned about their small seed size, seed coat color and texture, and pod shattering (Nkoana et al., 2019; Boukar et al., 2020). At the cellular level, three main processes govern plant growth and development: cell division (mitosis), cell expansion, and cell differentiation (Wang and Ruan, 2013). A cell population's mitotic index has long been considered a crucial characteristic for cell and tissue development and multiplication. One reason for mitotic indexing of species is to generate data for breeding purposes (Darbelley et al., 1989; Driss-Ecole et al., 1994). The ongoing provision of new germplasm material as a donor of numerous agronomically significant genes is a critical condition for future development of cowpea cultivars, particularly given concerns that yield peaks in key crop species, including cowpea, have been achieved. Crossings among different cowpea varieties can surely help to the development of germplasm pools (Willie and Aikpokpodion, 2015). Better genotype development and selection is a crucial long-term method to fighting the problem of low yield in arid areas such as Egypt and other countries. One of the most efficient

Abbreviations: 'AI', Asian Introduction; 'Col', Colossus cv.; 'Com1', Commercial 1; 'Cr7', Cream 7 cv.; 'D331', Dokki 331 cv.; 'PV', Phenotypic Variance; 'GV', Genotypic Variance; 'PCV', Phenotypic Coefficient of Variance; 'GCV', Genotypic Coefficient of Variance; 'H', Heritability; 'GAM', Genetic Advance Mean; 'MI', Mitotic Index.

traditional breeding approaches is systematic germplasm development and evaluation of promising genotypes for adaptability and production stability. The goal of this study was to develop and evaluate new high-yielding cowpea crosses objectively. Meanwhile, the aim was to determine the genotypic and phenotypic variance and heritability of seeds yield and yield components in Egyptian developed cowpea crosses. The horticultural and cytological performance of five cowpea parental genotypes and six generated crosses were evaluated in F_1 and F_2 generations. The current study aimed to improve the characteristics of local commercial cultivars such Cream 7 cv. 'Cr7', Dokki 331 cv. 'D331', and Commercial 1 'Com1' by crossing them with introduced cultivars such Colossus cv. 'Col', and Asian Introduction 'AI'.

2 Materials and methods

2.1 Experimental site

The experiment was carried out at Minia University's Horticulture Department, Faculty of Agriculture, El-Minia, Egypt. The field site is located at latitude 28°7'N and longitude 30°43'E. The research was carried out over three summer seasons in 2016, 2017, and 2018. Physical and chemical analyses of soil collected from a depth of 0.0 to 30 cm were performed over the seasons, and the average results are shown in Supplementary Table 1.

2.2 Parental genotypes

The present investigation searched for five different cowpea genotypes (*Vigna unguiculata* L. Walp). Local commercial cultivars, i.e., Cream 7 cv. 'Cr7', Dokki 331 cv. 'D331', and introduced cultivars, i.e., Colossus cv. 'Col' and Asian Introduction 'AI'. Another cultivar, Commercial 1 'Com1', was collected from the local market in El-Minia governorate for its seed quality characteristics. Twenty-eight cowpea varieties, including the current genotypes, were investigated for over 8 seasons in a comprehensive study undertaken by the author. Meanwhile, these genotypes were selected because of genetic variation in morphological, floral, pod, and seed traits (Figure 1), and genotypes such as 'Cr7' and 'D331' are commonly farmed in Egypt. All genotypes were selected for evaluation and cross experiments in the F_1 and F_2 generations. Genotype seeds were obtained from the Horticulture Department of the Faculty of Agriculture at Minia University in El-Minia, Egypt. The list of the examined genotypes with flower color, seed coat color, seed eye color, source and desirable traits is shown in Table 1.

FIGURE 1
(A) flower traits, **(B)** peduncle length, **(C)** pod traits and **(D)** seed traits of five different cowpea parental genotypes examined under current investigation, i.e., and Asian Introduction, Colossus cv., Commercial 1, Cream 7 cv., and Dokki 331.

2.3 Crossing experiment and evaluation

In this experiment, crossing among five cowpea genotypes, i.e., 'Cr7', 'D331', 'Col', 'AI', and 'Com1', was done to produce six crosses, which were studied in advanced generations of F_1 and F_2. Cultivars, 'AI' and 'Com1' were used as pollen donor parents, whereas the other cultivars, 'Col', 'Cr7' and 'D331' were employed as maternal parents. Crossing was carried out in the early morning by removing all anthers to prevent self-pollination and cutting all buds on the peduncle, followed by the application of pollen grains collected from donor plants to the pistil of the emasculated flowers. Wet cotton was used to cover the area of the removed buds, and paper bags were used to avoid cross-contamination with any foreign pollen grains. In 2016, through the crossing, F_1 seeds of six crosses were obtained. In the meantime, F_1 seeds were cultivated to produce F_2 plants. The six crossings were evaluated for their morphological, yield, and yield component traits. Genotypes and crosses were distributed in plots, and each plot consisted of five rows (4 m length x 0.7 m wide), whereas the plot area was 14 m^2. The inner three ridges were used for sampling, and the two outer ridges were left as guard ridges. The genotypes

TABLE 1 List of the tested parental genotypes with their flower color, seed color, source and desirable traits.

Parental genotypes*	Flower color	Seed coat color	Seed eye color	Source	Desirable traits
Paternal parents					
Asian Introduction 'AI'	Purple	Cupreous	Colorless	Bangladesh	Pod length and number of seeds/pod
Commercial 1 'Com1'	White	White	Brown	Egyptian Market	Pod diameter, number of seeds/pod, seeds weight/pod
Maternal parents					
Colossus 'Col'	Purple	Brown	Colorless	USA	Yield of seeds/plant
Cream 7 'Cr7'	White	Cream	Colorless	Egypt	number of pods/plant
Dokki 331 'D331'	White	White	Black	Egypt	Less number of aborted ovules/pod

*The study included five parental genotypes; local commercial cultivars, i.e., Cream 7 'Cr7', Dokki 331 'D331', and introduced cultivars, i.e., Colossus 'Col' and Asian Introduction 'AI'. Another cultivar, Commercial 1 'Com1', was collected from the local market for its seed's quality characteristics.

and crosses were organized using a Randomized Complete Block Design (RCBD). Each genotype was represented by a single plot, which was repeated three times. During the three seasons, seeds were sown at a rate of 2 seeds per hole, with a spacing of 25 cm between holes. During harvesting, crosses plants were sampled, and data was collected. All other agricultural practices were in accordance with commercial production guidelines.

2.4 Data collection and analysis

At harvesting time, plants and pods were chosen at random, and the morphological, floral, pods, seeds, and yield traits were examined.

2.4.1 Morphological traits

Thirty plants were randomly selected from each plot of each genotype/cross, and the average shoot length, number of branches/plant, and stem diameter were investigated.

2.4.2 Flower traits

The average length and diameter and number of peduncles/plant and flower length were examined in thirty plants which were randomly harvested from each genotype/cross.

2.4.3 Pod and seed traits

The average number of pods/plant, seed weight, seed width, and seed length were calculated for thirty plants for each genotype/cross. Meanwhile, the average pod length, pod diameter, pod weight, number of seeds/pod, and number of aborted ovules/pod of ten pods were measured.

2.4.4 Yield

The weight of the seeds per plant were measured in thirty plants randomly selected from each plot for each parental genotype and cross.

2.4.5 Genetic analysis

The genotypic and phenotypic variance were estimated by using the following equations:

a. Genotypic variance

$$\sigma^2 g = \frac{MSt - MSe}{r}$$

Whereas:
MSt = Mean sum of squares for trait of genotype/cross.
MSe = Mean sum of squares for error of genotype/cross.
r = Number of replications.

b. Phenotypic variance

$$\sigma^2 p = \sigma^2 g + \sigma^2 e$$

Whereas:
σ^2p = Phenotypic variance for each trait of genotype/cross.
σ^2g = Genotypic variance for each trait of genotype/cross.
σ^2e = Environmental variation among the tested traits of genotype/cross.

c. Phenotypic coefficient of variance and genotypic coefficient of variance

The PCV and GCV expressed as percentages were calculated as suggested by Burton and Vane (1953). In the meantime, PCV and GCV were classified into three classes; less than 10% (Low), 10 – 20% (Moderate) and more than 20% (High).

$$\text{Phenotypic Coefficient of Variance (PCV)} = \frac{\sqrt{Phenotypic\ Variance}}{Mean} \times 100$$

$$\text{Genotypic Coefficient of Variance (GCV)} = \frac{\sqrt{Genotypic\ Variance}}{Mean} \times 100$$

d. Heritability in broad sense

Heritability in broad sense meaning was estimated as the ratio of genetic variance to the phenotypic variance as reported by Burton (1952) as follow:

$$\text{Heritability (road sense)} = \frac{Genotypic\ \ Variance}{Phenotypic\ \ Variance} \times 100$$

It was categorized according to Robinson et al. (1949) to three classes: 0.0-30% (Low), 31-60% (Medium) and more than 60% (High).

e. Genetic advance

Genetic advance as a percent of mean (GAM) was estimated and categorized as reported by Johnson et al. (1955) by the following formula:

$$\text{GAM\%} = \frac{K * H * p}{Mean} \times 100$$

Whereas:

K = 2.06 at 5% selection intensity.

H = Heritability.

P = Phenotypic standard deviation.

Meanwhile, GAM was categorized to three classes: less than 10% (Low), 10-20% (Moderate) and more than 20% (High).

2.4.6 Cytological analysis

Cowpea seeds from the studied plant materials which included five parental genotypes and six crossings in F_1 and F_2 were germinated in Petri dishes with two layers of moist filter paper at room temperature for 48 hrs. Roots with a length of 1-2 cm were cut and fixed for 24 hrs. in a newly produced farmer's fixative solution (absolute ethyl alcohol: glacial acetic acid, 3:1 v/v). The fixed roots were stored in 70% ethyl alcohol in the refrigerator at 4°C until analysis. The fixed roots were rinsed with distilled water, then hydrolyzed in 1 N HCl at 60°C for 10 min before being washed again. Mitotic investigations were conducted using the aceto-carmine squash preparation. For each genotype, almost 1000 cells were investigated (consisting of ten seeds). Images were captured with an Olympus BX51 microscope and a C-4040 zoom digital camera whenever possible. Mitotic index, phase index, and chromosomal aberrations were recorded for each genotype, and mitotic index was computed using Racuciu (2009) formula:

$$\text{Mitotic index} = \frac{Total\ \ number\ \ of\ \ divided\ \ cells}{Total\ \ number\ \ of\ \ examined\ \ cells} \times 100$$

Percentage of abnormality of each stage of mitosis was counted for each slide.

$$\text{Percentage of abnormality} = \frac{Total\ \ number\ \ of\ \ abnormal\ \ cells}{Total\ \ number\ \ of\ \ examined\ \ cells} \times 100$$

2.5. Statistical analysis

Data obtained from this study were subjected to analysis using SAS, version 9.3 (Cary, NC). Differences among cowpea genotypes were tested by an analysis of variance (ANOVA), and mean significant differences were tested by the Least Significant Difference (LSD) test at the 0.05 level of significance.

3 Results

3.1 Morphological traits

Morphological variability in shoot length, number of branches/plant, and stem diameter was noted among F_1 and F_2 generation crosses, as shown in Table 2. The shoot lengths of the parental genotypes ranged from 106.5 cm for 'D331' cv. with a range of 82-137 cm to 184.7 cm for 'AI' with a range of 167-210 cm. All six developed crossings produced F_1 plants that were taller than F_2 plants. F_1 of Cr7 x AI cross had the shortest shoot length with a mean value of 122.7 cm, while D331 x AI cross had the largest shoot length with a mean value of 234.0 cm and exceeded the better parent. On the other hand, D331 x AI cross had the shortest shoot length of all F_2 plants (133.3 cm), while Col x Com1 cross was superior and had the longest shoot (193.7 cm). Furthermore, the parental cultivar 'D331' had the highest number of branches per plant (7.4 branches) with a range of 4.0–12 branches, whilst 'AI' was the least branching with a mean value of 2.9 and a range of 2.0–4.0 branches. In comparison, F_1 of Col x Com1 cross and F_2 of Cr7 x AI cross produced the highest values of 9.0 and 5.5, with a range of 7.0–11 and 3–9 branches, respectively (Table 2). In the meantime, the mean stem diameter varied among the parental genotypes and crosses. 'Cr7' cv. produced a thicker diameter (1.9 cm) than other parental genotypes and crosses, with the exception of Col x Com1 cross in F_1 plants, which had the largest stem diameter (2.0 cm). At the same time, the stem diameter of the obtained crosses ranged from 0.6 cm for cross Cr7 x AI with an average of 0.7 cm to 2.1 cm for cross Col x Com1 with an average of 2.0 cm in F_1s, when it varied from 0.9 cm for Cr 7 x AI, Cr7 x Com 1, Col x AI, and D331 x AI to 2.4 cm for D331 x Com1 with an average of 1.6 cm in F_2s. The data obtained, as given in Table 2, demonstrated that none of the developed crossings surpassed the diameter of the F_1 cross's Col x Com1 (2.0 cm).

3.2 Flower traits

In terms of peduncle length, F_1 and F_2 of Cr7 x Com1 cross outperformed the other crosses, with a mean of 51.9 and 40.2 cm and a range of 31.0-66.7 cm and 20.4-62.6 cm, respectively. Nevertheless, this cross (Cr7 x Com1) produced a taller peduncle than the better parent, 'D331' cv., which had a peduncle length of 33.2 cm. The peduncle diameter of the parental genotypes, on the other hand, ranged from 0.3 to 0.4 cm, and better parents were 'Col' cv. and 'Com1' cv. The peduncle diameter of F_1 plants ranged from 0.2 cm in Cr7 x AI cross to 1.1

TABLE 2 Mean and range of shoot length (cm), number of branches, stem diameter (cm), peduncle length (cm), peduncle diameter (mm), number of peduncles/plant, and flower length (cm) of six cowpea crosses produced from five different genotypes in F_1, F_2 generations.

Genotypes/Crosses	Shoot length (cm)		Number of branches/ plant		Stem diameter (cm)		Peduncle length (cm)		Peduncle diameter (mm)		Number of Peduncles/ plant		Flower length (cm)	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Parental genotypes*														
Asian Introduction (AI)	184.7	167-210	2.9	2-4	1.1	1.0-1.3	24.4	17.3-31.3	0.3	0.3-0.4	18.2	10-23	2.7	2.5-2.9
Colossus (Col)	134.0	100-157	5.5	4-9	1.7	1.5-2.3	25.8	19.5-34.5	0.4	0.4-0.6	30.0	22-43	2.1	2.0-2.3
Commercial 1 (Com1)	139.2	115-190	4.1	3-6	1.3	1.0-1.8	29.1	21.0-40.5	0.4	0.3-0.6	27.8	20-62	2.3	2.0-2.5
Cream 7 (Cr7)	133.2	87-172	6.4	5-8	1.9	1.5-2.2	32.0	23.0-40.0	0.3	0.3-0.5	48.9	30-74	2.0	2.0-2.2
Dokki 331 (D331)	106.5	82-137	7.4	4-12	1.7	1.2-2.5	33.2	21.5-47.0	0.3	0.3-0.5	39.0	23-79	2.0	1.8-2.2
Crosses (F_1)														
Col x AI	185.1	165-206	6.1	5-8	1.4	1.2-1.8	30.0	21.0-44.0	0.4	0.3-0.6	39.3	35-43	2.3	1.9-3.1
Col x Com1	220.5	189-252	9.0	7-11	2.0	1.9-2.1	42.0	41.0-43.0	1.1	0.6-1.7	55.0	48-62	2.0	1.7-2.1
Cr7 x AI	122.7	100-145	5.0	4-7	0.7	0.6-1.0	25.8	18.0-34.0	0.2	0.2-0.4	30.1	24-44	2.1	1.3-2.8
Cr7 x Com1	158.4	80-300	7.8	4-11	1.5	0.9-2.0	51.9	31.0-66.7	0.5	0.4-0.7	63.4	24-167	2.1	1.5-2.9
D331 x AI	234.0	233-235	8.5	6-11	1.3	1.0-1.6	24.4	10.1-30.0	0.4	0.4-0.5	49.4	32-69	1.9	1.7-2.2
D331 x Com1	146.4	125-212	6.8	5-9	1.6	1.2-1.9	37.5	35.5-39.5	0.3	0.3-0.4	41.5	33-50	2.1	1.8-2.9
Crosses (F_2)														
Col x AI	183.0	125-264	5.0	3-8	1.2	0.9-2.1	37.2	23.2-49.2	0.3	0.3-0.5	28.3	15 -47	2.6	2.3-3.0
Col x Com1	193.7	120-271	5.0	3-8	1.4	1.1-1.8	36.6	23.1-51.8	0.5	0.4-0.7	29.8	19-41	2.5	2.2-3.1
Cr7 x AI	163.1	92-231	5.5	3-9	1.2	0.9-2.0	37.2	23.2-49.2	0.4	0.3-0.6	36.4	22-61	2.5	2.2-3.0
Cr7 x Com1	153.5	88-223	5.2	3-9	1.4	0.9-2.0	40.2	20.4-62.6	0.4	0.3-0.8	34.0	18-82	2.6	2.3-3.0
D331 x AI	133.3	86-198	4.8	3-7	1.3	0.9-2.0	34.4	23.4-58.8	0.4	0.3-0.8	27.8	17-41	2.5	2.0-3.0
D331 x Com1	157.4	105-289	5.3	3-8	1.6	1.1-2.4	35.5	19.4-50.7	0.4	0.3-0.6	34.2	22-51	2.5	2.3-3.0

*The study was conducted over three summer seasons in 2016, 2017, and 2018. F_1 seeds from six crosses were obtained through crossing in 2016 meanwhile, F_1 seeds were grown in 2017 to produce F_2 seeds in 2018. The mean and range of the tested parents' scores, as well as statistics from the three growing seasons.

cm in Col x Com1 cross. F_2 of Col x Com1 cross produced the largest peduncle diameter, with an average of 0.5 cm and a range of 0.4 to 0.7 cm. These crosses outperformed the better parents, as indicated in Table 2.

The better parental genotype, 'Cr7' cv., produced the most peduncles (48.9) with a range of 30–74, whereas 'AI' produced the fewest peduncles (18.2) with a range of 10–23. In terms of the developed six crosses, the crosses in F_1s had more peduncles than the crosses in F_2s. F_1 of Cr7 x Com1 cross produced the most peduncles (63.4) with a range of 24–167, outnumbering all other crosses and parental genotypes, while F_2 of D331 x AI cross produced the fewest, still more than the least parent 'AI'. Flower length also differed between parental genotypes and the six produced crosses in F_1 and F_2 generations. In general, when compared to other genotypes and crosses, 'AI' had the longest flower length. In contrast to the previously mentioned characteristics, F_2 crosses had longer flower lengths than F_1 crosses. The mean flower length in F_1 and F_2 crosses varied widely, from 2.3 cm for Col x AI cross to 2.6 cm for Cr7 x Com1 and Col x AI crosses. None of the tested crosses outperformed the flower length of the better genotype 'AI' (2.7 cm). On the other hand, F_1 cross between the parental genotypes 'D331' and 'Com1' had the shortest flower length (1.9 cm) as shown in Table 2.

3.3 Pods and seed traits

Parental genotypes and all crosses in F_1 and F_2 generations showed considerable variations in pods and seed traits (Figures 2, 3). The mean number of pods per plant of the developed crosses in this experiment differed from 21.5 for the cross Cr7 x AI to 71.0 for the cross D331 x AI in F_1s, and from 38.7 for the cross Cr7 x Com1 to 48.2 for the cross Cr7 x AI in F_2s. As reported in Table 3, the F_1 of the crossing between 'D331' cv. and 'AI' outperformed the better parent, 'Cr7' cv. (58.9). The paternal parents, 'AI' and 'Com1' improved the pod traits in the six crosses. According to the approximated pod length values, F_1 of the crossing between 'D331' cv. and 'AI' and F_2 of the crossing between 'Col' and 'AI' showed the longest pod length (21.0 and 21.6 cm, respectively). However, none of the crosses exceeded the better parent, 'AI', which had the tallest pod with an average of 34.2 cm and a range of 24.0 to 42.0 cm (Table 3). All produced crossings in F_1s and F_2s had thicker pods than the parental genotypes in terms of pod diameter, which was particularly noticeable in F_1 crosses. Pod diameter ranged from 0.3 to 0.5 cm in the parental genotypes, while it ranged from 0.6 to 0.8 cm in F_1 crossings and from 0.6 to 0.7 cm in F_2 crosses. The maximum pod diameter was found in F_1 crosses of Cr7 x AI, Col x Com1, and Col x AI (Table 3).

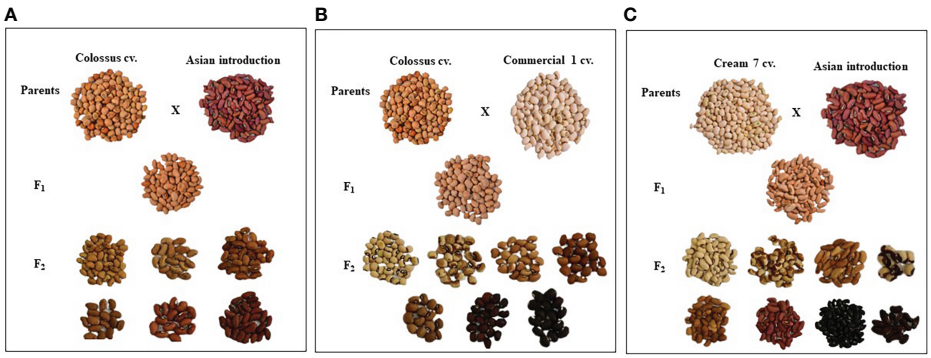


FIGURE 2
The diversification of seed characteristics of the parental genotypes, F_1 and F_2 of the crosses, **(A)** Colossus cv. x Asian Introduction, **(B)** Colossus cv. x Commercial 1and **(C)** Cream 7 cv. x Asian Introduction.

When it came to fresh pod weight, there were substantial differences among the tested genotypes and the six developed crosses in F_1 and F_2. The best parent was 'Com1' with an average of 3.7 g per pod, followed by 'Col' cv. with an average of 3.3 g per pod, whilst 'D331' cv. and 'Cr7' cv. were the least parents with an average of 2.2 and 1.6 g per pod, respectively. Meanwhile, 'Com1' cv. had the maximum average pod weight, followed by 'Col' cv., with no significant differences however, 'Cr7' cv. had the lowest pod weight values across the three seasons (Table 3). In F_1 crossings, pod weights varied from 1.6 g for Cr7 x AI cross to 2.7 g for Col x Com 1 cross, whereas, in F_2 crossings, pod weights ranged from 1.3 g for D331 x Com 1 cross to 3.0 g for Col x Com 1 cross. 'Com 1' was the better parent, with the greatest pod weight of 3.7 g on average and a range of 2.5–4.7 g. Meanwhile, as seen in Table 3, none of the crosses outperformed the superior parent.

In parental genotypes, the number of aborted ovules ranged from 4.1 to 8.6 ovules/pod, while in F_1 and F_2 crossings, the number ranged from 4.1 to 7.8 ovules/pod. 'D331' cv. yielded the greatest results and had the fewest aborted ovules/pod (4.1 ovules/pod), while the F_2 cross of Col x Com1 yielded the same result. The largest number of aborted ovules were found in 'AI' and F_1 of Cr7 x AI cross, with 8.6 and 7.8, respectively (Table 3).

Table 4 reveals that there were significant differences in seed length, seed width, and seed weight/pod between genotypes and crosses in F_1 and F_2.The parental genotype 'AI' and F_1 of D331 x AI cross had the greatest value (1.0 cm in seed length). On the other hand, 'Cr7' cv. and F_2 of Cr7 x Com1 cross had the lowest value (0.7 cm in seed length). At the same time in this trial, the mean seed width of the six crossings produced increased from 0.4 to 0.5 in F_1s and F_2s. In comparison to the other crosses, F_1 and F_2 of Col x Com1 cross generated the greatest values of 0.5 cm, but the crosses did not surpass the better parent, 'Col' cv. which had 0.6 cm in seed width (Table 4). Crossing among the parents improved the number of seeds/pod trait and it was considerably clear in the crosses in F_2s (Table 4). The paternal

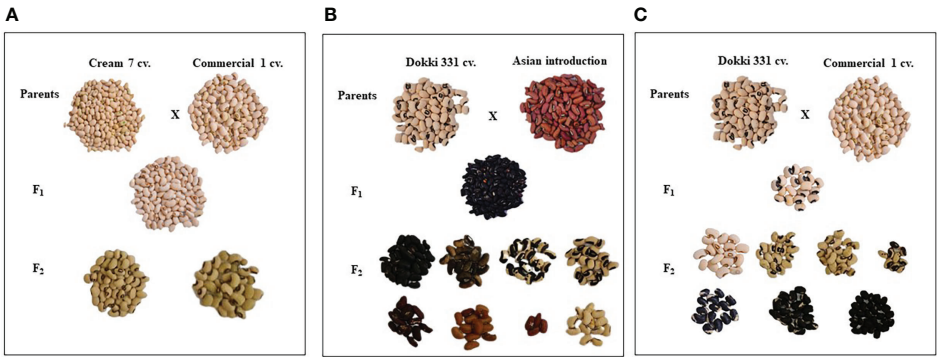


FIGURE 3
The diversification of seed characteristics of the parental genotypes, F_1 and F_2 of the crosses, **(A)** Cream 7 cv. x Commercial 1, **(B)** Dokki 331 cv. x Asian Introduction and **(C)** Dokki 331 cv. x Commercial 1.

TABLE 3 Mean and range of number of pods/plant, pod length (cm), pod diameter (cm), fresh pod weight (g), and number of aborted ovules/pod of six cowpea crosses produced from five different genotypes in F_1, F_2 generations.

Genotypes/Crosses	Number of pods/plant		Pod length (cm)		Pod diameter (cm)		Fresh pod weight (g)		Number of aborted ovules /pod	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Parental genotypes*										
Asian Introduction (AI)	18.3	10.0-27.0	34.2	24.0-42.0	0.3	0.3-0.5	2.3	0.7-3.4	8.6	3.0-13.0
Colossus (Col)	27.0	15.0-42.0	13.2	11.0-15.1	0.5	0.5-0.6	3.3	1.5-4.5	5.6	3.0-11.0
Commercial 1 (Com1)	30.1	16.0-90.0	19.4	14.7-23.5	0.5	0.5-0.6	3.7	2.5-4.7	5.1	2.0-9.0
Cream 7 (Cr7)	58.9	21.0-94.0	13.6	9.7-17.0	0.4	0.4-0.5	1.6	0.7-2.7	6.3	3.0-10.0
Dokki 331 (D331)	35.3	17.0-74.0	13.7	10.1-16.0	0.4	0.4-0.5	2.2	1.1-2.9	4.1	2.0-8.0
Crosses (F_1)										
Col x AI	41.1	37.0-55.0	19.4	8.6-29.5	0.8	0.6-0.9	2.1	0.6-3.7	7.3	3.0-12.0
Col x Com1	30.0	22.0-38.0	16.1	10.2-21.5	0.8	0.7-1.0	2.7	0.6-4.7	5.2	2.0-13.0
Cr7 x AI	21.5	14.0-34.0	19.6	13.3-29.0	0.8	0.7-1.0	1.6	0.9-2.7	7.8	5.0-12.0
Cr7 x Com1	55.5	33.0-88.0	17.6	12.4-20.3	0.7	0.6-0.8	2.4	1.0-3.2	4.5	2.0-7.0
D331 x AI	71.0	55.0-87.0	21.0	16.6-25.8	0.7	0.7-0.8	2.6	1.8-3.6	4.6	2.0-8.0
D331 x Com1	49.8	42.0-58.0	13.3	10.1-15.7	0.6	0.6-0.8	1.7	1.0-2.7	4.8	2.0-10.0
Crosses (F_2)										
Col x AI	40.3	23.0-96.0	21.6	11.5-44.3	0.6	0.5-0.9	2.3	0.6-4.1	5.7	2.0-11.0
Col x Com1	38.8	28.0-53.0	18.0	11.2-22.8	0.7	0.6-0.9	3.0	1.6-5.4	4.1	0.0-11.0
Cr7 x AI	48.2	23.0-69.0	19.1	7.4-37.3	0.6	0.5-1.0	1.5	0.3-3.2	5.6	2.0-13.0
Cr7 x Com1	38.7	15.0-74.0	15.5	9.4-20.5	0.6	0.5-1.1	1.9	0.4-3.2	4.7	1.0-14.0
D331 x AI	40.8	28.0-61.0	20.9	13.2-26.3	0.7	0.5-0.9	2.8	1.5-4.2	5.1	0.0-9.0
D331 x Com1	39.7	27.0-58.0	13.8	8.9-22.7	0.6	0.4-1.0	1.3	0.6-2.7	5.3	2.0-9.0

*The study was conducted over three summer seasons in 2016, 2017, and 2018. F^1 seeds from six crosses were obtained through crossing in 2016 meanwhile, F_1 seeds were grown in 2017 to produce F_2 seeds in 2018. The mean and range of the tested parents' scores, as well as statistics from the three growing seasons.

genotype, 'AI' was the better parent for the number of seeds per pod, with an average of 11.0 seeds and a range of 3.0-15.0 seeds. Four of the developed crosses, D331 x AI and Cr7 x Com1 in F_1 and Col x Com1 and D331 x AI in F_2 crosses, were higher than the better parent. Nevertheless, F_2 of Col x Com1 cross, followed by F_1 and F_2 of D331 x AI cross, produced the highest number of seeds/pod (11.9 and 11.7 seeds/pod, respectively). Figures 2, 3 illustrated the seed traits of the six produced crosses. Parental genotypes and developed crosses had a considerable variation in seed color, seed coat color and size. Furthermore, the seed weight/pod of the parental genotypes varied from 1.3 g for 'AI' to 2.7 g for 'Com1'. For crosses between parental genotypes, they ranged from 0.9 g for F_1 of Cr7 x AI cross to 2.0 g for F_1 of Col x Com1 cross. However, in F_2, the seed weight ranged from 0.9 g for D331 x Com1 cross to 2.2 g for Col x Com1 cross. Meanwhile, Col x Com1 cross was superior in the two generations but still weighted less than the better parents, 'Com1' cv. (2.7 g) and 'Col' cv. (2.6 g) as clear in Table 4.

3.4 Seed yield

The seed weight per plant produced by the F_1 and F_2 generations of crossings, as well as the parental genotypes, indicated significant variation. In parental genotypes, 'Com1' and 'Col' cv. yielded higher seed weight than the others, with mean values of 24.7 and 23.1 g/plant, respectively, while 'AI' produced the least of all genotypes and examined crosses. Crosses in F_1 had a higher seed weight than crosses in F_2, with the exception of Cr7 x AI cross, which had the lowest seed weight of all the crosses. D331 x AI cross exceeded the parental genotypes as well as the other crosses in F_1 with an average of 80.7 g/plant and a range of 55.4-106.1 g/plant. Col x AI had heavier seeds in F_2s than other crosses with an average of 27.7 g/plant and a range from 10.2-59.0 g/plant, which was still greater than the better parents (Table 4).

3.5 Genetic parameters analysis

Tables 5, 6 present the findings of the genetic investigation among the parental genotypes, as well as the crosses seen in F_1 and F_2. The successful crosses were accomplished using 'Cr7' cv., 'Col' cv. and 'D331' cv. as female parents and 'AI' and 'Com1' genotypes as male parents. Meanwhile, phenotypic variance (PV), genotypic variance (GV), phenotypic coefficient of variance (PCV %), genotypic coefficient of variance (GCV %), heritability (H %), and genetic advance mean (GAM) of number

TABLE 4 Mean and range of seed length (mm), seed width (mm), number of seeds/pod, seeds weight/pod (g) and seeds weight/plant (g) of six cowpea crosses of F_1, F_2 generations which produced from five different genotypes.

Genotypes/Crosses	Seed length (mm)		Seed width (mm)		Number of seeds/pod		Seeds weight/pod (g)		Seeds weight/plant (g)	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Parental genotypes*										
Asian Introduction (AI)	1.0	0.9-1.1	0.3	0.3-0.4	11.0	3.0-15.0	1.3	0.1-2.3	13.3	10.1-18.5
Colossus (Col)	0.8	0.7-0.9	0.6	0.6-0.7	10.4	5.0-15.0	2.6	1.0-4.0	23.1	17.2-27.2
Commercial 1 (Com1)	0.9	0.8-1.2	0.5	0.5-0.6	10.3	6.0-16.0	2.7	1.7-3.7	24.7	18.6-31.6
Cream 7 (Cr7)	0.7	0.7-0.8	0.4	0.4-0.6	8.1	3.0-11.0	1.0	0.3-1.9	16.2	11.8-21.4
Dokki 331 (D331)	0.9	0.8-1.2	0.5	0.4-0.6	8.6	4.0-12.0	1.8	0.8-2.4	17.7	12.0-23.0
Crosses (F_1)										
Col x AI	0.9	0.8-1.2	0.4	0.4-0.6	7.7	2.0-12.0	1.3	0.3-2.6	43.6	38.1-55.9
Col x Com1	0.8	0.8-1.0	0.5	0.4-0.6	9.9	3.0-16.0	2.0	0.7-2.1	46.1	27.9-46.3
Cr7 x AI	0.9	0.8-1.1	0.4	0.3-0.5	7.6	5.0-12.0	0.9	0.5-1.4	15.4	8.3-26.6
Cr7 x Com1	0.8	0.8-1.0	0.4	0.4-0.5	11.3	5.0-15.0	1.7	0.7-2.5	67.2	35.5-106.6
D331 x AI	1.0	0.8-1.1	0.4	0.4-0.5	11.7	7.0-15.0	1.8	0.8-2.8	80.7	55.4-106.1
D331 x Com1	0.8	0.7-1.1	0.4	0.4-0.5	7.5	4.0-10.0	1.3	2.1-0.7	53.7	47.7-61.9
Crosses (F_2)										
Col x AI	0.9	0.8-1.3	0.4	0.4-0.6	11.0	3.0-18.0	1.7	0.3-3.1	27.7	10.2-59.0
Col x Com1	0.8	0.7-0.9	0.5	0.4-0.6	11.9	6.0-20.0	2.2	0.9-4.4	22.9	11.1-51.7
Cr7 x AI	0.8	0.7-1.2	0.4	0.3-0.5	8.8	3.0-16.0	1.0	0.1-2.4	23.9	11.7-42.5
Cr7 x Com1	0.7	0.6-1.0	0.4	0.4-0.5	10.0	2.0-16.0	1.4	0.2-2.6	17.5	5.0-39.0
D331 x AI	0.9	0.9-1.1	0.4	0.4-0.5	11.7	7.0-17.0	2.1	0.9-3.4	23.1	12.1-40.1
D331 x Com1	0.8	0.6-1.1	0.4	0.3-0.7	7.2	2.0-13.0	0.9	0.2-2.0	19.3	7.6-33.8

*The study was conducted over three summer seasons in 2016, 2017, and 2018. F^1 seeds from six crosses were obtained through crossing in 2016 meanwhile, F_1 seeds were grown in 2017 to produce F_2 seeds in 2018. The mean and range of the tested parents' scores, as well as statistics from the three growing seasons.

of pods/plant, pod length, number of seeds/pod, number of aborted ovules/pod, fresh pod weight, seeds weight/pod and seeds yield/plant of the five parental genotypes and six crosses in F_1 and F_2 were evaluated. Other morphological and floral parameters, such as shoot length, number of branches/plant, peduncle length, and number of peduncles/plant, were also genetically analyzed for parental genotypes and crosses as clear in Supplementary Tables 2, 3, respectively.

The genetic analysis revealed significant differences between the parental genotypes (Table 5) and obtained crosses (Table 6). For the parental genotypes, high and moderate genotypic and phenotypic coefficients of variance (PCV %) were obtained for the number of pods/plant. In F_1 and F_2, the genotypic variance (GV) was greater than the phenotypic variance (PV) in the five parents and the six crossings. Simultaneously, GCV % had greater impacts than PCV % in the studied characteristics. All parents and six crossings in F_1 and F_2 had high heritability values. GAM ranges were found to be broad for all parents and crosses in F_1 and F_2. Heritability values for the parental genotypes and F_1 and F_2 crossings ranged from 76.68% to 96.96%, with a high H% for pod length. With the exception of 'D331' cv., which showed low values for PCV %, GCV %, and GAM, all parental genotypes had moderate PCV % and GCV %, whereas the six crossings in F_1 and F_2 had high GAM values for the trait, as shown in Tables 5, 6.

The H % and GAM % of the number of seeds/pod were both high in all parents and the resulting F_1 and F_2 crosses (Tables 5 and 6). The estimations of PCV % were lower in all genotypes than the estimations of GCV %, indicating that this trait is influenced by the environment. Furthermore, the PCV %, GCV %, H %, and GAM of the number of aborted ovules per pod were all high in the five parents (Table 5). When compared to the other parental genotypes, 'AI' and 'Col' cv. demonstrated greater values. All F_1 and F_2 crossings yielded high values for the genetic characteristics tested for the number of aborted ovules/pod. F_1 of Col x AI cross and F_2 of Cr7 x Com1 cross had the highest H % with 91.71% and 94.49%, respectively (Table 6).

PCV %, GCV %, H %, and GAM of fresh pod weight were all high for all parental genotypes excluding the 'D331' cv., which had low values for PCV %, H %, and GAM, 8.44%, 18.07%, and 6.70%, respectively. All crosses in F_1 and F_2 had high H % values, other than F_1 crossings D331 x Com1 and Cr7 x AI, which had moderate H % values of 45.18% and 56.87%, respectively. GAM of all F_1 and F_2 crosses revealed elevated fresh pod weight values. At the same time, all F_1 and F_2 crosses showed high PCV % and GCV %, apart from D331 x Com1 and Cr7 x Com1, which had

TABLE 5 Genotypic variance (GV), phenotypic variance (PV), genotypic coefficient of variance (GCV %), phenotypic coefficient of variance (PCV %), heritability (H%), and genetic advance mean (GAM) of pods, seeds traits and yield of the five parental genotypes; local commercial cultivars, i.e., Cream 7 'Cr7', Dokki 331 'D331', and introduced cultivars, i.e., Colossus 'Col' and Asian Introduction 'AI'. Another cultivar, Commercial 1 'Com1', was collected from the local market in El-Minia governorate for its seed's quality characteristics.

Parental genotypes	Traits*	GV	PV	GCV %	PCV %	H %	GAM
AI	NPo/P	32.011	30.22	30.91	30.04	94.41	60.13
	PoL	26.3	24.7	14.98	14.51	93.84	28.97
	NS/Po	19.5	18.1	40.20	38.73	92.85	76.89
	NAO/Po	13.6	12.4	42.88	41.00	91.42	80.76
	FPoW	0.62	0.37	33.70	26.13	60.12	41.76
	SW/Po	0.57	0.33	55.64	42.49	58.33	66.86
	SY/P	103.91	100.69	32.83	32.31	96.89	65.53
Col cv.	NPo/P	498.32	491.26	74.16	73.63	98.58	150.61
	PoL	5.44	4.71	12.01	11.17	86.45	21.39
	NS/Po	8.23	7.32	27.85	26.27	88.98	51.05
	NAO/Po	5.21	4.48	44.76	41.54	86.14	79.43
	FPoW	0.69	0.42	22.13	17.42	61.99	28.27
	SW/Po	0.50	0.27	25.48	18.95	55.35	29.06
	SY/P	123.06	119.55	35.13	34.62	97.14	70.30
Com1	NPo/P	66.44	63.86	30.19	29.59	96.12	59.77
	PoL	1.84	1.41	10.27	8.99	76.68	16.22
	NS/Po	11.8	10.7	33.06	31.50	90.80	61.83
	NAO/Po	5.37	4.64	41.41	38.48	86.37	73.67
	FPoW	1.35	0.98	34.94	29.80	72.76	52.36
	SW/Po	1.14	0.80	39.87	33.45	70.40	57.84
	SY/P	1206.4	1195.4	71.13	70.81	99.08	145.2
Cr7 cv.	NPo/P	275.12	269.87	46.98	46.53	98.09	94.95
	PoL	4.34	3.68	15.13	13.93	84.82	26.44
	NS/Po	8.71	7.77	34.31	32.42	89.28	63.11
	NAO/Po	3.65	3.05	46.63	42.60	83.45	80.16
	FPoW	0.38	0.18	26.93	18.79	48.69	27.00
	SW/Po	0.27	0.10	28.92	18.08	39.11	23.27
	SY/P	131.93	128.30	41.25	40.68	97.24	82.64
D331 cv.	NPo/P	744.3	735.6	46.31	46.05	98.84	94.31
	PoL	3.62	3.02	13.99	12.77	83.37	24.03
	NS/Po	7.43	6.57	33.65	31.64	88.40	61.28
	NAO/Po	6.90	6.06	41.69	39.10	87.95	75.55
	FPoW	0.30	0.13	33.83	22.09	42.62	29.69
	SW/Po	0.19	0.05	41.38	22.34	29.14	24.84
	SY/P	216.02	211.38	52.57	52.00	97.84	105.9

*Number of pods/plant (NPo/P), pod length (PoL), number of seeds/pod (NS/Po), number of aborted ovules /pod (NAO/Po), fresh pod weight (FPoW), seeds weight/pod (SW/Po) and seeds yield/plant (SY/P).

moderate PCV % values of 16.01% and 19.75%, respectively (Tables 5, 6).

The H % of seeds weight per pod was low, medium, and high for the parental genotypes studied. The GAM of seeds weight/pod values was high in 'D331' cv., 'Cr7' cv., 'Com1', 'Col' cv., and 'AI'. The PCV and GCV % of the parents examined suggested moderate and high levels of the analyzed trait. In F_1 crossings, seeds weight heritability varied from 32.11% to 76.46%. Similarly, the proportion of F_2 crossings varied between 59.06 and 81.67%. Meanwhile, F_1 and F_2 crossings produced moderate and high PCV % and GCV % of seeds weight/pod values, respectively. Seeds weight per plant of the five parents and six crossings in F_1 and F_2 showed greater GCV % values than PCV %, indicating that genotypes interact with environmental variables to influence the expression of this trait which was found to have a high H % and GAM. In general, genetic diversity

TABLE 6 Genotypic variance (GV), phenotypic variance (PV), genotypic coefficient of variance (GCV %), phenotypic coefficient of variance (PCV %), heritability (H%), and genetic advance mean (GAM) of pods, seeds traits and yield of the six crosses in F_1 and F_2 produced from crossing between five parental genotypes; Cream 7 cv. (Cr7), Colossus (Col) cv., and Dokki 331 cv. (D331) genotypes as female parents and Asian Introduction (AI) and Commercial 1 (Com1) genotypes as male parents.

Crosses	Traits*	GV		PV		GCV %		PCV %		H %		GAM	
		F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2
Col x AI	NPo/P	57.2	92.54	53.81	91.02	18.37	23.29	17.82	23.1	94.08	98.35	35.61	47.19
	PoL	21.1	49.4	20.1	48.3	23.71	29.18	23.13	28.85	95.13	97.75	46.47	58.76
	NS/Po	7.90	18.7	7.27	18.0	36.51	35.19	35.03	34.54	92.04	96.34	69.23	69.85
	NAO/Po	7.27	5.60	6.67	5.22	36.94	46.40	35.38	44.82	91.71	93.32	69.79	89.18
	FPoW	0.90	1.09	0.69	0.93	45.15	38.79	39.49	35.73	76.49	84.86	71.14	67.79
	SW/Po	0.55	0.68	0.38	0.55	55.44	44.29	46.36	39.83	69.92	80.90	79.86	73.79
	SY/P	52.30	175.71	49.07	173.6	16.55	43.82	16.03	43.55	93.81	98.80	31.99	89.19
Col x Com1	NPo/P	128.0	54.64	120.0	53.47	37.71	19.25	36.51	19.04	93.75	97.86	72.83	38.80
	PoL	14.3	8.98	13.4	8.51	23.52	16.56	22.81	16.11	94.09	94.72	45.59	32.32
	NS/Po	14.3	13.5	13.4	12.9	38.20	30.90	37.05	30.23	94.08	95.69	74.04	60.93
	NAO/Po	7.03	5.68	6.44	5.30	51.02	59.59	48.82	57.58	91.57	93.36	96.24	114.6
	FPoW	1.29	0.94	1.03	0.79	40.87	32.51	36.63	29.74	80.32	83.70	67.62	56.05
	SW/Po	0.90	0.74	0.69	0.6	46.18	39.87	40.38	36.03	76,46	81.67	72.74	67.06
	SY/P	662.8	74.1	644.63	72.74	55.79	40.20	55.02	39.83	97.25	98.16	111.7	81.31
Cr7 x AI	NPo/P	33.35	274.5	31.68	271.9	26.86	32.48	26.18	32.33	95.00	99.04	52.57	66.29
	PoL	15.5	49.2	14.6	48.1	20.09	31.75	19.51	31.39	94.31	97.74	39.03	63.95
	NS/Po	5.18	9.85	4.67	9.36	29.96	36.93	28.45	35.99	90.18	94.96	55.67	72.24
	NAO/Po	3.29	6.02	2.88	5.64	23.26	40.91	21.77	39.58	87.66	93.56	42.00	78.86
	FPoW	0.26	0.53	0.15	0.42	31.24	44.37	23.56	39.27	56.87	78.35	36.60	71.61
	SW/Po	0.10	0.32	0.03	0.23	33.89	51.82	19.20	44.06	32.11	72.30	22.42	77.18
	SY/P	43.44	58.13	41.54	56.93	42.71	30.06	41.77	29.75	95.61	97.92	84.14	60.65
Cr7 x Com1	NPo/P	427.3	150.0	420.8	148.0	37.24	30.77	36.96	30.57	98.47	98.70	75.56	62.57
	PoL	4.81	6.79	4.32	6.37	12.46	16.18	11.81	15.68	89.81	93.93	23.05	31.32
	NS/Po	8.01	12.9	7.37	12.3	25.04	34.53	24.03	33.76	92.09	95.59	47.52	68.01
	NAO/Po	2.47	8.24	2.11	7.78	34.93	62.40	32.34	60.65	85.75	94.49	61.70	121.4
	FPoW	0.36	0.42	0.22	0.31	24.92	31.80	19.75	27.64	62.77	75.53	32.23	49.50
	SW/Po	0.29	0.31	0.17	0.22	30.94	37.42	23.74	31.69	58.84	71.70	37.51	55.21
	SY/P	550.5	77.17	543.08	75.78	34.89	49.35	34.65	48.90	98.65	98.2	70.91	43.01
D331 x AI	NPo/P	512.0	74.98	496.0	73.61	31.86	20.71	31.36	20.52	96.87	98.17	63.59	41.89
	PoL	6.31	8.61	5.75	8.14	11.96	14.38	11.42	13.99	91.10	94.61	22.45	28.03
	NS/Po	6.82	42.3	6.24	41.2	22.33	26.43	21.35	26.11	91.44	97.56	42.07	42.80
	NAO/Po	2.34	4.91	2.00	4.56	33.28	41.04	30.76	39.55	85.41	92.87	58.57	78.51
	FPoW	0.34	0.60	0.21	0.47	21.70	29.47	17.03	26.28	61.58	79.53	27.53	48.27
	SW/Po	0.24	0.5	0.13	0.39	26.43	35.63	19.61	31.42	55.06	77.75	29.98	57.02
	SY/P	1284.2	42.30	1258.8	41.27	44.36	26.43	43.92	26.11	98.02	97.56	89.58	53.14
D331 x Com1	NPo/P	54.20	70.56	50.90	69.23	14.78	21.59	14.32	21.39	93.92	98.11	28.60	43.64
	PoL	2.71	6.82	2.34	6.4	12.37	19.2	11.50	18.61	86.42	93.94	22.03	37.16
	NS/Po	4.68	5.83	4.20	5.45	28.85	34.01	27.32	32.88	89.66	93.45	53.30	65.48
	NAO/Po	6.23	2.94	5.68	2.67	52.03	31.76	49.69	30.27	91.04	90.79	97.59	59.43
	FPoW	0.16	0.21	0.07	0.14	23.82	32.80	16.01	26.63	45.18	65.89	22.17	44.54
	SW/Po	0.12	0.14	0.04	0.08	26.75	38.21	15.95	29.37	35.53	59.06	19.58	46.49
	SY/P	44.89	46.84	41.90	45.76	12.46	35.17	12.04	34.76	93.32	97.69	23.96	70.78

*Number of pods/plant (NPo/P), pod length (PoL), number of seeds/pod (NS/Po), number of aborted ovules /pod (NAO/Po), fresh pod weight (FPoW), seeds weight/pod (SW/Po) and seeds yield/plant (SY/P).

and the heritability of desired trait control the overall performance of crop development and breeding (Tables 5, 6).

3.6 Cytological analysis of parental genotypes and crossings

The mean proportions of phase and mitotic index were measured in root meristem cells from the five parents and six F_1 and F_2 crossings (Table 7). The mean percentages of mitotic index (MI) ranged from 2.56 to 4.53% in parental genotypes, from 2.63% to 4.53% in F_1, and from 2.39% to 4.31% in F_2. The maximum amount of MI (4.53%) was obtained from Col x AI cross in F_1 and 'AI' genotype, whereas the lowest value (2.39%) was obtained from Cr7 x Com1 cross. Except for the Col x AI cross in F_1 and the 'AI' genotype, which had the same value (4.53%), and Col x AI cross in F_2 (4.31%), the percentages of MI were almost same with in all genotypes/crosses. On the other hand, prophase index data revealed a significant difference across all studied parents and crossings. Furthermore, 'Col' cv. exhibited the greatest percentage of prophase index (44.60%), with a substantial rise above almost all other genotypes. 'Com 1' cv., on the other hand, had the lowest proportion (18.34%). Nevertheless, as shown in Table 7, Cr7 x AI cross, followed by Col x AI cross in F_1, exhibited the greatest prophase index of 36.84% and 33.46%, respectively, when compared to the other produced crossings in F_1 and F_2. Metaphase index values were often greater than previous phases. Among all genotypes, D331 x AI F_1 cross had the highest metaphase index (66.23%), whereas 'Col' cv. produced the lowest (39.84%). Despite the fact that Col x AI cross in F_2 (26.22%), which was greater than its F_1 value (19.16%), the anaphase and telophase indexes appeared at a low frequency compared to other phases.

Table 8 demonstrated many forms of mitotic abnormalities such as lagging chromosomes, chromosomal bridges, outside chromosomes, stickiness, and micronuclei. When compared to all other crosses and parental genotypes, D331 x Com1 cross in F_1 and 'Cr7' cv. provided the highest values of total mitotic abnormalities (6.61% and 5.77%, respectively), while the two crosses in F_2, Col x AI and D331 x AI recorded the lowest values (0.93% and 1.01%, respectively). The percentage of remained genotypes with abnormalities varied from 1.43% to 5.57%. Table 8 also showed that the two genotypes, 'Col' cv. and Cr7 x Com1 cross in F_2, had the greatest frequencies of chromosomal bridges (2.06% and 1.96%, respectively) compared to the other parental genotypes and crosses studied. Almost all cowpea genotypes tested positive for the outside chromosome. It was found in high frequency in D331 x Com1 cross in F_1 and 'Cr7' cv. (2.66% and 2.16%, respectively) with considerable differences with all other genotypes. In terms of chromosomal stickiness %, the aforementioned genotypes 'Cr7' cv. followed by D331 x Com1 cross in F_1 displayed the greatest values (3.61% and 2.1%,

TABLE 7 Proportions of phase and mitotic index (MI) derived from the root tips of six crosses in F_1 and F_2 produced from crossing between five parental genotypes; Cream 7 cv. 'Cr7', Colossus cv. 'Col', and Dokki 331 cv. 'D331' genotypes as female parents and Asian Introduction 'AI' and Commercial 1 'Com1' genotypes as male parents.

Genotypes	Number of examined cells	Prophase %	Metaphase %	Ana & telophase %	Mitotic index %
Parental parents					
Asian Introduction (AI)	1165	27.80	52.11	20.09	4.53
Colossus (Col)	1108	44.60	39.84	15.56	2.56
Commercial 1 (Com1)	1153	18.34	57.04	24.61	2.63
Cream 7 (Cr7)	1046	19.22	55.07	25.70	2.75
Dokki 331 (D331)	1171	18.46	61.33	20.21	2.96
Crosses (F_1)					
Col x AI	1172	33.46	47.38	19.16	4.53
Col x Com1	1234	24.97	54.19	20.84	3.80
Cr7 x AI	1151	36.84	40.24	22.92	3.41
Cr7 x Com1	1220	30.22	50.91	18.86	2.63
D331 x AI	1182	18.53	66.23	15.24	2.75
D331 x Com1	1210	23.79	60.66	15.55	3.33
Crosses (F_2)					
Col x AI	1364	29.94	43.84	26.22	4.31
Col x Com1	1267	26.63	54.09	19.28	3.53
Cr7 x AI	1359	25.46	55.92	18.62	2.84
Cr7 x Com1	1455	23.25	60.63	16.13	2.39
D331 x AI	1347	25.64	53.57	20.80	2.51
D331 x Com1	1383	29.80	50.34	19.87	2.75
$LSD_{0.05}$		11.83	10.77	11.39	1.23

TABLE 8 Proportions of total mitotic abnormalities derived from the root tips of six crosses in F_1 and F_2 produced from crossing between five parental genotypes; Cream 7 cv. 'Cr7', Colossus cv. 'Col', and Dokki 331 cv. 'D331' genotypes as female parents and Asian Introduction 'AI' and Commercial 1 'Com1' genotypes as male parents.

Genotypes	Number of examined cells	Bridges %	Outside %	Stickiness %	Laggards %	Micro nuclei %	Total Abnormalities %
Parental parents							
Asian Introduction (AI)	1165	0.00	1.19	0.00	0.74	0.00	1.93
Colossus (Col)	1108	1.96	0.00	0.00	0.00	0.00	1.96
Commercial 1 (Com1)	1153	0.00	0.88	1.01	1.01	0.12	2.90
Cream 7 (Cr7)	1046	0.00	2.16	3.61	0.00	0.00	5.77
Dokki 331 (D331)	1171	0.00	0.76	1.95	0.00	0.00	2.71
Crosses (F_1)							
Col x AI	1172	0.00	0.72	0.71	0.00	0.00	1.43
Col x Com1	1234	0.98	0.98	0.48	0.00	0.00	2.44
Cr7 x AI	1151	1.11	1.68	1.68	1.11	0.14	5.57
Cr7 x Com1	1220	1.15	0.00	0.68	0.00	0.00	1.83
D331 x AI	1182	0.81	0.90	0.00	0.00	0.00	1.71
D331 x Com1	1210	0.93	2.66	2.10	0.93	0.00	6.61
Crosses (F_2)							
Col x AI	1364	0.51	0.42	0.00	0.00	0.00	0.93
Col x Com1	1267	0.00	1.21	0.85	0.00	0.00	2.07
Cr7 x AI	1359	1.90	0.95	0.85	0.00	0.00	3.71
Cr7 x Com1	1455	2.06	0.00	1.96	0.98	0.00	5.00
D331 x AI	1347	0.00	0.00	1.01	0.00	0.00	1.01
D331 x Com1	1383	1.83	0.00	0.00	0.00	0.00	1.83
$LSD_{0.05}$		2.44	2.87	2.79	1.14	0.02	5.01

respectively). Laggard chromosomes were only detected in five genotypes with low frequencies: 'AI', Cr7 x AI F^1 cross, 'Com1' cv., D331 x Com1 F_1 cross, and Cr7 x Com1 F_2 cross. Micronuclei had the lowest frequency of all mitotic aberrations detected in this experiment. It was found in only Cr7 x AI F_1 cross and 'Com1' cv., and at extremely low levels (0.14% and 0.12%, respectively).

Discussion

Cowpea's narrow base of genetic diversity can be attributed to its self-pollinating nature, evolution from limited wild germplasm, and extremely minimal gene transfer between wild and cultivated varieties. Better variety breeding and selection is an important long-term technique to fighting the problem of low yield in arid or semi-arid regions (Zaki and Radwan, 2022). One of the most efficient traditional breeding methods is germplasm development and evaluation of promising varieties for adaptation and production stability (Tarawali et al., 2002; Singh, 2012). Breeding for consistent production would also require testing crop varieties in a wide range of environments both within and outside of regions, to identify superior genotypes with broad or specialized adaptation due to genotype x environment interactions (Nwosu and Awa, 2013). Summer season is especially challenging in Egypt and other African countries because of high temperatures combined with drought and other stressors. As a result, it will have a detrimental impact on vegetable productivity, quality, and production costs. Cowpea, being a major food in these regions, is thus introduced, particularly during the hot summer season, to provide fresh green vegetables all year (Reda et al., 2016). Introducing cultivars from other countries and planting them for evaluation to select superior cultivars is one of the necessary steps involved in the breeding for a new desired plant species (Animasaun et al., 2015; EL-Ameen, 2018). Varietal adaptation may vary dramatically among environments (Kaya et al., 2002; Eric et al., 2018). Meanwhile, the present study aims, *via* its crossings with introduced cultivars, such as 'Col' cv. and 'AI', to improve the performance of local commercial cultivars, such as 'Cr7' cv., 'D331' cv., and 'Com1'. Plant breeders attempt to develop varieties that reduce the genotype's adverse climatic interactions, varieties that can control their developmental processes so that high yields of high-quality food are produced (Singh et al., 1997; Animasaun et al., 2015).

The purpose of the current study was to determine the diversity of morphological, floral, pod, seed, and yield characteristics, as well as cytological analyses, among five

distinct cowpea genotypes and six developed crosses from two generations apart, F_1 and F_2. The chosen genotypes and resulting crosses were very variable. Govindaraj et al. (2015) underlined the significance of genetic diversity as the lifeline of genetic improvement, whereas Meena et al. (2017) stated that the degree of genetic variability in the breeding population is dependent on the development of high yielding varieties.

In terms of the examined morphological, floral, pods and seeds traits, there were statistically significant variations amongst the genotypes for the most of traits. 'AI' had the longest pod length, while 'D331' cv. had the smallest length. A plant's seed yield is significantly connected to the number of pods, the number of seeds per pod, and the weight of the pod (Oladejo et al., 2011). The findings of Jaydeep and Srinivasan (2011) demonstrated that the length of pods is a changeable trait that may be entirely or partially controlled by plant breeding. In general, heritability values are high for pod length (Diriba et al., 2014a and Diriba et al., 2014b; Sabale et al., 2018). F_1 of D331 x AI cross produced the tallest shoot and the largest pod length, whereas F_2 of Col x AI cross produced the longest pod length. Col x Com1 cross had the greatest peduncle diameter, and the most seeds per pod in F_1 and F_2. Cr7 x Com1 cross in F_1 and F_2 offspring surpassed the other crosses by having a taller peduncle than the better-parent and producing the most peduncle. F_2 of Col x Com1 cross generated the best results and had the fewest ovules that were abortive per pod. In regards of seeds weight/pod, Col x Com1 cross was also superior in both generations. For seeds weight per plant, F_1 of D331 x AI cross had the highest value and outperformed the parental genotypes as well as the other crosses in F_1. That to say, crossing with 'Col', 'AI' and 'Com1' genotypes generally helped to improve the performance of the local varieties, 'D331' and 'Cr7'. This demonstrates that this population might be exploited to develop a breeding program that would result in more productive progeny with more seeds, longer pods, and taller plants (Ajayi et al., 2014a and Ajayi et al., 2014b; Arup et al., 2014). This research has significance for the possibility for genetic development in the breeding program. The capability to select superior genotypes is entirely dependent on the genetic diversity of the collection of varieties, which is a function of additive variance. Important for the selection of possible genotypes is the presence of genetic variation in progenitors (Krause et al., 2012; Sabale et al., 2018).

Furthermore, the current study found a larger percentage of genotypic variance (GV) to total phenotypic variance (PV) in cowpea for selected agronomically significant variables. Other studies (Allen and Allen, 1981; Singh and Rachie, 1985; Damarany, 1994a; Ishiyaku et al., 2005) found a substantial fraction of GV in cowpea that contributed significantly to PV. The GV in F_1 and F_2 was greater than the PV in the five parents for the number of pods/plant which showed high H % values. Omoigui et al. (2006) found a H of 20% for the number of pods/plant, which was very low compared to the current estimate. In contrast to these findings, Singh and Rachie (1985) and Damarany (1994b) calculated the H of the number of pods/plant to be 53% and 86%, respectively. Parental genotypes, as well as F_1 and F_2 crosses, had high H % and genetic advance mean (GAM) for pod length and number of seeds/pod. Nevertheless, the five parents showed high genotypic coefficient of variance (GCV %), phenotypic coefficient of variance (PCV %), H %, and GAM of the number of aborted ovules/pod. Except for the F_1 crosses, D331 x Com1 and Cr7 x AI, which had moderate values for H % for pod weight, all crossings in F_1 and F_2 exhibited high values for H % for the trait. All F_1 and F_2 crosses, on the other hand, produced moderate to high PCV % and GCV % of seeds weight/pod. These differences show variability as a result of additive and non-additive gene effects, emphasizing the possibility of developing novel varieties or hybrids (Silva et al., 2004; Adewale et al., 2011; Manggoel et al., 2012; Ogunkanmi et al., 2013; Lal et al., 2014). Carvalho et al. (2012) discovered similar results in cowpea for 100-seed weight, number of seeds/pod, and yield, whereas Kimani and Derera (2009) reported the same flowering time, number of seeds/pod, and 100-seed weight in common bean, indicating the presence of selection improvements.

In addition, the existing research investigated the cytological performance of produced crosses and parental genotypes to identify new prospective candidates for cowpea breeding programs. The same normal chromosomal number (2n = 22) was found in all cowpea genotypes and crosses studied. The majority of the genotypes/crosses had significantly different mitotic index (MI) overall numbers. Few genotypes showed a high prevalence of total mitotic chromosomal abnormalities. Mitotic abnormalities such as lagging chromosomes, chromosomal bridges, outside chromosomes, stickiness, and micronuclei were seen in all genotypes. The chromosomal number found in this study agrees with previous findings of 2n = 22 for *V. unguiculata* and some allied wild species (Damayanti et al., 2010; Shambhu, 2013). By investigating mitotic characteristics like as mitotic index and mitotic abnormalities, a tentative portrait of cytogenetic differences among genotypes/crosses of interest might be produced. The mitotic index is a variable that may be used to assess the frequency of cellular division (Marcano et al., 2004; Leme and Marin-Morales, 2009). The means of the mitotic index (MI) were substantially different amongst genotypes, owing to large variability in the percentage values of mitotic phases. This might be due to changes in mitotic genetic regulatory systems (cell cycle programs) and/or the number of somatic mutations (Yasuhara and Shibaoka, 2000). This study demonstrated that assessing cowpea genotypes and produced crosses under realistic growth circumstances may objectively reveal the differential contributions of genotype through selection. In addition, the approach provides adequate estimates of variance components and heritability of certain traits, highlighting the significant impact of genotypic variation on traditional breeding. There is sufficient evidence to imply a substantial relationship between the ranks of the variance components' magnitudes.

Conclusion

Six crosses were developed in two generations apart using five different parental genotypes: 'AI', 'Col', 'Com1', 'Cr7', and 'D331'. When examined under different crossings in F_1 and F_2, the study indicated that there is a substantial degree of genotypic variability across the variables investigated in cowpea parental genotypes. Genotypic variance was greater than phenotypic variance for the number of pods/plant, pod length, number of seeds/pod, number of aborted ovules/pod, fresh pod weight, seeds weight/pod, and seeds yield/plant. Except for seeds weight/pod in 'D331' cv. and F_2 of Col x Com1, genotypes/crosses showed high H % for all variables tested. F_2 plants had a greater H % than F_1 plants and the genotypes of their parents. The produced crosses, Cr7 x Com1 and D331 x AI, have the potential for future genetic breeding research and are especially promising for yield and yield component selection. This might allow cowpea producers in Egypt and other similar regions to engage in strategic breeding and breeding trait modification.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

HZ: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualization, writing - original draft, and writing - review and editing. KR: data curation, formal analysis, investigation, methodology, visualization, writing - original draft, and writing - review and editing. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to express their gratitude to the Department of Genetics, Faculty of Agriculture, Minia University, El-Minia, Egypt, for technical assistance with cytological examination of cowpea genotypes.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.987985/full#supplementary-material

References

Adewale, B. D., Adeigbe, O. O., and Aremu, C. O. (2011). Genetic distance and diversity among some cowpea (Vigna unguiculata l. walp) genotypes. *Int. J. Res. Plant Sci.* 1 (2), 9–14. doi: 10.1155/2016/8956412

Ajayi, A. T., Adekola, M. O., Taiwo, B. H., and Azuh, V. O. (2014a). Character expression and differences in yield potential of ten genotypes of cowpea (*Vigna unguiculata* l. *Walp). Int. J. Plant Res.* 4 (3), 63–71. doi: 0.5923/j.plant.20140403.01

Ajayi, A. T., Osekita, O. S., and Osaghae, S. O. (2014b). Variability among some accessions of cowpea (Vigna unguiculata l. walp) in Nigeria on the expression of quantitative traits. *FUTA. J. Res. Sci.* 1), 26–34. doi: 10.5923/j.plant.20140403.01

Allen, O. N., and Allen, E. K. (1981). *The leguminosae: a source book of characteristics, uses and nodulation* (Madison: The University of Wisconsin Press).

Animasaun, D. A., Oyedeji, S., Azeez, Y. K., Mustapha, O. T., and Azeez, M. A. (2015). Genetic variability study among ten cultivars of cowpea (*Vigna unguiculata* l walp) using morpho-agronomic traits and nutritional composition. *J. Agric. Sci.* 10, 119–130. doi: 10.4038/jas.v10i2.8057

Arup, C., Pandia, R. N., Tania, S., Sibsankar, D., Soumitra, C., and Subrata, D. (2014). Identification of selection indices and choosing of parents for vegetable cowpea (*Vigna unguiculata CV-gr.sesquipedalis*) breeding programme. *Legume. Res.* 37 (1), 19–25. doi: 10.5958/J.0976-0571.37.1.003

Baudoin, J. P., and Marechal, R. (2001). "Genetic diversity in í vigna," in *Cowpea research, production and utilization*. Eds. S. R. Singh and K. O. Rachie (Chichester: John Wiley & Sons), 3–11.

Boukar, O, Abberton, M, Oyatomi, O, Togola, A, Tripathi, L, Fatokun, C, et al (2020). Introgression breeding in (*Vigna unguiculata* (L.) walp.). *Front. Plant Sci.* 11, 567425. doi: 10.3389/fpls.2020.567425

Burton, G. W. (1952). Quantitative inheritance in grasses. *Proc. 6th. Int. Grasslands. Cong. J.* 1, 227–283.

Burton, G. W., and De Vane, E. H. (1953). Estimating heritability in tall fescue (*Fistvea arundiancea*) from replicated clonal material. *Agric. J.* 45, 284–291. doi: 10.2134/agronj1953.00021962004500100005x

Carvalho, L. C. B., Silva, K. J. D., Rocha, M. M., Sousa, M. B., Pires, C. J., and Nunes, J. A. R. (2012). Phenotypic correlations between combining abilities of F_2 cowpea populations. *Crop Breed. Appl. Biotechnol.* 12, 211–214. doi: 10.1590/S1984-70332012000300008

Damarany, A. M. (1994a). Estimates of genotypic and phenotypic correlation, heritability and potency of gene set in cowpea (*Vigna unguiculata* [L.] walp). *Assiut. J. Agric. Sci.* 25, 1–8.

Damarany, A. M. (1994b). Testing and screening of some cowpea (*Vigna unguiculata* (L) walp) genotypes under assiut condition. *Assiut. J. Agric. Sci.* 25 (4), 9–19.

Damayanti, F., Lawn, R. J., and Bielig, L. M. (2010). Genetic compatibility among domesticated and wild accessions of the tropical tuberous legume *Vigna vexillata* (L.) a. *Rich. Crop Pasture Sci.* 61 (10), 785–797. doi: 10.1071/CP10060

Darbelley, N. D., Driss-Ecole, D., and Perbal, G. (1989). Elongation of mitotic activity of lentil roots grown in micro- gravity. *Plant Physiol. Biochem.* 27, 341–347.

Diriba, S., Mebeaselassie, A., and Habtamu, Z. (2014a). Genetic variability and heritability of yield and related characters in cowpea (*Vigna unguiculata l. walp.*). *Res. Plant Biol.* 4 (2), 21–26.

Diriba, S., Mebeasellasie, A., and Habtamu, Z. (2014b). Interrelationship and path coefficient analysis of some growth and yield characteristics in cowpea (*Vigna unguiculata l.* walp) genotypes. *J. Plant Sci.* 2 (2), 97–101.

Driss-Ecole, D., Schoevaert, D. D., Noin, M., and Perbal, G. (1994). Densitometric analysis of nuclear DNA content in lentil roots grown in space. *Cell* 81, 59–64. doi: 10.1016/0248-4900(94)90055-8

EL-Ameen, T. M. (2018). Genetic components of some economic traits in cowpea (*Vigna unguiculata*). *J. Agric. Sci.* 33, 135–149. doi: 10.21608/jpp.2008.126209" 10.21608/JPP.2008.126209

Eric, B. K., Toscani, N. D., Mariette, A., and Remy, S. P. (2018). Preliminary study on character associations phenotypic and genotypic divergence for yield and related quantitative traits among cowpea landraces (*Vigna unguiculata*) from the Western highland region of Cameroon. *Open Agric.* 3, 84–97. doi: 10.1515/opag-2018-0009

FAOSTAT (2019). *Agriculture production* (Viale delle Terme di Caracalla, Rome, Italy: Food and Agriculture organization of the United Nations). Available at: http://www.faostat.fao.org.

FAOSTAT (2020). *Agriculture production* (Viale delle Terme di Caracalla, Rome, Italy: Food and Agriculture organization of the United Nations). Available at: http://www.faostat.fao.org.

Govindaraj, M., Vetriventhan, M., and Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genet. Res. Int.* 2015, 431–487. doi: 10.1155/2015/431487

Hall, A. E. (2004). Breeding for adaptation to drought and heat in cowpea. *Europe. J. Argon.* 21, 447–454. doi: 10.1016/j.eja.2004.07.005

Ishiyaku, M. F., Singh, B. B., and Craufurd, P. Q. (2005). Inheritance of time to flowering in cowpea (*Vigna unguiculata* (L.) walp.). *Euphytica* 142, 291–300. doi: 10.1007/s10681-005-2435-0

Jaydeep, H., and Srinivasan, S. (2011). Varietal screening and role of morphological factors on distribution and abundance of spotted pod borer, *Maruca vitrata* (Geyer) on cowpea. *Ann. Plant Prot. Sci.* 19 (1), 71–74.

John, M., and Clabe, W. (2016). Genetic diversity of cowpea (*Vigna unguiculata* (L.) walp.) accession in Kenya gene bank based on simple sequence repeat markers. *Int. J. Genomics* 2016, 1–5. doi: 10.1155/2016/8956412

Johnson, H. W., Robinson, H. F., and Comstock, R. W. (1955). Estimates of genetic and environment. *Agron. J.* 47 (7), 314–318. doi: 10.2134/agronj1955.00021962004700070009x

Kaya, Y., Palta, Ç., and Taner, S. (2002). Additive main effects and multiplicative interactions analysis of yield performances in bread wheat genotypes across environments. *Turk. J. Agric.* 26, 275–279.

Kimani, J., and Derera, J. (2009). Combining ability analysis across environments for some traits in dry bean (*Phaseolus vulgaris* l.) under low and high soil phosphorus conditions. *Euphytica* 166, 1–13. doi: 10.1007/s10681-008-9795-1

Krause, W., Souza, R. S., Neves, L. G., Carvalho, M. L. S., Viana, A. P., and Faleiro, F. G. (2012). Selection gain in intrapopulation genetic improvement of yellow passion fruit. *Braz. Agric. Res.* 47, 51–57.

Kumar, L. S. (1999). DNA Markers in plant improvement: an overview. *Biotechnol. Adv.* 17, 143–182. doi: 10.1016/S0734-9750(98)00018-4

Lal, H., Bhardwaj, D. R., Vishwa, N., and Rameshwar, S. (2014). Character association and genetic divergence studies in cowpea [*Vigna unguiculata (L.) walp.*]. *Ann. Plant Soil Res.* 16 (3), 186–191.

Leme, D. M., and Marin-Morales, M. A. (2009). *Allium cepa* test in environmental monitoring: A review on its application. *Mutat. Res.* 682 (1), 71–81. doi: 10.1016/j.mrrev.2009.06.002

Manggoel, W., Uguru, M. I., Ndam, O. N., and Dasbak, M. A. (2012). Genetic variability, correlation and path coefficient analysis of some yield components of ten cowpea [*Vigna unguiculata (L.)* walp] accessions. *J. Plant Breed. Crop Sci.* 4 (5), 80–86. doi: 10.5897/JPBCS12.007

Marcano, L., Carruyo, I., Del Campo, A., and Montiel, X. (2004). Cytotoxicity and mode of action of maleic hydrazide in root tips of *Allium cepa* l. *Environ. Res.* 94, 221–226. doi: 10.1016/S0013-9351(03)00121-X

Meena, S. S., Verma, S. K., Choudhary, R., Nehra, M., Panwar, R. K., and Singh, J. P. (2017). Interpretation of genotype x environment interaction and grain yield stability in some advance lines of pigeon pea (*Cajanus cajan* (L.) millsp.) grown at different altitudes. *Chem. Sci. Rev. Lett.* 6, 1113–1119.

Nameirakpam, B., and Khanna, V. K. (2018). Studies on crossability and genetic diversity in cowpea (*Vigna unguiculata* l. walp). *Int. J. Environ. Sci. Natural Resour.* 13 (1), 8–16. doi: 10.19080/IJESNR.2018.13.555852

Nkoana, D. K., Abe Shegro, G., and Gwata, E. T. (2019). Agronomic performance and genetic variability of cowpea (*Vigna unguiculata*) accessions. *Legume. Res. Int. J.* 42 (6), 757–762. doi: 10.18805/LR-450

Nwosu, D. J. J., and Awa, E. N. (2013). Cross-compatibility between some cultivated cowpea varieties and a wild relative (*Subsp. dekindtiana var pubescens*). *J. Sci. Res.* 5 (1), 195–200. doi: 10.3329/jsr.v5i1.10761

Ogunkanmi, L. A., Fakoya, A. O., and Adekoya, K. O. (2013). Breeding for long pod trait in cowpea (*Vigna unguiculata l.* walp.). *J. Agric. Res.* 14, 9–14.

Oladejo, A. S., AkinwaleI, R. O., and Obisesan, I. O. (2011). Interrelationships between grain yield and other physiological traits of cowpea cultivars. *Afr. Crop Sci. J.* 19 (3), 189–200.

Omoigui, L. O., Ishiyaku, M. F., Kamara, A. Y., Alabi, S. O., and Mohammed, S. G. (2006). Genetic variability and heritability studies of some reproductive traits in cowpea (*Vigna unguiculate* (L.) walp.). *Afr. J. Biotechnol.* 5, 1191–1195.

Patel, U. V., Parmar, V. K., Patel, P. B., and Malviya, A. V. (2016). Correlation and path analysis study in cowpea (*Vigna unguiculata (L.)* walp.). *Int. J. Sci. Environ. Technol.* 5 (6), 3897–3904.

Patil, D. M., Sawardekar, S. V., and Gokhale, N. B. (2013). Genetic diversity analysis in cowpea [*VignaUnguiculata* (L.) walp.] by using RAPD markers. *Int. J. Innovat. Biotechnol. Biochem.* 1), 15–23.

Pavadai, P., Girija, M., and Dhanavel, D. (2009). Effectiveness, efficiency and biochemical content of physical and chemical mutagens in soybeans (*Glycine max* (L.) merr.). *J. Phytol.* 1 (6), 444–447.

Racuciu, M. (2009). Effect of radiofrequency radiation on root tip cells of *Zea mays*. *Romanin. Biotechnol. Lett.* 14 (3), 4365–4369.

Reda, M. G., Marwa, H., and Abdelfattah, B. (2016). Seed coat color, weight and eye pattern inheritance in gamma-rays induced cowpea M2-mutant line. *J. Genet. Eng. Biotechnol.* 14, 61–68. doi: 10.1016/j.jgeb.2015.12.005

Robinson, H. F., Comstock, R. E., and Harvey, P. H. (1949). Breeding procedure designed to make maximum use of both general and specific combining ability. *Agron. J.* 41, 360–367. doi: 10.2134/agronj1949.00021962004100080006x

Sabale, G. R., Bhave, S. G., Desai, S. S., Dalvi, M. B., and Pawar, P. R. (2018). Variability, heritability and genetic advance studies in F_2 generation of cowpea (*Vigna unguiculata sub* sp. *unguiculata*). *Int. J. Curr. Microbiol. Apply. Sci.* 7 (9), 3314–3320. doi: 10.20546/ijcmas.2018.709.411

Shambhu, B. (2013). *Studies on flower visitors of field bean lablab purpureus (L.) sweet and their role in pollination and pod set (Doctoral dissertation)*. University of Agricultural Sciences, GKVK.

Silva, M. P., Amaral Junior, A. T., Rodrigues, R., Daher, R. F., Leal, N. R., and Schuelter, A. R. (2004). Diallel analysis of combining ability in snap bean. *Braz. Horticult.* 22), 277–280. doi: 10.1590/S0102-05362004000200022

Singh, B. B. (2012). "Breeding diverse, durable, and diet-plus cowpea varieties for increased cowpea production and enhanced nutrition and health in the 21st century," in *Innovative research along the cowpea value chain. proceedings of the fifth world cowpea conference on improving livelihoods in the cowpea value chain through advancement in science, saly, 27 September -1 October 2010*. Eds. O. Boukar, O. Coulibaly, C. A. Fatokun, K. Lopez and M. Tamo (Ibadan: International Institute of Tropical Agriculture (IITA), 17–29.

Singh, B. B., Chambliss, O. L., and Sharma, B. (1997). Recent advances in cowpea breeding. *Biblio1.iita.org* 3, 30–49.

Singh, B. B., Ehlers, J. D., Sharma, B., and Freire Filho, F. R. (2002). *Challenges and opportunities for enhancing sustainable cowpea production* (Ibadan: International Institute of Agriculture), 4–8.

Singh, S. R., and Rachie, K. O. (1985). *Cowpea research, production and utilization* (Chichester: Wiley Interscience).

Tarawali, S. A., Singh, B. B., Gupta, S. C., Tabo, R., Harris, F., Nokoe, S., et al. (2002). *Cowpea as a key factor for a new approach to integrated crop livestock systems research in the dry savannas of West Africa* (Ibadan, Nigeria: International Institute of Tropical Agriculture), 233–251.

Timko, M. P., and Singh, B. B. (2008). "Cowpea, a multi-functional legume," in *Genomics of tropical crop plants*. Eds. P. H. Moore and R. Ming (New York: Springer), 227–256.

Wang, L., and Ruan, Y. L. (2013). Regulation of cell division and expansion by sugar and auxin signaling. *Front. Plant Sci.* 4, 163. doi: 10.3389/fpls.2013.00163

Willie, P. O., and Aikpokpodion, P. O. (2015). Mitotic activity in cowpea (*Vigna unguiculata* (L.) land race "Olaudi" walp)in Nigeria. *Am. J. Plant Sci.* 6, 1201–1205. doi: 10.4236/ajps.2015.68123

Yasuhara, H., and Shibaoka, H. (2000). Inhibition of cell-plate formation by brefeldin inhibited the depolymerization of microtubules in the central region of the phragmoplast. *Plant Cell Physiol.* 41, 300–310. doi: 10.1093/pcp/41.3.300

Zaki, H. E. M., and Radwan, K. S. A. (2022). Response of potato (*Solanum tuberosum* l.) cultivars to drought stress under *in vitro* and field conditions. *Chem. Biol. Technol. Agric.* 9, 1. doi: 10.1186/s40538-021-00266-z

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fpls.2022.970865

Check for updates

OPEN ACCESS

EDITED BY
Eric Von Wettberg,
University of Vermont, United States

REVIEWED BY
Ahmed Sallam,
Assiut University, Egypt
Debjyoti Sen Gupta,
Indian Institute of Pulses Research (ICAR), India

*CORRESPONDENCE
Estefanía Carrillo-Perdomo
estefania.carrillo-perdomo@inrae.fr

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 16 June 2022
ACCEPTED 28 September 2022
PUBLISHED 19 October 2022

CITATION
Carrillo-Perdomo E, Magnin-Robert J-B, Raffiot B, Deulvot C, Floriot M, Lejeune-Hénaut I, Marget P, Burstin J, Tayeh N and Aubert G (2022) A QTL approach in faba bean highlights the conservation of genetic control of frost tolerance among legume species. *Front. Plant Sci.* 13:970865. doi: 10.3389/fpls.2022.970865

COPYRIGHT
© 2022 Carrillo-Perdomo, Magnin-Robert, Raffiot, Deulvot, Floriot, Lejeune-Hénaut, Marget, Burstin, Tayeh and Aubert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A QTL approach in faba bean highlights the conservation of genetic control of frost tolerance among legume species

Estefanía Carrillo-Perdomo[1,2*], Jean-Bernard Magnin-Robert[1], Blandine Raffiot[1,3], Chrystel Deulvot[1], Matthieu Floriot[4], Isabelle Lejeune-Hénaut[5], Pascal Marget[1,6], Judith Burstin[1], Nadim Tayeh[1] and Grégoire Aubert[1]

[1]Agroécologie, INRAE, Institut Agro, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, Dijon, France, [2]UMR AGAP Institut, Univ. Montpellier, CIRAD, INRAE, Institut Agro, San Giuliano, France, [3]Terres Inovia, Thiverval-Grignon, France, [4]Agri Obtentions, Ferme de Gauvilliers, Orsonville, France, [5]Département de génétique et protection des cultures, BioEcoAgro Joint Research Unit, INRAE, Université de Lille, Université de Liège, Université de Picardie Jules Verne, Estrées-Mons, France, [6]INRAE, UE115 Domaine Expérimental d'Epoisses, Dijon, France

Frost is a major abiotic stress of winter type faba beans (*Vica faba* L.) and has adverse effects on crop yield. Climate change, far from reducing the incidence of frost events, is making these phenomena more and more common, severe, and prolonged. Despite the important interaction that the environment has in the tolerance of faba bean to frost, this trait seems to have good levels of heritability. Several QTLs for frost tolerance have already been reported, however, a more robust identification is needed to more precisely identify the genomic regions involved in faba bean tolerance to sub-zero temperatures. Several pea (*Pisum sativum* L.) and barrel medic (*Medicago truncatula* L.) frost tolerance QTLs appear to be conserved between these two species, furthering the hypothesis that the genetic control of frost tolerance in legume species might be more generally conserved. In this work, the QTL mapping in two faba bean recombinant inbred line (RIL) populations connected by a common winter-type parent has led to the identification of five genomic regions involved in the control of frost tolerance on linkage groups I, III, IV, and V. Among them, a major and robust QTL of great interest for marker-assisted selection was identified on the lower part of the long-arm of LGI. The synteny between the faba bean frost tolerance QTLs and those previously identified in other legume species such as barrel medic, pea or soybean highlighted at least partial conservation of the genetic control of frost tolerance among different faba bean genetic pools and legume species. Four novel RILs showing high and stable levels of tolerance and the ability to recover from freezing temperatures by accumulating frost tolerance QTLs are now available for breeding programs.

KEYWORDS

Frost tolerance, faba bean (Vicia faba L.), genetic linkage mapping, recombinant inbred line (RIL) population, quantitative trait loci (QTL), synteny, barrel medic (Medicago truncatula G.), pea (Pisum sativum L.)

Introduction

Faba bean (*Vicia faba* L.) is a cool season rain-fed legume and as such a key component of sustainable agriculture and food security due to its high nutritional value and the ecological services it provides. Faba bean seeds are rich in proteins, carbohydrates, dietary fibers, and micronutrients (Mulualem et al., 2012). They are important components of human food and animal feed in many countries. The crop improves soil fertility through symbiosis with nitrogen-fixing bacteria and is capable of increasing the yield of other crops such as cereals when included in rotations (Duc, 1997; Crépon et al., 2010; Maqbool et al., 2010). Faba bean cultivars are divided into spring or winter types. Their sowing season (spring/autumn/winter) varies according to the climate including the risk of frost in the cropping region (Flores et al., 2013). In this sense, there are three cultivation arrangements: i) in the Mediterranean areas, cultivars should be adapted to autumn sowing; (ii) in oceanic or fairly mild-winter continental climate areas, cultivars should be of winter-type and adapted to autumn sowing and; (iii) in cold regions of northern or continental Europe, cultivars should be of spring-type, adapted to early-spring sowing (Annicchiarico et al., 2015). When sown in autumn or early winter, the crop has to withstand low temperatures that can freeze its aerial organs and its roots, from the early stages of its development.

Frost is a major abiotic stress for winter-type faba beans and has adverse effects on crop production (Sallam et al., 2016a; Sallam et al., 2016b). The combination of climatic factors such as the rate of temperature decrease, the intensity and the duration of the freezing period, the intensity of the wind, the daily insulation, and the absence/presence of snow cover are determinants in the survival of faba bean plants (Duc et al., 2011). Low temperatures threaten plants by inducing the formation of ice crystals in plant tissues (Ambroise et al., 2020) and causing dehydration (Ambroise et al., 2020). Frost-damaged faba bean plants have leaves with a water-soaked appearance that wilt and turn dark brown or black. Although frost damage is less common on the roots thanks to soil insulation, below-ground freezing temperatures can seriously impact both the thin roots that absorb water and minerals and the symbiosis with nitrogen-fixing bacteria, thus affecting nitrogen fixation (Stoddard et al., 2006; Ambroise et al., 2020). After a frost event plants still have to de-acclimate to survive thawing. The thawing rate plays a key role in survival chances since water must be absorbed faster than it is released (Scarth, 1944). Particularly harmful are late frost events that occur once plants have already begun to mobilize their reserves to resume sprouting or have started the reproduction phase (Maqbool et al., 2010; Duc et al., 2011). Late frosts can cause the abortion of buds and flowers and affect pod formation and seed filling (Duc et al., 2011). In addition, frost damages make the plants more vulnerable to pests and pathogens (Maqbool et al., 2010).

However, cold acclimation occurs naturally in faba bean fields during the course of fall to winter. As described in other plants (Thomashow, 2010; Guy, 2011; Liu et al., 2019), faba bean has the capacity to improve its response to freezing after going through a cold acclimation period. During this period, the plants are exposed to successive days of progressively decreasing positive temperatures. This is usually correlated with an increasingly shorter photoperiod. During cold acclimation, changes in gene expression occur that lead to biochemical and physiological adaptations reinforcing the plant's ability to subsequently cope with intense frosts.

Frost tolerant faba bean varieties can survive frost or freezing temperatures. However, surviving plants (tolerant plants) need to invest time and resources to recover and regenerate. It is achieved at the cost of compromising their potential productivity. Different strategies can be activated simultaneously by the plant to tolerate and survive face frost: (i) limitation of the formation of ice crystals and prevention of freezing if water is kept liquid below 0°C (supercooling), (ii) insulation of cells from low temperatures by modifying the lipid bilayer of the cell wall and; (iii) tolerance to the presence of ice crystals and dehydration (Ambroise et al., 2020).

The warming caused by climate change, far from reducing the frequency, severity and duration of frost events, is making these phenomena increasingly problematic and common in many regions of the world (Gu et al., 2008). The less extreme temperatures that are being reached in winter, destabilize the climatic conditions of the poles and cause an increase in extreme meteorological events such as early and late frosts (Francis and Skific, 2015). Milder autumn temperatures could affect the efficiency of faba bean cold acclimation, as is the case for pea, another grain legume species (Castel et al, 2017). Furthermore, early and late frosts affect plants during the most critical stages of their life cycles (seedling, sprouting and, reproduction), making the development of frost-tolerant varieties strategic to ensure the future of the crop.

Within legumes, the genetic determinism of frost tolerance has been studied in the model plant *Medicago truncatula* (Avia et al., 2013; Tayeh et al., 2013a), in the forage crop *Medicago sativa* (Brouwer et al., 2000; Li et al., 2015; Adhikari et al., 2018, 2021) and in important grain legume crops such as pea (Dumont et al., 2009; Zhang et al., 2011; Klein et al., 2014), faba bean (Arbaoui et al., 2008; Ali et al., 2016; Sallam et al., 2016a; Sallam et al., 2016b), lentil (Kahraman et al., 2015) and chickpea (Kahraman et al., 2004). In all these species, the trait appears to be polygenic and some results have highlighted syntenic QTLs between pea (Avia et al., 2013;

Tayeh et al., 2013a) and *M. truncatula* and between faba bean and *M. truncatula* (Ali et al., 2016). In faba bean, several cultivars and breeding lines capable of withstanding sub-zero temperatures have been identified (Stoddard et al., 2006; Arbaoui et al., 2008b; Duc et al., 2011; Annicchiarico et al., 2015). Despite important environmental effects on frost tolerance (Duc et al., 2011; Flores et al., 2013; Sallam et al., 2016a), this trait appears highly heritable and polygenic, with large additive effects (Duc and Petitjean, 1995). Several quantitative trait loci (QTL) related to frost tolerance, winter hardiness, fatty acid content (FAC), and yield have been reported for this species (Arbaoui and Link, 2008; Arbaoui et al., 2008a; Sallam et al., 2015; Sallam et al., 2016b). Although several QTLs have already been identified, it would be necessary to carry out QTL analyses on more saturated genetic linkage maps that allow reducing the confidence interval of the QTLs to facilitate the identification of molecular markers useful in marker-assisted selection (MAS).

Considering that the faba bean genome sequence was not available at the time this study was conducted, the strategy to identify QTLs for frost tolerance in faba bean was to perform QTL mapping on two bi-parental populations with diverse genetic backgrounds. We identified genomic regions that control frost tolerance in a field trial network, using two faba bean recombinant inbred line (RIL) populations connected by a common winter-type parent. We used the densest faba bean consensus map (Carrillo-Perdomo et al., 2020) to date to study the synteny between the frost tolerance QTLs identified in faba bean and those previously identified in other legume species to highlight genomic regions useful for marker-assisted selection.

Materials and methods

Plant material

Two populations of connected recombinant inbred line (RIL) families *POP2* and *POP3* described by Carrillo-Perdomo et al. (2020) have been used in this study (Figure S1). Both populations share Hiverna (IVIP102372, accession code of the INRAE faba bean collection) as a common female parent, and Silian (IVIP100340) and Quasar (IVIP102378) are the respective male parents for *POP2* and *POP3*. Cultivar Hiverna (Germany) is known to be well adapted to continental climates (Link et al., 2010) while Quasar (United Kingdom) is a winter cultivar bred in England (Ondrej and Hunady, 2007) and thus adapted to oceanic climates (cool winters with abundant rainfall). Silian is a *V. faba* spp. *minor* landrace from Northern Sudan susceptible to frost. RILs used for phenotyping were $F_{5:6}$ families produced by single seed descent (SSD) until the F3 generation. *POP2* and *POP3* consisted of 145 and 96 RIL families respectively. Genotyping data and genetic maps described by Carrillo-Perdomo et al. (2020) were used for further QTL analyses.

Multi-year and multi-location field experiments: Frost damage assessments

Three experiments were carried out: i) a first trial *(B_2016-17)*, at the Epoisses experimental farm of INRAE in Bretenière (France) (Latitude 47° 23 '70 ′ ′ N; Longitude 5° 09 ′ 80 ′ ′ E; Altitude 210 m), was sown on November 3, 2016; ii) a second trial *(B_2017-18)*, at the Epoisses experimental farm, was sown on October 16, 2017; and iii) a third trial (*O_2017-18*), in Orsonville (France) (Latitude 48° 47 '77' 'N, 1° 83' 53 " E) was sown on November 3, 2017. Climatic parameters recorded at nearby weather stations are available in Table S1 and Figure S1 for each environment. Each trial consisted of a complete randomized two-block design. For each RIL and parent of the two populations, twenty seeds were sown in rows of 2.5 m with a 1 m distance between rows. The trials were chemically protected against weeds, fungi, and pests, when necessary. Freezing events were defined as periods of at least five successive days with negative temperatures measured under shelter. Frost damage (*FD*) was scored 15 to 30 days after freezing events. Therefore, the number of evaluations depended on the number of frost events that occurred at each location. A scale of 0 to 5 (Figure S2) was used to evaluate the symptoms of frost damage (*FD*) in plants: zero corresponded to no visual damage and five corresponded to dead plant. The area under the symptoms progress curve (*AUSPC*) (Wilcoxson et al., 1975) was calculated in each environment when more than one evaluation of frost damage was available to evaluate its progress over time. The equation used was: $\sum_{i=1}^{n-1} \frac{y_i + y_{i+1}}{2} x (t_{i+1} - t_i)$; where y_i is the assessment of frost damage at the ith observation, t_i is time in days at the ith observation, and n is the total number of observations. Additionally, the number of plants per accession and block was counted from the emergence to the end of the winter season. Survival rate (*SR*) was assessed once at the end of the winter season by counting the percentage of surviving plants for each accession.

Statistical analysis

Analysis of variance (ANOVA) for a complete block randomized design was conducted to test significant differences among accessions and environments for *FD*, *AUSPC*, and *SR* with accession (G) and environment (E) as fixed factors in order to determine the genotypic (G) and the genotype × environment (G×E) interaction effects for frost tolerance in *POP2* and *POP3*. We used the statistical model: $y_{ijkl} = \mu + E_i + R(E)_{j(i)} + G_k + GE_{jk} + \epsilon_{ijkl}$ where, i = 1,…,145 for *POP2* and i = 1,…,96 for *POP3*; j = 1,…,3; k = 1, 2; and yijk denotes the response variable of genotype i, in environment j, block k; μ is the mean of all plots; E_i is the environmental main

effect of trial i; $R(E)_{j(i)}$ stands for the effect of the block j in trial l; G_k is the genetic effect of genotype k; GE_{jk} is the genotype-by-environment interaction effect for genotype k in trial i, and ϵ_{ijkl} is the plot residual term. The underlined terms were considered random effects, which were assumed to be normally and independently distributed. Three environments *(B_2016-17*, *B_2017-18*, *and O_2017-18*) were defined as the combination of the sowing season (2016-2017 or 2017-2018) and the trial location (Bretenière or Orsonville). The normality of residuals and homogeneity of variances were checked using Shapiro–Wilk and Bartlett's test ($P \geq 0.05$). Broad-sense heritability (H^2) across environments was calculated from ANOVA by $H^2 = \delta_g^2/(\delta_g^2 + (\delta_{ge}^2/ne)$, where δ_g^2 is the genotypic variance, δ_{ge}^2 is the genotype x environment interaction variance, and n_e is the number of environments. Analyses were performed in IBM® SPSS Statistics version 26.0.0.0. (https://www.ibm.com/fr-fr/analytics/spss-statistics-software). Principal component analysis (PCA) and Pearson correlation coefficients between traits were calculated using the software Past4.07b (https://www.nhm.uio.no/english/research/infrastructure/past/).

QTL mapping

Individual genetic maps for *POP2* and *POP3* and the faba bean consensus map described by Carrillo-Perdomo et al. (2020) were used for Quantitative trait loci (QTL) detection. QTLs were identified using MCQTL software v5.2.4 (Jourjon et al., 2005). A composite interval mapping (CIM) and an iterative QTL mapping (iQTL) were carried out. Marker cofactor selection by the forward stepwise method was implemented as well as computation of threshold test values of LOD-scores by 1000 permutations. The ProbaPop component computed QTL genotype probabilities giving marker information at each linkage group for each population (*POP2* and *POP3*). MultiPop detection was performed using the faba bean consensus map developed by Carrillo-Perdomo et al. (2020). Model additive and Interpop connected were used in the MultiPop function to build a pooled QTL model fitting the observations on genotype probabilities for *Pop2* and *Pop3*. MultiPop provided threshold computation, QTL detection methods, and model estimation. The logarithm of the odds (LOD) score, global R^2, individual R^2, confidence interval, and allelic effect at each QTL were estimated for each trait and population.

Comparison of identified QTLs with previous studies in faba bean and other legume species

In order to compare the positions of frost tolerance-related QTLs from this study with those described earlier in faba bean, and in absence of a reference genome sequence for this crop, we took advantage of the bridge markers between the map published by Webb et al. (2016) and our map (Carrillo-Perdomo et al, 2020) as Sallam and Ul-Allah (2019) report QTLs and association peaks anchored to the Webb et al. map. In addition, we used the syntenic relations with *Medicago truncatula* to compare the QTL regions on both maps as both sets of markers have been blasted on the chromosomes of the Mt4.0 model legume genome sequence (Pecrix et al., 2018), https://phytozome-next.jgi.doe.gov/info/Mtruncatula_Mt4_0v1).

Similar to the work done on *M. truncatula*, and in order to search for potential syntenic QTLs from other legume species, these markers were also positioned on the *P. sativum* genome assembly version v.1a (Kreplak et al., 2019) (https://urgi.versailles.inra.fr/Species/Pisum/Pea-Genome-project), on *Cicer arietinum* L. v1 assembly (Varshney et al., 2014) (https://www.ncbi.nlm.nih.gov/assembly/GCF_000331145.1/) and on *Glycine max* L. assembly v2.0 (Schmutz et al., 2010) (http://www.plantgdb.org/GmGDB/) by applying BLASTn searches with marker context sequences and identifying the best match(es) for each sequence in each genome.

Results

Faba bean response to freezing temperatures

Two bi-parental RIL populations, namely *POP2* and *POP3*, were evaluated for frost tolerance in three environments (*B_2016-17*, *B_2017-18*, *and O_2017-18*). As expected, the parental lines showed contrasting responses to frost both when considering frost damage (*FD*) and survival rate (*SR*) scores (Table 1). The damage caused by up to three frost events (*FD1*, *FD2*, and *FD3*) was scored and the *SR* was noted in each of the three environments and for the two populations (*POP2* and *POP3*) studied. *AUSPC* was calculated to study the evolution of frost damage over time. Environment *B_2016-17* displayed the most extreme conditions (Figure S1, S3). A major freezing event took place in January 2017, leading to sixteen successive days of frost with a minimum temperature under shelter close to -10°C and eleven days of ground frost. During the 2017-2018 cropping season, frost events were shorter in duration, both at Bretenière and Orsonville (Figures S1, S3). The minimum temperature recorded over the growing period was -11.6°C for the environment *B_2017-18* and -9.5°C in *O_2017-18*, both in February. No negative underground temperatures were recorded for these trials.

In general, the susceptible parent Silian showed higher *FDs* and *AUSPC* scores and a lower *SR* than Hiverna and Quasar. Thus, Silian seems to present a higher sensitivity to frost damage as compared to Hiverna and Quasar. No significant differences (P>0.05) were found between Hiverna and Quasar in the

TABLE 1 Scores of parental accessions and recombinant inbred lines (RILs) ($F_{5:6}$) from *POP2* (Hiverna x Silian) and *POP3* (Hiverna x Quasar) for the parameters related to faba bean response to freezing temperatures.

Environment[a]	Trait[b]	*Parents*			*POP2*					*POP3*				
		Silian	Hiverna	Quasar	Mean	SD	Min.-Max.	Kurtosis	Skewness	Mean	SD	Min.-Max.	Kurtosis	Skewness
B_2016-17	**FD1**	5.00	0.00	1.50	1.82	1.30	0.00-5.00	-1.16	0.21	0.56	0.73	0.00-3.00	2.56	2.64
	FD2	5.00	1.50	4.00	3.99	0.82	1.00-5-00	-0.27	-0.43	2.93	0.88	0.00-5.00	0.39	0.39
	FD3	5.00	2.00	4.50	4.30	0.79	1.00-5.00	0.36	-0.91	3.21	0.98	1.00-5.00	-0.50	-0.39
	AUSPC	422.50	140.00	351.25	323.54	58.50	84.00-422.00	0.71	-0.67	237.45	72.92	53.00-401.00	-0.32	0.81
	SR	0.00	59.50	7.00	14.95	20.08	0-00-94.00	2.03	1.52	44.71	26.09	0.00-100.00	-0.87	-0.70
B_2017-18	**FD1**	3.50	0.00	0.00	0.21	0.25	0.00-0.50	-1.90	5.14	0.09	0.19	0.00-0.50	0.80	1.69
	FD2	5.00	1.00	1.50	1.91	0.71	0.00-4.00	-0.67	10.63	1.43	0.53	0.00-2.00	-1.27	-0.01
	AUSPC	288.00	24.00	36.00	54.25	38.50	0.00-552.00	94.38	6.51	34.21	17.06	0.00-72.00	-0.78	0.09
	SR	0.00	87.00	100.00	78.46	19.65	5.00-100.00	0.97	-1.27	92.57	7.75	67.00-100.00	0.28	-0.95
O_2017-18	**FD1**	4.00	2.50	2.00	2.84	0.79	1.00-5.00	-0.12	-0.09	2.27	0.71	1.00-4.00	-0.21	0.12
	SR	0.00	100.00	94.00	91.34	11.48	42.00-100.00	3.76	-3.01	97.06	6.87	43.00-100.00	23.67	-4.12

[a]B_2016-17, Bretenière in 2016-2017; B_2017-18, Bretenière in 2017-2018; O_2017-18, Orsonville in 2017-2018.
[b]FD1, first rating of symptoms of frost damage (0-5); FD2, second rating of symptoms of frost damage (0-5); FD3, third rating of symptoms of frost damage (0-5); AUSPC, area under the symptoms progress curve calculated from the ratings of FD1, FD2 and FD3 and; SR, survival rate after the frost events.

environment *O_2017-18*. Hiverna presented a higher tolerance to frost and ability to survive than Quasar during the 2016-2017 crop season in Bretenière (*FD2*, *FD3*, and *PS* in *B_2016-17* and *FD1* in *B_2017-18*) (Table 1). Nonetheless, the extreme frost event of January 2017 reduced the survival rate of the tolerant parents Hiverna and Quasar in the environment *B_2016-17*.

The frequency distribution of *FD1*, *FD2, FD3*, and *SR* in the RIL populations *POP2* and *POP3* showed a continuous non-normal distribution in the three environments (Table 1; Figure S3). Several transgressive RILs had higher tolerance to frost and survival rates than the tolerant parents Hiverna and Quasar. One line of *POP2* (VFP-2-0734) and three lines of *POP3* (VFP-3-0307, -1211, -1281) presented higher or similar levels of frost tolerance and survival regarding the tolerant parent Hiverna. These four lines presented stable levels of tolerance to frost and survival throughout the three environments. No lines more sensitive to frost than the susceptible parent Silian were identified.

Broad-sense heritability ranged from 0.51 (*FD1*) to 0.85 (*FD2*) in *POP2* and from 0.24 (*FD3*) to 0.78 (*FD2)* in *POP3* (Table 2). Pearson correlations between the traits, by environment and by population, are shown in Figures 1 and S3. In general, we observed positive correlations among different frost damage scores and negative correlations between frost damage scores and plant survival rate. This trend was confirmed by the Principal Component (PC) Analysis (Figure 2). PC1 captured 67.1% of the variation, allowing for differentiation between frost tolerant and susceptible accessions (Figure 2); while PC2 captured 9.2% of the variation, allowing for differentiation of the plant responses to frost in the different environments studied during the 2016-2017 and 2017-2018 cropping seasons.

Faba bean QTLs for frost damage and survival rate

QTL analysis using CIM and iQTL mapping revealed consistent genomic regions associated with frost damage, i.e. *FD1, FD2, FD3*, and *AUSPC*, and plant survival (*SR*) in the three environments *(B_2016-17, B_2017-18*, and *O_2017-18).*

TABLE 2 Broad-sense heritability (H^2) of the traits scored in the recombinant inbred lines (RILs) ($F_{5:6}$) from *POP2* (Hiverna x Silian) and *POP3* (Hiverna x Quasar).

Trait[a]	*POP2*	*POP3*
FD1	0.51	0.31
FD2	0.85	0.78
FD3	0.65	0.24
AUSPC	0.77	0.44
SR	0.77	0.69

[a]FD1, first rating of symptoms of frost damage; FD2, second rating of symptoms of frost damage; FD3, third rating of symptoms of frost damage; AUSPC area under the symptoms progress curve calculated from the ratings of FD1, FD2 and FD3 and; SR survival rate after the frost events.

(Table 3; Figures 3, S4). Five QTLs were identified along LG I, III, IV, and V. On one hand, the QTLs involved in plant tolerance to the damage caused by frost events explained individually from 8.37 to 27.60% of the total phenotypic variance, depending on the trait scored. They explained together from 12.96 to 58.26% of the phenotypic variance. On the other hand, the QTLs for plant survival rate explained from 17.22 to 24.37% of the phenotypic variance. Linkage group I contained a robust and consistent meta-QTL detected in *POP2* in all environments for frost damage scores and survival rate. It accounts for 12.96 to 27.60% of the phenotypic variance of frost damage scores and for 17.22 to 24.37% of survival rate variance. The meta-QTL is located between the gene-based SNP markers "dn_rep_c2728_1677" and "dn_rep_c30_291" and individual QTLs peaked around 378.31 to 392.55 cM with LOD values from 7.05 to 16.41, depending on the trait. QTLs identified in LGIII, LGIV and LGV were only detected in the two trials located in Bretenière. In LGIII, two QTLs associated to reduced damage after the first frost event were identified in both *POP2 and POP3*: the QTL *FD_III.1_FD1_B_17-18* identified during the 2017-2018 cropping season and the QTL *FD_III.2_FD1_B_16-17* during 2016-2017. *FD_III.1_FD1_B_17-18* accounts for 8.37% of phenotypic variance and is located between "dn_rep_c879_1279" and "dn_rep_c3378_1198" markers with a peak at 101.62 cM (LOD score of 4.33). *FD_III.2_FD1_B_16-17* accounts for 22.59% of phenotypic variance and is located between "dn_rep_c125_455" and "dn_rep_c3943_1691" markers with a peak at 196.16 cM (LOD score of 12.89). In LGIV, the QTL *FD_IV.1_FD1_B_17-18* is associated with reduced damage after the first frost event occurring during the 2017-2018 cropping season. It was only detected in *POP3* and accounts for 9.22% of the phenotypic variance. It is located between "dn_rep_c3237_978" and "dn_rep_c6181_713" markers with a peak at 117.47 cM (LOD score of 4.79). Finally, the QTL *FD_V.1 FD1_B_17-18* was identified on LGV in *POP2* and *POP3*. It accounts for 12.99% of the phenotypic variance associated with reduced frost damage during the first frost event. It is located between "dn_rep_c3712_1645" and "dn_rep_c958_552" markers with a peak at 93.08 cM and a LOD score of 6.89.

The desirable alleles for coping with frost tolerance are mainly derived from the parents Hiverna and Quasar (Table 3).

Co-localization of frost tolerance QTLs in different faba bean genetic pools

Previous studies have reported the identification of faba bean QTLs for frost tolerance (Arbaoui et al., 2008a; Sallam

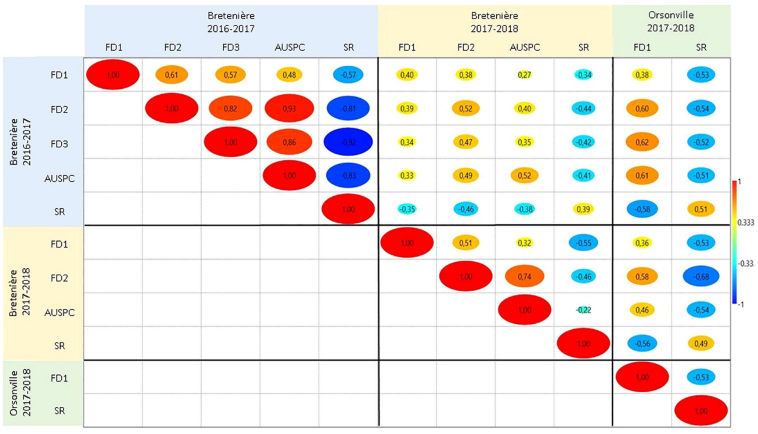


FIGURE 1
Pearson correlation between the traits scored in the two faba bean (Vicia faba) recombinant inbred line (RIL) populations *POP2* (Hiverna x Silian) and *POP3* (Hiverna x Quasar) in the three environments studied: Bretenière during the 2016-2017 cropping season, Bretenière during the 2017-2018 cropping season and Orsonville during the 2017-2018 cropping season. Cold colors indicate negative correlations while warm colors indicate positive correlations between traits. FD1 stands for plant damage caused by the first frost event. FD2 refers to plant damage after the second frost event. FD3 corresponds to plant damage after the third frost event. AUSPC is the area under the symptoms progress curve after the frost period. SR stands for survival rate after the frost period.

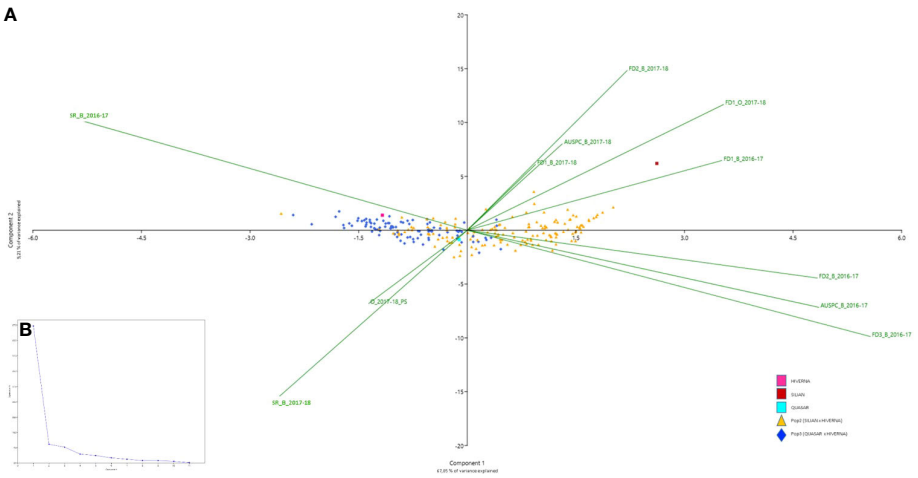


FIGURE 2
Principal component analysis (PCA) of the individuals of the two faba bean recombinant inbred line (RIL) populations *POP2* (Hiverna x Silian) and *POP3* (Hiverna x Quasar) and the parental lines Hiverna, Silian and Quasar and the scored variables: FD1, plant damage caused by the first frost event, FD2, plant damage after the second frost event FD3, plant damage after the third frost event, AUSPC, area under the symptoms progress curve after the frost period and SR, survival rate after the frost period; in the different environments studied (B_2016-17, Bretenière during the 2016-2017 cropping season, B_2017-18 Bretenière during the 2017-2018 cropping season and O_2017-18, Orsonville during the 2017-2018 cropping season). **(A)** Biplot of the first (x-axis) and second (y-axis) principal components (PCs) that shows the PCA scores of the explanatory variables as vectors (in green) and individuals (*POP2* in orange, *POP3* in dark blue, Hiverna in pink, Silian in burgundy and Quasar in light blue). Individuals on the same side as a given variable should be interpreted as having a high contribution on it. The magnitude of the vectors (lines) shows the strength of their contribution to each PC. Vectors pointing in similar directions indicate positively correlated variables, vectors pointing in opposite directions indicate negatively correlated variables, and vectors at proximately right angles indicate low or no correlation. **(B)** Scree plot of eigenvalues after PCA.

and Martsch, 2015; Sallam et al., 2016a; Sallam et al., 2016b). Among them, Sallam et al. (Sallam and Martsch, 2015; Sallam et al., 2015; Sallam et al., 2016b) used the SNP-based consensus map developed by Webb et al. (Webb et al., 2016) for QTL and association mapping, making possible the comparison with our results. Sallam and Martsch (2015) and Sallam et al. (2016b) also identified a QTL region associated with frost tolerance and fatty acid composition at the bottom of LGI (QTL region flanked by the markers VF_Mt5g026780 and VF_Mt2g027240). This QTL seems to co-localize with the major frost tolerance QTL (*FD_I.1* and *SR_I.1*) identified in the present study. Moreover, the environment-specific QTLs *FD_III.1*, *FD_III.2*, *FD_IV.1* and *FD_V.1* (identified in LGIII, IV, and V, respectively) seem to be located in between QTLs involved in frost tolerance and fatty acid composition previously identified on LGIII around 78 cM (Sallam and Martsch, 2015) and 136 cM (Sallam and Martsch, 2015; Sallam and Martsch, 2016; Sallam et al., 2016b), with a frost tolerance QTL of LGIV with a peak around 129 cM (Sallam and Martsch, 2015) and in between proline and fatty acid composition QTLs at 57 (Sallam et al., 2016b), 67 (Sallam and Martsch, 2016; Sallam et al., 2016b) and 73 (Sallam and Martsch, 2016; Sallam et al., 2016b) cM on LGV [see Sallam and Ul-Allah (2019)].

Synteny between genomic regions harboring QTLs for frost tolerance in faba bean and in close legume species

Taking advantage of the syntenic relationship between legume species, the positions of the QTLs identified in this study were compared with those QTLs of frost damage previously reported in other legume species (Kahraman et al., 2004; Lejeune-Hénaut et al., 2008; Dumont et al., 2009; Zhang et al., 2011; Zhang et al., 2012; Avia et al., 2013; Tayeh et al., 2013a; Kahraman et al., 2015; Ates et al., 2018; Deokar et al., 2019; Mugabe et al., 2019; Beji et al., 2020). The faba bean major QTL on LGI (*FD_I.1* and *SR_I.1*) is syntenic to the major freezing tolerance QTLs of chromosome 6 of *Medicago truncatula* (Tayeh et al., 2013a) and of LGVI of *Pisum sativum* (Lejeune-Hénaut et al., 2008) (Table S1). Interestingly, although the QTLs *FD_IV.1* (LGIV) and *FD_V.1* (LGV) are environment-specific, they seem to correspond to frost tolerance QTLs previously reported in other legume species. So, the QTL *FD_IV.1* appeared to be syntenic to *Sat_249* (Zhang et al., 2011) of LGJ of soybean (*Glycine max* L.) while *FD_V.1* is potentially syntenic to QTL *V.2* reported on LGV of pea (Dumont et al., 2009; Klein et al., 2014; Beji et al., 2020) and to *Sat_126* (Zhang et al., 2012) (LGK) of soybean (Table S1).

TABLE 3 Quantitative trait loci (QTL) for frost tolerance.

Trait[a]	LG[b]	QTL name	Peak position[c]	LOD[d]	Inferior marker[e]	Peak marker[f]	Superior marker[g]	*POP2* Hiverna x Silian	*POP3* Hiverna x Quasar	Add[he] Silian	Add[if] Hiverna	Add[jg] Quasar	R[2kh]
FD1_B_2016-17	I	*FD_I.1_FD1_B_16-17*	382.31	13.41	c16955_467	c16955_467	dn_rep_c9956_437	X		0.36	-0.24	-0.12	0.23
FD1_B_2016-17	III	*FD_III.2_FD1_B_16-17*	196.16	12.90	dn_rep_c125_455	dn_rep_c2278_599	dn_rep_c3943_1691	X	X	0.31	-0.19	-0.12	0.23
Total													*0.46*
FD1_B_2017-18	I	*FD_I.1_FD1_B_17-18*	378.31	12.85	c16955_467	c16955_467	dn_rep_c14390_1441	X		0.09	-0.05	-0.04	0.23
FD1_B_2017-18	III	*FD_III.1_FD1_B_17-18*	101.62	4.33	dn_rep_c879_1279	rep_c11396_777	dn_rep_c3378_1198	X	X	0.04	-0.02	-0.02	0.08
FD1_B_2017-18	IV	*FD_IV.1_FD1_B_17-18*	117.47	4.79	dn_rep_c3237_978	c17139_182	dn_rep_c6181_713		X	0.02	0.03	-0.05	0.09
FD1_B_2017-18	V	*FD_V.1_FD1_B_17-18*	93.08	6.89	dn_rep_c3712_1645	rep_c10474_1278	dn_rep_c958_552	X	X	0.05	-0.02	-0.03	0.13
Total													*0.53*
FD1_O_2017-18	I	*FD_I.1_FD1_O_17-18*	378.31	7.05	dn_rep_c2728_1677	c16955_467	dn_rep_c3119_1300	X		0.23	-0.17	-0.05	0.13
Total													*0.13*
FD2_B_2016-17	I	*FD_I.1_FD2_B_16-17*	392.55	9.64	c16955_467	dn_rep_c8507_205	dn_rep_c30_291	X		0.20	-0.16	-0.04	0.17
FD2_B_2017-18	I	*FD_I.1_FD2_B_17-18*	382.31	12.04	c16955_467	c16955_467	dn_rep_c3119_1300	X		0.28	-0.14	-0.14	0.21
Total													*0.38*
FD3_B_2016-17	I	*FD_I.1_FD3_B_16-17*	391.94	9.83	dn_rep_c14390_1441	dn_rep_c4551_1098	dn_rep_c30_291	X		0.17	-0.17	-0.01	0.18
Total													*0.18*
AUSPC_B_2016-17	I	*FD_I.1_AUSPC_B_16-17*	392.55	8.45	dn_rep_c14390_1441	dn_rep_c8507_205	dn_rep_c30_291	X		12.97	-11.29	-1.67	0.15
Total													*0.15*
AUSPC_B_2017-18	I	*FD_I.1_AUSPC_B_17-18*	380.31	16.41	c16955_467	c16955_467	dn_rep_c14390_1441	X		9.17	-4.47	-4.70	0.28
Total													*0.28*
SR_B_2016-17	I	*SR_I.1_PS_B_16-17*	392.55	9.60	dn_rep_c14390_1441	dn_rep_c8507_205	dn_rep_c30_291	X		-5.70	3.84	1.86	0.17
Total													*0.17*
SR_B_2017-18	I	*SR_I.1_PS_B_17-18*	378.31	14.19	c16955_467	c16955_467	c16955_467	X		-7.91	3.62	4.29	0.24
Total													*0.24*

[a]FD1_B_2016-17 first rating of symptoms of frost damage (FD1) in the trial of the Époisses experimental farm in Bretenière during 2016-2017 sowing season (B_2016-17), FD1_B_2017-18 first rating of symptoms of frost damage (FD1) in the trial at Bretenière during 2017-2018 sowing season (B_2017-18), FD1_O_2017-18 first rating of symptoms of frost damage (FD1) in the trial at Orsonville during 2017-2018 sowing season (O_2017-18); FD2_B_2016-17 second rating of symptoms of frost damage (FD2) in the trial at Bretenière during 2016-2017 sowing season (B_2016-17), FD2_B_2017-18 second rating of symptoms of frost damage (FD2) in the trial at Bretenière during 2017-2018 sowing season (B_2017-18), FD3_B_2016-17 third rating of symptoms of frost damage (FD3) in the trial at Bretenière during 2016-2017 sowing season (B_2016-17), AUSPC_B_2016-17 area under the symptoms progress curve calculated from the ratings of FD1, FD2 and FD3 in the trial of Bretenière during 2016-2017 sowing season (B_2016-17), AUSPC_B_2017-18 area under the symptoms progress curve calculated from the ratings of FD1 and FD2 in the trial at Bretenière during 2017-2018 sowing season (B_2017-18), SR_B_2016-17 survival rate (SR) in the trial at Bretenière during 2016-2017 sowing season (B_2016-17) and SR_B_2017-18, SR in the trial of Bretenière during 2017-2018 sowing season (B_2017-18),

[b]LG, linkage group.

[c]Peak position of the QTL (cM).

[d]LOD the peak LOD score.

[e]Inferior marker, molecular marker that coincides with the lowest position of the confidence interval of the QTL.

[f]Peak marker, molecular marker that coincides with the position in the confidence interval of the QTL at which the LOD score value is maximum.

[g]Superior marker, molecular marker that coincides with the uppest position of the confidence interval of the QTL.

[h]Add Silian the additive effect of the multicross Silian-Hiverna-QUASAR regarding the parent Silian.

[i]Add Hiverna the additive effect of the multicross Silian-HIVERNA-QUASAR regarding the parent HIVERNA.

[j]Add Quasar the additive effect of the multicross Silian-HIVERNA-QUASAR regarding the parent QUASAR.

[k]R^2 proportion of the phenotypic variance explained by the respective QTL.

10.3389/fpls.2022.970865

10.3389/fpls.2022.970865

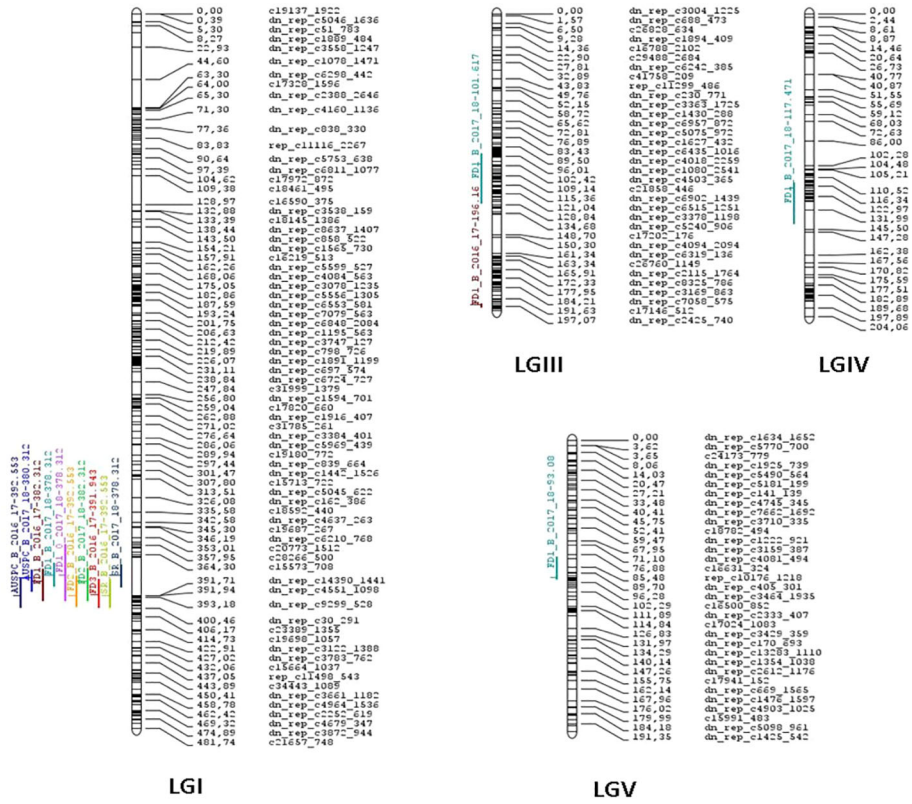


FIGURE 3
Faba bean (Vicia faba) consensus genetic linkage map constructed from two populations formed by 145 F5:6 recombinant inbred lines (RILs) derived from the cross between Hiverna x Silian and 96 RILs F5:6 derived from the cross Hiverna x Quasar. Colored lines on the left of chromosome bars indicate the locations of quantitative trait loci (QTLs). The lengths of the lines are proportional to the lengths of the respective confidence intervals.

Discussion

Progress towards improving frost tolerance in faba bean

Autumn-sown faba beans are known to be more productive than spring types (Sallam et al., 2016b; Landry and Hu, 2019). However, yield stability across environments continues to be a limitation that needs to be tackled. The development of frost-tolerant varieties with stable yields is especially critical to ensure productivity, so that, even if the temperatures seriously decrease, the physiological impact on the plants does not translate into a significant reduction in their yield. This is especially important when dealing with climate change since the incidence of extreme climatic events is becoming more frequent. Our results, together with other efforts to decipher frost tolerance genetics in winter faba bean (Arbaoui et al., 2008a; Sallam and Martsch, 2015; Sallam et al., 2015; Sallam et al., 2016), open avenues for breeding improved varieties. As expected, the level of tolerance and survival to frost along the RILs families of both populations was affected by the environmental conditions. This has also been previously reported in other faba been genetic pools (Arbaoui et al., 2008b). Nonetheless, four RILs [one line of *POP2* (VFP-2-0734) and three of *POP3* (VFP-3-0307, -1211, -1281)] presented high and stable levels of tolerance to frost. These lines are available for faba bean breeding programs.) Moreover, the identification of a major QTL controlling frost tolerance on the LGI (*FD_I.1* and *SR_I.1*) holds promise for MAS development (Table 3). Favorable alleles for frost tolerance mainly came from Hiverna and Quasar (Table 3). Interestingly, the recombinant inbred lines VFP-2-0734, VFP-3-0307, VFP-3-1211 and VFP-3-1281 cumulated several frost damage and/or plant survival QTLs, which makes them good genitors for faba bean breeding programs. Pyramiding favorable alleles from Quasar and Hiverna should increase the chances of coping with sub-zero temperatures by accumulating genes involved in different mechanisms for acclimatizing and de-acclimating to sub-zero temperatures, as suggested by progenies showing transgressive performances.

Some genomic regions of frost tolerance are common among different faba bean genetic backgrounds

The major QTL on LGI harboring *FD_I.1* and *SR_I.1*, co-localizes with the genomic region involved in frost tolerance and fatty acid composition previously described in Sallam and Martsch (2015) and in Sallam et al. (2016a). Sallam et al. (2016b) confirmed several frost tolerance QTLs (Sallam et al., 2015) by association mapping using 156 gene-based SNP markers (Webb et al., 2016), confirming their stability in different environments and genetic backgrounds. The authors used the synteny between the faba bean genetic map and the genome sequence of the model legume *Medicago truncatula* L. (Webb et al., 2016) to investigate the genomic regions involved in the faba bean response to frost. Thus, *FD_I.1* and *SR_I.1* are robust QTLs of utility for MAS. We also identified (only in Bretenière) QTLs on LGIII (*FD_III.1* and *FTIII.2*), LGIV (*FD_IV.1*), and LGV (*FD_V.1*) (Table 3). These QTLs co-localize with faba bean frost tolerance QTLs previously reported on LGIII (Sallam and Martsch, 2015; Sallam and Martsch, 2016; Sallam et al., 2016b), LGIV (Sallam and Martsch, 2015), and LGV (Sallam et al., 2016b) from different populations and different trials.

Some frost tolerance genomic regions are conserved among legume species

The solid blocks of macrosynteny between faba bean and closely-related legume species[22,23] (i.e. pea, barrel medic, or chickpea (Webb et al., 2016; Carrillo-Perdomo et al., 2020), but also with divergent legume species (Carrillo-Perdomo et al., 2020) [i.e. common bean (*Phaseoulus vulgaris* L.), soybean or birdsfoot trefoil (*Lotus japonicus* L.)] facilitates comparative genomic studies that allow the exploration of gene and QTL conservation among legumes. Genomic regions harboring syntenic QTLs between faba bean and other legume species have been reported for flowering and pod setting (Cruz-Izquierdo et al., 2012), where QTL were conserved among faba bean, *M. truncatula*, pea, lupine (*Lupinus albus* L.), chickpea (*C. arietinum*) and birdsfoot trefoil; or for resistance to *Ascochyta fabae* Speg. and *Orobanche crenata* Forssk., where QTLs (Gutierrez and Torres, 2021) were syntenic between faba bean and barrel medic. The comparison of the positions of QTLs reported in this study and those of previously reported QTLs for frost tolerance in other legume species (Kahraman et al., 2004; Lejeune-Hénaut et al., 2008; Dumont et al., 2009; Zhang et al., 2011; Zhang et al., 2012; Avia et al., 2013; Tayeh et al., 2013b; Kahraman et al., 2015; Ates et al., 2018; Deokar et al., 2019; Mugabe et al., 2019; Beji et al., 2020), showed some conserved genomic regions. The faba bean major QTL *FD_I.1* (LGI) is syntenic with the major freezing tolerance QTLs reported in barrel medic (Tayeh et al., 2013b) on chromosome 6 and in pea on LGVI (Lejeune-Hénaut et al., 2008). Moreover, *FD_IV.1 QTL* (LGIV) seems to be syntenic to *Sat_249* (Zhang et al., 2011) described in LGJ of soybean and *FD_V.1* (LGV) is syntenic to the frost tolerance QTLs *V.2* reported in pea on LGV (Dumont et al., 2009; Klein et al., 2014; Beji et al., 2020) and with *Sat_126* present on LGK of soybean (Zhang et al., 2012). These results further support the functional conservation of key genetic determinants for agronomic traits between legumes. Exploiting this conservation could make the identification of potential candidate genes more agile.

The region of chromosome 6 of barrel medic (Tayeh et al., 2013b) named Mt-FTQTL6 syntenic to the faba bean QTL *FD_I.1* contains twelve C-repeat binding factor (CBF)/dehydration-responsive element binding factor 1 (DREB1) genes organized in a tandem array. This QTL region is also collinear with a frost damage QTL of LGVI of pea (Tayeh et al., 2013a) and was in addition revealed by a Genome-Wide approach, which pinpointed CBF genes at the vicinity of frost-damage association peaks (Dumont et al., 2009). The CBF/DREB1 transcription factors have been shown to play a major role in plant cold acclimation in the model plant *Arabidopsis thaliana* (Alonso-Blanco et al., 2005; Thomashow, 2010) and to co-localize with freezing tolerance QTLs in other species (Vágújfalvi et al., 2003; Knox et al., 2008; Sandve et al., 2011). In durum wheat, (Sieber et al., 2016) showed that frost tolerance is mainly controlled by a *CBF* copy number variation at the *Fr* locus, a main frost resistance QTL of the *Poaceae*. In *Triticum monococcum*, studies of lines carrying recombination events within the homologous CBF cluster evidenced a non-functional CBF copy in a more sensitive line and three copies allowing to cold acclimate at milder temperatures in a more tolerant line (Knox et al., 2008).In addition, mutational approaches have evidenced the role of CBF regulons in cold acclimation in *Arabidopsis thaliana* (Jia et al., 2016; Zhao et al., 2016). The ongoing genomic and pan genomic approaches in faba bean (Khazaei et al., 2015) will allow establishing the presence of a CBF regulon in the QTL *FD_I.1* region and will enable the study of the organization and polymorphism of the locus in faba bean accessions with contrasted phenotypes for frost tolerance. The same synteny-based approach could be applied to explore the gene content of the other QTLs evidenced in this study.

In conclusion, four novel RILs showing high levels of tolerance and ability to recover from freezing temperatures by accumulating frost tolerance QTLs are now available for breeding programs. This work has demonstrated the suitability of high-density gene-based genetic maps (Carrillo-Perdomo et al., 2020) to identify QTLs with relatively small confidence intervals and exploit syntenic relationships between legumes in order to validate genomic regions involved in frost tolerance. It also highlighted the conservation of the genetic control of frost tolerance among different faba bean genetic pools and legume species, showing the relevance of translational approaches. The development of pan

genomic sequencing in legumes will open new perspectives to tackle the understanding of traits as complex as frost tolerance.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

EC-P analyzed the data, performed the QTL and synteny analysis and wrote the manuscript; JM-R, MF, BR, and CD performed the phenotyping; PM developed the recombinant populations and conceived the experiment; JB, IL-H, and NT edited and critically reviewed the manuscript; GA conceived and supervised the experiments, analyzed the data, performed the synteny analysis, edited and reviewed the manuscript. All authors read and approved the manuscript.

Funding

This research was supported by the PeaMUST project, which was funded by the French government through the Investment for the Future program (project ANR-11-BTBR-0002) and Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation.

Acknowledgments

Authors thank the field and laboratory staff at INRAE and Agri Obtentions for technical assistance and for their participation in the production of the recombinant inbred line populations and phenotyping.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.970865/full#supplementary-material

References

Adhikari, L., Lindstrom, OM, and Markham and Missaoui, J AM. (2018). Dissecting key adaptation traits in the polyploid perennial medicago sativa using GBS-SNP mapping. *Front. Plant Sci.* 9. doi: 10.3389/FPLS.2018.00934

Adhikari, L., Makaju, SO, Lindstrom, O. M., and Missaoui, A. M.. (2021). Mapping freezing tolerance QTL in alfalfa: based on indoor phenotyping. *BMC Plant Biol.* 21 (1), 1–13. doi: 10.1186/S12870-021-03182-4

Ali, M. B. M. , Welna, G. C., Sallam, A., Martsch, R., Balko, C., Gebser, B., et al. (2016). Association analyses to genetically improve drought and freezing tolerance of faba bean (Vicia faba l.). *Crop Sci.* 56 (3), 1036–1048. doi: 10.2135/CROPSCI2015.08.0503

Alonso-Blanco, C., Gomez-Mena, C., Llorente, F., Koornneef, M., Salinas, J., Martínez-Zapater, J. M., et al. (2005). Genetic and molecular analyses of natural variation indicate CBF2 as a candidate gene for underlying a freezing tolerance quantitative trait locus in arabidopsis. *Plant Physiol.* 139 (3), 1304–1312. doi: 10.1104/PP.105.068510

Ambroise, V., Legay, S., Guerriero, G., Hausman, J. F., Cuypers, A., and Sergeant, K.. (2020). The roots of plant frost hardiness and tolerance. *Plant Cell Physiol.* 61 (1), 3–20. doi: 10.1093/pcp/pcz196

Annicchiarico, P., Barrett, B., Brummer, E. C., Julier, B., Marshall, AH., et al. (2015). Achievements and challenges in improving temperate perennial forage legumes. *Crit. Rev. Plant Sci.* 34 (1–3), 327–380. doi: 10.1080/07352689.2014.898462

Arbaoui, M., Balko, C., and Link, W. (2008b). Study of faba bean (Vicia faba l.) winter-hardiness and development of screening methods. *Field Crops Res.* 106 (1), 60–67. doi: 10.1016/j.fcr.2007.10.015

Arbaoui, M., Link, W., Satovic, Z., and Torres, A-M. M.. (2008a). Quantitative trait loci of frost tolerance and physiologically related trait in faba bean (Vicia faba l.). *Euphytica* 164 (1), 93–104. doi: 10.1007/s10681-008-9654-0

Arbaoui, M., and Link, W. (2008). Effect of hardening on frost tolerance and fatty acid composition of leaves and stems of a set of faba bean (Vicia faba l.) genotypes. *Euphytica* 162 (2), 211–219. doi: 10.1007/s10681-007-9521-4

Ates, D., Aldemir, S., Alsaleh, A., Erdogmus, S., Nemli, S., Kahriman, A., et al. (2018). A consensus linkage map of lentil based on DArT markers from three RIL mapping populations. *PloS One* 13 (1), 1–13. doi: 10.1371/journal.pone.0191375

Avia, K., Pilet-Nayel, M. L., Bahrman, N., Baranger, A., Delbreil, B., Fontaine, V., et al. (2013). Genetic variability and QTL mapping of freezing tolerance and related traits in medicago truncatula. *Theor. Appl. Genet. 2013 126:9* 126 (9), 2353–2366. doi: 10.1007/S00122-013-2140-7

Beji, S., Fontaine, V, Devaux, R, Thomas, M, Negro, SS, Bahrman, N, et al. (2020). Genome-wide association study identifies favorable SNP alleles and candidate genes for frost tolerance in pea. *BMC Genomics* 21 (1), 1–21. doi: 10.1186/S12864-020-06928-W

Brouwer, D. J., Duke, S. H., and Osborn, T. C. (2000). Mapping genetic factors associated with winter hardiness, fall growth, and freezing injury in autotetraploid alfalfa. *Crop Sci.* 40 (5), 1387–1396. doi: 10.2135/CROPSCI2000.4051387X

Carrillo-Perdomo, E., Vidal, A., Kreplak, J., Duborjal, H., Leveugle, M., Duarte, J., et al. (2020). Development of new genetic resources for faba bean (Vicia faba l.) breeding through the discovery of gene-based SNP markers and the construction of

a high-density consensus map. *Sci. Rep.* 10 (1), 1–14. doi: 10.1038/s41598-020-63664-7

Castel, T., Lecomte, C., Richard, Y., Lejeune-Hénaut, I., and Larmure, A. (2017). Frost stress evolution and winter pea ideotype in the context of climate warming at a regional scale. *OCL*, 24 (1), D106.

Crépon, K., Marget, P., Peyronnet, C., Carrouée, B., Arese, P., and Duc, G.. (2010). Nutritional value of faba bean (Vicia faba l.) seeds for feed and food. *Field Crops Res.* 115 (3), 329–339. doi: 10.1016/j.fcr.2009.09.016

Cruz-Izquierdo, S., Avila, C. M., Satovic, Z., Palomino, C., Gutierrez, N., Ellwood, S. R., et al. (2012). Comparative genomics to bridge vicia faba with model and closely-related legume species: Stability of QTLs for flowering and yield-related traits. *Theor. Appl. Genet.* 125 (8), 1767–1782. doi: 10.1007/s00122-012-1952-1

Deokar, A., Sagi, M., and Tar'an, B. (2019). Genome-wide SNP discovery for development of high-density genetic map and QTL mapping of ascochyta blight resistance in chickpea (Cicer arietinum l.). *Theor. Appl. Genet.* 132 (6), 1861–1872. doi: 10.1007/S00122-019-03322-3

Duc, G., Link, W., Marget, P., Redden, R. J., Stoddard, F. L., Torres, A. M., Cubero, J. I., et al (2011). Genetic adjustment to changing climates: faba bean. *Crop adaptation to climate change*, 269–86.

Duc, G. (1997). 'Faba bean(vicia faba l.). *Field Crops Res.* 53, 99–109. doi: 10.1016/S0378-4290(97)00025-7

Duc, G., and Petitjean, F. (1995). "Study on the inheritance of freezing tolerance in vicia faba l," in *European Conference on Grain Legumes, Copenhagen* 130–1.

Dumont, E., Fontaine, V., Vuylsteker, C., Sellier, H., Bodèle, S., Voedts, N., et al. (2009). Association of sugar content QTL and PQL with physiological traits relevant to frost damage resistance in pea under field and controlled conditions. *Theor. Appl. Genet. 2009 118:8* 118 (8), 1561–1571. doi: 10.1007/S00122-009-1004-7

Flores, F., Hybl, M., Knudsen, J. C., Marget, P., Muel, F., Nadal, S., et al. (2013). Adaptation of spring faba bean types across European climates. *Field Crops Res.* 145, 1–9. doi: 10.1016/j.fcr.2013.01.022

Francis, J., and Skific, N. (2015). 'Evidence linking rapid Arctic warming to mid-latitude weather patterns'. *Philos. Trans. R. Soc. A* 373, 20140170. doi: 10.1098/rsta.2014.0170

Hanson, P. J., Post, W. M., Kaiser, D. P., Yang, B., and Nemani, R.. (2008). The 2007 eastern US spring freeze: increased cold damage in a warming world? *BioScience* 58 (3), 253–262. doi: 10.1641/B580311

Gutierrez, N., and Torres, A. M. (2021). QTL dissection and mining of candidate genes for ascochyta fabae and orobanche crenata resistance in faba bean (Vicia faba l.). *BMC Plant Biol.* 21 (1), 1–12. doi: 10.1186/S12870-021-03335-5/TABLES/6

Guy, C. L. (2011). Freezing tolerance of plants: current understanding and selected emerging concepts. *Can. J. Bot.* 81 (12), 1216–1223. doi: 10.1139/B03-130

Jia, Y., Ding, Y., Shi, Y., Zhang, X., Gong, Z., and Yang, S.. (2016). The cbfs triple mutants reveal the essential functions of CBFs in cold acclimation and allow the definition of CBF regulons in arabidopsis. *New Phytol.* 212 (2), 345–353. doi: 10.1111/NPH.14088

Jourjon, M.-F., Jasson, S., Marcel, J., Ngom, B., Mangin, B., et al. (2005). MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics* 21 (1), 128–130. doi: 10.1093/bioinformatics/bth481

Kahraman, A., Kusmenoglu, I., Aydin, N., Aydogan, A., Erskine, W., Muehlbauer, F. J., et al. (2004). QTL mapping of winter hardiness genes in lentil. *Crop Sci.* 44 (1), 13–22. doi: 10.2135/CROPSCI2004.1300

Kahraman, A., Demirel, U., Ozden, M., Muehlbauer, F. J., et al. (2015). *Afr. J. Biotechnol.* 9 (50), 8515–8519. doi: 10.4314/ajb.v9i50

Khazaei, H., O'Sullivan, D. M., Jones, H., Pitts, N., Sillanpää, M. J., Pärssinen, P., et al. (2015). Flanking SNP markers for vicine–convicine concentration in faba bean (Vicia faba l.). *Mol. Breed.* 35 (1), 38. doi: 10.1007/s11032-015-0214-8

Klein, A., Houtin, H., Rond, C., Marget, P., Jacquin, F., Boucherot, K., et al. (2014). QTL analysis of frost damage in pea suggests different mechanisms involved in frost tolerance. *Theor. Appl. Genet.* 127 (6), 1319–1330. doi: 10.1007/s00122-014-2299-6

Knox, A. K., Li, C., Vágújfalvi,A., Galiba, G., Stockinger, E. J., and Dubcovsky, J.. (2008). Identification of candidate CBF genes for the frost tolerance locus fr-a m 2 in triticum monococcum. *Plant Mol. Biol. 2008 67:3* 67 (3), 257–270. doi: 10.1007/S11103-008-9316-6

Kreplak, J., Madoui, M-A., Cápal, P., Novák, P., Labadie, K., Aubert, G., et al. (2019). A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* 51 (9), 1411–1422. doi: 10.1038/s41588-019-0480-1

Landry, E. J., and Hu, J. (2019). Increasing pre-acclimation temperature reduces the freezing tolerance of winter-type faba bean (*Vicia faba* l.). *J. Agron. Crop Sci.* 205 (1), 46–53. doi: 10.1111/jac.12289

Lejeune-Hénaut, I., Hanocq, E., Béthencourt, L., Fontaine, V., Delbreil, B., Morin, J., et al. (2008). The flowering locus hr colocalizes with a major QTL affecting winter frost tolerance in pisum sativum l. *Theor. Appl. Genet. 2008 116:8* 116 (8), 1105–1116. doi: 10.1007/S00122-008-0739-X

Li, X., Alarcón-Zúñiga, B., Kang, J., Tahir, M. H.N., Jiang, Q., Wei, Y., et al. (2015). Mapping fall dormancy and winter injury in tetraploid alfalfa. *Crop Sci.* 55 (5), 1995–2011. doi: 10.2135/CROPSCI2014.12.0834

Link, W., Balko, C., and Stoddard, F. L. (2010). Winter hardiness in faba bean: Physiology and breeding. *Field Crops Res.* 115 (3), 287–296. doi: 10.1016/J.FCR.2008.08.004

Liu, Y., Dang, P., Liu, L., and He, C.. (2019). Cold acclimation by the CBF–COR pathway in a changing climate: Lessons from arabidopsis thaliana. *Plant Cell Rep.* 38 (5), 511–519. doi: 10.1007/S00299-019-02376-3

Maqbool, A., Shafiq, S., and Lake, L. (2010). Radiant frost tolerance in pulse crops-a review. *Euphytica. Springer* pp, 1–12. doi: 10.1007/s10681-009-0031-4

Mugabe, D., Coyne, C. J., Piaskowski, J., Zheng, P., Ma, Y., Landry, E., et al. (2019). Quantitative trait loci for cold tolerance in chickpea. *Crop Sci.* 59 (2), 573–582. doi: 10.2135/CROPSCI2018.08.0504

Mulualem, T., Dessalegn, T., and Dessalegn, Y. (2012). Participatory varietal selection of faba bean (*Vicia faba* l.) for yield and yield components in dabat district, Ethiopia. *Wudpecker J. Agric. Res.* 7, 270–274.

Ondrej, M., and Hunady, I. (2007). Faba bean (*Vicia faba* l.) breeding for resistance to anthracnose (*Ascochyta fabae* speg.) in the Czech republic. *Czech J. Genet. Plant Breed.* 43 (2), 61.

Pecrix, Y., Staton, S. E., Sallet, E., Lelandais-Brière, C., Moreau, S., Carrère, S., et al. (2018). Whole-genome landscape of medicago truncatula symbiotic genes. *Nat. Plants* 4 (12), 1017–1025. doi: 10.1038/s41477-018-0286-7

Sallam, A., et al. (2016a). Association mapping of winter hardiness and yield traits in faba bean (*Vicia faba* l.). *Crop Pasture Sci.* 67 (1), 55–68. doi: 10.1071/CP15200

Sallam, A., et al. (2016b). Identification and verification of QTL associated with frost tolerance using linkage mapping and GWAS in winter faba bean. *Front. Plant Sci.* 7 (1098). doi: 10.3389/fpls.2016.01098

Sallam, A., and Martsch, R. (2015). Association mapping for frost tolerance using multi-parent advanced generation inter-cross (MAGIC) population in faba bean (Vicia faba l.). *Genetica* 143, 501–514. doi: 10.1007/s10709-015-9848-z

Sallam, A., and Martsch, R. (2016). Validation of RAPD markers associated with frost tolerance in winter faba bean (Vicia faba l.). *Turkish J. Bot.* 40, 488–495. doi: 10.3906/bot-1508-21

Sallam, A., Martsch, R., and Moursi, Y. S. (2015). Genetic variation in morpho-physiological traits associated with frost tolerance in faba bean (Vicia faba l.). *Euphytica* 205 (2), 395–408. doi: 10.1007/s10681-015-1395-2

Sallam, A., and Ul-Allah, S. (2019). Genomics-aided breeding for climate-smart traits in faba bean. *Genomic Designing Climate-Smart Pulse Crops*, 359–395. doi: 10.1007/978-3-319-96932-9_7

Sandve, S. R., et al. (2011). Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates. *Plant Sci.* 180 (1), 69–77. doi: 10.1016/J.PLANTSCI.2010.07.011

Scarth, G. W. (1944). Cell physiological studies of frost resistance: A review. *New Phytol.* 43 (1), 1–12. doi: 10.1111/j.1469-8137.1944.tb04997.x

Schmutz, J., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463 (7278), 178–183. doi: 10.1038/nature08670

Sieber, A. N., et al. (2016). Copy number variation of CBF-A14 at the fr-A2 locus determines frost tolerance in winter durum wheat. *Theor. Appl. Genet. 2016 129:6* 129 (6), 1087–1097. doi: 10.1007/S00122-016-2685-3

Stoddard, F. L., et al. (2006). Screening techniques and sources of resistance to abiotic stresses in cool-season food legumes. *Euphytica* 147 (1–2), 167–186. doi: 10.1007/s10681-006-4723-8

Tayeh, N., Bahrman, N., Devaux, R., et al. (2013a). A high-density genetic map of the medicago truncatula major freezing tolerance QTL on chromosome 6 reveals colinearity with a QTL related to freezing damage on pisum sativum linkage group VI. *Mol. Breed. 2013 32:2* 32 (2), 279–289. doi: 10.1007/S11032-013-9869-1

Tayeh, N., Bahrman, N., Sellier, H., et al. (2013b). A tandem array of CBF/DREB1 genes is located in a major freezing tolerance QTL region on medicago truncatula chromosome 6. *BMC Genomics* 14 (1), 1–16. doi: 10.1186/1471-2164-14-814/FIGURES/4

Thomashow, M. F. (2010). Molecular basis of plant cold acclimation: Insights gained from studying the CBF cold response pathway. *Plant Physiol.* 154 (2), 571–577. doi: 10.1104/PP.110.161794

Vágújfalvi, A., et al. (2003). The cold-regulated transcriptional activator Cbf3 is linked to the frost-tolerance locus fr-A2 on wheat chromosome 5A. *Mol. Genet. Genomics 2003 269:1* 269 (1), 60–67. doi: 10.1007/S00438-003-0806-6

Varshney, R. K., et al. (2014). Genetic dissection of drought tolerance in chickpea (Cicer arietinum l.). *Theor. Appl. Genet.* 127 (2), 445–462. doi: 10.1007/s00122-013-2230-6

Webb, A., et al. (2016). A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (Vicia faba l.). *Plant Biotechnol. J.* 14 (1), 177–185. doi: 10.1111/pbi.12371

Wilcoxson, R. D., Skovmand, B., and Atif, A. H. (1975). Evaluation of wheat cultivars for ability to retard development of stem rust. *Ann. Appl. Biol.* 80, 275–281. doi: 10.1111/j.1744-7348.1975.tb01633.x

Zhang, W. B., et al. (2011). Dissection of genetic overlap of drought and low-temperature tolerance QTLs at the germination stage using backcross introgression lines in soybean. *Mol. Biol. Rep. 2011 39:5* 39 (5), 6087–6094. doi: 10.1007/S11033-011-1423-9

Zhang, W. B., et al. (2012). Genetic overlap of QTL associated with low-temperature tolerance at germination and seedling stage using BILs in soybean. *Can. J. Plant Sci.* 92 (7), 1381–1388. doi: 10.4141/CJPS2011-098

Zhao, C., et al. (2016). Mutational evidence for the critical role of CBF transcription factors in cold acclimation in arabidopsis. *Plant Physiol.* 171 (4), 2744–2759. doi: 10.1104/PP.16.00533

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 November 2022
DOI 10.3389/fpls.2022.950992

Check for updates

OPEN ACCESS

EDITED BY
Jose C. Jimenez-Lopez,
Estación Experimental del Zaidin, (CSIC), Spain

REVIEWED BY
Naiyin Xu,
Jiangsu Academy of Agricultural Sciences (JAAS), China
Ayda Krisnawati,
Research Organization for Agriculture and Food, National Research and Innovation Agency (BRIN), Indonesia
Jan Bocianowski,
Poznan University of Life Sciences, Poland

*CORRESPONDENCE
Mesfin Hailemariam Habtegebriel
mesfin.hailemariam@aau.edu.et

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 23 May 2022
ACCEPTED 20 September 2022
PUBLISHED 23 November 2022

CITATION
Hailemariam Habtegebriel M (2022) Adaptability and stability for soybean yield by AMMI and GGE models in Ethiopia. *Front. Plant Sci.* 13:950992. doi: 10.3389/fpls.2022.950992

COPYRIGHT
© 2022 Hailemariam Habtegebriel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Adaptability and stability for soybean yield by AMMI and GGE models in Ethiopia

Mesfin Hailemariam Habtegebriel [1,2*]

[1]Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, Addis Ababa, Ethiopia, [2]Ethiopian Institute of Agricultural Research (EIAR), Addis Ababa, Ethiopia

Genotype by environment interaction (GEI) is a phenomenon that occurs in heterogeneous environments that slows breeding progress by preventing the selection of superior cultivars for breeding and commercialization. Therefore, the objectives of this study were to find out how GEI impacts soybean output and to identify the most adapted and stable genotypes. Moreover, to look at the possibility of other mega environments for testing in the future. The experiments were grown for two years in a four-replicated randomized block design at each environment. Over the course of several harvests, yield components, days to flowering, days to maturity, plant height, the number of pods per plants, the number of seeds per plant, hundred seed weight and grain yield per hectare were evaluated in the main for 2018 and 2019.To analyze the stability performance of the genotypes, general linear method, GGE and Additive main effect and multiplicative interaction effects analysis (AMMI) and ASV rank analysis were applied. The GGE biplot revealed that the GGE biplots explained 74.29% of the total variation distributed as,56.69% and 17.62% of sum of squares between principal component PC1 and PC2, respectively whereas, AMMI model, the first two interaction principal component axes (IPCA1 and IPCA2) explained 47.74% and 26.62% of the variation due to GEI, respectively, exposed genotypes identified the five as best performer. The results from the four distinct stability statistics AMMI biplot (G8, G2, G1, G11), ASV (G1, G11; (GSI; G9, G1, G11) and (GGE: G2, G8, G9) are taken into account together with the genotypes` grand mean. The genotypes JM-CLK/CRFD-15-SD (G8) and 5002T (G1), which rank among the best and have the highest seed output, are suitable for hybridization as a parent and commercial production. Therefore, genotypes JM-CLK/CRFD-15-SD (G8) and 5002T(G1) have the highest seed output were among the best and thus could be recommended for release as a new soybean varieties cultivation across.

KEYWORDS

genotype by environment interaction, *Glycine max*., predictability and performance trials, medium set, mega environment

1 Introduction

Cultivated soybean [*Glycine max (L.) Merrill*] is a self-pollinating and an important leguminous source of food, feed, fuel and oil crop with high commercial value widely cultivated globally (Bocianowski et al., 2019; Sritongtae et al., 2021 and Abebe et al., 2021). First, the annual wild soybean (*Glycine soja*), the kindred ancestor of the current cultivated soybean (*Glycine max*), originated and is found throughout China (Qiu and Chang, 2010). In Ethiopia, it is commonly known "*akuri ater*" in *Amharic, "akuri bonsitu"* in Oromic, imported for tackling for multi purposes problems in early 1950's. The crop also thrives in places as low as 500 m.a.s.l and as high as 1900 m.a.s.l. that get a well-distributed average rainfall of 550 to 700 mm over the growing year.

In addition, a lack of soybean genotypes suitable for growth in diverse conditions as soybean breeders have developed their own varieties for specific purposes (Sritongtae et al., 2021). Similarly, Oladosu et al. (2017) and Mushoriwa et al. (2022) also reported that the GEI reveals the challenges that plant breeders have, when identifying a superior genotype for commercial farming before releasing it as a variety. Due to the fact that analysis of variance (ANOVA) cannot adequately shed light on the genotypes or environments that contribute to the interaction (Oladosu et al., 2017; Khan et al., 2021). In general, the knowledge and extent of the GEI provided by this study would aid breeders in allocating limited resources to the appropriate varietal development cycle, as a phenotype is a combination of genotype (G) and environment (E) components, as well as a GEI. Plant breeders conduct multi-environment experiments (MET) to examine the yield stability performance of genetic resources under diverse environmental circumstances (Yan and Hunt, 2002; Mwiinga et al., 2020).

Many researchers and breeders have investigated and reported grain yield stability using a variety of soybean genotypes in various agro-climatic zones (Samyuktha et al., 2020; Mwiinga et al., 2020; Happ et al., 2021), are useful for evaluating crop yield stability across varied environmental circumstances and determining acceptable habitats for all genotypes studied. Many stability methods, such as regression, have been devised to uncover the GEI pattern (Eberhart and Russell, 1966), but AMMI and GGE were by far providing the better interpretation of GEI data and widely used than the rest of stability procedures (Bhartiya et al., 2017; Khan et al., 2021).The AMMI technique combines ANOVA and principal component analysis with a biplot displaying the genotype means and their relationship to the first PCA as a key output (Mitrović et al., 2012; Bocianowski et al., 2019). The GGE biplot will aid researchers in better understanding complicated GEIs in multi-environment breeding line trials and agronomic investigations (Luo et al., 2015).

Phenotypic selection, one of the most critical processes in genetic improvement, is the foundation of traditional soybean breeding. According to Mederos-Ramirez et al. (2021), GEI in the improvement programs has a determining influence on the new cultivar obtaining and more stable and adapted genotypes. However, several studies on the adaptation and stability of soybean genotypes (Gurmu et al., 2009; Krisnawati et al., 2018a; Happ et al., 2021; Mushoriwa et al., 2022). So far, several statistical models have been proposed to analyze feature stability and GEIs through traditional analysis. The AMMI and the genotype and genotype x environmental interaction effect (GGE) model can detect GEIs in terms of crossover effects due to significant changes in ranking. Therefore, it is widely used. The result is the surrounding genotype. In addition, AMMI and GGE have proven to be particularly useful in visualizing GEI effects in graphic representations. Therefore, the objectives of this study were (i). to investigate how GEI affects in the soybean yield using AMMI and GGE analysis and (ii). identification of most adapted and stable, best-performing genotypes that can recommend as a variety to potential customers. Furthermore, to investigate the existence of other mega-environments in order to determine the optimum test environments for future testing.

2 Materials and methods

2.1 Plant genetic materials, trial planning and execution

Nine elite soybean lines developed by Jimma Agricultural Research Centre, the national soybean program, and SCS-1 and Nyala were used as control (Table 1) were used in the study. Over two years, eleven environments, experiments with genotypes were performed in a randomized complete block design (RCBD) with four replicates was used to conduct the genotype experiment over each year. The plot size was 4 x 1.6 m, there were four rows and the spacing between rows was 60 cm and between plants was maintained at 5 cm. Two seeds per hill were planted, and after 2-3 weeks, the hill was reduced to one plant per hill. The NPS and Urea fertilizer was applied at the recommended amount of 122 kg ha^{-1} and 50 kg ha^{-1} during sowing and flowering and respectively. Weeds were controlled three times, remainder of the agronomic management practices like pest or diseases control was done as required.

2.2 Study area

The trials were conducted over two years in 2018 and 2019 at Jimma (L1), Metu (L2), Asosa (L3), Shire (L4), Gondar (L5), Jinka (L6), and Tepi (L7). The trials were planted throughout two years in 2018 and 2019, but at Gondar, Jinka, and Tepi, the

TABLE 1 Description of the eleven soybean genotypes used in this study.

Genotype	Pedigree	Source	Maturity	Remark	Growth Habit	Generation
G1	5002T	USA introduction	Medium	Promising line	Indeterminate	F8
G2	SCS-1	CIMMYT-Zimbabwe	Medium	Released	Indeterminate	Control
G3	Ozark	USA – introduction	Medium	Promising line	Indeterminate	F8
G4	KS4895	USA –introduction	Medium	Promising line	Indeterminate	F8
G5	Harber	USA – introduction	Medium	Promising line	Indeterminate	F8
G6	JM-PR142/CLK-15-SE	Ethiopia- Cross	Medium	Promising line	Indeterminate	F8
G7	Hs93-4118	USA –introduction	Medium	Promising line	Indeterminate	F8
G8	JM-CLK/CRFD-15-SD	Ethiopia- Cross	Medium	Promising line	Indeterminate	F8
G9	PI471904	USA- introduction	Medium	Promising line	Indeterminate	F8
G10	PI417089A	USA- introduction	Medium	Promising line	Indeterminate	F8
G11	Nyala	Ethiopia –released	Medium	Released	Indeterminate	Control

trials were only planted in 2019 (See Figure 1) with the total descriptions below (Table 2).

2.3 Evaluation of morphological traits and method of phenotypic data collection

All phenotypic data were collected from 10 randomly selected and tagged plants in each plot in each replication at various growth stages in the field and the field crops laboratory after harvest according to the soybean descriptor (1986). Days to flowering, days to maturity, plant height (cm), the number of pods per plant, the number of seeds per plant, and hundred seed weight (gm) were all reported as agronomic parameters.

2.4 Statistical analysis

2.4.1 Analysis of variance

Prior to proceeding with the analysis of variance (ANOVA), a homogeneity test was conducted, and then all the data

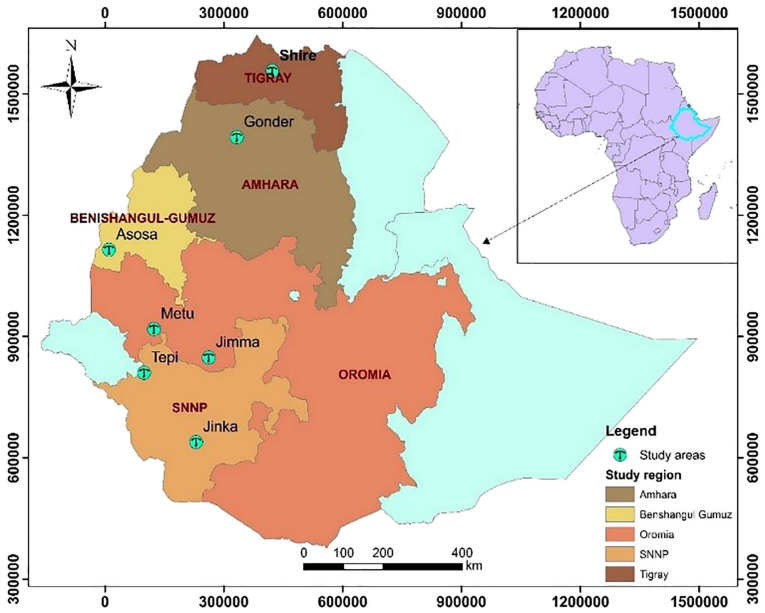


FIGURE 1
Geographical map (constructed using geographic information system (ArcGIS) showing testing environments: Jimma and Metu are situated Oromia regional state, Shire is located in Tigray regional state, Gondar situated Amhara regional state, Jinka and Tepi are located in southern nation, nationalities and peoples Regional State, Asosa (2018-2019) is situated Benishangul Gumez regional states.

TABLE 2 The eleven sites and soil attributes used for evaluation across the country are described.

Year Designations	Environment	Year	Latitude (N)	Longitude (E)	Elevation (m.a.s.l)	Rainfall (mm)	Sunshine hour (hr.)	Average temperature (o c)	R.H (%)	Soil type
Y1L1	Jimma	2018	$36^0$82'	$7^0$67'	1775	1561	6.8	9-28	65.1	Chromic Nitosol and Combisol
Y1L2	Metu	2018	$35^0$57'	$8^0$28'	1550	1835	–	12.5-28.6	82	Reddish brown
Y1L3	Asosa	2018	$34^0$52'	$10^0$00'	1580	1,236	7	14- 39	58	Dystric Nitosols
Y1L4	Shire	2018	$38^0$29'	$14^0$10'	1871	905	7.8	17.57	52.6	Cambisols
Y2L1	Jimma	2019	$36^0$82'	$7^0$67'	1775	1561	5.6	9-28	74.9	Chromic Nitosol and Combisol
Y2L2	Metu	2019	$35^0$57'	$8^0$28'	1550	1835	–	12.5-28.6	82	Reddish brown
Y2L3	Asosa	2019	$34^0$52'	$10^0$00'	1580	1,236	8	16.5- 27.5	38	Dystric Nitosols
Y2L4	Shire	2019	$38^0$29'	$14^0$10'	1871	905	7.8	12.5-28.2	46	Cambisols
Y2L5	Gondar	2019	$37^0$43'	$12^0$52'	1973	912	7.7	19.8-26.1	56.1	Vertisol
Y2L6	Jinka	2019	$36^0$55'	$5^0$77'	1375	162.9	5.5	15.56-29.6	67.3	Nitosol
Y2L7	Tepi	2019	$35^0$44'	$7^0$20'	1208	1559	–	15.5 -29.7	78	Nitosol

m.a.s. l, meter above sea level; mm, millimetres; hr, hours; o ^{c}, degree centigrade; %, percent; R.H, Relative humidity.

considered were subjected to a combined analysis of variance over their environment, which was performed using the SAS program software with the following the combined ANOVA model equation:

$$Y_{ijkl} = \mu + G_i + E_j + Rk_{(j)} + Bl(jk) + GE_{ij} + \epsilon_{ijkl} \quad \text{(Combined ANOVA model)}$$

Where Y_{ijkl} is the response of the ith genotype in jth environment and k^{th} replication within environment and l^{th} block within replication; μ is the grand mean, G_i is the genotype effect i; E_j is the environment effect j; Rk_j is the replication within environment effect k;Bl_{jk} is the block within replication effect l; GE_{ij} is the genotype × environment interaction effect; and ε_{ijkl} is the random error.

2.4.2 Additive main effects and multiplicative interaction (AMMI) analysis and AMMI's stability value (ASV)

AMMI analysis was carried out using the software GEA-R version 4.1 (Angela et al., 2015) with the model equation:

$$Y_{ger} = \mu + \alpha_g + \beta_e + \sum_{n=1}^{N} \lambda\varsigma^{gn}\eta_{en} + \rho_{ge} + \epsilon_{ger} \quad \text{(AMMI model)}$$

where:

$\mathbf{Y_{ger}}$=the grain yield level for genotypes g in environment e for replicate r

μ = the grand mean

$\boldsymbol{\alpha_g}$= genotype mean deviations (mean minus the grand mean)

$\boldsymbol{\beta_e}$=the environmentsmean mean deviations

N=the number of singular value decompostion(SVD) axes retained in the model

$\boldsymbol{\lambda_n}$= the singular value for SVD axis

$\boldsymbol{\eta\varsigma_{gn}}$ = the genotype of singular vector values for SVD axis $\boldsymbol{n}$

$\boldsymbol{\theta_{ge}}$ =the interaction residuals

$\boldsymbol{\rho_{ge}}$ =The AMMI residuals

$\boldsymbol{\varepsilon_{ger}}$ =Error term mean

$\sum_{n=1}^{N}\lambda\varsigma^{gn}\eta_{en}+\rho_{ge}+\epsilon_{ger}$ is equivalent to the interaction term in the ANOVA model.

The AMMI stability value (ASV) as described by Purchase et al. (2000), was calculated as follows: $ASV = \sqrt{\left[\frac{(SS_{IPCA1})}{(SS_{IPCA2})}(IPCA_1)\right]^2 + (IPCA_2)^2}$, where, SS is the sum of squares, IPCA1 and IPCA2 are the first and second interaction principal component axes, respectively; and the IPCA1 and IPCA2 scores were the genotypic scores in the AMMI model. ASV is the distance from zero in a two-dimensional scatterplot of IPCA1 scores against IPCA2 scores.

Another parameter that we can calculate in AMMI analysis is the genotype selection index (GSI). Selection for stability will not necessarily lead to the best genotype. Thus, the GSI for each genotype was computed *via* the sum of the rank of the genotype grain yield (RYi) and the rank of the genotype ASV (RASVi). The most stable genotype was suggested to be the one with the lowest GSI (Farshadfar et al., 2012).

2.4.3 GGE analysis

If there is a significant difference in genotype–environment interaction, the GGE biplot method will be employed to analyse and assess the interaction and yield stability (Olanrewaju et al., 2021). A general model of GGE biplot based on environment-centric or environment-standardized singular value decomposition (SVD) can be written as:

$$Y_{ij} = \mu + \beta_j + \lambda_1\xi_{i1}\eta_{j1} + \lambda_2\xi_{i2}\eta_{j2} + \epsilon_{ij}$$

Where, Y_{ij}, is the trait mean for genotype i in environment j; μ is the grand mean; β_j is the main effect of environment j; μ + β_j

is the mean yield across all gen'genotypes in environment j; λ1 and λ2 are the singular values (SV) for the first and second principal components), respectively; ξ_{i1} and ξ_{i2} are eigenvectors of genotype *i*for PC1 and PC2, respectively; η1j and η2j are eigenvectors of environment j for PC1 and PC2; and ξ_{ij} is the residual associated with genotype *i* in environment *j*. In GGE biplot analysis, scores of PC1 were plotted against PC2.

3 Results

3.1 Combined analysis of variance for yield and related traits

For the majority of the factors, the results of the analysis of variance show that there is a substantial difference between genotypes, environment and genotypes x environment (P<0.001) (Table 3). The yield production of the examined soybean genotypes ranged from 1.70 t ha^{-1} to 2.84 t ha^{-1}, according to the combined data analysis results. The average grain yield of the control variety was 2.62 t ha^{-1}. The genotypes JM-CLK/CRFD-15-SD had a maximum grain yield of 2.85 t ha^{-1}, while the genotypes Hs93-4118 had the lowest grain yield of 1.70 t ha^{-1}, respectively. The genotypes 5002T and SCS-1 yielded more than the control variety.

To explain the main effect and quantify the interactions among and within the sources of variability, a combined analysis of variance was performed. The findings of the pooled analysis of variance are shown in Table 3. Significant variations (p≤ 0.01, p ≤0.05) were found in the mean square of environments, genotypes, and GEI for DTF, DTM, and NSP. For genotypes by environment interaction (GEI), all but one variable had significant variance, while the trait had non-significant variation for genotypes by environment interactions (GEI). Large differences in environments, and genotypes could be attributed to differences in environmental conditions and genetic makeup that differ from one place to the next. Table 3 shows the percentages of variation for all attributes when the proportion of GEI (% of GEI) is partitioned using the whole sum of the squares. Other attributes exhibited a large range of variance related to environment, ranging from 0.61 to 86.92%, with the exception of the total number of pods (NPP). The majority of genotype performance variance is attributable to a greater disparity in genotype means between sites. Genotype by environment variation was found to be lower, ranging from 0.42 to 2.13%. Days to flowering (15.08%), plant height (42.59%), and hundred seed weight (24.28%) all contributed nearly the same amount to genotype effect, whereas days to maturity (4.24%), the number of pods per plant (6.76%), number of seeds per pod (0.923%), and yield per hectare (4.73%) all contributed very little. The genotype by environment interaction (GEI) is significant for almost all variables, including days to flowering, days to maturity, plant height, number of pods per plant, number of seeds per plant, and grain yield per hectare, but hundred seed weight.

3.1.1 Mean performance and comparison of genotypes

The means comparison and average performance of eleven soybean genotypes over eleven environments were listed in Table 4. Over the environment, all the genotypes showed significant variation for days to flowering, days to maturity, plant height, the number of pods per plant, the number of seeds per plant, hundred seed weight (HSW), and yield per hectare. The days to flowering ranged from 51.4 (5002T) to 63.3 (PI471904 and SCS-1) with an average of 55.9 days. The genotype PI471904 produced the highest number of days to maturity (122.5) followed by SCS-1 (119.71), JM-CLK/CRFD-15-SD (117.9), and JM-PR142/CLK-15-SE (117.7) though across the genotype it was 106 while the lowest was 51. The highest plant height was 86.8 cm for genotype PI471904, followed by PI417089A (81.5 cm), and the lowest was 44.4 cm for the genotype Ozark. The greater value of hundred seed weight was accounted for 21.6 g (5002T) followed by 21.0 g (PI417089A)

TABLE 3 Estimation of significant level for yield and yield contributed traits for eleven soybean genotypes revealed by ANOVA.

SOV	df	DTF(Days)		DTM (Days)		PH (cm)		NPP		NSP		HSW (gm)		Yield (t ha^{-1})	
		MS	TSS (%)	MS	TSS (%)	MS	TSS (%)	MS	TSS (%)	MS	TSS (%)	MS	TSS (%)	MS	TSS (%)
Replications (R)	3	58.5ns	0.71	18.1ns	0.13	57.9 ns	0.34	23.3 ns	0.20	362.5ns	0.080	1.9 ns	0.38	0.5 ns	0.79
Locations (L)	6	2804.4**	34.06	11980.6**	86.92	5162.5**	30.75	69.7**	0.61	126832.9**	27.889	238**	47.96	11.7**	18.46
Genotypes (G)	10	1241.7**	15.08	584.0**	4.24	7151**	42.59	768.7**	6.76	4198.1**	0.923	120.5**	24.28	3.0**	4.73
GEI	30	175.3**	2.13	95.7**	0.69	180.3**	1.07	182.1**	1.60	1914.7**	0.421	7.2ns	1.45	1.3**	2.05
Error	261	54.6**	0.66	20.7**	0.15	34.9**	0.21	61.3**	0.54	451.6**	0.099	2.2**	0.44	0.20**	0.32

SOV, Source of variations; df, degree of freedom; TSS (%), Total sum of squares; DTF, Days to flowering; DTM, Days to maturity; NPP, Number of pods per plant; NSP, Number of seeds per plant; HSW, Hundred seed weight; **, Significant at 0.01.

TABLE 4 Mean performance days to flowering, days to maturity, plant height, number of pods per plant and, number of seeds per, branch per plant early maturing across years and environments in 2018 and 2019.

N°	Genotype	Yield (t ha^{-1})												Combined value of yield related traits				
		Y1L1	Y1L2	Y1L3	Y1L4	Y2L7	Y2L1	Y2L2	Y2L3	Y2L4	Y2L5	Y2L6	Overall	DF	DM	PH (cm)	NPP	NSP	HSW (gram)
1.	5002T	2.97	2.61	1.64	3.72	2.24	2.82	2.90	3.70	2.89	2.18	2.40	2.73	51.4	114.6	64.1	29.6	70.5	20.6
2.	SCS-1	2.95	1.97	1.31	4.72	2.26	1.92	2.09	4.00	3.03	1.92	2.12	2.57	63.3	119.7	64.7	38.8	87.0	17.1
3.	Ozark	2.58	2.24	0.97	3.07	1.76	2.87	2.97	2.26	2.75	2.02	2.54	2.37	50.7	112.3	44.4	28.3	59.1	18.3
4.	KS4895	2.22	2.08	1.28	3.02	1.97	2.37	2.21	2.45	2.80	1.52	2.80	2.25	51.7	110.5	46.7	29.2	65.0	16.9
5.	Harber	2.31	1.70	1.17	2.65	1.95	2.58	2.30	2.42	2.39	1.32	2.21	2.09	51.2	111.6	45.3	30.9	69.3	18.8
6.	JM-PR142/CLK-15-SE	2.33	2.39	1.36	3.29	2.03	1.36	3.05	2.66	2.71	1.91	2.25	2.30	60.9	117.7	62.3	37.2	76.9	18.8
7.	Hs93-4118	1.40	1.19	0.40	2.30	1.59	2.39	1.84	1.31	2.97	1.28	2.04	1.70	45.2	108.1	42.4	27.6	62.3	18.6
8.	JM-CLK/CRFD-15-SD	3.05	2.72	1.91	3.77	2.47	2.37	2.98	4.27	2.91	2.17	2.62	2.84	62.8	117.9	64.5	36.7	78.8	18.4
9.	PI471904	1.76	2.60	1.53	3.74	3.24	1.53	2.41	3.06	3.18	2.96	3.03	2.64	63.3	122.5	86.8	42.2	96.2	14.1
10.	PI417089A	2.47	2.26	1.50	2.62	1.82	2.66	2.61	3.22	2.60	1.79	1.80	2.30	56.5	114.0	81.5	35.1	78.8	21.0
11.	Nyala(C1)	2.56	2.44	1.70	2.58	2.44	1.71	2.32	3.71	1.83	1.80	2.71	2.35	57.5	115.9	53.4	30.9	62.8	20.2
Mean		2.4	2.2	1.3	3.2	2.2	2.2	2.5	3.0	2.7	1.9	2.4	2.4	55.9	115.0	59.6	33.3	73.3	18.4
	Min. across environments	1.4	1.2	0.4	2.3	1.6	1.4	1.8	1.3	1.8	1.3	1.8	1.7	45.2	108.1	42.4	27.6	59.1	14.1
	Max. across environments	3.1	2.7	1.9	4.7	3.2	2.9	3.1	4.3	3.2	3.0	3.0	2.8	63.3	122.5	86.8	42.2	96.2	21.0
CV (%)		12.83	16.92	19.66	20.28	19.83	12.46	17.47	19.19	14.37	30.27	34.18	18.8	13.3	4.0	9.9	23.5	29.0	8.1
LSD $_{0.05}$		0.41	0.61	0.38	0.79	0.61	0.43	0.63	0.83	0.66	0.82	1.18	0.21	3.6	2.2	2.9	3.9	10.5	0.7

Jimma (L1), Metu (L2), Asosa (L3), Shire(L4), Gonder(L5), Jinka(L6), and Tepi(L7), DF, Days to flowering; DM, Days to maturity; Plant Height(cm), NPP, Number pf pods per plant; NSP, Number of seeds per plant; HSW, Hundred seed weight(gm).

10.3389/fpls.2022.950992

though the lower value was 14.1 g for PI471904. The average weight of hundred seeds was 18.4 g with a range of 14.1 g to 21.0 g over tested environments. In terms of yield per hectare, genotype JM-CLK/CRFD-15-SD had the highest yield (2.84 t ha^{-1}), followed by genotypes 5002T (2.73 t ha^{-1}) and PI471904 (2.64 t ha^{-1}), while genotype Hs93-4118 had the lowest yield of 1.7 t ha^{-1}. However, yield per hectare varied from 1.70 to 2.84 t ha^{-1}, with an average of 2.4 t ha^{-1} throughout the studied environments.

3.2 AMMI model analysis for grain yield

According to the results of the AMMI model's analysis of variance, genotype, environment, and GEI all had a significant ($p<0.001$) impact on grain yield, explaining 16.18%, 49.42%, and 33.91% of the variation, respectively. It also revealed three PCs with highly significant differences ($p<0.001$) and the first three interaction principal component of AMMI, explaining 47%, 26.62%, and 17.88% of the GEI with 19, 17, and 15 degrees of freedom (df), respectively, as well as the fourth PC with significant differences ($p<0.05$) explaining 7.77% of the interaction with the degree with total cumulative of nearly 100%. The environment explains a major amount of the yield differential, showing that the environments were different, according to AMMI.

3.3 IPCA scores, AMMI stability values and mean yields

According to this parameter, lines G6, G1, G10, G5 and G11 had lowest stability values of ASV, identified as stable lines. On contrary, genotypes G8, G7, G2, G4 and G3 had the highest stability value hence are less stable. GSI integrates both yield and stability across environments. Genotypes with lower GSI (G9, G1 and G11) were desirable since they combine high mean yield performance with stability. Three out of three had higher stability than the average site of 2.4 t ha^{-1}; G9 had a mean performance ranking of 2.64 t ha^{-1}, followed by G1 and G11, which have 2.53 t ha^{-1} and 2.48 t ha^{-1}, respectively. Since these three genotypes had above the average yield, we can consider as the stable and adapted genotypes for the wider productions. And, similarly for the environments Asosa19, Shire18, Shire19, Jimma18 and Jinka19 were stable places, which is less discriminatory. In addition, Shire19 which had the highest yield of 3.24 t ha^{-1} (Table 5).

TABLE 5 Average grain yield, IPCA score, ASV and GSI Soybean genotypes across sites.

Genotypes Name	Code	Mean	RY_i	IPCA1	IPCA2	ASV	ASV Rank(RASVi)	GSI
5002T	G1	2.75	4	0.37	-0.36	0.53	2	6
PI417089A	G10	2.30	7	-0.09	-0.53	0.53	3	10
Nyala (C1)	G11	2.35	5	0.56	-0.29	1.11	5	10
SCS-1	G2	2.56	3	0.74	0.11	5.12	9	12
Ozark	G3	2.37	6	-0.61	-0.26	1.43	7	13
KS4895	G4	2.24	9	-0.39	0.03	4.82	8	17
Harber	G5	2.09	10	-0.39	-0.33	0.56	4	14
JM-PR142/CLK-15-SE	G6	2.30	7	0.09	0.26	0.26	1	8
Hs93-4118	G7	1.70	11	-0.98	0.19	5.12	10	21
JM-CLK/CRFD-15-SD	G8	2.75	1	0.64	-0.02	19.96	11	12
PI471904	G9	2.64	2	0.06	1.22	1.22	6	8
Environments Name								
Y1L1	Jimma18	2.14	8	2.14	-0.96	4.84	1	9
Y1L2	Mettu18	2.52	4	2.52	-0.22	29.25	8	12
Y1L3	Asosa18	1.30	11	1.30	0.11	15.53	5	16
Y1L4	Shire18	2.72	3	2.72	-0.52	14.32	3	6
Y2 L1	Jimma19	2.45	5	2.45	0.31	19.47	6	11
Y2 L2	Mettu19	2.17	7	2.17	0.08	61.81	10	17
Y2 L3	Asosa19	3.03	2	3.03	1.15	8.03	2	4
Y2 L4	Shire19	3.24	1	3.24	0.47	22.19	7	8
Y2 L5	Gonder19	1.87	10	1.87	-0.07	48.11	9	19
Y2 L6	Jinka 19	2.38	6	2.38	-0.37	15.48	4	10
Y2 L7	Tepi19	2.13	9	2.13	0.02	293.91	11	20

ASV, AMMI stability value; IPCA1 and IPCA2 are the first and second interaction principal component axes; respectively and GSI, Genotype selection index; RYi, the sum of the rank of the genotype grain yield and RASVi , the rank of the genotype ASV (RASVi).

3.4 AMMI biplot analysis

The AMMI model analysis was used in this study to provide a substantial and informative summary of GEIs, as well as to examine the correlations and differences in genotype performance across diverse testing conditions. In the AMMI 1 biplot, the usual interpretation of the biplot is that the displacements along the abscissa indicate differences in main (additive) effects, whereas displacements along the ordinate indicate differences in interaction effects. The first two principal components, IPCA1 and IPCA2, explained 69.37% of the total GEI variation (Figure 2). In terms of the grain yield feature, G8 > G2 > G1 > G11 and G5 showed the minimum interplay between genotypes and environments.

Genotypes having a zero IPCA 1 score are less influenced by the environment and better adapted to all environments. Since the IPCA 1 scores of varieties G2, G6, G8, G4 and G9 were close to zero, they were the most adapted and stable genotypes across these environments (Figure 3). However, the mean yield of genotype G8 was higher than the remaining genotypes, so it is preferable since it had a mean yield above average, while the other two genotypes, G6 and G4, had a mean below average. In summary, an adapted and stable variety might not be the highest yielding.

The environmental score is connected to the origin by side lines. Short arrow does not provide powerful interactive power. The long arrow has strong interactions. Genotypes near the coordinates represent general adaptations, and further genotypes represent more specific adaptations to the environment (Gauch and Zobel, 1996). The scatter plot of grain yield vs. IPCA1 (Figure 3) illustrates that the superior genotype had a higher agricultural yield and IPCA1, which is shown on the vertical axis, had a minimum value and was near zero. The superior genotypes were G8< G2< G9. Whereas, it is located on the right side of the graph and close to zero in terms of the IPCA1 axis. The most unstable genotypes and the lowest grain yield among the genotypes belonged to G5 and G10, G7, G4, G6, and G8 had the least GEI. Y1L1, Y2L1, Y2L7, and Y2L5 were the unstable environments.

4 GGE biplot analysis

The genotype + genotype × environment (GGE) biplot analysis result of the eleven soybean genotypes evaluated across environments with respect to yield. The GGE biplots explained 74.29% of the total variation distributed as 56.69% and 17.62% of sum of squares between principal component PC1 and PC2, respectively (Figure 4).

4.1 Which-won-where polygon view of GGE biplot

The polygon is formed by connecting the signs of the genotypes that are farthest away from the biplot origin, such that all other genotypes are contained in the polygon. In this case, the polygon

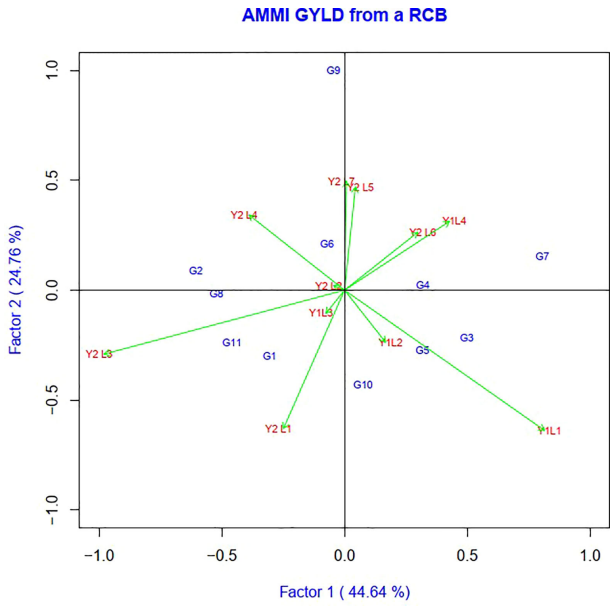


FIGURE 2
The biplot 'AMMI 2' illustrated the first two principal components (PC1 and PC2) effects of genotype plus GE interaction effect of 11 soybean genotypes under two years (11environemnets) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

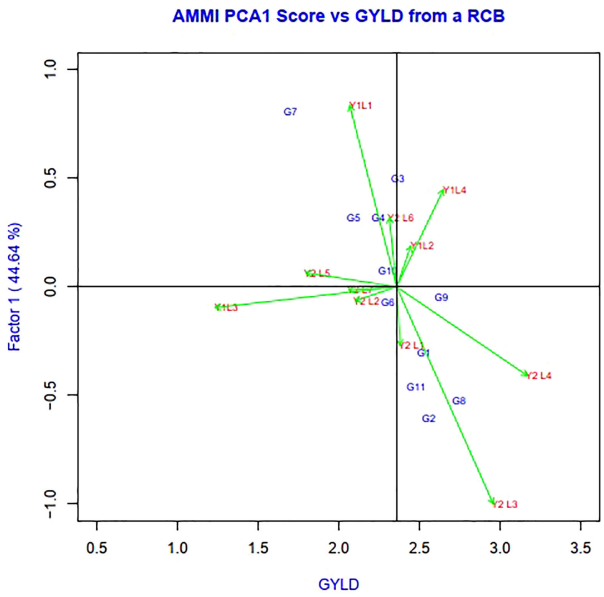


FIGURE 3
The biplot 'AMMI 1' illustrated the trait main effect and first interaction principal component (IPC1) effects of both genotype and environment of 11 soybean genotypes under two years (11 environments) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

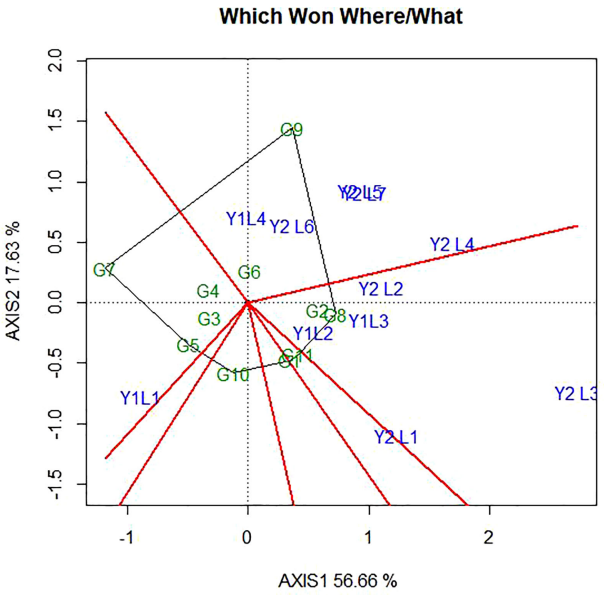


FIGURE 4
"Which-won where" pattern of GGE biplot polygon view displaying the genotype main effect plus GEI effect of 11 soybean genotypes under two years (11 environments) for grain yield per hectare. The biplots were based on Centering =0, SVP=2, Scaling=2.

connects all the farthest genotypes and perpendicular lines divide the polygon into sectors. Sectors help to visualize the mega-environments. This means that winning genotypes for each sector are placed at the vertex. The pattern on the environment in the biplot suggests that the existence of seven sector and five different mega-environments (Figure 4). The vertex genotype of each sector is the one that gave the highest yield for the environments which fall within that sector. The vertex genotypes in this study were G2, G11, G7, G8, G10, G1 and G9. According to Yan and Tinker (2006), the vertex genotypes were the most responsive genotypes, as they have the longest distance from the origin in their direction. From this figure, G9 was the best performer at Y2L5, Y2L7, Y1L1, Y2L6 and Y2L4 on the first environment; the second environment containing the higher yielding environment Y2L2, Y1L3, Y1L2 and Y2L3 with a winner genotypes G8 and G2. The third environment include Y2L1 with a vertex genotype G11 and G1 while the fourth environment include only Y1L1 with the winner genotype G7. From the figure 4, G10 was on the vertex with no environment. However, genotypes within the polygon, particularly those located near the biplot origin were less responsive than the genotypes on the vertices and the ideal genotype would be one closest to the origin.

4.2 Relationship among environments

In the present study, the relationships among the test environments are presented in Figure 5. The lowest angle was observed between Y2L5 and Y2L5, Y1L3 and Y2L3 followed by those between Y1L2 and Y2L1indicating the existence of high correlation between them. This shows that genotypes performing best at Y2L5 can repeat the same performance at Y2L5, and vice versa. The angles between Y2L1and Y2L5 on the other hand, were closer to 90° showing that they have no correlation or indicating that each environment has independent genotypic performance. Furthermore, the angle between Y2L1 and Y2L6 is greater than 90^0 showing that they have negative correlations. Hence, genotypes performing best at Y2L1 does not repeat the same performance at Y2L6, and vice versa. Based on the representativeness and discriminating ability of the study environment, Yan et al. (2007), classified the environment into three major types. Type 1 environments are characterized by short vectors, providing little or no information genotypes; thus, they are inappropriate as test environments. Type 2 environments are characterized by long vectors, forming smaller angles with the AEC abscissa, and are useful for selecting superior genotypes. Type 3 environments with long vectors form large angles with the AEC abscissa; therefore, they are inappropriate for the selection of superior genotypes. Despite their limitation, Type 3 environments can be useful in culling unstable genotype. The eleven test environments in the present study were categorized into three based on the relationship among environments. Thus, category-I contained of Y2L5, Y2L7, Y1L1, Y2L6, Y2L3, Y2L4 and Y2L2; category-II comprised Y2L3, Y1L1 and Y2L1, whereas the category-III consisted of Y1L1 only. The categorizing was based on relationship among environments is in line with the environmental categorizing of the polygon view.

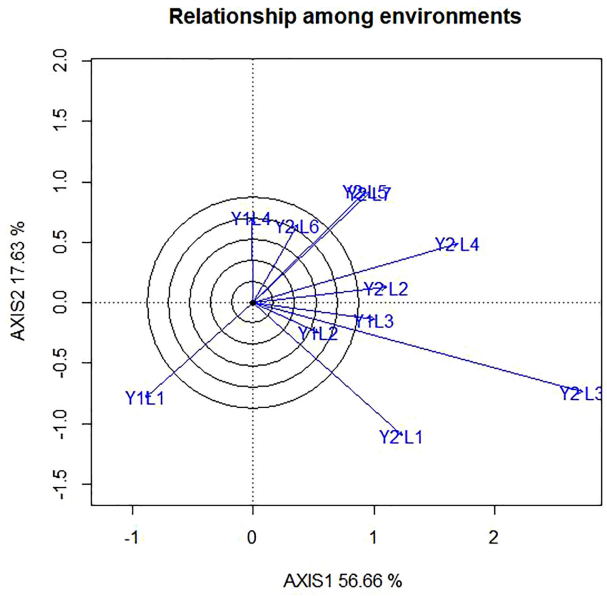


FIGURE 5
The vector view of GGE biplot showing the relationship among environment (tested environment with the ideal environment) of 11 soybean genotypes two years (11 environments) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

4.3 Discriminativeness vs representativeness

It is known that the GGE biplot is also useful to assess how much a test environment is capable of generating unique information about the differences among genotypes and how representative the mega-environment is. In this study, based on the length of vector the different environments (Figure 6), Y2L3, Y2L1, Y1L1, Y2L4, Y2L5 and Y2L7 had the longest vector length and, therefore, are the most discriminating environments in the present study. Similarly, Y1L2, Y2L2 and Y1L3 had the shortest vector length which indicates that they are the least discriminating of all the test environments.

4.4 Genotype ranking based on their mean performance and stability

The ranking of the genotypes based on their mean performance and stability presented in Figure 7. It has been established that if the PC1 of a GGE biplot approximates the genotype main effects, PC2 must approximate the GE effects associated with each genotype, which is a measure of instability (Yan et al., 2000; Yan, 2001). The line passing through the biplot origin and the average environment indicated by a circle is called the average environment coordinate (AEC) axis, which is defined by the average PC1 and PC2 scores of all the environments. By using the average principal components in all environments, the AEC method was employed to evaluate the yield stability of genotypes. In our study, the "mean vs. stability" pattern of GGE biplot revealed 74.29% for yield per hectare of G + G E variation (Figure 7). A line drawn through the average environment and the biplot origin, having one direction pointed to a greater genotype main effect. Moving in either direction away from AEC ordinate and from the biplot origin indicates the greater GEI effect and reduced stability. The genotypes with below-average means and those with above-average means are divided by the AEC ordinate. As a result, genotypes G8, G11, G1, G2, and G9 showed yield performances that were higher than the average yield in this study (Figure 7). The genotypes G7, G10, G5, G4 and G3 had lesser yielding performance in comparison to the mean, while genotypes G8, G2 and G4 are more adapted and stable. This is because the genotype on the right of the ordinate line had yield less than the average mean yield. The least stable genotypes are G9 and G10, on the other hand. Stability can be identified based on concentric circles and also ideal genotypes are on the centre of concentric circles i.e., high mean and stable. Beside this, good genotypes are close to ideal genotypes.

4.5 Evaluation of genotypes relative to ideal genotypes

The ideal genotype is the one that with the highest mean performance and absolutely stable (Yan and Kang, 2002). This is assumed to be in the centre of the concentric circles is an ideal

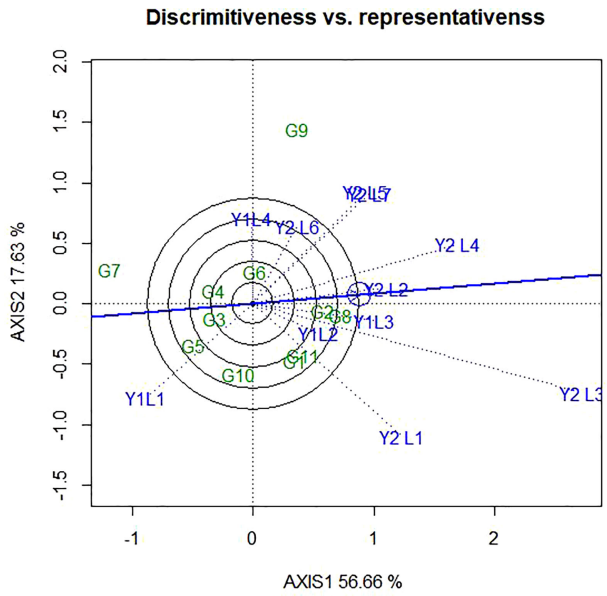


FIGURE 6
The GGE biplot 'Discriminativeness vs. Representativeness' pattern for genotype comparison with ideal genotype showing G + G × E interaction effect of 11 soybean genotypes under two years (11 environments) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

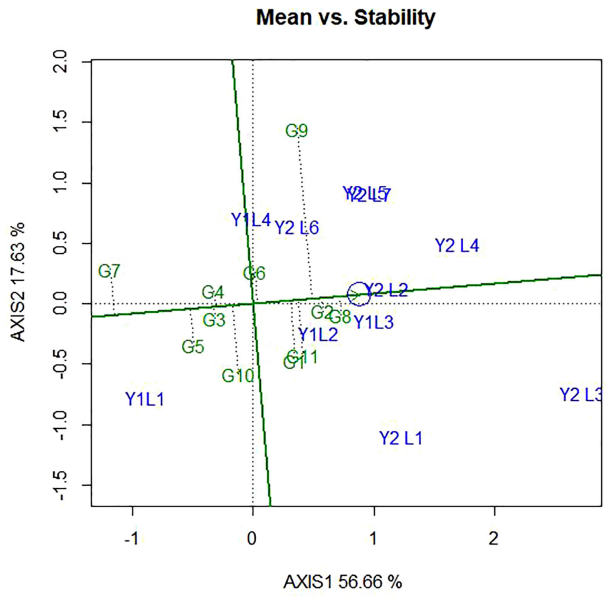


FIGURE 7
The "Mean vs. stability" pattern of GGE biplot illustrating interaction effect of 11 soybean genotypes under two years (11environemnets) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

genotype across the tested environment. It is more desirable for a genotype to be located closer to the ideal genotype. Hence, the GGE biplots (Figure 8) shows that G8 and G2 were ideal in terms of higher-yielding ability and stability as compared to the other genotypes. While genotypes G7, G5, G9, G10 and G5, were unfavourable since there are too far from the ideal genotypes. Some of the environments located close to the ideal genotypes (Y1L2, Y1L3, and Y2L2) and most (Y2L3, Y1L1, Y2L1, and Y2L4) are located far from the ideal variety.

4.6 Evaluation of environments relative to ideal environments

Ideal environments had the longest vector with small IPCA, which fell into the centre of concentric circles. Hence, from this study Y2L4 and Y2L3 are ideal environment. Ideal environment is the most representative of the overall environments and the most powerful to discriminate genotypes. Likewise, Y2L2 and Y1L3 were closer to the ideal environment and considered as second powerful to discriminate genotypes. On the other hand, environments Y1L1 and Y1L4 were found far from the ideal environment and considered as less powerful to discriminate genotypes (Figure 9).

5 Discussions

Ethiopia still imports processed soybean products such as soya edible oil and other oil crops like Sesame and Niger seed. Soybeans generated $376 million (USDA, 2021) and around 1.4 million farmers produced oilseed crops in the country (CSA, 2020). The mean square of environments, genotypes, and genotypes by environments interaction showed very highly significant differences ($p \leq 0.001$) for yield per hectare. This is an indication that the response of soybean genotypes is highly dependent on the site where the genotypes are grown and annual variation of temperature and rainfall. The results also indicate that environmental factors have a significant impact on the performance of different soybean genotypes, indicating that different soybean genotypes should be tested over time and in different locations. The highest grain yield in tone was recorded at Y1L4 (3.2), followed by Y2L3 (3.0) and lastly, Y2L4 (2.7). The influence of genotype, environment, and GEI in the soybean for different traits has been reported by (Gurmu et al., 2009; Agoyi et al., 2017; Mwiinga et al., 2020).

GGE biplot can be graphically detect the genotype by environment interaction pattern, identify winning genotype, and delineating mega environments among the testing environments (Yan et al., 2007). So, this potential tool has been employed to analyze the multienvironment data of grain soybean trials (Mwiinga et al., 2020). In our study, environment contributed 49.42% of the variation in the data, while the contribution of genotype and their interaction with environment was around 33.91%. Gauch and Zobel (1997) reported that normally in MET data, the environment accounts for about 80% of the total variation. Zamalotshwa et al. (2020) found that 89.65 and 89.06% of the difference in Glu

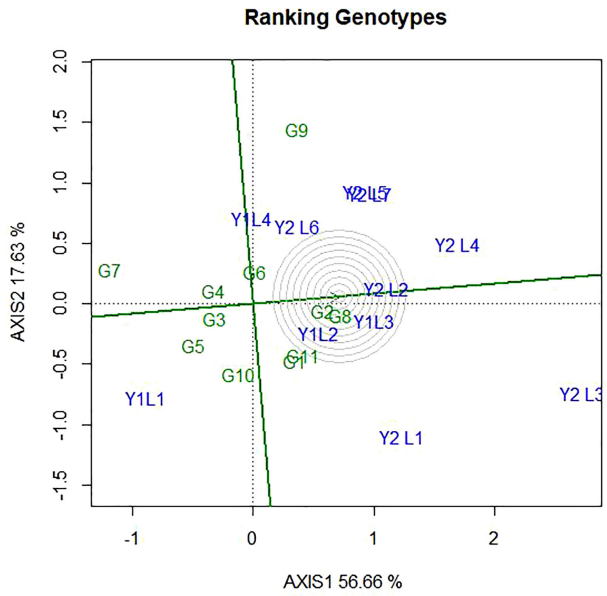


FIGURE 8
The GGE biplot 'Genotype ranking' pattern for environment comparison with ideal environment showing G + G × E interaction effect of 11 soybean genotypes under two years (11environemnets) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

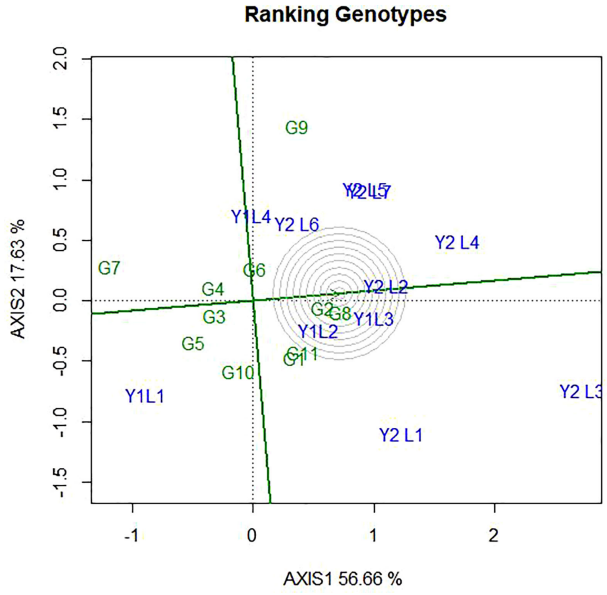


FIGURE 9
The GGE biplot 'Environment ranking' pattern for environment comparison with ideal environment showing G + G × E interaction effect of 11soybean genotypes under two years (11environemnets) for grain yield per hectare. The biplots were created based on Centering = 0, SVP = 2, Scaling = 0.

and Glut was explained by the environment in soybean-MET data. Agoyi et al. (2017) reported a similar trend for soybean dry weight of nodules at 83.79% and the percentage of effective nodules was 12.98%. However, Agoyi et al. (2017) reported a slightly lower (12.98%) contribution of environment to the total variation in soybean yield number of nodules from Uganda. In a two-year study of eleven diverse lines across seven environments, Mwiinga et al. (2020) attributed 21.04% of overall variation to environment, 31.59% to genotype main effect, and 47.36% to GEI in a two-year study of eleven diverse lines. In our study, GEI explained a higher proportion of the variation than genotype alone. The higher proportion of GEI as compared to genotype is indicative of the possible existence of different mega-environments in testing environments (Yan, 2002; Yan and Kang, 2002). This could be true not only in Ethiopia, but also in other soya-growing regions. But in the presence of the GEI effect could complicate the selection process of superior genotypes and also may reduce the selection efficiency in breeding programs (Gauch, 2006; Kumar et al., 2014). As a result, when breeding soybeans in their individual environments, soybean breeders must consider this fact.

5.1 Mean performance and comparison of genotypes

Plant height, number of seeds per plant, number of pods per plant, and weight of 100 seeds are important agriculture for soybean producers to develop excellent varieties of soybean. It will be the standard. A genotype that is adapted and stable to yield in a variety of environments, on the other hand, is highly valued by researchers in a breeding program to limit the risk of yield loss due to climatic circumstances. The results also demonstrated that the control Nyala (2.35 t ha^{-1}) gave the sixth and unreleased genotypes, which come from the modified single descent method from the soybean breeding, namely JM-CLK/CRFD-15-SD (2.84 t ha^{-1}), and the remaining three genotypes through direct introduction from various sources, namely 5002T (2.73 t ha^{-1}) and PI471904 (2.64 t ha^{-1}) and SCS-1 (2.57 t ha^{-1} were among the top performers with respect to yield and stability. The genotype Nyala, which was released in 2014, was among the top five ranked genotypes. When genotype performance varies in a variety of environments, Gurmu et al. (2009) and Asrat et al. (2009), recommends studying GEI followed by stability analysis. Similarly, studying on seven soybean promising lines and two check varieties, on eight locations of soybean production, TGx-1835-10F and TGx-1876-4E exhibit higher stability than other genotypes (Asrat et al. 2009). In evaluating the performance of five soybean genotypes, Al-Assily et al. (2002) found that three cultivars consistently produced mean yields that were higher than the trial mean. In a similar regard, Cucolotto et al. (2007) discovered that out of 30 cultivars, four had good adaptation and stability. In contrast to our findings, Asrat et al. (2009) did not report any genotype with widespread adaption.

5.2 The AMMI model analysis

The application of AMMI model for partitioning of GEI showed that the first four (IPCA1-4) multiplicative terms of AMMI were significant (Table 6) and this implies that the interaction of 11 genotypes of soybean with eleven environments was predicted by the first four principal components of genotypes and environments, which is in agreement with the recommendation of Behrouz et al. (2018) and Goa et al. (2022). This is also following the results of Mwiinga et al. (2020) and Mushoriwa et al. (2022), whereas much as the first five and six IPCAs were significant respectively. However, this contradicted with the findings of Zobel et al. (1988) and Abiriga et al. (2020), which recommended that the most accurate model for AMMI, can be predicted using the first two IPCAs.

Environments accounted for the largest share (49.42%) of the overall treatment sum of squares in this study, followed by GEI (33.91%) and genotypes (16.18%), which supports the necessity for multi-environmental trials conducted throughout several seasons in this nation. The environment was revealed to be the main source of variance in additional studies by Gurmu et al. (2009); Rakshit et al. (2012); Mushoriwa et al. (2022), and Vaezi et al. (2018). In contrast, studies by Mulugeta et al. (2013), Mwiinga et al. (2020), and Bhartiya et al. (2017) likewise found that the GEI contributed more to overall variation, with values of 60%, 47.36%, and 41%, respectively.

AMMI 1 identified the genotypes as the most stable and high producing genotypes based on how well they performed in various situations (G8, G2, and G9). The highly interacting environment (Y2L3) and genotypes (G9 and G7) were found in the AMMI 2 biplot. Biplot has also been used to illustrate how precisely genotypes adapt to the right environment. The genotype G9 benefited more from the use of Y2L7.The genotype G3 interacted favorably with Y1L1, while the genotype G7 expressed the high yielding potential in Y1L4. Similar positive interactions were seen between the genotypes G5, G6, G1, G6, G11, G2, G6, and G4 and the locations Y1L2, Y1L3, Y2L1, Y2L2, Y2L3, Y2L4, and Y2L5.

According to similar approaches applied, Mwiinga et al. (2020), the genotypes and the environments average was 2.4 t ha^{-1}.The precise adaptation of genotypes to the appropriate environment was also visualized using biplots (Samyuktha et al., 2020).

The AMMI stability values (ASV), IPCA and GSI scores were used to classify the genotypes according to stability. With consideration, according to ASV, the genotypes close to 0 (G6, G1, G10 and G5) were highly adapted and stable, whereas the genotypes having high value (G8, G7, G2, and G4) were found to

TABLE 6 Grain yield AMMI investigation of variance across seven locations in eleven environments.

Sources of Variations	df	SS	MS	Total Variation (%)	GEI Explained (%)	GEI Cumulative (%)
Block, Env.	33	15.54	0.471			
Genotypes, G	10	36.61	3.661***	16.18		
Environments, E	10	126.65	12.665***	49.42	49.42	
Interactions, GEI	100	83.56	0.836***	33.91		
IPCA 1	19	37.29	1.963***		47.74	47.74
IPCA 2	17	20.67	1.216***		26.62	74.36
IPCA 3	15	13.78	0.919***		17.88	92.23
IPCA 4	13	5.93	0.456*		7.77	100.00
Error	330	82.15	0.249			
Total	483	344.51	0.713			

GEI, Genotype by Environment interaction; df, degree of freedom; SS, Sum Square; MS, Mean Square; IPCA, Interaction principal component axis.
***Significant at $P < 0.001$, *Significant at $P < 0.05$.

be highly unstable. From the ASV, both the top and the lowest yielder *viz*., G8 and G7 categorized under unstable genotypes. Based on GSI, the following genotypes *viz*., G9 (PI471904), G1 (5002T), and G11(Nyala) were identified as stable and high yielding across all the test environments under study. In the study, Mushoriwa et al. (2022), identified both the IPCA scores and GSI identified genotype G25 was the most superior followed by genotype G1 while genotype G16 performed the least among the 42 studied genotypes. At the same time these authors further explained that IPCA and GSI stability parameters could be used for simultaneously selection for high yield and stability. Therefore, the two stability parameters could be used for simultaneously selection for high yield and stability. From the results of Mwiinga et al. (2020) shows that, the same results and reason were obtained in studying on twenty elite soybean lines and five commercial checks across six environments across four Southern Africa countries, in six environments. In addition, Mushoriwa et al. (2022), considering the two analytical methods (AMMI and GSI), two genotypes (Sovreign and SC Siesta) were among the best and thus could be recommended for cultivation across the three countries and utilization as breeding stocks in programs that aim to improve both stability and productivity of soybean.

5.3 GGE biplot analysis

In the GGE bi-plot analysis, the variation 74.29% for the first two principal components was higher than the ideal limit (66%) (Yan et al., 2000). Similarly, this result is lower than that obtained by Arega et al. (2020) (93.59%), Amira et al. (2013) (86.6%); Carvalho et al. (2021) (80%); Mulugeta et al. (2013) (63.4%), Krisnawati and Adie (2018a) (60.87%) and Bhartiya et al. (2017) (74.40%) in contrary to that found by Kocaturk et al. (2019) (51.90%).

GGE biplot ('Which-won-where' pattern) - According to Yan and Tinker (2006), the vertex genotypes were the most responsive genotypes, as they have the longest distance from the origin in their direction. In this study since there are five mega environments, but from the Yan et al. (2007), infer that the repeatable environment grouping is necessary, but not sufficient, for declaring different mega-environments. The necessary and sufficient condition for mega-environment division is a repeatable which-won-where pattern rather than merely a repeatable environment-grouping pattern (Yan and Rajcan, 2002; Yan and Kang, 2002). This indicates that genotypes in vertex without environment performed poorly in all the sites (Alake and Ariyo, 2012). Similarly, Mulugeta et al. (2013) found that from the polygon view of this biplot, test environments and genotypes fell into three and four sectors, respectively, with three sectors having no test environment.

5.4 Relationship among environments

GGE biplot analysis helped us to understand the utility of different sites in terms of relative discrimination between genotypes and the relationship between them for different traits. The grouping based on relationship among environments is in line with the environmental grouping of the polygon view. According to Yang et al. (2009), an ideal environment should have a high PC1 score and zero scores for PC2. Similar results to our findings were stated by Zhang et al. (2006) and Bhartiya et al. (2017). In the study by Mulugeta *et al.* (2013), identified three environments namely "Dibate_2008", "Bullen_2007" and "Pawe_2007" were the most discriminating of the genotypes whereas "Dibate_2007" provided very little information about genotypic differences in the north western regions of Ethiopia.

5.5 Discriminativeness vs representativeness

According to Yan et al. (2007), described the discriminating power vs. representativeness view of the GGE biplot is an effective tool for test-environment evaluation, which can lead to the identification of a minimum set of discriminating and representative test environments. Test-environment evaluation has not been a research topic in AMMI analysis. The Y1L2, Y2L2 and Y1L3 had the shortest vector length which indicates that they are the least discriminating of all the test environments. Hashim et al. (2021), reported one environment is suitable for genotype selection considering yield per hectare among the tested four environments.

5.6 Genotype ranking based on their mean performance and stability

The present study showed that there were a number of genotypes that performed better than the control Nyala (G11) the most farmer-preferred and high yielding varieties in Ethiopia. This study also showed that JM-CLK/CRFD-15-SD (G8) derived through across modified single seed decent (MSSD) plant selection Clarck -63K and Crowford had the highest yield level compared to all the other genotypes. Stability can be identified based on concentric circles and also ideal genotypes are on the centre of concentric circles.

5.7 Evaluation of genotypes relative to ideal genotypes

According to Yan et al. (2007), genotype evaluation is meaningful only for a specific mega-environment, and an ideal genotype should have both high mean performance and high stability within a mega-environment. The ideal genotype is the one that with the highest mean performance and absolutely stable (Yan and Kang, 2002). According to Yan and Tinker (2006), the most responsive genotypes were located at the vertexes of the polygon, since they have the longest distance from the origin in their direction. Oladosu et al. (2017) finds out similar results observation.

5.8 Evaluation of environments relative to ideal environments

According to the biplot graph, the biplot accounted for 56.66% (PC1) and 17.63% (PC2) of the variability in grain yield of G+GEI across the studied environment. Mushoriwa et al. (2022) found that the majority of the settings had the fewest interaction effects, indicating that they were perfect for evaluation and selection since the genotypes' performance could be consistent in them. In this study, eleven test environments in were categorized into three based on the relationship among environments and thus, category-I contained of Y2L5, Y2L7, Y1L1, Y2L6, Y2L3, Y2L4 and Y2L2; category -II comprised Y2L3, Y1L1 and Y2L1, whereas the category -III consisted of Y1L1 only. In a similar study by Khan et al. (2021), the tested environments were categorized in to three groups for different agronomic traits. In a similar study, Arega et al. (2020) identified the test location (Billo and Gute) as the most discriminating environment, while Bako and Uke were identified as the least discriminating testing environments. Chewaka was identified as the least discriminating environment among the testing environments. "Bullen_2007", "Dibate_2007", and "Manbuk_2008" were found to be more representative environments for soybean regional trials by Mulugeta et al. (2013).

6 Conclusion

This research evaluated the stability and adaptability of soybean genotypes based on combined ANOVA and, AMMI, GGE biplot analysis. Analysis of variance for every location and combined over locations showed significant differences among genotypes, environments and GEI for grain yield, as well as most of the yield-related traits of soybean. According to AMMI and GGE analyses to generate adapted, stable and high-yielding soybean genotypes across several environments would be beneficial. Five different mega-environments and seven different sectors are indicated by the biplot's environment pattern. The vertex genotypes used in this study were G2, G11, G7, G8, G10, G1, and G9. In the first environment's Y2L5, Y2L7, Y1L1, Y2L6, and Y2L4 circumstances, the genotype G9 performed best. In the second environment's Y2L2, Y1L3, Y1L2, and Y2L3 conditions, the winning genotypes G8 and G2 had greater yields. The third environment also comprises Y2L1 with vertex genotypes G11 and G1, but the fourth environment only has Y1L1 with the winning genotype G7. Unusually, the genotype G10 vertex lacked an environment.

The yields of genotypes G8 and G1 ranged from 2.85 to 2.75 t ha^{-1} and were significantly greater than those of other genotypes and the control variety. The results from the four distinct stability statistics AMMI biplot (G8, G2, G1, G11, ASV: G1, G11; GSI; G9, G1, G11 and GGE: G2, G8, G9 are taken into account together with the genotypes' grand mean. The genotypes JM-CLK/CRFD-15-SD (G8) and 5002T (G1), which rank among the best and have the highest seed output, are suitable for hybridization as a parent and commercial production. The highly high grain output may be due to genotype variations in relative performance across environments and their changing

sensitivity to a variety of abiotic variables seen in a typical Ethiopian soybean growing environment.

Data availability statement

The raw data supporting the conclusions of this article willbe made available by the authors, without undue reservation.

Author contributions

MH took part in the conception and design of the study, data collection, paper writing, statistical analysis, and interpretation of the data, as well as article revision for important intellectual components. The final text has been reviewed and approved by the author.

Funding

Ethiopian Institute of Agricultural Research (EIAR), Ethiopia.

Acknowledgments

The authors appreciate the financing provided by the Soybean Innovation Laboratory (SIL), as well as the experimental lines and other resources provided by EIAR. In addition, the Ethiopian Institute of Agricultural Research [EIAR-Soybean (1-2-15)] generously funded this research.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.950992/full#supplementary-material

References

Abebe, A. T., Kolawole, A. O., Unachukwu, N., Chigeza, G., Tefera, H., and Gedil, M. (2021). Assessment of diversity in tropical soybean (Glycine max (L.) merr.) varieties and elite breeding lines using single nucleotide polymorphism markers. *Plant Genet. Resour.* 19 (1), 20–28. doi: 10.1017/S1479262121000034

Abiriga, F., Ongom, P. O., Rubaihayo, P. R., Edema, R., Gibson, P. T., Dramadri, I., et al (2020). Harnessing genotype-by-environment interaction to determine adaptability of advanced cowpea lines to multiple environments in Uganda *J. Plant Breed. Crop Sci.* 12(2), 131–145. doi: 10.5897/JPBCS2020.0891

Adane, A., Alemayehu, D., Meseret, T., and Chala, D. (2018). Genotype and genotype by environment interaction and grain yield stability of medium maturity groups of soybeans [*Glycine max* (L.) Merrill] varieties in Western oromia, Ethiopia. *Afr. J. Plant Sci.* 12 (9), 227–237. doi: 10.5897/AJPS2018.1674

Agahi, K., Ahmadi, J., Oghan, H. A., Fotokian, M. H., and Orang, S. F. (2020). Analysis of genotype× environment interaction for seed yield in spring oilseed rape using the AMMI model. *Crop Breed. Appl. Biotechnol.* 20, 7–15. doi: 10.1590/1984-70332020v20n1a2

Agoyi, E. E., Odong, T. L., Tumuhairwe, J. B., Chigeza, G., Diers, B. W., and Tukamuhabwa, P. (2017). Genotype by environment effects on promiscuous nodulation in soybean (*Glycine max* L. Merrill). *Agric. Food Secur.* 6 (1), 1–14. doi: 10.1186/s40066-017-0107-7

Alake, C. O., and Ariyo, O. J. (2012). Comparative analysis of genotype x environment interaction techniques in West African okra, (Abelmoschus caillei, a. chev stevels). *J. Agric. Sci.* 4 (4), 135. doi: 10.5539/jas.v4n4p135

Al-Assily, K. A., Saleeb, S. R., Mansour, S. H., and Mohamed, M. S. (2002). Stability parameters for soybean genotypes as criteria for response to environmental conditions. *Minufia J. Agric. Res.* 27 (2).

Amira, J. O., Ojo, D. K., Ariyo, O. J., Oduwaye, O. A., and Ayo-Vaughan, M. A. (2013). Relative discriminating powers of GGE and AMMI models in the selection of tropical soybean genotypes. *Afr. Crop Sci. J.* 21 (1), 67–73.

Angela, P., Mateo, V., Gregorio, A., Francisco, R., Jose, C., and Juan, B. (2015). "GEA-r (Genotype × environment analysis with r for windows) version 4.1," in *CIMMYT research data and software repository network* (Mexico: ORACLE AMERICA, INC).

Arega, A., Dabessa, A., and Bekele, S. (2020). Yield stability analysis of soybean varieties (early set) in Western oromia, Ethiopia. *Ethiopian J. Crop Sci.* 8 (2), 67–77.

Asrat, A., Fistum, A., Fekadu, G., and Mulugeta, A. (2009). AMMI and SREG GGE biplot analysis for matching varieties onto soybean production environments in Ethiopia. *Sci. Res. Essays* 4 (11), 1322–1330.

Bhartiya, A., Aditya, J. P., Kumari, V., Kishore, N., Purwar, J. P., Agrawal, A., et al. (2017). GGE biplot & ammi analysis of yield stability in multi-environment trial of soybean [Glycine max (L.) Merrill] genotypes under rainfed condition of north western Himalayan hills. *J. Anim. Plant Sci.* 27 (1), 227–238.

Bocianowski, J., Warzecha, T., Nowosad, K., and Bathelt, R. (2019). Genotype by environment interaction using AMMI model and estimation of additive and epistasis gene effects for 1000-kernel weight in spring barley (Hordeum vulgare l.). *J. Appl. Genet.* 60 (2), 127–135. doi: 10.1007/s13353-019-00490-2

Carvalho, M. P., Nunes, J. A. R., Carmo, E. L. D., Simon, G. A., and Moraes, R. N. O. (2021). Adaptability and stability of conventional soybean by GGE biplot analysis. *Pesquisa Agropecuária Trop.* 51, 1–10. doi: 10.1590/1983-40632021v5167995

CSA (2020). *Agricultural sample survey report on area, production and yields of major crops* (Addis Ababa, Ethiopia: Central statistical Authority).

Cucolotto, M., Pípolo, V. C., Garbuglio, D. D., Junior, N. D. S. F., Destro, D., and Kamikoga, M. K. (2007). Genotype x environment interaction in soybean: evaluation through three methodologies. *Crop Breed. Appl. Biotechnol.* 7 (3), 270–277. doi: 10.12702/1984-7033.v07n03a06

Ebdon, J. S., and Gauch, J. H.G. (2002). Additive main effect and multiplicative interaction analysis of national turfgrass performance trials: I. interpretation of

genotype× environment interaction. *Crop Sci.* 42 (2), 489–496. doi: 10.2135/cropsci2002.4890

Eberhart, S. T., and Russell, W. A. (1966). Stability parameters for comparing varieties 1. *Crop Sci.* 6 (1), 36–40. doi: 10.2135/cropsci1966.0011183X000600010011x

Eltaher, S., Baenziger, P. S., Belamkar, V., Emara, H. A., Nower, A. A., Salem, K. F., et al. (2021). GWAS revealed effect of genotype× environment interactions for grain yield of Nebraska winter wheat. *BMC Genomics* 22 (1), 1–14. doi: 10.1186/s12864-020-07308-0

Farshadfar, E., Mahamodi, N., and Yaghotipoor, A. (2012). AMMI stability value and simultaneous estimation of yield and yield stability in bread wheat (Triticum aestivum l.). *Aust. J. Crop Sci.* 5, 1837–1844.

Gauch, J. H.G. (2006). Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46 (4), 1488–1500. doi: 10.2135/cropsci2005.07-0193

Gauch, H. G., and Zobel, R. W. (1996). "AMMI analysis of yield trials," in *Genotype-by-Environment interaction*. Eds. M. S. Kang and H. G. Gauch (Boca Raton: CRC Press), 85–122. doi: 10.1201/9781420049374.ch4

Gauch, J. H.G., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37 (2), 311–326. doi: 10.2135/cropsci1997.0011183X003700020002x

Goa, Y., Mohammed, H., Worku, W., and Urage, E. (2022). Genotype by environment interaction and yield stability of cowpea (Vigna unguiculata (L.) walp.) genotypes in moisture limited areas of southern Ethiopia. *Heliyon* 8 (3), e09013. doi: 10.1016/j.heliyon. 2022.e09013

Gurmu, F., Mohammed, H., and Alemaw, G. (2009). Genotype x environment interactions and stability of soybean for grain yield and nutrition quality. *Afr. Crop Sci. J.* 17 (2), 87–99.

Happ, M. M., Graef, G. L., Wang, H., Howard, R., Posadas, L., and Hyten, D. L. (2021). Comparing a mixed model approach to traditional stability estimators for mapping genotype by environment interactions and yield stability in soybean [Glycine max (L.) merr.]. *Front. Plant Sci.* 12, 630175. doi: 10.3389/fpls.2021.630175

Hashim, N., Rafii, M. Y., Oladosu, Y., Ismail, M. R., Ramli, A., Arolu, F., et al. (2021). Integrating multivariate and univariate statistical models to investigate genotype–environment interaction of advanced fragrant rice genotypes under rainfed condition. *Sustainability* 13 (8), 4555. doi: 10.3390/su13084555

Hongyu, K., García-Peña, M., de Araújo, L. B., and dos Santos Dias, C. T. (2014). Statistical analysis of yield trials by AMMI analysis of genotype× environment interaction. *Biometr. Lett.* 51 (2), 89–102. doi: 10.2478/bile-2014-0007

IBPGR (1984). "Descriptors for soybean" in *International board for plant genetic resources* (Rome, Italy: IBPGR Secretariat), 19–38.

Kandus, M., Almorza, D., Boggio Ronceros, R., and Salerno, J. C. (2010). Statistical models for evaluating the genotype-environment interaction in maize (*Zea mays* L.). *Phyton-Revista Internacional Botanica Exp.* 79, 39–46.

Kaya, M. D., Okçu, G., Atak, M., Cıkılı, Y., and Kolsarıcı, Ö. (2006). Seed treatments to overcome salt and drought stress during germination in sunflower (Helianthus annuus l.). *Eur. J. Agron.* 24 (4), 291–295. doi: 10.1016/j.eja.2005.08.001

Khan, M. M. H., Rafii, M. Y., Ramlee, S. I., Jusoh, M., and Al Mamun, M. (2021). AMMI and GGE biplot analysis for yield performance and stability assessment of selected bambara groundnut (Vigna subterranea l. verdc.) genotypes under the multi-environmental trials (METs). *Sci. Rep.* 11 (1), 1–17. doi: 10.1038/s41598-021-01411-2

Kocaturk, M., CUBUKCU, P., Goksoy, A. T., Sincik, M., Ilker, E., Kadiroglu, A., et al. (2019). GGE biplot analysis of genotype× environment interaction in soybean grown as a second crop. *Turkish J. Field Crops* 24 (2), 145–154. doi: 10.17557/tjfc.615175

Krisnawati, A., and Adie, M. M. (2015). Selection of soybean genotypes by seed size and its prospects for industrial raw material in Indonesia. *Proc. Food Sci.* 3, 355–363. doi: 10.1016/j.profoo.2015.01.039

Krisnawati, A., and Adie, M. M. (2018a). Genotype by environment interaction and yield stability of soybean genotypes. *Indonesian J. Agric. Sci.* 19 (1), 25–32. doi: 10.21082/ijas.v19n1.2018.p25-32

Krisnawati, A., and Adie, M. M. (2018b). "Yield stability of soybean promising lines across environments," in *IOP Conference Series: Earth and Environmental Science* (IOP Publishing). 102 (1), 012044. doi: 10.1088/1755-1315/102/1/012044

Kumar, A., Kumar, S., Kapoor, C., Bhagawati, R., Pandey, A., and Pattanayak, A. (2014). GGE biplot analysis of genotype× environment interaction in soybean grown in NEH regions of India. *Environ. Ecol.* 32 (3A), 1047–1050. doi: 10.17557/tjfc.615175

Li, W., Yan, Z. H., Wei, Y. M., Lan, X. J., and Zheng, Y. L. (2006). Evaluation of genotype× environment interactions in Chinese spring wheat by the AMMI model, correlation and path analysis. *J. Agron. Crop Sci.* 192 (3), 221–227. doi: 10.1111/j.1439-037X.2006.00200.x

Luo, J., Pan, Y. B., Que, Y., Zhang, H., Grisham, M. P., and Xu, L. (2015). Biplot evaluation of test environments and identification of mega-environment for sugarcane cultivars in China. *Sci. Rep.* 5, 15505. doi: 10.1038/srep15505

Mederos-Ramirez, A., Ortiz-Perez, R., and Postal, G. (2021). Analysis of the genotype environment interaction in the of soybean crop (Glycine max (L) Merrill). *Cultivos Tropicales* 42 (1), 1–10. doi: 10.1016/j.fcr.2020.107922

Mitrović, B., Stanisavljević, D., Treskić, S., Stojaković, M., Ivanović, M., Bekavac, G., et al. (2012). Evaluation of experimental maize hybrids tested in multi-location trials using AMMI and GGE biplot analyses. *Turkish J. Field Crops* 17 (1), 35–40.

Mulugeta, A., Sisay, K., Seltene, A., and Zelalem, F. (2013). GGE biplots to analyze soybean multi-environment yield trial data in north Western Ethiopia. *J. Plant Breed. Crop Sci.* 5 (12), 245–254. doi: 10.5897/JPBCS13.0403

Mushoriwa, H., Mathew, I., Gwata, E. T., Tongoona, P., and Derera, J. (2022). Grain yield potential and stability of soybean genotypes of different ages across diverse environments in southern Africa. *Agronomy* 12 (5), 1147. doi: 10.3390/agronomy12051147

Mwiinga, B., Sibiya, J., Kondwakwenda, A., Musvosvi, C., and Chigeza, G. (2020). Genotype x environment interaction analysis of soybean (*Glycine max* (L.) Merrill) grain yield across production environments in southern Africa. *Field Crops Res.* 256, 107922. doi: 10.1016/j.fcr.2020.107922

Oladosu, Y., Rafii, M. Y., Abdullah, N., Magaji, U., Miah, G., Hussin, G., et al. (2017). Genotype× environment interaction and stability analyses of yield and yield components of established and mutant rice genotypes tested in multiple locations in Malaysia. *Acta Agricult. Scandinavica Sect. B—Soil Plant Sci.* 67 (7), 590–606. 10.1080/09064710.2017.1321138

Olanrewaju, O. S., Oyatomi, O., Babalola, O. O., and Abberton, M. (2021). GGE biplot analysis of genotype× environment interaction and yield stability in bambara groundnut. *Agronomy* 11 (9), 1839. doi: 10.3390/agronomy11091839

Purchase, J. L., Hatting, H., and Van Deventer, C. S. (2000). Genotype× environment interaction of winter wheat (Triticum aestivum l.) in south Africa: II. stability analysis of yield performance. *South Afr. J. Plant Soil* 17 (3), 101–107. doi: 10.1080/02571862.2000.10634877

Qiu, L. J., and Chang, R. Z. (2010). "The origin and history of soybean," in *The soybean: botany, production and uses* (Wallingford UK: CABI), 1–23.

Rakshit, S., Ganapathy, K. N., Gomashe, S. S., Rathore, A., Ghorade, R. B., Kumar, M. V., et al. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* 185 (3), 465–479. doi: 10.1007/s10681-012-0648-6

Samyuktha, S. M., Malarvizhi, D., Karthikeyan, A., Dhasarathan, M., Hemavathy, A. T., Vanniarajan, C., and Senthil, N. (2020) Delineation of genotype× environment interaction for identification of stable genotypes to grain yield in mungbean. *Front. Agron.* 2, 577911. doi: 10.3389/fagro.2020.577911

Sritongtae, C., Monkham, T., Sanitchon, J., Lodthong, S., Srisawangwong, S., and Chankaew, S. (2021). Identification of superior soybean cultivars through the indication of specific adaptabilities within duo-environments for year-round soybean production in northeast Thailand. *Agronomy* 11 (3), 585. doi: 10.3390/agronomy11030585

Thungo, Z., Shimelis, H., Odindo, A., and Mashilo, J. (2020). Genotype-by-environment effects on grain quality among heat and drought tolerant bread wheat (*Triticum aestivum* l.) genotypes. *J. Plant Interact.* 15 (1), 83–92. doi: 10.1080/17429145.2020.1748732

USDA (2021). *Ethiopia Oilseeds Report Annual. Report Number: ET2021 0006.2021*. Available at: https://www.google.com/search?client=firefoxbd&q=Ethiopia+still+imports+processed+soybean+products+such+as+soya+edible+oil+and+other+oil+crops+like+Sesame+and+Niger+see.

Vaezi, B., Pour-Aboughadareh, A., Mehraban, A., Hossein-Pour, T., Mohammadi, R., Armion, M., et al. (2018). The use of parametric and non-parametric measures for selecting stable and adapted barley lines. *Arch. Agron. Soil Sci.* 64 (5), 597–611. doi: 10.1080/03650340.2017.1369529

Yan, W. (2001). GGEbiplot–a windows application for graphical analysis of multi environment trial data and other types of two-way data. *Agron. J.* 93 (5), 1111–1118. doi: 10.2134/agronj2001.9351111x

Yan, W. (2002). Singular-value partitioning in biplot analysis of multienvironment trial data. *Agron. J.* 94 (5), 990–996. doi: 10.2134/agronj2002.9900

Yan, W., and Hunt, L. A. (2002). Biplot analysis of diallel data. *Crop Sci.* 42 (1), 21–30. doi: 10.2135/cropsci2002.0021

Yan, W., Hunt, L. A., Sheng, Q., and Szlavnics, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40 (3), 597–605. doi: 10.2135/cropsci2000.403597x

Yan, W., and Kang, M. S. (2002). *GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists* (CRC press: Boca Raton).

Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 47 (2), 643–653. doi: 10.2135/cropsci2006.06.0374

Yan, W., and Rajcan, I. (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42 (1), 11–20. doi: 10.2135/cropsci2002.1100

Yang, W., Shen, C., Ji, Q., An, H., Wang, J., Liu, Q., et al (2009). Food storage material silver nanoparticles interfere with DNA replication fidelity and bind with DNA. *Nanotechnology* 20 (8), 085102. doi: 10.1088/0957-4484/20/8/085102

Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Can. J. Plant Sci.* 86 (3), 623–645. doi: 10.4141/P05-169

Zhang, S., Dailey, G. M., Kwan, E., Glasheen, B. M., Sroga, G. E., and Page-McCaw, A. (2006). An MMP liberates the ninjurin a ectodomain to signal a loss of cell adhesion. *Genes Dev.* 20 (14), 1899–1910. doi: 10.1101/gad.1426906

Zobel, R. W., Wright, M. J., and Gauch, J. H.G. (1988). Statistical analysis of a yield trial. *Agron. J.* 80 (3), 388–393. doi: 10.2134/agronj1988.00021962008000030002x

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 25 January 2023
DOI 10.3389/fpls.2022.1089720

Check for updates

OPEN ACCESS

EDITED BY
Petr Smýkal,
Palacký University in Olomouc,
Czechia

REVIEWED BY
Renato De Mello Prado,
São Paulo State University, Brazil
Rohit Joshi,
Institute of Himalayan Bioresource Technology (CSIR), India

*CORRESPONDENCE
Marion Prudent
marion.prudent@inrae.fr

SPECIALTY SECTION
This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 04 November 2022
ACCEPTED 20 December 2022
PUBLISHED 25 January 2023

CITATION
Jacques C, Girodet S, Leroy F, Pluchon S, Salon C and Prudent M (2023) Memory or acclimation of water stress in pea rely on root system's plasticity and plant's ionome modulation. *Front. Plant Sci.* 13:1089720. doi: 10.3389/fpls.2022.1089720

COPYRIGHT
© 2023 Jacques, Girodet, Leroy, Pluchon, Salon and Prudent. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Memory or acclimation of water stress in pea rely on root system's plasticity and plant's ionome modulation

Cécile Jacques[1], Sylvie Girodet[1], Fanny Leroy[2], Sylvain Pluchon[3], Christophe Salon[1] and Marion Prudent[1]*

[1]Agroécologie, INRAE, Institut Agro, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, Dijon, France, [2]Plateforme PLATIN', US EMerode, Normandie Université, Unicaen, Caen, France, [3]Laboratoire de Nutrition Végétale, Centre Mondial de l'Innovation Roullier, TIMAC AGRO, Saint Malo, France

Introduction: Peas, as legume crops, could play a major role in the future of food security in the context of worldwide human nutrient deficiencies coupled with the growing need to reduce consumption of animal products. However, pea yields, in terms of quantity and quality (i.e. grain content), are both susceptible to climate change, and more specifically to water deficits, which nowadays occur more frequently during crop growth cycles and tend to last longer. The impact of soil water stress on plant development and plant growth is complex, as its impact varies depending on soil water availability (through the modulation of elements available in the soil), and by the plant's ability to acclimate to continuous stress or to memorize previous stress events.

Method: To identify the strategies underlying these plant responses to water stress events, pea plants were grown in controlled conditions under optimal water treatment and different types of water stress; transient (during vegetative or reproductive periods), recurrent, and continuous (throughout the plant growth cycle). Traits related to water, carbon, and ionome uptake and uses were measured and allowed the identification typical plant strategies to cope with water stress.

Conclusion: Our results highlighted (i) the common responses to the three types of water stress in shoots, involving manganese (Mn) in particular, (ii) the potential implications of boron (B) for root architecture modification under continuous stress, and (iii) the establishment of an "ecophysiological imprint" in the root system *via* an increase in nodule numbers during the recovery period.

KEYWORDS

legume, water deficit, recurrent stress, nutrient uptake, agroecology

1 Introduction

Drought is one of the most impacful environmental factors that impairs plant growth, development, and, finally, plant yields (Prasad et al., 2008; Lipiec et al., 2013; Conforti et al., 2018). In the context of climate change, crops face frequent and severe stressful conditions during their growth cycle; among these is the higher frequency of drought events. Nowadays, drought events tend to occur earlier in the crop growth cycle, in the spring period, whereas in previous years they tended to occur at later stages of the growth cycle (Hänsel et al., 2019).

Drought decreases photosynthetic activity (Zhou et al., 2007) either by way of stomatal closure (Chaves, 1991; Flexas et al., 2004) or directly, negatively impacting metabolic activities (Farquhar et al., 1989; Parry, 2002; Bota et al., 2004). In parallel, a decrease of carbon dioxide (CO_2) uptake or internal resistance to CO_2 diffusion can be responsible for an increase in photorespiration (Wingler et al., 2000; Wingler et al., 2002; Prasad et al., 2008). Drought also induces morphological changes; these include a reduction in leaf expansion and size (Alves and Setter, 2004), which result from both a lower number of cells and smaller cells, which generates in turn a lower transpiration area (Randall and Sinclair, 1988; Tardieu et al., 2000; Basu et al., 2016). Drought-induced stomatal closure also lowers plant transpiration and water loss through leaves (Sobeih et al., 2004; Liu et al., 2006; Chai et al., 2016). These responses depend upon the extent of drought. In mild drought, both the number of leaves and the rate of leaf expansion are reduced; in severe drought, leaf growth cessation may occur (Prasad et al., 2008), and a prolongated period of drought may even accelerate leaf senescence (De Souza et al., 1997; Jagadish et al., 2015). Furthermore, roots, being the first organs perceiving soil water stress (Brunner et al., 2015; Weemstra et al., 2016), are also greatly impacted by water availability (Chaves, 1991; Zhou et al., 2007; Prasad et al., 2008; Chapman et al., 2012); under moderate water stress, carbohydrates partitioning to roots is maintained or increased, promoting their growth, but this is, by contrast, reduced by severe drought (Sharp and Davies, 1989; Wu and Cosgrove, 2000; Prasad et al., 2008; Prudent et al., 2016).

Lower soil water availability also reduces the nutrient availability for plants, which in turn limits their nutrient uptake and assimilation (Prasad et al., 2008; Gregory et al., 2012; Barzana et al., 2021). Moreover, water stress negatively impacts the transport of elements within plants owing to the decrease of transpiration, transporter activity, and cell membrane permeability (Viets, 1972; Hsiao, 1973; Kramer and Boyer, 1995; Sánchez-Rodríguez et al., 2010; Ahanger et al., 2016). Thus, under water stress, a nutritional imbalance of the plant ionome can be observed, which results from lower concentrations of essential elements for plant development and growth: macronutrients [nitrogen (N), sulfur (S), phosphorous (P), magnesium (Mg), calcium (Ca), and potassium (K)] or micronutrients [iron (Fe), manganese (Mn), copper (Cu), zinc (Zn), nickel (Ni), molybdenum (Mo), boron (B), and chlorine (Cl)] (Kirkby, 2012). Moreover, the availability of elements considered to be "beneficial" [(cobalt (Co), sodium (Na), vanadium (V), aluminium (Al), selenium (Se), and silicon (Si)] because they can improve plant development and growth and plant response to abiotic stresses may be reduced within water-stressed soils. However, interactions among ionome mineral elements complexifies the assessment of which ones are first impacted by water stress (Baxter et al., 2008; Salt et al., 2008). For example, the reduced availability of Fe can increase the uptake of other metal cations (Ni, Cu, Mn, and Zn), as these elements share the common transporter Iron-Regulated Transporter 1 (IRT1) (Pilon et al., 2009). If rewetting occurs after water stress, such a "recovery period" could trigger a higher level of soil mineralization that enhances nutrient availability for plants (Austin et al., 2004; He and Dijkstra, 2014; Wang et al., 2017).

Under low soil mineral nitrogen availability, legumes such as peas (*Pisum sativum* L.) mainly rely for their nitrogen nutrition on the symbiotic atmospheric N_2 fixation, which is decreased by water stress *via* its negative impact on both structural and functional components (Zahran, 1999; Salon et al., 2001; Prudent et al., 2016). Indeed, drought decreases nodulation initiation, nodule growth, and nodule specific activity (King and Purcell, 2001; Serraj et al., 2001; Streeter, 2003; Larrainzar et al., 2009; Mahieu et al., 2009; Prudent et al., 2016). Altogether, this leads to a decrease in plant nitrogen levels and overall plant growth.

During long periods of continuous stress, plant acclimation includes reprogramming of development, physiology, and metabolism to improve plant functioning and promote better plant health (Lichtenthaler, 1996; Bohnert and Sheveleva, 1998; Orcutt et al., 2000; Wilson and Franklin, 2002; Sanchez et al., 2011). During their acclimation to periods of continuous stress, plants decrease their organ growth. Although early responses to water stress inlcude an induction of ABA signaling leading to stomatal closure (Harb et al., 2010; Pirasteh-Anosheh et al., 2016), the plant response evolves after some days, entering an intermediate stage that includes the regulation of cell wall properties (Neumann, 1995; Moore et al., 2008; Harb et al., 2010). Changes in cell walls, which are essential for plant adaptation (Bacon, 1999), seem to be linked to an early induction of expansin genes that are repressed during continuous stress (Jones and McQueen-Mason, 2004; Harb et al., 2010). Finally, during the later stages of periods of continuous moderate water stress, plant growth is so reduced that it enables the maintenance of plant metabolism and physiology at the same level as control plants as a result of a reduced energy demand (Harb et al., 2010). Indeed, acclimation was shown to maintain stomatal conductance and photosynthesis similarly to well-watered plants (Harb et al., 2010). Moreover, negative regulation of jasmonic acid biosynthesis during the later stages of continuous water stress

could be beneficial to minimize the negative effects of this hormone on plant growth (Ellis et al., 2002; Harb et al., 2010). Thus, despite smaller plant size, their reproductive ability may be maintained in a stressful environment (Chaves et al., 2002).

Several studies have highlighted the importance of the recovery period following drought for plant resilience (Yaqoobbr et al., 2013; Yousfi et al., 2015; Chen et al., 2016; Sun et al., 2016), and notably for pea plants (Prudent et al., 2016; Couchoud et al., 2020). This period enables mitigation of the negative impact of water stress on ecophysiological processes related to plant nutrition. However, the efficiency of recovery highly depends on plant genotypic strategies for restoring nutrition. For instance, in pea, the genotype Kayanne (that is investigated in the present study) was shown to totally recover, through the fine tuning of its nitrogen nutrition status and yield after a drought event during its vegetative growth period, unlike the genotype Puget, for which a drop in yield was observed (Couchoud et al., 2020).

Drought events can be recurrent and a recovery period can be followed by a new period of water stress. In such a case, plant "memory" from the previous stress event has been pointed out as another mechanism, part of resilience-associated processes (Trewavas, 2003; Bruce et al., 2007; Jacques et al., 2021). Plants memorizing their first period of water stress can modify their response to subsequent stress, during which they can display a faster and/or stronger response. The first stress period (priming) induces various molecular and physiological responses that generate a stress imprint (Hilker and Schmülling, 2019). These impacts of recurrent water stresses on different physiological variables linked to overall growth and yield, on physiological processes such as photosynthetic activity, transpiration, osmotic regulation, and antioxidant system, have been studied in several crop species, including wheat (Wang et al., 2014; Abid et al., 2016; Abid et al., 2017; Abid et al., 2018), corn (Virlouvet et al., 2018), rice (Auler et al., 2017; Aihemaiti et al., 2019; Auler et al., 2021), sugarcane (Marcos et al., 2018), sugarbeet (Leufen et al., 2016), and potato (Ramírez et al., 2015), but not in legumes. Priming enables the maintenance of growth rate through greater leaf water potential, leaf water status, and photosynthetic efficiency, and antioxidant system efficiency, whereas non-primed plants are less likely to maintain their growth rate (Abid et al., 2016; Abid et al., 2018; Auler et al., 2021). However, stress memory has been studied mainly at the molecular level, and particularly in leaves; it is also essential to understand if and how memory of water stress takes place in roots for potentially enhancing plant hydro mineral nutrition.

To our knowledge, studies dealing with pea responses to water stress mainly concern transient stresses; plant memory and acclimation to water stress have rarely been studied in this species. We can take into consideration a study highlighting the impact of recurrent stress on the pea shoot metabolome (Lahuta et al., 2022). Taking into consideration what has been described in the literature for several species, we can hypothesize that pea memory and acclimation to water stress involve different physiological changes in plant adaptation. To test this hypothesis, several ecophysiological variables reflecting water and carbon assimilation and allocation have been either measured or calculated and integrated in an ecophysiological framework, enriched by an analysis of the ionome. Altogether these closely connected variables allow us to decipher and distinguish underlying processes involved in (i) plant response to transient water stress, (ii) stress memory during recurrent water stresses, and (iii) stress acclimation during continuous water stress.

2 Materials and methods

2.1 Plant growth conditions

Pea seeds (*Pisum sativum* L. cv. Kayanne, obtained from KWS Momont, Mons-en-Pévèle, France) were calibrated, their surfaces sterilized by exposure to 70% ethanol for 1 minute followed by 0.6% sodium hypochlorite for 10 minutes, and then imbibed in distilled water for 2 hours. Seeds were then pre-germinated in germination boxes for 3 days in the dark at 22°C in a Fitoclima S600 germinator (Aralab, Rio de Mouro, Portugal). Seedlings were then transferred to the 4PMI high-throughput phenotyping platform (Dijon, France) in two litters pots filled with a mix of sand and soil harvested in Aiserey, France (Figure 1). Soil characteristics are given in Supplementary Table S1. Greenhouse environmental conditions were 21.0 ± 1.5°C during the day and 16.6 ± 1.0°C at night, with a photoperiod of 16 h. Supplementary artificial lighting was supplied with sodium lamps (MACS 400W; Mazda, Dijon, France), allowing a mean incident photosynthetically active radiation (PAR) of 280 $\mu E.m^{2}.s^{-1}$. Plants were automatically watered three times per day with N_7-free nutrient solution [see Voisin et al. (2003) for the composition].

2.2 Experimental design and measurements

During the first 2 weeks of growth, plants were watered to reach 90% of substrate water-holding capacity, corresponding to well-watered (WW) conditions. To determine the quantity of the nutrient solution to deliver at each watering, each pot was weighed, and the solution was added to compensate the weight lost due to evapotranspiration. After 2 weeks (t_0, Figure 1), pots were randomly divided into five groups corresponding to the five watering treatments. The first group hereafter called "WW_c" corresponded to plants that were well watered throughout the experiment. The group called "WS_1" corresponded to plants that encountered a short period of water stress (2 weeks) during their vegetative stage (transient vegetative stress), the group called

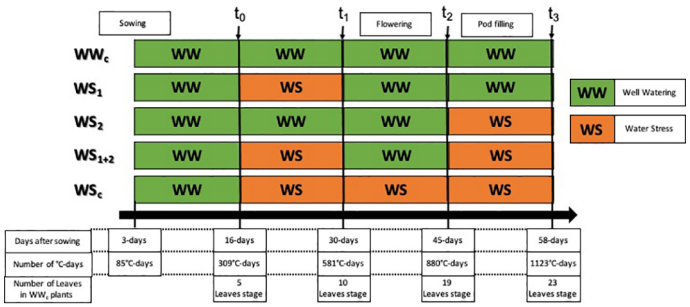


FIGURE 1
Experimental design used to characterize the response of pea plants to four types of water stress. WS_1, transient at vegetative stage; WS_2, transient at the reproductive stage; WS_{1+2}, recurrent stage; WS_c, continuous stress. The control treatment corresponds to WW_c. Well-watered (WW) corresponds to a well-watered period and WS corresponds to a period of water deficit.

"WS_2" corresponded to plants that encountered a short period of water stress (2 weeks) during their reproductive stage (transient reproductive stress), the group called "WS_{1+2}" corresponded to plant that encountered two recurrent water stresses spaced by a re-watering period, and the group called "WS_c" corresponded to plants that encountered a period of continuous water stress (6 weeks).

After the first 2 weeks of growth (t_0, Figure 1), the watering of WS_c, WS_1, and WS_{1+2} plants were stopped to reach 40% of substrate water-holding capacity and was maintained at this level until the end of the water stress period (t_1, Figure 1). After their first water stress period, WS_1 and WS_{1+2} plants were re-watered for 2 weeks to reach 90% of substrate water-holding capacity (WW), whereas WS_c plants watering was maintained at 40% of substrate water-holding capacity for 2 more weeks (t_2, Figure 1). After recovery (for WS_{1+2} plants) or water stress (for WSc plants), WS_c plant watering was maintained at 40% of substrate water-holding capacity, whereas for WS_2 and WS_{1+2} plants, watering was stopped to reach 40% of substrate water-holding capacity for 2 weeks (t_3, Figure 1).

For each treatment (WW_c, WS_1, WS_2, WS_{1+2}, WS_c) five plants were harvested at t_0, t_1, t_2, and t_3. Before each harvest, stomatal conductance (expressed in $mol.m^{-2}.s^{-1}$) was measured on the last fully opened stipule using a porometer (AP4 porometer; Delta T device, Cambridge, United Kingdom). Harvested plants were divided into four samples corresponding to leaves (including stipules and tendrils), stem, roots, and nodules. Leaf and stem areas ($Leaves_A$ and $Stem_A$, in cm^2) were measured using an area meter (LI 3100c area meter; LI-COR, Inc., Lincoln, NE, USA). The root system was thoroughly washed with demineralized water to ensure that the entire root system was free from any soil particles. The root system morphometry was characterized by measuring the taproot length ($Taproot_{LEN}$, in cm), the primary lateral root number (Lateral $root_{Nb}$), the number of nodules localized either on the taproot (Taproot $Nodule_{Nb}$) or on lateral roots (Lateral root $Nodule_{Nb}$). Following each harvest, tissue samples (leaves, stem, roots, and nodules) were dried at 80°C for 48 h. Each organ was weighed to obtain the dry weight of the leaves ($Leaves_{DW}$), stem ($Stem_{DW}$), roots ($Root_{DW}$), and nodules ($Nodule_{DW}$) prior to elemental analyses. Throughout the experiment, each pot was weighed before and after each watering; this allowed us to calculate the plant water uptake (W_{Qty}, in g).

2.3 Elemental analysis of plant tissues

Each sample was ground to fine powder using MM400 vibratory mixer mill (Retch, France). For each sample, carbon (C) and nitrogen (N) concentrations were measured from 5 mg of ground tissue using the Dumas procedure (Thermo Electron NC 2500 Elemental Analyzer, Courtaboeuf, France).

The other element concentrations (S, P, K, Ca, Mg, Cu, Ni, Mn, B, Zn, Fe, V, Co, Na, and Mo) were measured with a high-resolution, inductively coupled plasma mass spectrometer (HR ICP-MS, Element2, Thermo Fisher) using the PLATIN' Platform (Caen, France), following methods previously described in Lurthy et al. (2020). For this, a sample of 40 mg of dry powder was submitted to acid digestion composed of 1 mL of nitric acid (HNO_3), 250 µL of hydrogen peroxide (H_2O_2), 900 µL of ultrapure water and 10 µL of internal standard solution of gallium and rhodium before to be diluted to 50 mL with ultrapure water. Then, the solution obtain was filtered with a 0.45 µm Teflon filter. Finally, quantification was obtained by correction of calibration curves by subtracting blank and using internal standards [gallium (Ga) and rhodium Rh)], and evaluated by certified reference plant material (Citrus leaves, CRM NCS ZC73018, Skylab, Metz, France).

10.3389/fpls.2022.1089720

2.4 Calculated variables and statistical analyses

From measured variables, several other variables were calculated to build an ecophysiological framework originally linking C and N in the plant (adapted from Couchoud et al., 2020) and to calculate water and element uptake and use efficiencies.

The shoot dry weight ($Shoot_{DW}$; *expressed in g*) was calculated as:

$$Shoot_{DW} = Leaves_{DW} + Stem_{DW}$$

Where $Leaves_{DW}$ is the dry weight of leaves (in g) and $Stem_{DW}$ is the dry weight of stems (in g).

The root system dry weight ($RootSystem_{DW}$; *expressed in g*) was calculated as:

$$RootSystem_{DW} = Root_{DW} + Nodule_{DW}$$

Where $Root_{DW}$ is the dry weight of root (in g) and $Nodule_{DW}$ is the dry weight of nodule (in g).

The plant dry weight ($Plant_{DW}$; *expressed in g*) was calculated as:

$$Plant_{DW} = Shoot_{DW} + RootSystem_{DW}$$

The ratio of shoot dry weight (DW) over root system (root and nodule) DW ($R_{Shoot:RootSystem}$) was calculated as:

$$R_{Shoot:RootSystem} = \frac{Shoot_{DW}}{RootSystem_{DW}}$$

The ratio of nodule dry weight over root dry weight (R_{Nodule}) was calculated as:

$$R_{Nodule} = \frac{Nodule_{DW}}{Root_{DW}}$$

Where $Nodule_{DW}$ is the dry weight of nodules (in g).

The element quantity in different organs (leaves, stem, root, and nodule) ($Organ_{Element_{Qty}}$; expressed in g) was calculated as:

$$Organ_{Element_{Qty}} = Organ_{DW} \times Organ_{[Element]}$$

Where $Organ_{[Element]}$ is the element concentration in organ (in g/g$Organ_{DW}$).

And the element quantity in plants ($Plant_{Element_{Qty}}$; expressed in g) was calculated as:

$$Plant_{Element_{Qty}} = Leaves_{Element_{Qty}} + Stem_{Element_{Qty}} + Root_{Element_{Qty}} + Nodule_{Element_{Qty}}$$

That enabled us to calculate the element concentration in plant components (leaves, stem, root, and nodule) ($Plant_{[Element]}$; expressed in g/g$Plant_{DW}$) as follows:

$$Plant_{[Element]} = \frac{Plant_{Element_{Qty}}}{Plant_{DW}}$$

The relative variation of each variable between time t_i and time t_{i+1} ($\Delta Var_{ti \to ti+1}$ such as the dry weight, $Element_{Qty}$ (quantity of element accumulated in plant, expressed in g) or W_{Qty} (water quantity uptake by root, expressed in g) between two harvests was calculated as:

$$\Delta Var_{ti \to ti+1} = \frac{Var_{ti+1} - Var_{ti}}{Var_{ti}} \times 100$$

Where Var_{ti} is the variable value at the harvest time $t_i$and Var_{ti+1} the variable value at the following harvest time t_{i+1}.

The specific leaf area (SLA; expressed in cm^2/g) between two harvests (t_i and t_{i+1}) was calculated as:

$$SLA = \frac{Leaves_{A\ t_{i+1}} - Leaves_{A\ t_i}}{Leaves_{DW\ t_{i+1}} - Leaves_{DW\ t_i}}$$

Where $Leaves_{DW}$ is the dry weight of leaves (in g) and $Leave_A$ is the projected area of leaves ($in\ cm^2$).

The radiation use efficiency (RUE, expressed in g/cm^2) between two harvests (t_i and t_{i+1}) was calculated as:

$$RUE = \frac{Plant_{DW\ t_{i+1}} - Plant_{DW\ t_i}}{\int_{ti}^{ti+1} Leaves_{A\ t_{i+1}} . dt}$$

The water use efficiency (WUE; expressed in g/m^3) between two harvests (t_i and t_{i+1}) was calculated as:

$$WUE = \frac{Plant_{DW\ t_{i+1}} - Plant_{DW\ t_i}}{W_{Qty\ t_{i+1}} - W_{Qty\ t_i}}$$

The element use efficiency (EUE; expressed in g of $Plant_{DW}$/g) was calculated as:

$$EUE = \frac{Plant_{DW\ t_{i+1}} - Plant_{DW\ t_i}}{Element_{Qty\ t_{i+1}} - Element_{Qty\ t_i}}$$

The element uptake efficiency (EUpE; expressed g/g of *Root System*$_{DW}$) between two harvests (t_i and t_{i+1}) was calculated as:

$$EUpE = \frac{Element_{Qty\ t_{i+1}} - Element_{Qty\ t_i}}{\int_{ti}^{ti+1} RootSystem_{DW} . dt}$$

Statistical analyses were performed with R software (https://www.r-project.org/, v.3.5.2). At each harvest (t_1, t_2, t_3), multiple comparisons were performed between treatments (WS_2, WS_{1+2}, WS_c, and WW_c). Kruskal–Wallis tests and Wilcoxon tests were performed *via* Kruskal.test and wilcox.test functions, respectively, and results were displayed *via* ggboxplot function (package ggpubr). Effects were considered significant when the *p*-value was lower than 0.05.

10.3389/fpls.2022.1089720

3 Results

All primary data are given in Supplementary Tables S2-S4, and pictures of plants are given in Supplementary Figure S5.

3.1 Only single stress period improves water use efficiency

To characterize plant water management, several traits were measured and calculated for each water stress treatment (WS_1, WS_2, WS_{1+2} and WS_c) such as stomatal conductance (Figure 2), water uptake, and use efficiencies (Figures 3, 4). Under recurrent stress during the first stress period (WS_{1+2} at t_1), water uptake efficiency (WUpE) and stomatal conductance were decreased respectively by 22.8% and 31.7% (Figures 3, 2), whereas water use efficiency (WUE) was increased by 95.7% (Figure 4). However, during the recovery period (WS_{1+2} at t_2) all these variables returned to the values of control plants. Then during the second stress period (WS_{1+2} at t_3), stomatal conductance and WUE decreased. The impact on stomatal conductance was equivalent under transient reproductive stress (WS_2 at t_3), under which WUpE was decreased by 76.0%. Whereas, WUE decrease was equivalent for recurrent (WS_{1+2}) and single stress at vegetative stage (WS_1) during reproductive period (t_3). And under continuous stress of 6 weeks (WS_c at t_3), stomatal conductance was less decreased by 54.3% and WUpE was decreased by 77.5%.

3.2 Two strategies for soil prospection: Increased taproot growth under WS_2 and enhanced lateral root initiation under WS_c

The first water stress period at the vegetative stage (t_1) induced numerous changes related to plant morphology and carbon fluxes (Figure 5). First, carbon was preferentially allocated to the nodulated root system, at the expense of shoots. Indeed, the shoot dry weight accumulation (including both leaves and stem) of water-stressed plants was reduced by 9.25% when compared with well-watered plants, because of a decrease in leaf area of 8.49% and a decrease in radiation use efficiency (RUE) of 12.4%. At the same time, the dry weight of the root system increased by 32.1% because of the increase in taproot dry weight (+68.4%), the length of which also increased, and of the lateral roots' dry weight (+60.2%), which were were numerous in water-stressed plants than in well-watered plants. Moreover, the ratio of nodule dry weight to root dry weight (R_{Nodule}) was decreased (by 49.4%), indicating during this stress period a higher allocation of C to the root than to nodule. This explained the decrease of nodule dry weight accumulation and nodule initiation, respectively, by 46.6% and 33.6%. During the following recovery period (WS_{1+2} at t_2) the variables which are related to the photosynthetic C fixation and to the allocation of C in the shoot were almost unchanged, except for the specific leaf area (SLA), which decreased by 15.5% (Figure 5). On the contrary, opposite responses were observed between the end of the first water stress period and the end of the recovery period in below-ground parts. Indeed, root system DW

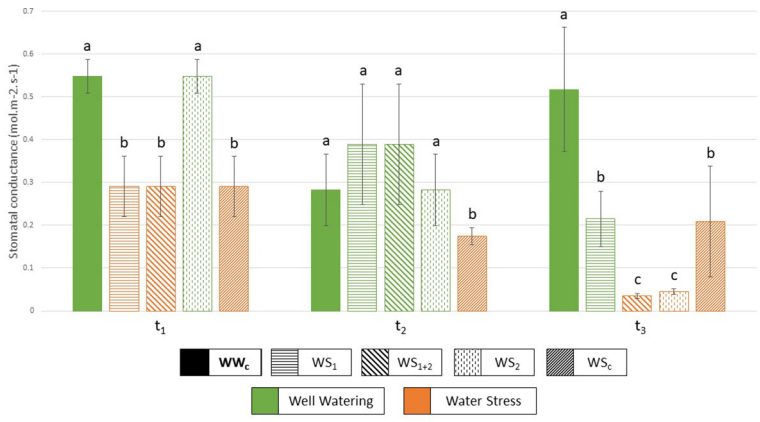


FIGURE 2
Stomatal conductance (mol.m^{-2}.s^{-1}) measured at each harvest (t_1, t_2, t_3), and for each treatment. WW_c, well-watered control plants; WS_1, transient vegetative stress; WS_{1+2}, recurrent stress, WS_2, transient reproductive stress, and WS_c, continuous stress. Green represents a well-watered period and orange, a water-stressed period. Values correspond to means ± SD (n = 4 or n = 5), letters regroup treatment with a non-significant difference for each harvest.

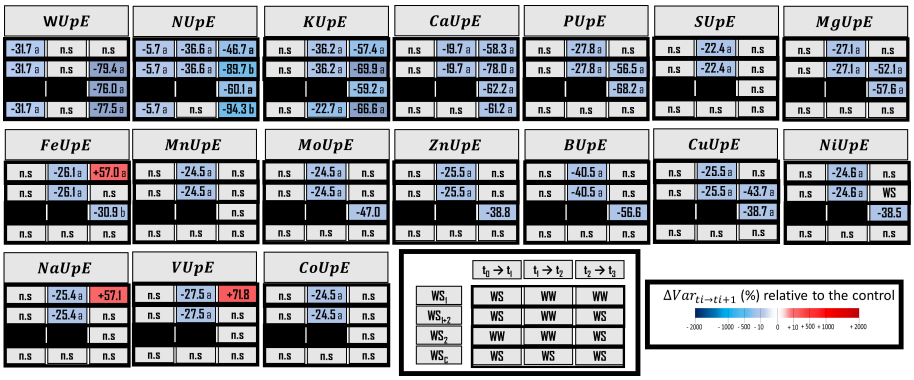


FIGURE 3
Water (WUpE) and element uptake efficiency (EUpE) for macronutrients, micronutrients and three beneficial elements. Elements (E) characterized are N, nitrogen; K, potassium; Ca, calcium; P, Phosphorus; S, Sulfur; Mg, Magnesium; Fe, Iron; Mn, Manganese; Mo, Molybdenum; Zn, Zinc; B, Boron; Cu, Copper; Ni, Nickel; Na, Sodium; V, Vanadium, and Co, Cobalt. The effect of each treatment (WS_1, transient vegetative stress; WS_{1+2}, recurrent stress; WS_2, transient reproductive stress; WS_c, continuous stress) was expressed relatively to control plants (WW_c) at each harvest (t_1, t_2, t_3) (ΔVar $t_i \rightarrow t_{i+1}$). Boxes were colored in red when the treatment induced an increase of the variable, and in blue when the treatment induced a decrease of the variable compared with conrol plants (WW_c); the intensity of the effect was characterized by the scale of color and the value (percent compared with WW_c) present in the box. Data are presented as a percentage relative to control plants (n = 4 or n = 5). Primary data are available in Supplementary Table S4. Letters regroup treatments with non-significant difference for each harvest.

accumulation was decreased, and nodule initiation increased, respectively by 37.6% and 177.2%.

During the reproductive stage (t_3), whatever type of water stress was considered (transient, WS_1, or WS_2; recurrent, WS_{1+2}, or continuous WS_c), we found that most of the traits reflecting photosynthetic carbon fixation and allocation were similarly impacted. However, the intensity of this impact depended on the type of stress. For instance, plant dry weight accumulation was more negatively impacted under recurrent stress (–80.1%) than under continuous stress (–73.3%), single reproductive stress (–59.8%), and single vegetative stress (–26.6%). On the other hand, traits reflecting root morphology were typically affected by each of the four water stress treatments. When compared with well-watered plants, taproot elongation was increased under transient reproductive stress (+1860%) but decreased under recurrent stress (–771.0%) and even more under continuous stress (–2118.2%). The number of lateral roots initiated during this period was not affected by transient reproductive stress or a recurrent stress but was greatly decreased (+766.7%) under continuous stress. Conversely, the

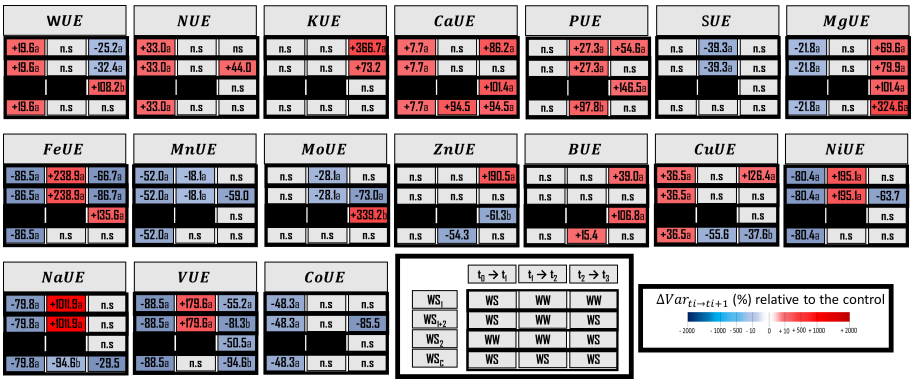


FIGURE 4
Water use efficiency (WUE) and element use efficiency (EUE) for macronutrients, micronutrients and three beneficial elements. Elements (E) characterized are N, Nitrogen; K, Potassium; Ca, Calcium; P, phosphorus; S, sulfur; Mg, magnesium; Fe, iron; Mn, manganese; Mo, molybdenum; Zn, zinc; B, boron; Cu, copper; Ni, nickel; Na, sodium; V, vanadium, and Co, cobalt. The effect of each treatment (WS_1, transient vegetative stress; WS_{1+2}, recurrent stress; WS_2, transient reproductive stress; WS_c, continuous stress) was expressed relatively to control plants (WW_c) at each harvest (t_1, t_2, t_3; n = 4 or n = 5) (ΔVar $t_i \rightarrow t_{i+1}$). Boxes were colored in red when the treatment induced an increase of the variable, and in blue when the treatment induced a decrease of the variable compared with conrol plants (WW_c); the intensity of the effect was characterized by the scale of color and the value (percent compared with WW_c) present in the box. Primary data are available in Supplementary Table S4. Letters regroup treatment with non-significant difference for each harvest.

10.3389/fpls.2022.1089720

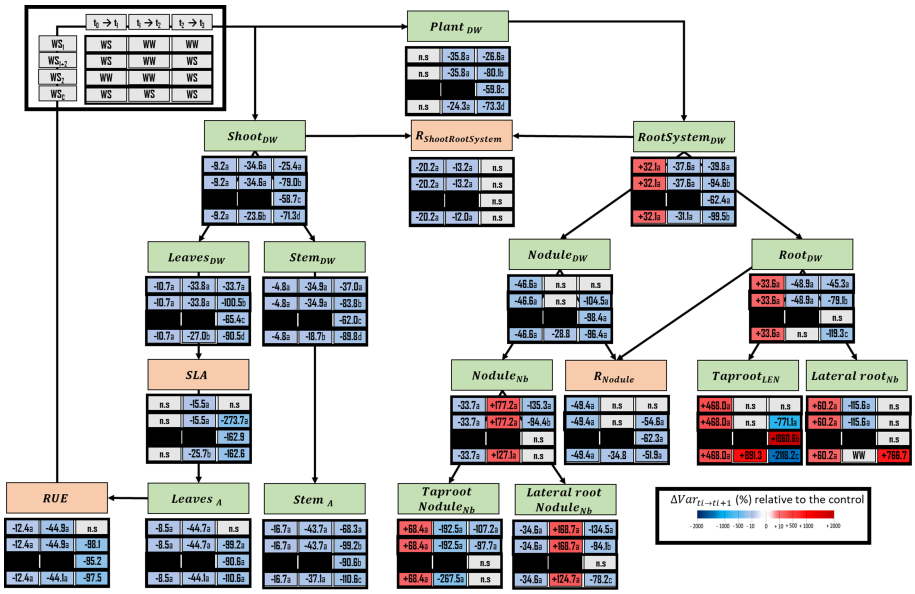


FIGURE 5
Ecophysiological framework highlighting several variables related to carbon fixation or allocation within plant, that have been either measured or calculated. Lateral root$_{Nb}$, lateral root number; Lateral root Nodule$_{Nb}$, number of nodules located on the lateral roots; Leaves$_{DW}$, leaves dry weight; Leaves$_{A}$, Leaves projected area; Nodule$_{DW}$, nodule dry weight; Nodule$_{Nb}$, total nodule number; Plant$_{DW}$, plant dry weight; R$_{Nodule}$, ratio of nodule dry weight over the root dry weight; R$_{Shoot\ :\ RootSystem}$, ratio of shoot dry weight over the whole root system dry weight; Root$_{DW}$, root dry weight; RootSystem$_{DW}$, root System (root and nodule) dry weight; RUE, radiation use efficiency; Shoot$_{DW}$, shoot dry weight; Stem$_{A}$, stem projected area; Stem$_{DW}$, stem dry weight; Taproot$_{LEN}$, taproot length; Taproot Nodule$_{Nb}$, number of nodules located on the taproot; SLA, specific leaf area. The effect of each treatment (WS$_1$, transient vegetative stress; WS$_{1+2}$, recurrent stress; WS$_2$, transient reproductive stress; WS$_c$, continuous stress) was expressed relatively to control plants (WW$_c$) at each harvest (t_1, t_2, t_3; n=4 or 5) (ΔVar $t_i \rightarrow t_{i+1}$). Boxes were colored in red when the treatment induced an increase of the variable, and in blue when the treatment induced a decrease of the variable; the intensity of the effect was characterized by the scale of color and the value (percent compared with WW$_c$) present in the box. Primary data are available in Supplementary Table S2. Letters regroup treatment with non-significant difference for each harvest.

number of nodules initiated during this period was decreased by 94.4% under recurrent stress and by 135.3% transient vegetative stress at reproductive period.

3.3 Element uptake efficiencies were more negatively impacted under continuous water stress and element use efficiencies were more negatively impacted under recurrent water stress

To characterize pea mineral nutrition during the four types of water stress [transient stress (vegetative, WS$_1$; or reproductive, WS$_2$), recurrent stress (WS$_{1+2}$), and continuous stress (WS$_c$)], an analysis of the plant ionome was performed in the different organs of the plant and allowed to characterize both the uptake efficiency and the use efficiency of 16 nutrients and beneficial elements.

When water stress occurred during the vegetative stage (WS$_{1+2}$ at t_1), 11 out of the 16 elements were impacted for their uptake efficiencies (UpE) or for their use efficiencies (UE), or for both (Figures 3 and 4). However, only N uptake efficiency (NUpE) and use efficiency (NUE) and use efficiencies responded in a compensatory way, with a lower level of uptake efficiency (–5.7%) that could be compensated for by a higher use efficiency of this element (+32.7%). The use efficiencies of Ca and Cu were increased (by 7.7% and 36.5%, respectively). Conversely, Mg, Fe, Mn, Ni, Na, V and Co use efficiencies were decreased (by 21.8%, 86.5%, 52.0%, 80.4%, 79.8%, 88.5%, and 48.28%, respectively). Under water stress at the vegetative stage, VUE (Vanadium Use Efficiency, -88.5%) and FeUE (Iron Use Efficiency, -86.5%) were the most negatively impacted. Interestingly, K, P, S, Mo and Zn and use efficiencies were not impacted by the water stress at vegetative stage. Similarly, to what was previously observed for variables related to root morphology, it appeared that for most of the elements (except for NUpE, KUE, ZnUE, and BUE) their uptake or use efficiencies were drastically changed during the recovery period following the water stress (WS$_{1+2}$ at t_2). Moreover, other elements were impacted by their uptake efficiencies (KupE, +22.3%; CaUpE, –19.7%; PUpE, –27.8%; SUpE, +81.0%; MgUpE, –24.5%; FeUpE, –26.1%; MnUpE, –24.5%; MoUpE, +80.9%; ZnUpZ, –25.5%; BUpE, –40.5%; CuUpE, –25.5%; NiUpE, –24.5%; NaUpE, –25.4%; VUpE, –27.5%; and CoUpE, –24.5%) or their use efficiencies (PUE, +27.3%; and SUE, +39.3%). During recovery, MnUE decreased by a smaller amount (by 18.4%) than in the first stress period (WS$_{1+2}$ at t_2).

During the reproductive stage (t_3), either uptake efficiencies, use efficiencies, or both were impacted. In terms of elements and the three types of stress, MgUpE was decreased and MgUE was increased, by −52.5% and +79.9%, respectively; for recurrent stress, by −57.9% and +181.9%, respectively, for transient stress; and only MgUE was increased, by +323.6%, for continuous stress. Moreover, CaUpE (Calcium Uptake Efficiency) and CaUE (Calcium Use Efficiency) were similarly regulated under transient vegetative or reproductive stress (−58.3% and +86.2%; −62.2% and +101.4%, respectively) and continuous stress (by −61.2% and +94.5%). In terms of transient reproductive stress (WS_2), the most negatively impacted uptake efficiency was PUpE (−68.2%) and the only use efficiency negatively impacted was VUE (−50.2%). In terms of recurrent stress, antagonistic regulations were observed, such as the decrease of NUpE (−86.7%) and the increase of NUE (+44.1%). NUpE was the uptake efficiency that was most decreased (−86.7%) and the FeUE use efficiency was most decreased (−86.7%). Moreover, NUpE and VUE were less impacted under transient vegetative stress (WS_1 at t_3) (respectively, −46.7% and −55.2%) than under recurrent stress (WS_{1+2} at t_3) (respectively, −86.7% and −81.4%). Whereas, these two types of stress impact KUpE, MgUE, and FeUE in a similar way. Under continuous stress, CuUE (−55.6%) and NaUE (−94.62%) were the only element use efficiencies that were decreased as soon as after 4 weeks of stress. However, the uptake efficiencies of K were negatively impacted after 4 weeks of stress. After 6 weeks of stress, three uptake efficiencies were negatively impacted (NUpE, −94.3%; KUpE, −66.6%; CaUpE, −61.2%) Finally, under continuous stress, NaUE had the most decreased use efficiency (−94.6%).

3.4 Water stress modulated plant element concentrations progressively from roots to shoot

A description of the plant ionome based on the element concentrations of each organ was performed for each water stress treatments (Figure 6). After the first stress period (WS_{1+2} at t_1), eight element concentrations were modulated: seven were increased (C, Mg, Fe, Mn, Ni, Na, V and Co) among which the Fe concentration was most improved (+157.3%), and the N concentration was the only negatively altered concentration (-18.8%) (Figure 6A). Moreover, roots were the organ most impacted by the six elements with increased concentrations (Fe, Mn, Mo, Ni, Na and V) among which Fe concentration (+166.6%) was the most increased, and Ca concentration was the only decreased concentration (−19.8%) (Figure 6D). Leaves were the organ least impacted by stress, with only two element concentrations that were increased, among which the Ni concentration (+85.3%) was most increased (Figure 6B). Moreover, in leaves, the Zn concentration was most increased (+44.7%) and the N concentration was most decreased (−64.1%). In nodules, the Ni concentration was also greatly increased (+125.9%) and the N concentration was similarly decreased (−63.7%) (Figure 6E).

During the recovery period after the first water stress period (WS_{1+2} at t_2), the Mn and Co concentrations continued to increase at the whole plant level (Figure 6A). Roots were still the most impacted organ during recovery (Figure 6D). The concentrations of Mn and Mo were enhanced in this organ, but the K concentration was most amplified (+293.2%). The organs least impacted were the nodules, in which the Na concentration was the only decreased concentration (−59.4%) (Figure 6E).

During the second stress period (WS_{1+2} at t_3), Mn and Co concentrations remained increased at the whole plant level, and the Fe concentration was most increased (+180.1%) (Figure 6A). However, although the Co concentration was equivalently impacted under transient vegetative stress (WS_1 at t_3), the Mn concentration was less impacted. Moreover, in root Mg and B concentrations were equivalently impacted by these two types of stresses (Figure 6D). Under recurrent stress, stem was the organ that was not affected while nodules were the most impacted (five element concentrations were enhanced). In nodules, the K concentration was the most increased (+413.7%) (Figure 6E).

After 4 weeks of continuous stress (WS_c at t_2), the P concentration was the only concentration that was decreased at the whole plant level (−35.5%) (Figure 6A). The leaves were the compartment most impacted by the four elements with increased concentrations (Ca, Mn, Ni and Co) and the seven elements with decreased concentrations (N, K, P, S, Mn, B and Na) (Figure 6B). In leaves, the element with the most increased concentration was Ni (+56.1%) while the element with the most reduced concentration was P (−34.7%). The stem was the least impacted organ, with only four elements whose concentrations were decreased (K, P, Cu and V) (Figure 6C). Then, after 6 weeks of continuous stress (WS_c at t_3), the P concentration was still decreased by 42.6% (as was as the N concentration) and Mn was the only element whose concentration increased (+75.2%) at the whole plant level. Moreover, leaves and stem were the most impacted components, with eight element concentrations that increased or decreased. In the stem, six element concentrations increased (K, S, Fe, Mn, B and Co), among which the Fe concentration was most increased (+193.5%) and the P and Na concentrations were most decreased (−50.4% and -43.4%, respectively). Nodules were only slightly affected by decreased Mn and V concentrations (Figure 6E).

Under single transient stress at the reproductive stage (WS_2 at t_3), no change in element concentration was observed at the whole plant level (Figure 6A). Moreover, roots were the organs that were the most impacted by the six elements with increased concentrations (K, Mg, Fe, Mn, Ni and V) among which K concentration was the most increased (+132.0%) and four element concentrations that decreased (Ca, Zn, Cu and Co), among which the Cu concentration was most decreased (−

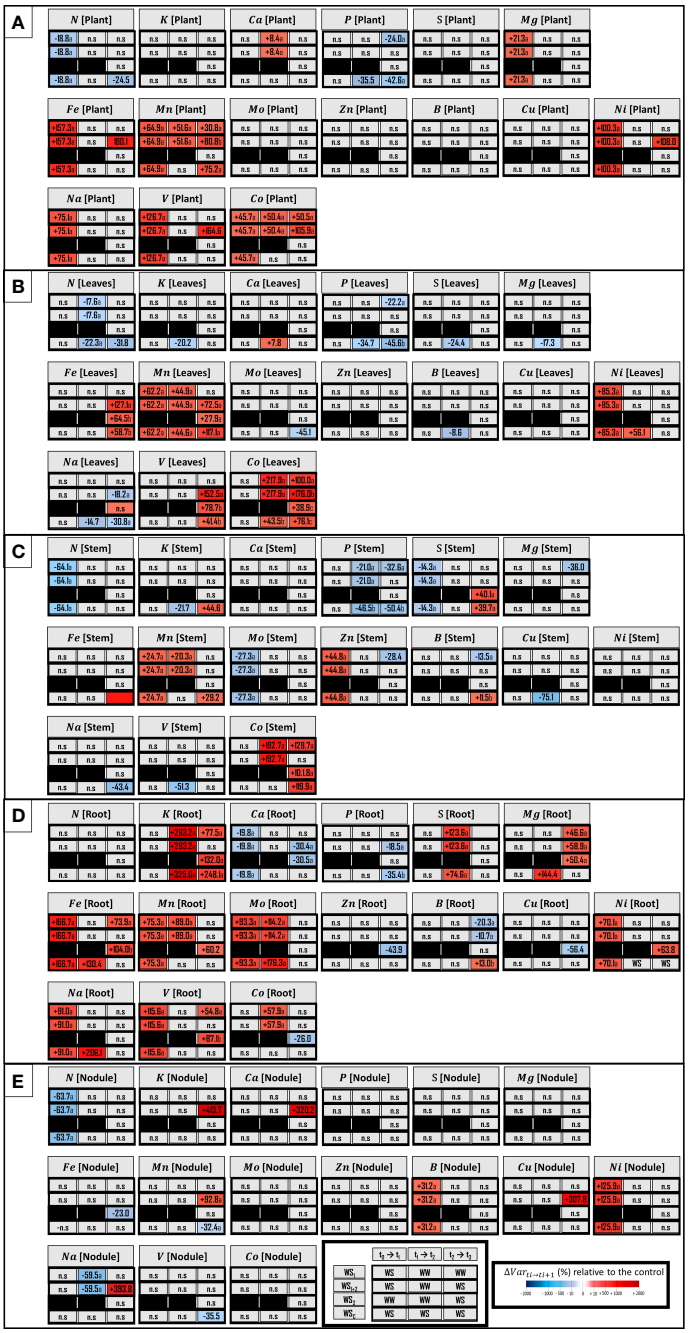


FIGURE 6
Concentrations of macro-nutrients, micro-nutrients and three beneficial elements in whole plant **(A)** and leaves **(B)**, stem **(C)**, roots **(D)**, nodules **(E)**. Elements (E) characterized are N, nitrogen; K, potassium; Ca, calcium; P, phosphorus; S, sulfur; Mg, magnesium; Fe, iron; Mn, manganese; Mo, molybdenum; Zn, zinc; B, boron; Cu, copper; Ni, nickel; Na, sodium; V, vanadium, and Co, cobalt. The effect of each treatment (WS_1, transient vegetative stress; WS_{1+2}, recurrent stresses; WS_2, transient reproductive stress; WS_c, continuous stress) was expressed relatively to control plants (WW_c) at each harvest (t_1, t_2, t_3; n=4 or 5) ($\Delta Var\ t_i \rightarrow t_{i+1}$). Boxes were colored in red when the treatment induced an increase of the variable, and in blue when the treatment induced a decrease of the variable; the intensity of the effect was characterized by the scale of color and the value (percent compared with WW_c) present in the box. Primary data are available in Supplementary Table S3. Letters regroup treatment with non-significant difference for each harvest.

56.3%) (Figure 6D). Finally, nodules and aerial parts were the organs least impacted (Figure 6E).

4 Discussion

In this study, plant strategies related to hydromineral nutrition status maintenance or restoration were characterized according to the different types of water stress: single transient stress at the vegetative stage (WS_1), single transient reproductive stress (WS_2), recurrent stress (WS_{1+2}), and continuous stress (WS_c). Thus, plant strategies (hereafter, we will refer to plant responses that may play a role in resilience mechanisms to water stress as "strategies") that could be involved in resilience to water stress present common and typical response among the different types of stress. Plants were described at the whole-plant level through ecophysiological variables related to water, soil mineral resource uptakes and uses, and, finally, their ionome composition.

4.1 The shoot compartment was negatively impacted, whatever type of water stress it encountered

Shoot compartment was negatively impacted whatever the type of stress encountered by the plant and the period in the growth cycle at which the stress occurred (Figure 5). However, the decrease of plant growth, and more specifically of its shoot, was greater under recurrent stress and during continuous stress than single reproductive stress. However, the impact on shoot growth of stress at reproductive stage during recurrent stress (WS_{1+2} at t_3, -79.0%) could represent cumulative effect of single vegetative stress (WS_1 at t_3, -25.4%) and single reproductive stress (WS_2 at t_3, -58.7). A general trend was observed for all stresses, whatever the period of the plant cycle (except WS_1 at t_3). Mn was the only element which concentration was systematically enhanced in leaves (Figure 6B). This element is known to be involved in several processes related to photosynthesis (Goussias et al., 2002; Schmidt et al., 2016), ATP synthesis (Pfeffer et al., 1986; Ahanger et al., 2016), RuBP carboxylase reaction (Houtz et al., 1988; Bloom and Lancaster, 2018), and biosynthesis of fatty acid, acyl lipid, and proteins (Ness and Woolhouse, 1980; Millaleo et al., 2010; Tripathi et al., 2015). Moreover, Mn plays a role in antioxidant response under water stress *via* its requirement for superoxide dismutase (Sevilla et al., 1980; Bridges and Salin, 1981; Sandmann and Böger, 1983; Polle et al., 1992). So, even if we have not demonstrated that plants experienced a Mn deficiency during drought, one could suggest that an increased concentration of Mn in leaves could improve carbon fixation and thus plant growth under water stress. Moreover, under continuous stress (WS_c at t_3), the increase of Mn concentration in leaves was amplified when compared with recurrent (WS_{1+2} at t_3) and transient reproductive stresses (WS_2 at t_3). This could explain the smaller reduction of plant and shoot growth under continuous stress as compared with recurrent stress. Moreover, some studies have highlighted the beneficial role of a foliar application of Mn for plants grown under water deficit, as it could improve growth parameters and N_2 fixation (Purcell et al., 2000; Karim et al., 2012; Ghorbani et al., 2019). Concomitantly, under a continuous stress, the higher stomatal conductance that was observed could be linked to a higher photosynthetic rate than during recurrent or transient stresses. It thus appears that pea plants, by reducing their growth, seem to induce an acclimation process leading to the maintenance of a physiological activity similar to that observed in control plants. This limitation of energy-consuming processes during the first weeks of drought reduced plant growth on the long term (Harb et al., 2010).

As with the concentration of Mn, concentrations of Fe, Co, and V were increased regardless of the type of water stress encountered at the reproductive stage (except for Fe and V, WS_1 at t_3), with higher increases observed under recurrent stress (Figure 6A). In terms of the Co concentration, its increase under recurrent stress (WS_{1+2} at t_3, +176.0%) seemed higher than the cumulative impact of single vegetative stress (WS_1 at t_3, +100.0%) and single reproductive stress (WS_2 at t_3, +38.9%). In the same way as for Mn, increases in the concentrations of Fe and V in leaves could also help the plant to maintain its growth under water stress at the reproductive stage. Indeed, Fe is known to be involved in chlorophyll biosynthesis, photosynthesis, and respiration (Halvin et al., 1999; Rehman et al., 2021), and the concentration of V can enhance plant growth *via* a higher tissue elasticity, thus enabling cell expansion (García-Jiménez et al., 2018).

On the other hand, Na concentration in leaves only increased under transient reproductive stress, but decreased under recurrent and continuous stress at the reproductive stage (Figure 6A). Therefore, the increased concentration of Na in leaves could be beneficial *via* its osmoticum role and its growth promotion under the non-limiting conditions of K (Lehr and Wybenga, 1955; Tinker, 1965; Milford et al., 1977; Pessarakli et al., 2015).

Although more focused investigations are needed to assess the different roles of Na, Mn, Fe and V, these elements may be beneficial for various plant biological processes despite a lower plant and shoot growth rate under water stress at the reproductive stage. For instance, fine measures of physiological parameters (such as photosynthesis) under different levels of Na, Mn, Fe and V could validate the implication of their concentration increase on the water stress tolerance process.

4.2 Within the root system, priming inhibited nodule initiation and acclimation modulated root architecture

The root system was impacted in a different manner by each of the three types of water stress applied at the reproductive

stage. Taproot elongation was promoted only under transient reproductive stress (WS_2 at t_3), whereas lateral root initiation increased only under continuous stress (WS_c at t_3). On the other hand, recurrent stress (WS_{1+2} at t_3) was the only stress applied at the reproductive stage that decreased nodule initiation, with nodulation most typically occurring on the taproot. However, the same trend was observed at this period in plants for which a single period of stress applied at the vegetative stage (WS_1 at t_3). Therefore, at the root system scale, different responses were induced under water stress during the reproductive period, depending on the type of stress.

4.2.1 After single reproductive transient stress, plant promoted taproot growth and maintained nodule initiation

The typical impact of single reproductive transient stress was mostly observed in roots whose taproot length was increased and whose nodule numbers remained stable (Figure 5). The increase of the conctration of V in roots is also an important element for increased taproot growth rate under single reproductive stress. Indeed, this element has been shown to improve taproot elongation *via* its role in the elasticity of tissues associated with cell expansion (García-Jiménez et al., 2018; Chen et al., 2021).

Moreover, the modification of root and nodule ionomes (Figures 6D, E) could in turn impact nodule growth, development, and activity. On one hand, the amount of N fixed by nodules depends on structural components such as nodule number and dry weight. On the other hand, it depends on functional component like N_2 fixation activity. Several mineral elements were important for both the establishment and the maintenance of these two components. Concerning the structural component of legume symbiosis, Fe impacts nodule formation (Su et al., 2008) and, as Cu is also critical for nodule dry weight accumulation (Hemantaranjan and Garg, 1986; Tang et al., 1990; O'Hara et al., 1993; Tang, 2001; Rotaru and Sinclair, 2009).

The larger concentration of Fe in roots observed in our study could explain the maintenance of nodule initiation during the reproductive stage (Figure 5). Similarly, the observed decrease of nodule dry weight accumulation could be linked to both the decrease of Fe concentration in nodules and the decrease of Cu concentration in roots (Figure 6D). Regarding the functional component of N_2 fixation, decreased concentrations of Fe in nodules and Cu in roots could negatively impact nodule activity. Indeed, Fe is an essential element for N_2 fixation, because of its role in bacteroid and nitrogenase activity (Kaczor et al., 1994; Rotaru and Sinclair, 2009) and Cu deficiency has been shown to negatively impact leghaemoglobin concentration (Snowball et al., 1980; Seliga, 1993; Seliga, 1998). Moreover, increased Ni and V concentrations in roots could enhance nodule activity. Indeed, Ni is essential to nitrogenase activity in many rhizobial bacteria (Albrecht et al., 1979; Brito et al., 1994; Cammack, 1995; Freitas et al., 2019).

Further studies should aim at understanding the most limiting elements for nitrogen fixation in peas among Fe, Cu, and Ni are, and confirm if an increase of these element concentrations in nodule or roots could directly positively impact the structural and functional components of N_2 fixation under water stress. This could help to identify which elements ought to be preventively supplemented to the crop before the appearance of potential transient water stress.

4.2.2 Priming limited plant growth and promoted element concentrations of tissues

Under recurrent stress, ecophysiological memory was characterized by analogy with memory characterization at the molecular scale (Ding et al., 2013; Ding et al, 2014; Jacques et al., 2021), by quantifying the changes induced by the second stress period regarding those induced by a single stress period (Supplementary Material S6). The typical impact of recurrent stress (WS_{1+2} at t_3) on the root system consisted in the decrease of the number of nodules initiated on both the taproot and the lateral roots during the stress such as observed during the reproductive period after a transient vegetative stress (WS_1 at t_3) (Figure 5). However, the decrease in the number of nodules initiated on lateral roots was lower under recurrent stress compared with the decrease that occurred after transient vegetative stress. This decrease of nodule initiation during the second stress period could be explained by the previous increase of nodule initiation during the recovery period following the first water stress period. Nodules were the only organs that displayed the most intensive ionome regulation under recurrent stress, as illustrated by increases of K, Ca, Cu and Na concentrations (Figure 6E), but were not impacted under transient vegetative stress. In faba bean or yellow lupin, Cu accumulation in nodules was shown to improve nodule dry weight accumulation and leghaemoglobin concentration (Seliga, 1998). Under recurrent water stress, an increase of Cu concentration in nodules may improve water stress tolerance in the pea plant, with a lower negative impact on nodule development and function. It seems consistent with the lower decrease of nodule dry weight accumulation under recurrent stress, regarding transient reproductive stress (Figure 5). Finally, although recurrent stress had a more negative impact on plant growth than transient reproductive stress, changes related to the nodule compartment (initiation, growth, elemental concentration) may be considered a ecophysiological imprint of previous water stress. Moreover, this imprint could be more beneficial if a second period of stress occurs. This highlights the importance of the trade-off between memory and the absence of memory if a second period of stress does not occur (Crisp et al., 2016; Jacques et al., 2021).

4.2.3 The acclimation of plants during continuous stress, enhanced plant soil exploration, and delayed nodule initiation

The typical impact of continuous stress on the root system consisted in an increase of lateral root number after 6 weeks of stress (Figure 5). This promotion of lateral root initiation could be putatively and partly explained by an increase of B concentration in roots that could also promote root hair formation, N_2 fixation, and ion uptake (Gupta and Solanki, 2013; Shireen et al., 2018). Because the limit between plant deficiency and excess for B has been reported to be quite low, the small increase in B concentration observed could limit toxic effect of this element (Dell and Huang, 1997). However, contrary to what has been previously observed in the literature, the increase in the concentration of B did not seem to be beneficial for root dry weight or elongation in pea roots (Hodecker et al., 2014; Shireen et al., 2018), whereas changes in root architecture could be beneficial for soil water and nutrient uptake by other plants (Vilches-Barro and Maizel, 2015). It will thus be necessary to determine the importance of B concentration to the root architecture in pea plants submitted to water stress under different B concentrations.

On the other hand, for nodules, a similar pattern to that observed under recurrent stress was highlighted under long-term continuous stress: although the initiation of nodules was promoted after 4 weeks of stress, it declined after 6 weeks. Stress memory and acclimation to water stress both seem to induce a delay of nodulation, as already observed during the recovery period following drought in pea (Couchoud et al., 2020). From a functional point of view, V concentrations were decreased under continuous stress. Yet, V is a beneficial element for nitrogenase activity that could be decreased under continuous stress (Imtiaz et al., 2015; Altaf et al., 2020). Finally, the acclimation of pea plants after 6 weeks of water stress reduced plant growth, with changes in root system architecture leading to extended soil exploration. It could be interesting to characterize the impact of V concentration under water stress on N_2 fixation, to confirm the potential negative impact of the decrease in the concentration of V on the functional component of nodulation.

5 Conclusion

Our results allowed the identification of different plant strategies related to acclimation or stress memory according to the type of water stress encountered by the plant during its growth cycle. The type of stress can be characterized in leaves or roots by ionomic imprint of each stress. The increased concentration of Mn in leaves seems to be a common response of pea plants to water stress, whatever the stage of the plant growth cycle or type of stress encountered. Moreover, under single transient reproductive stress, Na and Mn concentrations in leaves may be involved in the maintenance of shoot growth. On the other hand, the modification of root system seems to be a typical response to considered water stress. During acclimation to long-term water stress, soil prospection was promoted *via* a higher lateral root number that could also be related to high levels of B in this organ. Finally, under recurrent stress, the recovery period between the two stress periods was beneficial for N_2 fixation through the increase of nodule initiation that could be thus considered as an ecophysiological imprint of a first stress period.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

CS, SP, and MP conceived the project. CJ, SG, FL, and MP performed the experiments. CJ analyzed the data. All authors contributed to the article and approved the submitted version.

Funding

This work was conducted through the EAUPTIC project supported by the "Fond Unique Interministériel" (3870401/1), BPI France (0097244/00), the Regional Council of Burgundy (0133465/00), Dijon Metropole (2018–118–20180820), and the Fonds Européen de Développement Régional (2018–6200FEO003S01889). CJ obtained a Ph.D. grant from Université de Bourgogne.

Acknowledgments

We thank the greenhouse team for taking care of the plants during their growth on the 4PMI platform. We acknowledge Vincent Durey and Marion Forest for their excellent experiment monitoring and sample preparation, the PLATIN' (Plateau d'Isotopie de Normandie) core facility for the element analysis, and more particularly Sylain Diquélou, Frédérine Marie, and Jéhane Lepoittevin.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1089720/full#supplementary-material

References

Abid, M., Hakeem, A., Shao, Y., Liu, Y., Zahoor, R., Fan, Y., et al. (2018). Seed osmopriming invokes stress memory against post-germinative drought stress in wheat (triticum aestivum l.). *Environ. Exp. Bot.* 145, 12–20. doi: 10.1016/j.envexpbot.2017.10.002

Abid, M., Shao, Y., Liu, S., Wang, F., Gao, J., Jiang, D., et al. (2017). Pre-drought priming sustains grain development under post-anthesis drought stress by regulating the growth hormones in winter wheat (Triticum aestivum l.). *Planta* 246, 509–524. doi: 10.1007/s00425-017-2698-4

Abid, M., Tian, Z., Ata-Ul-Karim, S. T., Liu, Y., Cui, Y., Zahoor, R., et al. (2016). Improved tolerance to post-anthesis drought stress by pre-drought priming at vegetative stages in drought-tolerant and -sensitive wheat cultivars. *Plant Physiol. Biochem.* 106, 218–227. doi: 10.1016/j.plaphy.2016.05.003

Ahanger, M. A., Morad-Talab, N., Abd-Allah, E. F., Ahmad, P., and Hajiboland, R. (2016). "Plant growth under drought stress: Significance of mineral nutrients," in *Water stress and crop plants*. Ed. P. Ahmad (Chichester, UK: John Wiley & Sons, Ltd), 649–668. doi: 10.1002/9781119054450.ch37

Aihemaiti, A., Jiang, J., Gao, Y., Meng, Y., Zou, Q., Yang, M., et al. (2019). The effect of vanadium on essential element uptake of setaria viridis' seedlings. *J. Environ. Manage.* 237, 399–407. doi: 10.1016/j.jenvman.2019.02.054

Albrecht, S. L., Maier, R. J., Hanus, F. J., Russell, S. A., Emerich, D. W., and Evans, H. J. (1979). Hydrogenase in *Rhizobium japonicum* increases nitrogen fixation by nodulated soybeans. *Science* 203, 1255–1257. doi: 10.1126/science.203.4386.1255

Altaf, M. M., Diao, X., ur Rehman, A., Imtiaz, M., Shakoor, A., Altaf, M. A., et al. (2020). Effect of vanadium on growth, photosynthesis, reactive oxygen species, antioxidant enzymes, and cell death of rice. *J. Soil Sci. Plant Nutr.* 20, 2643–2656. doi: 10.1007/s42729-020-00330-x

Alves, A. A. C., and Setter, T. L. (2004). Response of cassava leaf area expansion to water deficit: Cell proliferation, cell expansion and delayed development. *Ann. Bot.* 94, 605–613. doi: 10.1093/aob/mch179

Auler, P. A., do Amaral, M. N., Rodrigues, G., dos, S., Benitez, L. C., da Maia, L. C., et al. (2017). Molecular responses to recurrent drought in two contrasting rice genotypes. *Planta* 246, 899–914. doi: 10.1007/s00425-017-2736-2

Auler, P. A., Souza, G. M., da Silva Engela, M. R. G., do Amaral, M. N., Rossatto, T., da Silva, M. G. Z., et al. (2021). Stress memory of physiological, biochemical and metabolomic responses in two different rice genotypes under drought stress: The scale matters. *Plant Sci.* 311, 110994. doi: 10.1016/j.plantsci.2021.110994

Austin, A. T., Yahdjian, L., Stark, J. M., Belnap, J., Porporato, A., Norton, U., et al. (2004). Water pulses and biogeochemical cycles in arid and semiarid ecosystems. *Oecologia* 141, 221–235. doi: 10.1007/s00442-004-1519-1

Bacon, M. A. (1999). The biochemical control of leaf expansion during drought. *Plant Growth Regul.* 29, 101–112. doi: 10.1023/A:1006299514732

Barzana, G., Rios, J. J., Lopez-Zaplana, A., Nicolas-Espinosa, J., Yepes-Molina, L., Garcia-Ibañez, P., et al. (2021). Interrelations of nutrient and water transporters in plants under abiotic stress. *Physiol. Plantarum* 171, 595–619. doi: 10.1111/ppl.13206

Basu, S., Ramegowda, V., Kumar, A., and Pereira, A. (2016). Plant adaptation to drought stress. *F1000Res* 5, 1554. doi: 10.12688/f1000research.7678.1

Baxter, I. R., Vitek, O., Lahner, B., Muthukumar, B., Borghi, M., Morrissey, J., et al. (2008). The leaf ionome as a multivariable system to detect a plant's physiological status. *Proc. Natl. Acad. Sci.* 105, 12081–12086. doi: 10.1073/pnas.0804175105

Bloom, A. J., and Lancaster, K. M. (2018). Manganese binding to rubisco could drive a photorespiratory pathway that increases the energy efficiency of photosynthesis. *Nat. Plants* 4, 414–422. doi: 10.1038/s41477-018-0191-0

Bohnert, H. J., and Sheveleva, E. (1998). Plant stress adaptations — making metabolism move. *Curr. Opin. Plant Biol.* 1, 267–274. doi: 10.1016/S1369-5266(98)80115-5

Bota, J., Medrano, H., and Flexas, J. (2004). Is photosynthesis limited by decreased rubisco activity and RuBP content under progressive water stress? *New Phytol.* 162, 671–681. doi: 10.1111/j.1469-8137.2004.01056.x

Bridges, S. M., and Salin, M. L. (1981). Distribution of iron-containing superoxide dismutase in vascular plants. *Plant Physiol.* 68, 275–278. doi: 10.1104/pp.68.2.275

Brito, B., Palacios, J. M., Hidalgo, E., Imperial, J., and Ruiz-Argüeso, T. (1994). Nickel availability to pea (Pisum sativum l.) plants limits hydrogenase activity of rhizobium leguminosarum bv. viciae bacteroids by affecting the processing of the hydrogenase structural subunits. *J. Bacteriol.* 176, 5297–5303. doi: 10.1128/jb.176.17.5297-5303.1994

Bruce, T. J. A., Matthes, M. C., Napier, J. A., and Pickett, J. A. (2007). Stressful "memories" of plants: Evidence and possible mechanisms. *Plant Sci.* 173, 603–608. doi: 10.1016/j.plantsci.2007.09.002

Brunner, I., Herzog, C., Dawes, M. A., Arend, M., and Sperisen, C. (2015). How tree roots respond to drought. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00547

Cammack, R. (1995). Splitting molecular hydrogen. *Nature* 373, 556–557. doi: 10.1038/373556a0

Chai, Q., Gan, Y., Zhao, C., Xu, H.-L., Waskom, R. M., Niu, Y., et al. (2016). Regulated deficit irrigation for crop production under drought stress. a review. *Agron. Sustain. Dev.* 36, 1–21. doi: 10.1007/s13593-015-0338-6

Chapman, N., Miller, A. J., Lindsey, K., and Whalley, W. R. (2012). Roots, water, and nutrient acquisition: let's get physical. *Trends Plant Sci.* 17, 701–710. doi: 10.1016/j.tplants.2012.08.001

Chaves, M. M. (1991). Effects of water deficits on carbon assimilation. *J. Exp. Bot.* 42, 1–16. doi: 10.1093/jxb/42.1.1

Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P., Osório, M. L., et al. (2002). How plants cope with water stress in the field? photosynthesis and growth. *Ann. Bot.* 89, 907–916. doi: 10.1093/aob/mcf105

Chen, L., Liu, J., Hu, W., Gao, J., and Yang, J. (2021). Vanadium in soil-plant system: Source, fate, toxicity, and bioremediation. *J. Hazard. Mater.* 405, 124200. doi: 10.1016/j.jhazmat.2020.124200

Chen, D., Wang, S., Cao, B., Cao, D., Leng, G., Li, H., et al. (2016). Genotypic variation in growth and physiological response to drought stress and re-watering reveals the critical role of recovery in drought adaptation in maize seedlings. *Front. Plant Sci.* 6, 1241. doi: 10.3389/fpls.2015.01241

Conforti, P., Ahmed, S., and Markova, G. (2018). *Impact of disasters and crises on agriculture and food security, 2017*. FAO

Couchoud, M., Salon, C., Girodet, S., Jeudy, C., Vernoud, V., and Prudent, M. (2020). Pea efficiency of post-drought recovery relies on the strategy to fine-tune nitrogen nutrition. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00204

Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., and Pogson, B. J. (2016). Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. *Sci. Adv.* 2, e1501340. doi: 10.1126/sciadv.1501340

Dell, B., and Huang, L. (1997). Physiological response of plants to low boron. *Plant Soil* 193, 103–120. doi: 10.1023/A:1004264009230

De Souza, P. I., Egli, D. B., and Bruening, W. P. (1997). Water stress during seed filling and leaf senescence in soybean. *Agron. J.* 89, 807–812. doi: 10.2134/agronj1997.00021962008900050015x

Ding, Y., Liu, N., Virlouvet, L., Riethoven, J., Fromm, M., and Avramova, Z. (2013). Four distinct types of dehydration stress memoy genes in Arabidopsis thaliana. *BMC Plant Biol.* 13, 229. doi: 10.1186/1471-2229-13-229

Ding, Y., Virlouvet, L., Liu, N., Riethoven, J.-J., Fromm, M., and Avramova, Z. (2014). Dehydration stress memory genes of zea mays; comparison with arabidopsis thaliana. *BMC Plant Biol.* 14, 141. doi: 10.1186/1471-2229-14-141

Ellis, C., Karafyllidis, I., Wasternack, C., and Turner, J. G. (2002). The arabidopsis mutant cev1 links cell wall signaling to jasmonate and ethylene responses. *Plant Cell* 14, 1557–1566. doi: 10.1105/tpc.002022

Farquhar, G. D., Hubick, K. T., Condon, A. G., and Richards, R. A. (1989). "Carbon isotope fractionation and plant water-use efficiency," in *Stable isotopes in ecological research*. Eds. P. W. Rundel, J. R. Ehleringer and K. A. Nagy (New York, NY: Springer New York), 21–40. doi: 10.1007/978-1-4612-3498-2_2

Flexas, J., Bota, J., Loreto, F., Cornic, G., and Sharkey, T. D. (2004). Diffusive and metabolic limitations to photosynthesis under drought and salinity in c_3 plants. *Plant Biol.* 6, 269–279. doi: 10.1055/s-2004-820867

Freitas, D. S., Rodak, B. W., Carneiro, M. A. C., and Guilherme, L. R. G. (2019). How does Ni fertilization affect a responsive soybean genotype? a dose study. *Plant Soil* 441, 567–586. doi: 10.1007/s11104-019-04146-2

García-Jiménez, A., Trejo-Téllez, L. I., Guillén-Sánchez, D., and Gómez-Merino, F. C. (2018). Vanadium stimulates pepper plant growth and flowering, increases concentrations of amino acids, sugars and chlorophylls, and modifies nutrient concentrations. *PloS One* 13, e0201908. doi: 10.1371/journal.pone.0201908

Ghorbani, P., Eshghi, S., Ershadi, A., Shekafandeh, A., and Razzaghi, F. (2019). The possible role of foliar application of manganese sulfate on mitigating adverse effects of water stress in grapevine. *Commun. Soil Sci. Plant Anal.* 50, 1550–1562. doi: 10.1080/00103624.2019.1626873

Goussias, C., Boussac, A., and Rutherford, A. W. (2002). Photosystem II and photosynthetic oxidation of water: an overview. *Phil. Trans. R. Soc Lond. B* 357, 1369–1381. doi: 10.1098/rstb.2002.1134

Gregory, P. J., Hester, R., and Harrison, R. (2012). "Soils and food security: challenges and opportunities," in *Soils and food security* (Cambridge:RSC Publishing London), 1–30.

Gupta, U., and Solanki, H. (2013). Impact of boron deficiency on plant growth. *Int. J. Bioassays* 2, 1048–1050.

Halvin, J. L., Beaton, J. D., and Tisdale, S. L. (1999). *Soil fertility and fertilizersp. 499, edition* (NJ, USA: Prentice Hall, Upper Saddle River).

Hänsel, S., Ustrnul, Z., Łupikasza, E., and Skalak, P. (2019). Assessing seasonal drought variations and trends over central Europe. *Adv. Water Resour.* 127, 53–75. doi: 10.1016/j.advwatres.2019.03.005

Harb, A., Krishnan, A., Ambavaram, M. M. R., and Pereira, A. (2010). Molecular and physiological analysis of drought stress in arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiol.* 154, 1254–1271. doi: 10.1104/pp.110.161752

He, M., and Dijkstra, F. A. (2014). Drought effect on plant nitrogen and phosphorus: a meta-analysis. *New Phytol.* 204, 924–931. doi: 10.1111/nph.12952

Hemantaranjan, A., and Garg, O. K. (1986). Introduction of nitrogen-fixing nodules through iron and zinc fertilization in the nonnodule-forming French bean (phaseolus vulgaris l.). *J. Plant Nutr.* 9, 281–288. doi: 10.1080/01904168609363444

Hilker, M., and Schmülling, T. (2019). Stress priming, memory, and signalling in plants: Stress priming, memory, and signalling in plants. *Plant Cell Environ.* 42, 753–761. doi: 10.1111/pce.13526

Hodecker, B. E. R., De Barros, N. F., Da Silva, I. R., Diola, V., Sarkis, J. E. S., and Loureiro, M. E. (2014). Boron delays dehydration and stimulates root growth in eucalyptus urophylla (Blake, S.T.) under osmotic stress. *Plant Soil* 384, 185–199. doi: 10.1007/s11104-014-2196-4

Houtz, R. L., Nable, R. O., and Cheniae, G. M. (1988). Evidence for effects on the *in vivo* activity of ribulose-bisphosphate Carboxylase/Oxygenase during development of Mn toxicity in tobacco. *Plant Physiol.* 86, 1143–1149. doi: 10.1104/pp.86.4.1143

Hsiao, T. C. (1973). Plant responses to water stress. *Annu. Rev. Plant Physiol.* 24, 519–570. doi: 10.1146/annurev.pp.24.060173.002511

Imtiaz, M., Rizwan, M. S., Xiong, S., Li, H., Ashraf, M., Shahzad, S. M., et al. (2015). Vanadium, recent advancements and research prospects: A review. *Environ. Int.* 80, 79–88. doi: 10.1016/j.envint.2015.03.018

Jacques, C., Salon, C., Barnard, R. L., Vernoud, V., and Prudent, M. (2021). Drought stress memory at the plant cycle level: A review. *Plants* 10, 1873. doi: 10.3390/plants10091873

Jagadish, K. S. V., Kavi Kishor, P. B., Bahuguna, R. N., von Wirén, N., and Sreenivasulu, N. (2015). Staying alive or going to die during terminal senescence–an enigma surrounding yield stability. *Front. Plant Sci.* 6, 1070. doi: 10.3389/fpls.2015.01070

Jones, L., and McQueen-Mason, S. (2004). A role for expansins in dehydration and rehydration of the resurrection plant craterostigma plantagineum. *FEBS Lett.* 559, 61–65. doi: 10.1016/S0014-5793(04)00023-7

Kaczor, C. M., Smith, M. W., Sangwan, I., and O'Brian, M. R. (1994). Plant [delta]-aminolevulinic acid dehydratase (Expression in soybean root nodules and evidence for a bacterial lineage of the alad gene). *Plant Physiol.* 104, 1411–1417. doi: 10.1104/pp.104.4.1411

Karim, M., Zhang, Y., Zhao, R., Chen, X., Zhang, F., and Zou, C. (2012). Alleviation of drought stress in winter wheat by late foliar application of zinc, boron, and manganese. *Z. Pflanzenernähr. Bodenk.* 175, 142–151. doi: 10.1002/jpln.201100141

King, C. A., and Purcell, L. C. (2001). Soybean nodule size and relationship to nitrogen fixation response to water deficit. *Crop Sci.* 41, 1099–1107. doi: 10.2135/cropsci2001.4141099x

Kirkby, E. (2012). "Introduction, definition and classification of nutrients," in *Marschner's mineral nutrition of higher plants* (San Diego:Academic Press), 3–5. doi: 10.1016/B978-0-12-384905-2.00001-7

Kramer, P. J., and Boyer, J. S. (1995). *Water relations of plants and soils* (San Diego:Academic press).

Lahuta, L. B., Szablińska-Piernik, J., and Horbowicz, M. (2022). Changes in metabolic profiles of pea (Pisum sativum l.) as a result of repeated short-term soil drought and subsequent re-watering. *IJMS* 23, 1704. doi: 10.3390/ijms23031704

Larrainzar, E., Wienkoop, S., Scherling, C., Kempa, S., Ladrera, R., Arrese-Igor, C., et al. (2009). Carbon metabolism and bacteroid functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. *MPMI* 22, 1565–1576. doi: 10.1094/MPMI-22-12-1565

Lehr, J. J., and Wybenga, J. M. (1955). Exploratory pot experiments on sensitiveness of different crops to sodium: C. flax. *Plant Soil* 6, 251–261. doi: 10.1007/BF01464240

Leufen, G., Noga, G., and Hunsche, M. (2016). Drought stress memory in sugar beet: Mismatch between biochemical and physiological parameters. *J. Plant Growth Regul.* 35, 680–689. doi: 10.1007/s00344-016-9571-8

Lichtenthaler, H. K. (1996). Vegetation stress: an introduction to the stress concept in plants. *J. Plant Physiol.* 148, 4–14. doi: 10.1016/S0176-1617(96)80287-2

Lipiec, J., Doussan, C., Nosalewicz, A., and Kondracka, K. (2013). Effect of drought and heat stresses on plant growth and yield: a review. *Int. Agrophysics* 27, 463–477. doi: 10.2478/intag-2013-0017

Liu, F., Shahnazari, A., Andersen, M. N., Jacobsen, S.-E., and Jensen, C. R. (2006). Physiological responses of potato (Solanum tuberosum l.) to partial root-zone drying: ABA signalling, leaf gas exchange, and water use efficiency. *J. Exp. Bot.* 57, 3727–3735. doi: 10.1093/jxb/erl131

Lurthy, T., Cantat, C., Jeudy, C., Declerck, P., Gallardo, K., Barraud, C., et al. (2020). Impact of bacterial siderophores on iron status and ionome in pea. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00730

Mahieu, S., Germon, F., Aveline, A., Hauggaard-Nielsen, H., Ambus, P., and Jensen, E. S. (2009). The influence of water stress on biomass and n accumulation, n partitioning between above and below ground parts and on n rhizodeposition during reproductive growth of pea (Pisum sativum l.). *Soil Biol. Biochem.* 41, 380–387. doi: 10.1016/j.soilbio.2008.11.021

Marcos, F. C. C., Silveira, N. M., Marchiori, P. E. R., Machado, E. C., Souza, G. M., Landell, M. G. A., et al. (2018). Drought tolerance of sugarcane propagules is improved when origin material faces water deficit. *PloS One* 13, e0206716. doi: 10.1371/journal.pone.0206716

Milford, G. F. J., Cormack, W. F., and Durrant, M. J. (1977). Effects of sodium chloride on water status and growth of sugar beet. *J. Exp. Bot.* 28, 1380–1388. doi: 10.1093/jxb/28.6.1380

Millaleo, R., Reyes- Diaz, M., Ivanov, A. G., Mora, M. L., and Alberdi, M. (2010). Manganese as essential and toxic element for plants: transport, accumulation and resistance mechanisms. *J. Soil Sci. Plant Nutr.* 10, 470–481. doi: 10.4067/S0718-95162010000200008

Moore, J. P., Vicré-Gibouin, M., Farrant, J. M., and Driouich, A. (2008). Adaptations of higher plant cell walls to water loss: drought vs desiccation. *Physiol. Plantarum* 134, 237–245. doi: 10.1111/j.1399-3054.2008.01134.x

Ness, P. J., and Woolhouse, H. W. (1980). RNA Synthesis in *Phaseolus* chloroplasts: I. ribonucleic acid synthesis in chloroplast preparation from *Phaseolus vulgaris* l. leaves and solubilization of the RNA polymerase. *J. Exp. Bot.* 31, 223–233. doi: 10.1093/jxb/31.1.223

Neumann, P. M. (1995). The role of cell wall adjustments in plant resistance to water deficits. *Crop Sci.* 35, 1258–1266. doi: 10.2135/cropsci1995.0011183X003500050002x

O'Hara, G. W., Hartzook, A., Bell, R. W., and Loneragan, J. F. (1993). Differences between bradyrhizobium strains NC92 and TAL1000 in their nodulation and nitrogen fixation with peanut in iron deficient soil. *Plant Soil* 155–156, 333–336. doi: 10.1007/BF00025050

Orcutt, D. M., Nilsen, E. T., and Hale, M. G. (2000). *The physiology of plants under stress* (Chichester:John Wiley & Sons).

Parry, M. A. J. (2002). Rubisco activity: Effects of drought stress. *Ann. Bot.* 89, 833–839. doi: 10.1093/aob/mcf103

Pessarakli, M., Haghighi, M., and Sheibanirad, A. (2015). Plant responses under environmental stress conditions. *Adv. Plants Agric. Res.* 2, 73. doi: 10.15406/apar.2015.02.00073

Pfeffer, P. E., Tu, S.-I., Gerasimowicz, W. V., and Cavanaugh, J. R. (1986). *In vivo* ^{31}$ p NMR studies of corn root tissue and its uptake of toxic metals. *Plant Physiol.* 80, 77–84. doi: 10.1104/pp.80.1.77

Pilon, M., Cohu, C. M., Ravet, K., Abdel-Ghany, S. E., and Gaymard, F. (2009). Essential transition metal homeostasis in plants. *Curr. Opin. Plant Biol.* 12, 347–357. doi: 10.1016/j.pbi.2009.04.011

Pirasteh-Anosheh, H., Saed-Moucheshi, A., Pakniyat, H., and Pessarakli, M. (2016). "Stomatal responses to drought stress," in *Water stress and crop plants*. Ed. P. Ahmad (Chichester, UK: John Wiley & Sons, Ltd), 24–40. doi: 10.1002/9781119054450.ch3

Polle, A., Chakrabarti, K., Chakrabarti, S., Seifert, F., Schramel, P., and Rennenberg, H. (1992). Antioxidants and manganese deficiency in needles of Norway spruce (*Picea abies* l.) trees. *Plant Physiol.* 99, 1084–1089. doi: 10.1104/pp.99.3.1084

Prasad, P. V. V., Staggenborg, S. A., Ristic, Z., Ahuja, L. R., Reddy, V. R., Saseendran, S. A., et al. (2008). "Impacts of drought and/or heat stress on physiological, developmental, growth, and yield processes of crop plants," in *Advances in agricultural systems modeling* (American Society of Agronomy, Crop Science Society of America, Soil Science Society of America), 301–355. doi: 10.2134/advagricsystmodel1.c11

Prudent, M., Vernoud, V., Girodet, S., and Salon, C. (2016). How nitrogen fixation is modulated in response to different water availability levels and during recovery: A structural and functional study at the whole plant level. *Plant Soil* 399, 1–12. doi: 10.1007/s11104-015-2674-3

Purcell, L. C., King, C. A., and Ball, R. A. (2000). Soybean cultivar differences in ureides and the relationship to drought tolerant nitrogen fixation and manganese nutrition. *Crop Sci.* 40, 1062–1070. doi: 10.2135/cropsci2000.4041062x

Ramírez, D. A., Rolando, J. L., Yactayo, W., Monneveux, P., Mares, V., and Quiroz, R. (2015). Improving potato drought tolerance through the induction of long-term water stress memory. *Plant Sci.* 238, 26–32. doi: 10.1016/j.plantsci.2015.05.016

Randall, H. C., and Sinclair, T. R. (1988). Sensitivity of soybean leaf development to water deficits. *Plant Cell Environ.* 11, 835–839. doi: 10.1111/j.1365-3040.1988.tb01909.x

Rehman, A., Masood, S., Khan, N. U., Abbasi, M. E., Hussain, Z., and Ali, I. (2021). Molecular basis of iron biofortification in crop plants; a step towards sustainability. *Plant Breed* 140, 12–22. doi: 10.1111/pbr.12886

Rotaru, V., and Sinclair, T. R. (2009). Influence of plant phosphorus and iron concentrations on growth of soybean. *J. Plant Nutr.* 32, 1513–1526. doi: 10.1080/01904160903093828

Salon, C., Munier-Jolain, N. G., Duc, G., Voisin, A.-S., Grandgirard, D., Larmure, A., et al. (2001). Grain legume seed filling in relation to nitrogen acquisition: A review and prospects with particular reference to pea. *Agronomie* 21, 539–552. doi: 10.1051/agro:2001143

Salt, D. E., Baxter, I., and Lahner, B. (2008). Ionomics and the study of the plant ionome. *Annu. Rev. Plant Biol.* 59, 709–733. doi: 10.1146/annurev.arplant.59.032607.092942

Sanchez, D. H., Pieckenstain, F. L., Escaray, F., Erban, A., Kraemer, U., Udvardi, M. K., et al. (2011). Comparative ionomics and metabolomics in extremophile and glycophytic lotus species under salt stress challenge the metabolic pre-adaptation hypothesis: Ionomic and metabolomic phenotypes of l. creticus under salt stress. *Plant Cell Environ.* 34, 605–617. doi: 10.1111/j.1365-3040.2010.02266.x

Sánchez-Rodríguez, E., del Mar Rubio-Wilhelmi, M., Cervilla, L. M., Blasco, B., Rios, J. J., Leyva, R., et al. (2010). Study of the ionome and uptake fluxes in cherry tomato plants under moderate water stress conditions. *Plant Soil* 335, 339–347. doi: 10.1007/s11104-010-0422-2

Sandmann, G., and Böger, P. (1983). Comparison of the bleaching activity of norflurazon and oxyfluorfen. *Weed Sci.* 31, 338–341. doi: 10.1017/S0043174500069125

Schmidt, S. B., Jensen, P. E., and Husted, S. (2016). Manganese deficiency in plants: The impact on photosystem II. *Trends Plant Sci.* 21, 622–632. doi: 10.1016/j.tplants.2016.03.001

Seliga, H. (1993). "The role of copper in nitrogen fixation in lupinus luteus l," in *Plant nutrition–from genetic engineering to field practice* (Dordrecht:Springer), 459–462. doi: 10.1007/978-94-011-1880-4_96

Seliga, H. (1998). *Nitrogen fixation in several grain legume species with contrasting sensitivities to copper nutrition* (Acta physiologia plantarum), Vol. 20. 263–267. doi: 10.1007/s11738-998-0057-6

Serraj, R., Vadez, V., and Sinclair, T. R. (2001). Feedback regulation of symbiotic n $_2$ fixation under drought stress. *Agronomie* 21, 621–626. doi: 10.1051/agro:2001153

Sevilla, F., López-Gorgé, J., Gómez, M., and del Río, L. A. (1980). Manganese superoxide dismutase from a higher plant: Purification of a new Mn-containing enzyme. *Planta* 150, 153–157. doi: 10.1007/BF00582359

Sharp, R. E., and Davies, W. J. (1989). A restricted supply of water. *Plants Under Stress: Biochem. Physiol. Ecol. Their Appl. To Plant Improvement* 39, 71.

Shireen, F., Nawaz, M., Chen, C., Zhang, Q., Zheng, Z., Sohail, H., et al. (2018). Boron: Functions and approaches to enhance its availability in plants for sustainable agriculture. *IJMS* 19, 1856. doi: 10.3390/ijms19071856

Snowball, K., Robson, A. D., and Loneragan, J. F. (1980). The effect of copper on nitrogen fixation in subterranean clover (trifolium subterraneum). *New Phytol.* 85, 63–72. doi: 10.1111/j.1469-8137.1980.tb04448.x

Sobeih, W. Y., Dodd, I. C., Bacon, M. A., Grierson, D., and Davies, W. J. (2004). Long-distance signals regulating stomatal conductance and leaf growth in tomato (Lycopersicon esculentum) plants subjected to partial root-zone drying. *J. Exp. Bot.* 55, 2353–2363. doi: 10.1093/jxb/erh204

Streeter, J. G. (2003). Effects of drought on nitrogen fixation in soybean root nodules: Effects of drought on root nodules. *Plant Cell Environ.* 26, 1199–1204. doi: 10.1046/j.1365-3040.2003.01041.x

Sun, C., Gao, X., Chen, X., Fu, J., and Zhang, Y. (2016). Metabolic and growth responses of maize to successive drought and re-watering cycles. *Agric. Water Manage.* 172, 62–73. doi: 10.1016/j.agwat.2016.04.016

Su, C. T., Tang, J. S., Shan, H. P., Chen, H. F., Xia, R. Y., and Zou, S. Z. (2008). Study on the geochemical properties of fe-Mn nodule in litang karst area. *Carsologica Sin.* 1, 43–49.

Tang, C. (2001). Phosphorus deficiency impairs early nodule functioning and enhances proton release in roots of medicago truncatula l. *Ann. Bot.* 88, 131–138. doi: 10.1006/anbo.2001.1440

Tang, C., Robson, A. D., and Dilworth, M. J. (1990). The role of iron in nodulation and nitrogen fixation in lupinus angustifolius l. *New Phytol.* 114, 173–182. doi: 10.1111/j.1469-8137.1990.tb00388.x

Tardieu, F., Reymond, M., Hamard, P., Granier, C., and Muller, B. (2000). Spatial distributions of expansion rate, cell division rate and cell size in maize leaves: a synthesis of the effects of soil water status, evaporative demand and temperature. *J. Exp. Bot.* 51, 1505–1514. doi: 10.1093/jexbot/51.350.1505

Tinker, P. B. H. (1965). The effects of nitrogen, potassium and sodium fertilizers on sugar beet. *J. Agric. Sci.* 65, 207–212. doi: 10.1017/S0021859600083933

Trewavas, A. (2003). Aspects of plant intelligence. *Ann. Bot.* 92, 1–20. doi: 10.1093/aob/mcg101

Tripathi, D. K., Singh, S., Singh, S., Mishra, S., Chauhan, D. K., and Dubey, N. K. (2015). Micronutrients and their diverse role in agricultural crops: advances and future prospective. *Acta Physiol. Plant* 37, 1–14. doi: 10.1007/s11738-015-1870-3

Viets, F. G. J. (1972). Water deficits and nutrient availability. *Water Deficits Plant Growth* 3, 217–240.

Vilches-Barro, A., and Maizel, A. (2015). Talking through walls: mechanisms of lateral root emergence in arabidopsis thaliana. *Curr. Opin. Plant Biol.* 23, 31–38. doi: 10.1016/j.pbi.2014.10.005

Virlouvet, L., Avenson, T. J., Du, Q., Zhang, C., Liu, N., Fromm, M., et al. (2018). Dehydration stress memory: Gene networks linked to physiological responses during repeated stresses of zea mays. *Front. Plant Sci.* 9, 1058. doi: 10.3389/fpls.2018.01058

Voisin, A. S., Salon, C., Jeudy, C., and Warembourg, F. R. (2003). Symbiotic N2 fixation activity in relation to c economy of pisum sativum l. as a function of plant phenology. *J. Exp. Bot.* 54, 2733–2744. doi: 10.1093/jxb/erg290

Wang, Y., Jensen, C. R., and Liu, F. (2017). Nutritional responses to soil drying and rewetting cycles under partial root-zone drying irrigation. *Agric. Water Manage.* 179, 254–259. doi: 10.1016/j.agwat.2016.04.015

Wang, X., Vignjevic, M., Jiang, D., Jacobsen, S., and Wollenweber, B. (2014). Improved tolerance to drought stress after anthesis due to priming before anthesis in wheat (Triticum aestivum l.) var. vinjett. *J. Exp. Bot.* 65, 6441–6456. doi: 10.1093/jxb/eru362

Weemstra, M., Mommer, L., Visser, E. J., van Ruijven, J., Kuyper, T. W., Mohren, G. M., et al. (2016). Towards a multidimensional root trait framework: a tree root review. *New Phytol.* 211, 1159–1169. doi: 10.1111/nph.14003

Wilson, R. S., and Franklin, C. E. (2002). Testing the beneficial acclimation hypothesis. *Trends Ecol. Evol.* 17, 66–70. doi: 10.1016/S0169-5347(01)02384-9

Wingler, A., Lea, P. J., Quick, W. P., and Leegood, R. C. (2000). Photorespiration: metabolic pathways and their role in stress protection. *Philos. Trans. R. Soc. London Ser. B: Biol. Sci.* 355, 1517–1529. doi: 10.1098/rstb.2000.0712

Wingler, A., Quick, W. P., Bungard, R. A., Bailey, K. J., Lea, P. J., and Leegood, R. C. (2002). The role of photorespiration during drought stress: an analysis utilizing barley mutants with reduced activities of photorespiratory enzymes. *Plant Cell Environ.* 22, 361–373. doi: 10.1046/j.1365-3040.1999.00410.x

Wu, Y., and Cosgrove, D. J. (2000). Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. *J. Exp. Bot.* 51, 1543–1553. doi: 10.1093/jexbot/51.350.1543

Yaqoobbr, M., Hollingtonbr, P., and Gurmani, A. (2013). Yield performance and responses studies of chickpea (Cicer arietinum l.) genotypes under drought stress. *Emir. J. Food Agric.* 25, 117–123. doi: 10.9755/ejfa.v25i2.10655

Yousfi, N., Saïdi, I., Slama, I., and Abdelly, C. (2015). Phenology, leaf gas exchange, growth and seed yield in medicago polymorpha l. populations affected by water deficit and subsequent recovery. *Flora - Morphol. Distribution Funct. Ecol. Plants* 214, 50–60. doi: 10.1016/j.flora.2015.05.007

Zahran, H. H. (1999). *Rhizobium* -legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* 63, 968–989. doi: 10.1128/MMBR.63.4.968-989.1999

Zhou, Y., Lam, H. M., and Zhang, J. (2007). Inhibition of photosynthesis and energy dissipation induced by water and high light stresses in rice. *J. Exp. Bot.* 58, 1207–1217. doi: 10.1093/jxb/erl291

frontiers | Frontiers in Plant Science

TYPE Mini Review
PUBLISHED 30 January 2023
DOI 10.3389/fpls.2023.1120342

Check for updates

OPEN ACCESS

EDITED BY
Jose C. Jimenez-Lopez,
Department of Biochemistry, Cell & Molecular Biology of Plants,
Spanish National Research Council (CSIC),
Spain

REVIEWED BY
Hokuto Nakayama,
The University of Tokyo, Japan

*CORRESPONDENCE
Edelín Roque
edroque@ibmcp.upv.es
Luis A. Cañas
lcanas@ibmcp.upv.es

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 09 December 2022
ACCEPTED 16 January 2023
PUBLISHED 30 January 2023

CITATION
Rodas AL, Roque E, Hamza R, Gómez-Mena C, Beltrán JP and Cañas LA (2023) *SUPERMAN* strikes again in legumes. *Front. Plant Sci.* 14:1120342. doi: 10.3389/fpls.2023.1120342

COPYRIGHT
© 2023 Rodas, Roque, Hamza, Gómez-Mena, Beltrán and Cañas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

SUPERMAN strikes again in legumes

Ana L. Rodas, Edelín Roque*, Rim Hamza, Concepción Gómez-Mena, José Pío Beltrán and Luis A. Cañas*

Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia), Ciudad Politécnica de la Innovación, Valencia, Spain

The *SUPERMAN* (*SUP*) gene was described in *Arabidopsis thaliana* over 30 years ago. *SUP* was classified as a cadastral gene required to maintain the boundaries between reproductive organs, thus controlling stamen and carpel number in flowers. We summarize the information on the characterization of *SUP* orthologs in plant species other than Arabidopsis, focusing on the findings for the *MtSUP*, the ortholog in the legume *Medicago truncatula*. *M. truncatula* has been widely used as a model system to study the distinctive developmental traits of this family of plants, such as the existence of compound inflorescence and complex floral development. *MtSUP* participates in the complex genetic network controlling these developmental processes in legumes, sharing conserved functions with *SUP*. However, transcriptional divergence between *SUP* and *MtSUP* provided context-specific novel functions for a *SUPERMAN* ortholog in a legume species. *MtSUP* controls the number of flowers per inflorescence and the number of petals, stamens and carpels regulating the determinacy of ephemeral meristems that are unique in legumes. Results obtained in *M. truncatula* provided new insights to the knowledge of compound inflorescence and flower development in legumes. Since legumes are valuable crop species worldwide, with high nutritional value and important roles in sustainable agriculture and food security, new information on the genetic control of their compound inflorescence and floral development could be used for plant breeding.

KEYWORDS

SUPERMAN, legumes, Medicago truncatula, MtSUP, compound inflorescence, flower development, flower number

Introduction

Most angiosperm flowers are organized in four concentric whorls: sepals (W1), petals (W2), stamens (W3) and carpel/s (W4), (Smyth et al., 1990). The number of floral organs and the placement of the organs within each whorl are genetically determined and MADS-box floral homeotic genes play a crucial role in the specification of floral organ identity (Bowman et al., 2012). Other classes of genes, the ones that determine the boundaries of different cell identities, are also crucial players during floral development (Yu and Huang, 2016). They were classified as "cadastral genes", to which the Arabidopsis *SUPERMAN* (*SUP*) gene was assigned (Bowman et al., 1992; Sakai et al., 1995).

10.3389/fpls.2023.1120342

SUP is a transcriptional repressor, extensively studied in *Arabidopsis thaliana*, that encodes a plant-specific EPF-like protein with one Cys_2-His_2 zinc finger DNA binding domain and a C-terminus EAR-like (DLELRL) motif (Sakai et al., 1995; Hiratsu et al., 2002; Hiratsu et al., 2003; Hiratsu et al., 2004). The specific expression of *SUP* at the boundary between W3 and W4 (Sakai et al., 1995; Prunet et al., 2017) led to its classification as a cadastral gene specifying the stamens-carpel boundary. The supernumerary male organs at the expense of the female one of the *sup-1* mutant (Figure 1A) (Bowman et al., 1992) was initially associated with the expansion of the MADS-box genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) expression closer to the centre of the floral

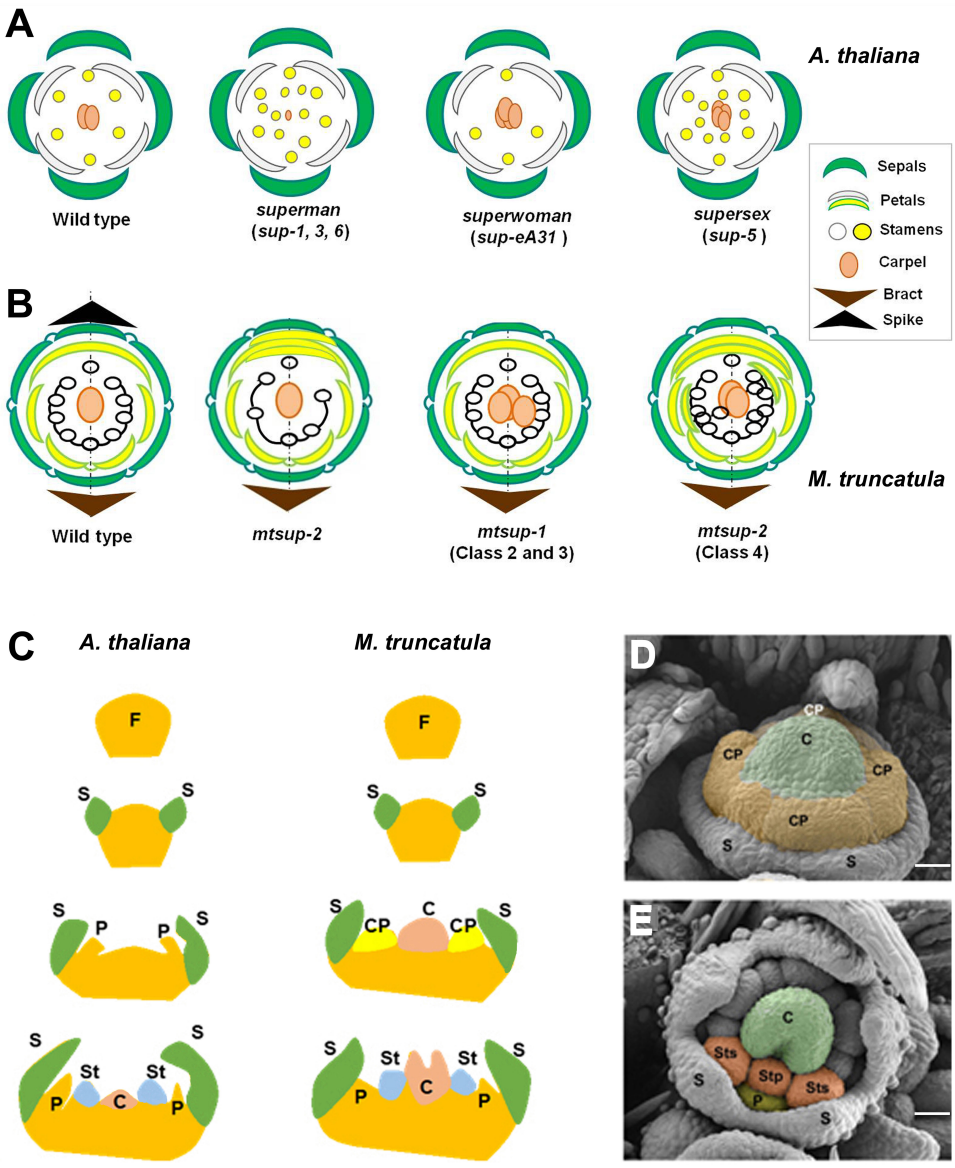


FIGURE 1
Comparative floral development of *Arabidopsis thaliana* and *Medicago truncatula*. **(A)** Comparative schematic representation among *superman* (*sup*) mutant alleles in Arabidopsis regarding floral organ number. Wild *type A. thaliana* flower: 4 sepals, 4 petals, 6 stamens and two fused carpels. **(B)** Comparative schematic representation among *superman* (*mtsup*) mutant alleles in Medicago regarding floral organ number. Wild type *M. truncatula* flower: 5 sepals, 5 petals, 10 stamens, 9 fused (staminal tube) and one free, and a single carpel. In Arabidopsis, the 'superman' class of mutants harbors supernumerary stamens and reduced or absent carpel, resembling *mtsup-2* showing additional petals at the expense of stamens. An increased number of carpels characterize the "superwoman" class. Similar phenotypes displayed the *mtsup-1* (class 2 and 3) alleles, with two or three carpels in *M. truncatula*. The "supersex" class, to which *sup-5* allele belong, produces more stamens and additional carpels. This phenotype is observed in *mtsup-1* allele (class 4). Also, additional petals are produced by this allele. **(C)** Left. In *A. thaliana* organ differentiation is centripetal and sequential. First are differentiated the sepal primordia, then the petal primordia, the stamen primordia and finally the carpel primordium. Right. In *M. truncatula*, the four common primordia differentiate petals and stamens in W2 and W3 respectively. **(D)** Floral meristem of *M. truncatula* showing the early carpel primordium (C, green) in the centre, the four common primordia (CP, orange) and the sepal primordia (S). **(E)** Each common primordium differentiates petals (P, yellow) in W2 and antepetal (Stp) and antesepal (Sts) stamens in W3 (orange). VM, vegetative meristem; FM, floral meristem; I1, primary inflorescence meristem; I2, secondary inflorescence meristem; spk, spike; S, sepal primordium; CP, common primordium; P, petal primordium; St, stamen primordium; C, carpel primordium. Scale bars, 25 µm in **(D, E)** Adapted from Benlloch et al., 2003; Breuil-Broyer et al., 2016; Prunet et al., 2017 and Rodas et al., 2021.

meristem (Sakai et al., 1995; Prunet et al., 2017). Initially, models to explain the *SUP* function were based on a single allele: *sup-1* (*flo-10*) (Bowman et al., 1992; Sakai et al., 1995; Prunet et al., 2017). However, the study of other *sup* alleles displaying phenotypes deviating from *sup-1* (Figure 1A) has shed light on the *SUP* functions (Breuil-Broyer et al., 2016; Prunet et al., 2017; Xu et al., 2018). *SUP* is a gene controlling the stamens-carpel boundary setting and is linked to floral meristem termination (FMT) at the early stages of flower development.

It has been described that *SUP* is required for the correct timing to turn off *WUSCHEL* (*WUS*) from the floral meristem centre (FMC), thus controlling the floral meristem termination. *WUS* activity is required for the stem cell division at the floral meristem centre, a prolonged expression of *WUS* would lead to a delayed floral meristem termination, and more floral organs could be produced. However, *SUP* and *WUS* do not show an overlapping spatial expression and, in *sup* mutants, *WUS* expression is prolonged (Prunet et al., 2017; Xu et al., 2018). Moreover, *SUP* contributes to carpel medial region formation and the tissues derived from this region (Breuil-Broyer et al., 2016).

These studies provided new information to generate different models to explain *SUP* functions. One of the models proposes that *SUP* indirectly promotes floral meristem termination by repressing B-class genes. This model explains the different *sup* alleles phenotypes (Figure 1A), showing an indistinct male-female boundary and a sporadic carpel development (*sup-1*). By contrast, the increased number of stamens and carpels in the *sup-5* mutant (Figure 1A) supports a second model that proposes that *SUP* controls the balance of cell proliferation and differentiation at W3 and W4 (Breuil-Broyer et al., 2016). To this regard, the effect of the overexpression of *SUP*-like genes supports the activity of *SUP* as a cell proliferation control gene (Nandi et al., 2000; Bereterbide et al., 2001; Hiratsu et al., 2002; Kazama et al., 2009; Nibau et al., 2011; Zhao et al., 2014). Recent studies demonstrated that *SUP* regulates both stem cell proliferation in the floral meristem and floral organogenesis through fine-tuning auxin biosynthesis (Xu et al., 2018). This mechanism might explain all *sup* mutant phenotypes (Figure 1A). Studies in *SUP* have shown its broad spectrum of action, highlighting how different are the floral phenotypes according to the type of mutation (Bowman et al, 2012; Breuil-Broyer et al, 2016).

Compound inflorescence and floral development in legumes. Distinctive traits

In addition to the well-known capacity to fix nitrogen symbiotically, some distinctive features of legumes are the presence of compound leaves and inflorescences and a complex floral development (Ferrándiz et al., 1999; Singer et al., 1999; Benlloch et al., 2003). All these traits make them of interest for the study of unique developmental processes (Hofer and Ellis, 2014; Cañas and Beltrán, 2018).

Most legumes show complex raceme inflorescences with more than one branching. In the model legume *Medicago truncatula* the primary inflorescence meristem (I1) differentiates a secondary inflorescence meristem (I2). The existence of the I2 is linked to the compound inflorescence development and is a distinctive feature compared to Arabidopsis, which produces a unique inflorescence meristem (IM) before differentiating the floral meristem (FM) (Tucker, 2003; Benlloch et al., 2007). The I2 is a transient meristem, and its identity is given by a genetic function of *VEGETATIVE1* (*VEG1*) in *Pisum sativum* and *MtFRUITFULLc* (*MtFULc*) in *M. truncatula* (Cheng et al., 2018). It has been proposed that this function was derived from the sub-functionalization of the *AGL79* MADS-box gene clade within the *AP1/SQUA/FUL* family (Berbel et al., 2012). The perpetual activity of the I2 meristem will define the number of flowers per inflorescence and its termination as a residual vegetative organ (stub or spike) in the legume compound inflorescences (Benlloch et al., 2003; Benlloch et al., 2015). In the model legume *M. truncatula*, the identity of the I1 and FM, also involved in this developmental process, are specified by *MtTERMINAL FLOWER1* (*MtTFL1*) *and MtAPETALA1* (*MtAP1*) or *MtPROLIFERATING INFLORESCENCE MERISTEM* (*MtPIM*), respectively (Benlloch et al., 2006; Cheng et al., 2018). Their spatial and temporal expression and mutual repression control the compound inflorescence development in *M. truncatula* (Cheng et al., 2018).

The wild type flower of *M. truncatula* (Figure 1B) displays pentamerous floral organs per whorl: five sepals in W1, five petals in W2 (a keel petal formed by two fused petals, two wing petals and one standard or vexillum), 10 stamens in W3 (nine fused in a staminal tube and one free) and a single carpel in W4 (Benlloch et al., 2003; Cañas and Beltrán, 2018). In contrast to Arabidopsis, organ differentiation shows a high degree of spatial and temporal overlapping. Even more characteristic is the presence of common primordia (CP), ephemeral meristems from which petals and stamens will differentiate, and the early carpel differentiation (Ferrándiz et al., 1999; Tucker, 2003; Benlloch et al., 2003; Roque et al., 2018). The model species *A. thaliana* (Figure 1C left) shows a centripetal and sequential organ differentiation. First, the sepal primordia are differentiated, then the petal primordia, followed by the stamen primordia, and finally, the carpel primordium. In contrast, *M. truncatula* (Figure 1C right) shows unidirectional differentiation of the organ primordia with a high degree of overlapping. Unique differences are the presence of four common primordia and the early carpel primordium differentiation (Figures 1D, E). Despite the functional divergence of the duplicated floral homeotic MADS-box genes in *M. truncatula*, the specification of the floral organs is conserved in this model legume (Roque et al., 2018).

MtSUPERMAN: Conserved and new functions controlling compound inflorescence and floral development

The *SUP* gene has been widely studied in *A. thaliana*. However, there is scant information on the role of *SUP* orthologs in other plant species. The petunia *PhSUP* gene (Nakagawa et al., 2004) had been the only *SUP* ortholog functionally characterized on its own species until it was studied in the model legume *M. truncatula* (Rodas et al., 2021).

Recently, the *SMALL REPRODUCTIVE ORGANS* (*SRO*) gene was described as the *SUP* ortholog in rice (Xu et al., 2022).

The results obtained from the functional characterization of *MtSUP* in Medicago (Rodas et al., 2021) uncovered new context-specific functions in a different plant species. This information may have changed not only the previously proposed idea of *SUPERMAN* as a boundary gene but also made *MtSUP* a key player of the complex regulatory network behind the compound inflorescence development, being an undescribed function for a *SUP* ortholog in eudicots. Nevertheless, there are also similarities between *MtSUP* and other *SUP* orthologs regarding flower development.

The floral phenotypes of *mtsup* mutants (Figure 1B) are different to Arabidopsis, Petunia and rice mutants in several respects. However, *sup*, *phsup* and *mtsup* mutants have in common the increase in the numbers of both stamens and carpels in their respective flowers. Thus, the early floral meristem function of *SUP* is conserved in these three species (Nakagawa et al., 2004; Rodas et al., 2021). However, the rice *SUP* ortholog controls the size of male and female organs but not their number (Xu et al., 2022).

MtSUP transcript is firstly detected in the I2 meristem and later in the FM (Figures 2A, B). The expression pattern of *MtSUP* during floral organogenesis (Rodas et al., 2021) showed that even before the carpel primordium is initiated *MtSUP* transcript is already detected in the floral meristem centre and later in the common primordia (Figure 2C). The proliferation of extra petals was a distinctive feature discovered for the *mtsup* mutants during floral development (Figure 1B). At the common primordia, the meristematic cells that will produce petals and stamens coexist, and a given number of meristematic cells will give place to the organ primordia (Bossinger and Smyth, 1996).

As a role already described for *SUP*, *MtSUP* might control cell proliferation in the common primordia. In other words, *MtSUP* is involved in the determinacy of the common primordia as prolonged maintenance of these meristematic cells can give rise to extra organs (Bowman et al., 1989; Bossinger and Smyth, 1996). The supernumerary petals, stamens and carpels in *mtsup* mutants (Figure 1B) might also be explained by a delayed floral meristem termination linked to *MtWUS* persistence, as also occurs for *WUS* in *sup* mutants of *A. thaliana* (Prunet et al., 2017; Xu et al., 2018). *MtWUS* expression is not detected in the wild type after the floral apex flattens (Rodas et al., 2021). This is consistent with the early carpel initiation in legumes (Ferrándiz et al., 1999), as floral meristem termination happens when the pool of stem cells of the floral meristem centre is set to a female fate (Prunet et al., 2009). In *mtsup-1*, the expression of *MtWUS* is prolonged, thus the pool of stem cells remains undifferentiated during more time at the floral meristem centre. Contrary to *SUP*, which is not expressed in the floral meristem centre and plays a non-cell-autonomous function there (Prunet et al., 2017). *MtSUP* and *MtWUS* expression overlaps, both spatially and temporally at the I2 and the floral meristem centre, which would allow them to interact physically (Rodas et al., 2021). *MtSUP* has a novel function in the common primordia determinacy and seems to conserve its cell antiproliferative role in this unique feature of legumes.

A proper carpel primordium formation requires a correct floral meristem termination (Sakai et al., 2000; Prunet et al., 2017). *MtSUP* is expressed in the carpel marginal tissue that will develop the parietal placenta. It agrees with the defects in the marginal derived tissues of the gynoecium in *mtsup* mutants. Defects in placenta morphogenesis were also observed in the Petunia *phsup* mutants (Nakagawa et al., 2004), and the strong *sup-5* mutant (Figure 1A) of *A. thaliana* (Gaiser et al., 1995). Therefore, the *SUP* orthologs *PhSUP* and *MtSUP* are required for proper floral meristem termination and the correct development of the carpel marginal tissues (Nakagawa et al., 2004; Rodas et al., 2021). Common aberrancies in the development of the placenta impacted ovule development in *mtsup* mutants, reducing fertility. Similar phenotypes were reported for Arabidopsis *sup-5* (Gaiser et al., 1995) and Petunia *phsup* mutants (Nakagawa et al., 2004). Thus, the late floral function of *SUP* controlling ovule development is conserved in these three species (Gaiser et al., 1995; Nakagawa et al., 2004; Breuil-Broyer et al., 2016; Rodas et al., 2021).

Unlike Arabidopsis, *MtSUP* was first detected in the whole I2 (Figures 2A, B), similar to the expression of the I2 identity gene *MtFULc* (Figures 2D, E). This expression matched with the multiflowered phenotype in *mtsup* mutants, assigning *MtSUP* a determinant role in controlling the maturation rate of the I2. This novel function has not been described for any *SUP*-like gene. In *M. truncatula* cv.R108, the I2 derived from the I1 divides to produce one or two floral meristems (Benlloch et al., 2003). The I2 is a transient state between the vegetative and the reproductive tissue that remains immature until the floral identity acquisition (Prusinkiewicz et al., 2007). After producing a floral meristem, the remaining cells of the I2 enter senescence and produce the spike (Tucker, 1989; Benlloch et al., 2007). In *mtsup* mutants, the I2 gives rise to more floral meristems than the wild type because the residual cells of the I2 terminate as a floral meristem instead of a spike (Figures 2F–J). The I2 determinacy could also be linked to the gradual turn-off of *MtWUS* in the I2. As *MtWUS* prolongs its expression, the I2 could extend its activity in *mtsup* mutants, or *MtSUP* could influence the I2 activity by controlling cell proliferation. *MtFULc* transcript occupies a broader area in *mtsup* mutants compared to the wild type (Figures 2D, E), and there might be more cells expressing *MtFULc*. According to the expression analysis, *MtSUP* also seems to restrict *MtPIM* expression to the floral meristem and this restriction could be considered another way to control the determinacy of the I2. In *mtsup* mutants, *MtPIM* invades the expression domain of *MtFULc* in the I2 and the remnant cells that lose their vegetative nature and acquire floral identity (Rodas et al., 2021).

SUPERMAN was classified as a "cadastral gene" after studying the *sup-1 (flo10)* allele (Bowman et al., 1992). However, the results obtained in Medicago might support that the conserved ancestral function of *SUP*-like genes is the control of cell proliferation rather than a cadastral function. *MtSUP* does not have a typical boundary expression pattern since it is expressed in the whole I2 and floral meristem. This transcript localization correlates with the floral and inflorescence phenotypes of *mtsup-1* mutant. Thus, the model that proposes the balance of cell proliferation explains *mtsup* mutants better than the model that considers that *SUP* is related to the repression of B-class MADS-box genes to the floral meristem centre. Indeed, *MtPI* expression (Roque et al., 2018) in *mtsup-1* expands towards the W1 instead of expanding to the floral meristem centre (Rodas et al., 2021). In both flower and inflorescence development, there is no need to invoke a boundary function to explain *mtsup* mutants. Certainly, the phenotypic

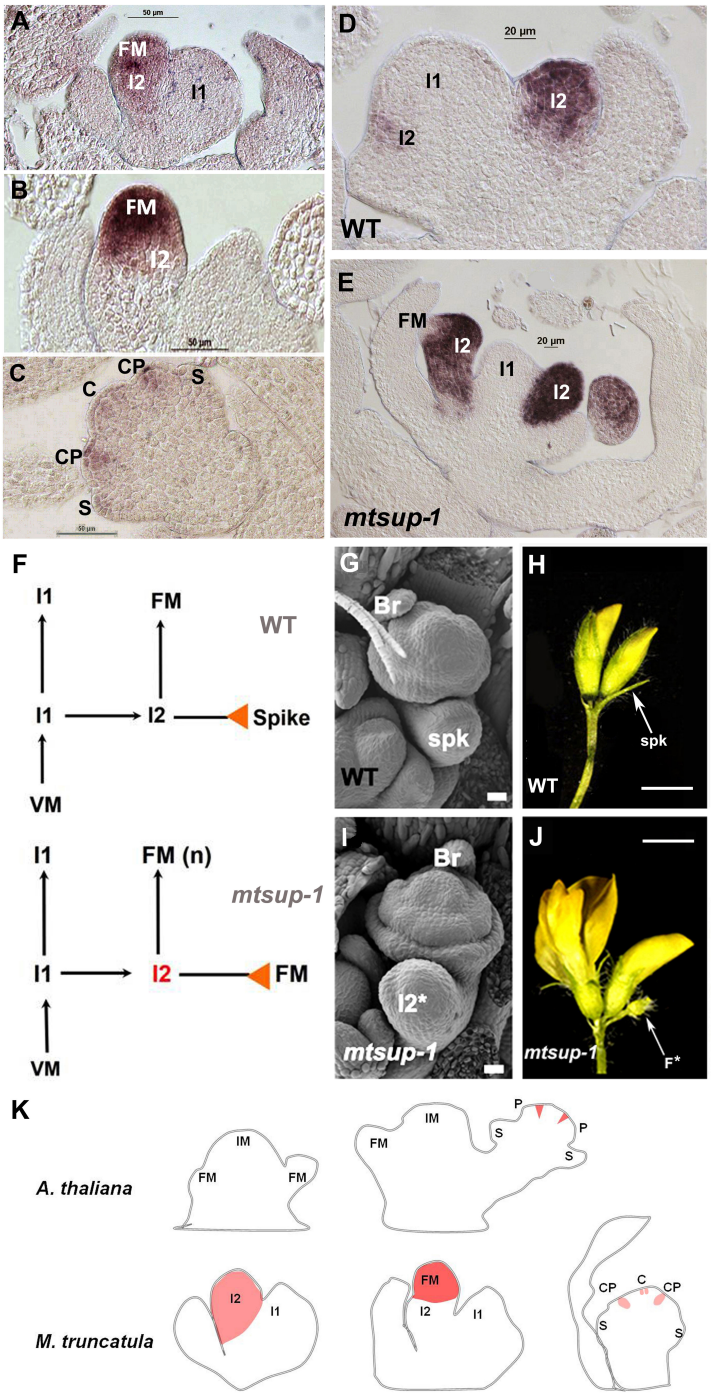


FIGURE 2
MtSUP controls compound inflorescence development in *M. truncatula*. **(A)** *MtSUP* transcript is firstly detected in the I2. **(B)** Later on *MtSUP* activity is detected in the FM. **(C)** During early floral development *MtSUP* expression is detected in the common primordia (CP). **(D)** *MtFULc* expression in the wild type (WT) flower. **(E)** *MtFULc* expression in the *mtsup-1* mutant. *MtFULc* transcript occupies a wider area in *MtSUP* mutants compared to the WT. **(F)** Schematic representation of the compound inflorescence development in Medicago with the formation of an I2 and the terminal spike in the WT and a new FM instead the spike in the *mtsup-1* mutant. **(G)** SEM image of a WT floral primordium with its respective bract and spike. **(H)** The WT of *M. truncatula* R108 produces one or two flowers per inflorescence and terminates in a spike. **(I)** In *mtsup-1*, the I2* (future spike) acquires floral identity. **(J)** In the *mtsup-1* mutant the residual cells of the I2 terminate as a new flower (F*) instead a spike. **(K)** Comparative schematic representation of *SUP* and *MtSUP* expression patterns during inflorescence and flower development in Arabidopsis and Medicago. *SUP* and *MtSUP* are orthologs that have functionally diverged through changes in their gene transcription patterns. IM, inflorescence meristem; I1, primary inflorescence meristem; I2, secondary inflorescence meristem; FM, floral meristem; CP, common primordium; S, sepal primordium; P, petal primordium; C, carpel primordium; Br, bract; Spk, spike; F*, new flower. Scale bars, 20 µm in **(G, I)**, and 2 mm in **(H, J)** Adapted from Rodas et al. (2021).

consequences of *sup* mutations in Arabidopsis are correlated to an over-proliferation of cells at W3 and the floral stem cells at the floral meristem centre (Prunet et al., 2009; Prunet et al., 2017; Xu et al., 2018). In addition, in *P. hybrida* the specific expression of *PhSUP* in the stamen primordia and the excessive proliferation of cells at the connective tissue in *phsup1* anthers suggest that the control of cell division and growth is the function of this *SUP* ortholog (Nakagawa et al., 2004). Similar conclusions were reached with the *SUP* ortholog in rice. The authors stated that *SRO* is not a cadastral gene based on its expression pattern and the mechanisms through which *SRO* regulates reproductive organ development (Xu et al., 2018).

From an evolutionary point of view, *SUP* and *MtSUP* are orthologs that have functionally diverged through changes in their gene transcription patterns while keeping some common functions (Figure 2K). Such changes can occur through transposition, rearrangement, duplication or point mutations in the regulatory regions (Carroll, 2005), which are frequent after whole-genome duplications (WGD) events, a common phenomenon in the evolution of angiosperms (Cui et al., 2006). The WGD event that pre-dated the speciation of legumes ~50–60 million years ago had an essential role in structuring the *M. truncatula* genome and in the success of papilionoid legumes (Cannon et al., 2006). However, these rounds of polyploidization have contributed mainly to the gradual decline in the conserved synteny between species, as is the case for Arabidopsis and *M. truncatula* (Young et al., 2011). An example is the absence of collinearity in the flanking regions of *MtSUP* and *SUP* in their respective genomes (Rodas et al., 2021).

The functional study of *SUP* orthologous genes in other legume species (alfalfa, common bean) or plants with complex inflorescence (*i.e.* tomato, mustard) could help to understand *SUP*-like genes implications in the development of higher order meristems (*i.e.* I2). Alternatively, they could show the emergence of new functions for transcription factors when they are expressed in species with different architectures. Recently, the *SINGLE FLOWER* (*SFL*) gene, a MYB transcription factor expressed in the I2, was shown to perform a similar role to *MtSUP* in chickpea (Caballo et al., 2022). Multiflowered pea and chickpea mutants have been reported for three decades (Murfet, 1985; Singer et al., 1999; Gaur and Gour, 2002; Srinivasan et al., 2006; Devi et al., 2018). However, the correlation between the multiflowered mutants and the genes responsible requires further studies. The genes involved in the specification and determinacy of inflorescence meristems could be used as bioengineering tools to optimize inflorescence traits (Wang et al., 2021). In line with this, *SUP*-like genes in other crops should be studied to determine their possible roles in inflorescence development.

Author contributions

ER, AR and LC wrote the manuscript and RH, CG-M and JB contributed with valuable comments during the manuscript writing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a grant from the Spanish Ministry of Science and Innovation (PID2019-106060RB-I00). AR acknowledges a Santiago Grisolía fellowship (GRISOLIA 2017/168) from the Generalitat Valenciana.

Acknowledgments

We want to thank John Wiley and Sons and The Plant Journal by the license (5444160987798) to reproduce some figure panels previously published in Rodas et al. (2021). We acknowledge the support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Benlloch, R., Berbel, A., Ali, L., Gohari, G., Millán, T., and Madueño, F. (2015). Genetic control of inflorescence architecture in legumes. *Front. Plant Sci.* 6, 1–14. doi: 10.3389/fpls.2015.00543

Benlloch, R., Berbel, A., Serrano-Mislata, A., and Madueño, F. (2007). Floral initiation and inflorescence architecture: A comparative view. *Ann. Bot.* 100, 659–676. doi: 10.1093/aob/mcm146

Benlloch, R., D'Erfurth, I., Ferrandiz, C., Cosson, V., Beltrán, J. P., Cañas, L. A., et al. (2006). Isolation of mtpim proves *Tnt1* a useful reverse genetics tool in *Medicago truncatula* and uncovers new aspects of *AP1*-like functions in legumes. *Plant Physiol.* 142, 972–983. doi: 10.1104/pp.106.083543

Benlloch, R., Navarro, C., Beltrán, J. P., and Cañas, L. A. (2003). Floral development of the model legume *Medicago truncatul*a: Ontogeny studies as a tool to better characterize homeotic mutations. *Sexual Plant Reprod.* 15, 231–241. doi: 10.1007/s00497-002-0157-1

Berbel, A., Ferrándiz, C., Hecht, V., Dalmais, M., Lund, O. S., Sussmilch, F. C., et al. (2012). *VEGETATIVE1* is essential for development of the compound inflorescence in pea. *Nat. Commun.* 3, 797. doi: 10.1038/ncomms1801

Bereterbide, A., Hernould, M., Castera, S., and Mouras, A. (2001). Inhibition of cell proliferation, cell expansion and differentiation by the arabidopsis *SUPERMAN* gene in transgenic tobacco plants. *Planta* 214, 22–29. doi: 10.1007/s004250100584

Bossinger, G., and Smyth, D. R. (1996). Initiation patterns of flower and floral organ development in *Arabidopsis thaliana*. *Development* 122, 1093–1102. doi: 10.1242/dev.122.4.1093

Bowman, J. L., Sakai, H., Jack, T., Weigel, D., Mayer, U., and Meyerowitz, E. M. (1992). *SUPERMAN*, a regulator of floral homeotic genes in arabidopsis. *Development* 114, 599–615. doi: 10.1242/dev.114.3.599

Bowman, J. L., Smyth, D. R., and Meyerowitz, E. M. (1989). Genes directing flower development in arabidopsis. *Plant Cell* 1, 37–52. doi: 10.1105/tpc.1.1.37

Bowman, J. L., Smyth, D. R., and Meyerowitz, E. M. (2012). The ABC model of flower development: then and now. *Development* 139 (22), 4095–4098. doi: 10.1242/dev.083972

Breuil-Broyer, S., Trehin, C., Morel, P., Boltz, V., Sun, B., Chambrier, P., et al. (2016). Analysis of the arabidopsis *superman* allelic series and the interactions with other genes demonstrate developmental robustness and joint specification of male-female boundary, flower meristem termination and carpel compartmentalization. *Ann. Bot.* 117, 905–923. doi: 10.1093/aob/mcw023

Caballo, C., Berbel, A., Ortega, R., Gil, J., Millán, T., Rubio, J., et al. (2022). The *SINGLE FLOWER* (*SFL*) gene encodes a MYB transcription factor that regulates the number of flowers produced by the inflorescence of chickpea. *New Phytol.* 234, 827–836. doi: 10.1111/nph.18019

Cañas, L. A., and Beltrán, J. P. (2018). "Chapter 2. model legumes: functional genomics tools in medicago truncatula," in *Functional genomics in medicago truncatula. methods and protocols. methods in molecular biology series n° 1822*. Eds. L. A. Cañas and JoséP. Beltrán (New York: Springer Nature, Humana Press), 11–38, ISBN: .

Cannon, S. B., Sterckc, L., Rombautsc, S., Sato, S., Cheung, F., Gouzy, J., et al. (2006). Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. *Proc. Natl. Acad. Sciences U.S.A.* 103, 14959–14964. doi: 10.1073/pnas.0603228103

Carroll, S. B. (2005). Evolution at two levels: On genes and form. *PloS Biol.* 3, 1159–1166. doi: 10.1371/journal.pbio.0030245

Cheng, X., Li, G., Tang, Y., and Wen, J. (2018). Dissection of genetic regulation of compound inflorescence development in *Medicago truncatula*. *Development* 145 (3), dev158766. doi: 10.1242/dev.158766

Cui, L., Wall, P. K., Leebens-Mack, J. H., Lindsay, B. G., Soltis, D. E., Doyle, J. J., et al. (2006). Widespread genome duplications throughout the history of flowering plants. *Genome Res.* 16, 738–749. doi: 10.1101/gr.4825606

Devi, J., Mishra, G. P., Sanwal, S. K., Dubey, R. K., Singh, P. M., and Singh, B. (2018). Development and characterization of penta-flowering and triple-flowering genotypes in garden pea (*Pisum sativum* l. var. hortense). *PloS One* 13, 1–15. doi: 10.1371/journal.pone.0201235

Ferrándiz, C., Navarro, C., Gómez, M. D., Cañas, L. A., and Beltrán, J. P. (1999). Flower development in *Pisum sativum*: From the war of the whorls to the battle of the common primordia. *Dev. Genet.* 25, 280–290. doi: 10.1002/(SICI)1520-6408(1999)25:3<280::AID-DVG10>3.0.CO;2-3

Gaiser, J. C., Robinson-Beers, K., and Gasser, C. S. (1995). The arabidopsis *SUPERMAN* gene mediates asymmetric growth of the outer integument of ovules. *Plant Cell* 7, 333–345. doi: 10.2307/3869855

Gaur, P. M., and Gour, V. K. (2002). A gene producing one to nine flowers per flowering node in chickpea. *Euphytica* 128, 231–235. doi: 10.1023/A:1020845815319

Hiratsu, K., Matsui, K., Koyama, T., and Ohme-Takagi, M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in arabidopsis. *Plant J.* 34, 733–739. doi: 10.1046/j.1365-313X.2003.01759.x

Hiratsu, K., Mitsuda, N., Matsui, K., and Ohme-Takagi, M. (2004). Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in arabidopsis. *Biochem. Biophys. Res. Commun.* 321, 172–178. doi: 10.1016/j.bbrc.2004.06.115

Hiratsu, K., Ohta, M., Matsui, K., and Ohme-Takagi, M. (2002). The SUPERMAN protein is an active repressor whose carboxy-terminal repression domain is required for the development of normal flowers. *FEBS Lett.* 514, 351–354. doi: 10.1016/S0014-5793(02)02435-3

Hofer, J. M. I., and Ellis, N. T. H. (2014). Developmental specialisations in the legume family. *Curr. Opin. Plant Biol.* 17, 153–158. doi: 10.1016/j.pbi.2013.11.014

Kazama, Y., Fujiwara, M. T., Koizumi, A., Nishihara, K., Kifune, E., Abe, T., et al. (2009). A *SUPERMAN*-like gene is exclusively expressed in female flowers of the dioecious plant silene latifolia. *Plant Cell Physiol.* 50, 1127–1141. doi: 10.1093/pcp/pcp064

Murfet, I. C. (1985). The influence of genes *ar* and *n* on senescence in *Pisum sativum* l. *Ann. Bot.* 55 (5), 675–683. doi: 10.1093/oxfordjournals.aob.a086946

Nakagawa, H., Ferrario, S., Angenent, G. C., Kobayashi, A., and Takatsuji, H. (2004). The petunia ortholog of arabidopsis *SUPERMAN* plays a distinct role in floral organ morphogenesis. *Plant Cell* 16, 920–932. doi: 10.1105/tpc.018838

Nandi, A. K., Kushalappa, K., Prasad, K., and Vijayraghavan, U. (2000). A conserved function for arabidopsis *SUPERMAN* in regulating floral-whorl cell proliferation in rice, a monocotyledonous plant. *Curr. Biol.* 10 (4), 215–218. doi: 10.1016/S0960-9822(00)00341-9

Nibau, C., Di Stilio, V. S., Wu, H. M., and Cheung, A. Y. (2011). Arabidopsis and tobacco *SUPERMAN* regulate hormone signalling and mediate cell proliferation and differentiation. *J. Exp. Bot.* 62, 949–961. doi: 10.1093/jxb/erq325

Prunet, N., Morel, P., Negrutiu, I., and Trehin, C. (2009). Time to stop: Flower meristem termination. *Plant Physiol.* 150, 1764–1772. doi: 10.1104/pp.109.141812

Prunet, N., Yang, W., Das, P., Meyerowitz, E. M., and Jack, T. P. (2017). *SUPERMAN* prevents class b gene expression and promotes stem cell termination in the fourth whorl of *Arabidopsis thaliana* flowers. *Proc. Natl. Acad. Sciences U.S.A.* 114, 7166–7171. doi: 10.1073/pnas.1705977114

Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L. D., and Coen, E. (2007). Evolution and development of inflorescence architectures. *Sci. (80-.)* 316, 1452–1456. doi: 10.1126/science.1140429

Rodas, A. L., Roque, E., Hamza, R., Gómez-Mena, C., Minguet, E. G., Wen, J., et al. (2021). *MtSUPERMAN* plays a key role in compound inflorescence and flower development in *Medicago truncatula*. *Plant J.* 105 (3), 816–830. doi: 10.1111/tpj.15075

Roque, E., Gómez-Mena, C., Ferrándiz, C., Beltrán, J. P., and Cañas, L. A. (2018). "Chapter 18. functional genomics and genetic control of floral and fruit development in medicago truncatula: an overview," in *Functional genomics in medicago truncatula. methods and protocols. methods in molecular biology series n° 1822*. Eds. L. A. Cañas and JoséP. Beltrán (New York: Springer Nature, Humana Press), 273–290, ISBN: .

Sakai, H., Krizek, B. A., Jacobsen, S. E., and Meyerowitz, E. M. (2000). Regulation of *SUP* expression identifies multiple regulators involved in arabidopsis floral meristem development. *Plant Cell* 12, 1607–1618. doi: 10.1105/tpc.12.9.1607

Sakai, H., Medrano, L. J., and Meyerowitz, E. M. (1995). Role of *SUPERMAN* in maintaining arabidopsis floral whorl boundaries. *Nature* 378, 199–203. doi: 10.1038/378199a0

Singer, S., Sollinger, J., Maki, S., Fishbach, J., Short, B., Reinke, C., et al. (1999). Inflorescence architecture: A developmental genetics approach. *Botanical Rev.* 65, 385–410. doi: 10.1007/BF02857756

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early flower development in arabidopsis. *Plant Cell* 2 (8), 75–67. doi: 10.1105/tpc.2.8.755

Srinivasan, S., Gaur, P. M., Chaturvedi, S. K., and Rao, B. V. (2006). Allelic relationships of genes controlling number of flowers per axis in chickpea. *Euphytica* 152, 331–337. doi: 10.1007/s10681-006-9219-z

Tucker, S. C. (1989). Overlapping organ initation and common primordia in flowers of *Pisum sativum* (Leguminosae : Papilionoideae). *Am. J. Bot.* 76, 714–729. doi: 10.1002/j.1537-2197.1989.tb11366.x

Tucker, S. C. (2003). Update on floral development floral development in legumes. *Plant Physiol.* 131, 911–926. doi: 10.1104/pp.102.017459

Wang, C., Yang, X., and Li, G. (2021). Molecular insights into inflorescence meristem specification for yield potential in cereal crops. *Int. J. Mol. Sci.* 22 (7), 3508. doi: 10.3390/ijms22073508

Xu, Y., Prunet, N., Gan, E.-S., Wang, Y., Stewart, D., Wellmer, F., et al. (2018). *SUPERMAN* regulates floral whorl boundaries through control of auxin biosynthesis. *EMBO J.* 37, e97499. doi: 10.15252/embj.201797499

Xu, W., Zhu, W., Yang, L., Liang, W., Li, H., Yang, L., et al. (2022). *SMALL REPRODUCTIVE ORGANS*, a *SUPERMAN*-like transcription factor, regulates stamen and pistil growth in rice. *New Phytol.* 233, 1701–1718. doi: 10.1111/nph.17849

Young, N. D., Debellé, F., Oldroyd, G. E. D., Cannon, S. B., Udvardi, M. K., et al. (2011). The medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524.

Yu, H., and Huang, T. (2016). Molecular mechanisms of floral boundary formation in arabidopsis. *Int. J. Mol. Sci.* 17 (3), 317. doi: 10.3390/ijms17030317

Zhao, J., Liu, M., Jiang, L., Ding, L., Yan, S. S., Zhang, J., et al. (2014). Cucumber *SUPERMAN* has conserved function in stamen and fruit development and a distinct role in floral patterning. *PloS One* 9, 1–9. doi: 10.1371/journal.pone.0086192

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 March 2023
DOI 10.3389/fpls.2023.1083086

Check for updates

OPEN ACCESS

EDITED BY
Karam B. Singh,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

REVIEWED BY
R. Varma Penmetsa,
University of California, Davis, CA, United States
Lars Kamphuis,
Curtin University, Australia

*CORRESPONDENCE
Thomas D. Warkentin
tom.warkentin@usask.ca

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 28 October 2022
ACCEPTED 14 February 2023
PUBLISHED 10 March 2023

CITATION
Zhou J, Wan Z, Gali KK, Jha AB, Nickerson MT, House JD, Tar'an B and Warkentin TD (2023) Quantitative trait loci associated with amino acid concentration and *in vitro* protein digestibility in pea (*Pisum sativum* L.).
Front. Plant Sci. 14:1083086.
doi: 10.3389/fpls.2023.1083086

COPYRIGHT
© 2023 Zhou, Wan, Gali, Jha, Nickerson, House, Tar'an and Warkentin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Quantitative trait loci associated with amino acid concentration and *in vitro* protein digestibility in pea (*Pisum sativum* L.)

Junsheng Zhou[1], Zhongyang Wan[2], Krishna Kishore Gali[1], Ambuj Bhushan Jha[1], Michael T. Nickerson[3], James D. House[2,4], Bunyamin Tar'an[1] and Thomas D. Warkentin[1]*

[1]Crop Development Centre, Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada, [2]Department of Food and Human Nutritional Sciences, University of Manitoba, Winnipeg, MB, Canada, [3]Department of Food and Bioproduct Science, University of Saskatchewan, Saskatoon, SK, Canada, [4]Richardson Centre for Food Technology and Research, University of Manitoba, Winnipeg, MB, Canada

With the expanding interest in plant-based proteins in the food industry, increasing emphasis is being placed on breeding for protein concentration and quality. Two protein quality traits i.e., amino acid profile and protein digestibility, were assessed in replicated, multi-location field trials from 2019 to 2021 in pea recombinant inbred line population PR-25. This RIL population was targeted specifically for the research of protein related traits and its parents, CDC Amarillo and CDC Limerick, had distinct variations in the concentration of several amino acids. Amino acid profile was determined using near infrared reflectance analysis, and protein digestibility was through an *in vitro* method. Several essential amino acids were selected for QTL analysis, including lysine, one of the most abundant essential amino acids in pea, and methionine, cysteine, and tryptophan, the limiting amino acids in pea. Based on phenotypic data of amino acid profiles and *in vitro* protein digestibility of PR-25 harvested in seven location-years, three QTLs were associated with methionine + cysteine concentration, among which, one was located on chromosome 2 (R^2 = 17%, indicates this QTL explained 17% phenotypic variation of methionine + cysteine concentration within PR-25), and two were located on chromosome 5 (R^2 = 11% and 16%). Four QTLs were associated with tryptophan concentration and are located on chromosome 1 (R^2 = 9%), chromosome 3 (R^2 = 9%), and chromosome 5 (R^2 = 8% and 13%). Three QTLs were associated with lysine concentration, among which, one was located on chromosome 3 (R^2 = 10%), the other two were located on chromosome 4 (R^2 = 15% and 21%). Two QTLs were associated with *in vitro* protein digestibility, one each located on chromosomes 1 (R^2 = 11%) and 2 (R^2 = 10%). QTLs associated with *in vitro* protein digestibility, and methionine + cysteine concentration on chromosome 2 were identified to be co-localized with known QTL for total seed protein concentration in PR-25. QTLs associated with tryptophan and

methionine + cysteine concentration co-localized on chromosome 5. The identification of QTLs associated with pea seed quality is an important step towards marker-assisted selection of breeding lines with improved nutritional quality, which will further boost the competitiveness of pea in plant-based protein markets.

KEYWORDS

pea, methionine, cysteine, tryptophan, lysine, protein digestibility, QTLs

Introduction

Despite the currently relatively high sales price of the plant-based protein products, their popularity is increasing in recent years (Van Loo et al., 2020; Estell et al., 2021). Trends toward sustainable agriculture, healthy diets, food security, expanding global population, animal welfare, and religious beliefs are contributing to the expansion of plant-based protein product launches to meet the diverse needs (Sabate & Soret, 2014; Alcorta et al., 2021; Aschemann-Witzel et al., 2021).

Pea protein, as one of the common ingredients in the fabrication of meat alternatives, is rich in lysine, but relatively low in tryptophan and sulfur containing amino acids, such as methionine and cysteine (Leterme et al., 1990; Pilorge et al., 2021). To provide a complete and balanced set of amino acids, food companies need to complement pea protein with cereal protein, since cereals tend to be rich in the sulfur amino acids, but limiting in lysine (Srivastava et al., 2015). Compared to soybean protein, which is also widely used in plant-based protein products, pea protein is less allergenic, and contains less offensive off-flavors, such as beany, chalky flavors, and hence require less processing (Hoffmann et al., 2017; Taylor et al., 2021). These advantages could contribute to the wider use of pea protein in the future, but the current limitation of sulfur amino acids could also be an obstacle to wider use. Until now, most of the emphasis related to pea breeding for protein has been related to protein quantity instead of quality. But as the demands of the plant-based protein industries is continuously growing, a more profound study of pea protein quality traits is necessary.

Protein digestibility corrected amino acid score (PDCAAS) is an index that has been widely adopted when evaluating the quality of a specific protein. PDCAAS is based on both the amino acids profile and the digestibility of the proteins (Schaafsma, 2005). Amino acid profile is typically assessed by HPLC analysis, (Dołowy & Pyka, 2014; Sharma et al., 2014). HPLC analysis produces accurate results but is expensive and low throughput. Amino acids have distinct attributes, which increases the complexity of enzyme digestion steps. Three different digestion steps are required in HPLC analysis to assess the complete amino acid profile, and long incubation periods makes it a cumbersome method for amino acid profiling on large numbers of samples typical of plant breeding programs (Kambhampati et al., 2019). The use of near-infrared spectroscopy (NIR) has been widely used to assess the protein concentration in different crops. Recently, this non-destructive approach has been extended to assess the amino acid profile of legume seed samples (Hang et al., 2022; Shi et al., 2022). Shi et al. (2022) developed a calibration curve to predict the concentration of individual amino acids in soybean based on the NIR scans. The development and enhancement of the curve relies on data derived from established and validated wet chemistry reference methods, including the use of liquid chromatographic methods. For most amino acids, relatively high correlations were detected between the reference chemistry method, namely HPLC, and the NIR analysis. The correlation coefficients fell within a range of 0.76 to 0.91. Even for the least abundant amino acids, for instance, methionine, the correlation coefficient was 0.77. The benefits of using NIR relate to high throughput, low cost of analysis, and the potential to use intact seeds for non-destructive analysis (Jiang, 2020). The latter is particularly important when dealing with precious seed quantities such as those derived from breeding programs.

Protein digestibility reflects the extent to which a protein is broken down into its constituent amino acids and the latter made available for absorption within the gastrointestinal tract. *In vivo* and *in vitro* methods have been used to measure protein digestibility. The *in vivo* method assesses true fecal digestion in a rodent model and is accurate, but it has the limitations of high cost and low throughput, due to the significant involvement of animals in the analysis (Tavano et al., 2016). In contrast, *in vitro* methods simulate conditions within the digestive tract, including enzyme digestion, with protein digestibility assessed by various methods, including the change of pH of samples over a defined time course. This reproducible approach provides a method with relatively high throughput compared to the *in vivo* method and has been adopted to assess digestibility of different plant proteins (Duodu et al., 2002; Bessada et al., 2019; Ketnawa & Ogawa, 2021).

With more diverse needs from both consumers and the food industry, it has become necessary to develop crop varieties with improved protein quality. Understanding the underlying genetic control of protein quality traits is critical to aid in breeding such varieties. In the current study, we used PR-25, a recombinant inbred line population, to understand the genetic basis of protein quality traits in pea. PR-25 is derived from the cross of a medium protein parent CDC Amarillo and a high protein parent CDC Limerick.

10.3389/fpls.2023.1083086

These two genotypes differed in individual amino acid concentrations from 10% to 27%, which makes PR-25 as a valuable source for the identification of QTL regions associated with protein quality traits. These QTL regions have the potential to be developed into breeder-friendly DNA markers, which could be used to improve the protein quality of pea and hence, increase the competitiveness of pea in plant-based protein markets.

Materials and methods

PR-25 mapping population

The PR-25 mapping population used in this study is a recombinant inbred line (RIL) population with 110 lines derived from the cross of CDC Amarillo x CDC Limerick. The mapping population was grown in seven locations during 2019 to 2021 in microplots (1 m^2) and the harvested seed from individual plots was used in the current study. Information about PR-25, including its seeding date, harvest date and crop management in all locations were detailed in Zhou et al. (2022). Briefly, PR-25 was grown at Sutherland nursery in 2019 with 2 biological replicates. It was grown at Sutherland, Rosthern and Lucky Lake nurseries in 2020 and Floral, Rosthern and Lucky Lake nurseries in 2021 with 3 biological replicates per nursery. All nurseries were in Saskatchewan, Canada and among which, Sutherland and Floral are located in the Dark Brown soil zone, Rosthern is located in the Black soil zone, and Lucky Lake is located in the Brown soil zone. Best management practices for field pea production in western Canada were used at each location.

Amino acid profiling and NIR calibration development

A total of 2359 whole seed samples, including 159 samples from 2017 GWAS (Gali et al., 2019) and 2200 samples from 2019-2021 PR-25, were used for calibration development and amino acid profile assessment (Table 1). They were stored at -20°C before any analysis. All samples were scanned in Dr. James House's lab (University of Manitoba, Winnipeg, MB, Canada) *via* NIR spectroscopy on a PerkinElmer DA 7250 diode array NIR system (PerkinElmer Health Sciences Canada Inc.) to obtain spectral data. A sub-sample of 339, including all 159 samples from 2017 GWAS and 180 samples from 2019-2021 PR-25, were selected from 2359 samples for development and improvement of the calibration formula. These sub-samples were analyzed *via* HPLC for amino acid concentration and the protocol of HPLC was detailed by Shi et al. (2022). A prediction model was created using the whole seed spectrum and chemical data of 2017 GWAS samples, the calibration was then applied to all remaining samples to estimate chemical compositions (Shi et al., 2022). Annual maintenance and improvement of prediction model were made based on the HPLC data of selected samples from PR-25. Sixty samples each were chosen from year 2019 to 2021 to improve the accuracy of calibration formula. Samples in each year were divided into quartiles of predicted crude protein content in descending order and within each quartile, 15 samples were drawn randomly.

All seed samples were stored in a 4°C walk-in cooler prior to NIR scanning to avoid protein denaturation. After scanning with NIRS, approximately 20 grams of each selected sample was ground with a Retsch ZM-200 grinder (Retsch, Haan, Germany) using a 0.75mm sieve and then stored at -20 °C before further analysis.

Detailed information about NIR analysis of whole seeds, as well as wet chemistry analysis of protein and amino acid concentrations was reported by Shi (2021). Detailed information about the development of an NIR calibration model for pea amino acid concentration was reported by Hang et al. (2022).

In vitro protein digestibility assessment

Sample preparations were conducted in the Grain Innovation Lab (Crop Development Centre, University of Saskatchewan). Five grams of pea seeds per line from 2019 PR-25 yield trials were ground into homogeneous powder using a cyclone sample mill model 3010-030 (UDY Corporation, USA). Protein digestibility was evaluated using an *in vitro* method. Detailed protocol of *in vitro*

TABLE 1 Information of whole seed pea samples used for near infrared reflectance (NIR) calibration development and amino acid profile assessment.

Year	Population	Location	Lines	Replicates	Total number of tested samples	Use
2017	GWAS	Sutherland	80	2	159 (1 missing)	Calibration development
2019	PR-25	Sutherland	110	2	220	Calibration improvement/ Amino acid profiling
2020	PR-25	Sutherland	110	3	330	
	PR-25	Rosthern	110	3	330	
	PR-25	Lucky Lake	110	3	330	
2021	PR-25	Floral	110	3	330	
	PR-25	Rosthern	110	3	330	
	PR-25	Lucky Lake	110	3	330	
			Total		2359	

protein digestibility (IVPD) determination was described by Cabuk et al. (2018).

Measurement of *in vitro* protein digestibility-corrected amino acid score

In vitro PDCAAS was calculated as follows based on Nosworthy et al. (2017):

$$AAS = \frac{\text{mg of most limiting amino acid in 1g of tested protein}}{\text{mg of this particular amino acid in 1g of reference protein}}$$

$$in \quad vitro \quad PDCAAS = AAS * IVPD$$

Statistical analysis

Two sample t-tests with equal variance were conducted on protein related traits, including the concentration of protein and each of 18 amino acids, for the parents of PR-25, CDC Amarillo and CDC Limerick (Table 2). Data of CDC Amarillo and CDC Limerick from all 7 station-years were used for the t-test. A Pearson correlation analysis was performed among protein quality traits, including total protein, protein digestibility (*in vitro*), PDCAAS and 18 amino acids found in pea seeds (Table 3), based on the average of 7 station-years data (2 biological replicates in 2019 Sutherland, 3 biological replicates in 2020 Sutherland, 2020 Rosthern, 2020 Lucky Lake, 2021 Floral, 2021 Rosthern, 2021 Lucky Lake) for each trait. Analysis of variance (ANOVA) was conducted for amino acids of interest, including methionine, cystine, tryptophan and lysine using 7 station-years data; ANOVA of multi-environment function was conducted in QTL IciMapping software (Table 4). A comparison of amino acid profiles was made among years 2019, 2020, 2021 and the average of all 7 station-years (Figure 1). A comparison of amino acid profiles was also made between CDC Amarillo and CDC Limerick based on their average of 7 station-years (Figure 2). Each amino acid profile represented the average percentage of each amino acid in pea seeds in the given year. Frequency distribution of methionine + cysteine concentration, tryptophan concentration, lysine concentration, protein digestibility (*in vitro*) and PDCAAS was made based on the average of biological replicates for each line in the given year-location (Figures 3–7).

Genotyping, linkage map construction and QTL identification

PR-25 population was genotyped using Axiom® 90K SNP array and a linkage map representing 901 unique loci and measuring 855.4 cM was reported by Zhou et al. (2022). This linkage map was used for QTL analysis in the current study. The nomenclature of the markers was in accordance with Axiom® 90K SNP array. To increase the accuracy of QTL identification, QTL analyses were first performed individually in each station-year. Data of some station-years, in which their individual analysis didn't include any significant QTL, were excluded and the rest were averaged and used for combined QTL analysis. Phenotypic traits, including methionine + cysteine concentration, tryptophan concentration, lysine concentration and *in vitro* protein digestibility, were assessed. QTL analysis was conducted *via* Windows QTL Cartographer (Wang et al., 2012), where cross type was set as Ri1, map function was set as Kosambi, analysis type was set as composite interval mapping, permutation time was set as 1000, significance level was set as 0.05, and walk speed was set as 1.0 Cm. Other detailed information of the QTL identification approach was described in Zhou et al., 2022.

Results

The results of two sample t-test between the parents of PR-25 population, CDC Amarillo and CDC Limerick, showed that they differed significantly in their concentration of seed protein and most of the amino acids, except for alanine (Table 2). The average proportion of each amino acid from field trials conducted over three seasons is summarized in Figure 1. The average of each amino acid content from the 7 station-years can be ranked from top to bottom as: Glutamine (18.3%), aspartic acid (12.4%), lysine (7.7%), leucine (7.5%), arginine (7.5%), phenylalanine (5.2%), serine (5.1%), valine (4.9%), glycine (4.5%), isoleucine (4.4%), proline (4.4%), alanine (4.3%), threonine (3.9%), tyrosine (3.3%), histidine (2.9%), cysteine (1.4%), methionine (1.2%), tryptophan (1.1%). The average proportion of each amino acid of CDC Amarillo and CDC Limerick from 7 station-years is summarized in Figure 2. There were positive correlations among most amino acids (Table 3). Variation in amino acid profile was detected across the station-years. Analysis of variance of methionine, cysteine, tryptophan, and lysine showed that the phenotypic variation can be attributed to the

TABLE 2 Two sample t-tests with equal variance between the parents of PR-25, CDC Amarillo and CDC Limerick, across 7 station-years.

AA	Ala	Met	Cys	His	Ser	Arg	Gly
p value	0.188ns	<0.001***	0.002**	<0.001***	0.001**	0.031*	0.036*
AA	Asp	Glu	Thr	Pro	Lys	Tyr	Val
p value	0.03*	<0.001***	0.039*	0.001**	<0.001***	0.004**	0.007**
AA	Ile	Leu	Phe	Tryp	Protein		
p value	0.010*	0.002**	0.010*	0.021*	0.004**		

Significance levels for the correlation coefficient (r) is denoted by the symbols *, **, ***, for P< 0.05, P< 0.01, P< 0.001 or not significant (ns), respectively.

TABLE 3 Correlation analysis among total protein, protein digestibility (*in vitro*), PDCAAS and 18 amino acids found in pea seed.

	Ala	Met	Cys	His	Ser	Arg	Gly	Asp	Glu	Thr	Pro	Lys	Tyr	Val	Ile	Leu	Phe	Trp	Protein	IVPD
Met	0.52***																			
Cys	0.25**	0.51***																		
His	0.64***	0.37***	0.18ns																	
Ser	0.52***	0.11ns	0.27**	0.73***																
Arg	0.76***	0.63***	0.51***	0.56***	0.56***															
Gly	0.80***	0.14ns	0.16ns	0.75***	0.82***	0.57***														
Asp	0.89***	0.51***	0.42***	0.70***	0.70***	0.87***	0.78***													
Glu	0.76***	0.37***	0.34***	0.86***	0.89***	0.72***	0.88***	0.89***												
Thr	0.84***	0.23*	0.15ns	0.77***	0.85***	0.64***	0.96***	0.84***	0.90***											
Pro	0.75***	0.24*	0.35***	0.78***	0.88***	0.74***	0.92***	0.89***	0.94***	0.91***										
Lys	0.68***	0.13ns	0.15ns	0.84***	0.86***	0.58***	0.89***	0.80***	0.91***	0.90***	0.93***									
Tyr	0.50***	0.21*	0.38***	0.66***	0.83***	0.65***	0.69***	0.74***	0.81***	0.74***	0.83***	0.76***								
Val	0.76***	0.21*	0.24*	0.79***	0.90***	0.70***	0.94***	0.87***	0.94***	0.96***	0.98***	0.94***	0.84***							
Ile	0.75***	0.24*	0.28**	0.75***	0.88***	0.73***	0.91***	0.90***	0.94***	0.92***	0.98***	0.91***	0.84***	0.98***						
Leu	0.79***	0.31**	0.30**	0.78***	0.89***	0.76***	0.90***	0.92***	0.97***	0.93***	0.97***	0.92***	0.83***	0.98***	0.99***					
Phe	0.49***	-0.02ns	0.17ns	0.73***	0.94***	0.51***	0.83***	0.70***	0.87***	0.83***	0.91***	0.92***	0.85***	0.92***	0.91***	0.90***				
Trp	0.73***	0.73***	0.52***	0.66***	0.58***	0.81***	0.6***	0.79***	0.74***	0.68***	0.68***	0.55***	0.68***	0.66***	0.68***	0.71***	0.46***			
Protein	0.74***	0.47***	0.49***	0.80***	0.83***	0.85***	0.78***	0.92***	0.94***	0.83***	0.92***	0.84***	0.84***	0.90***	0.91***	0.93***	0.81***	0.81***		
IVPD	0.12ns	0.22ns	0.11ns	0.07ns	-0.04ns	0.1ns	-0.02ns	0.13ns	0.05ns	0.04ns	0.03ns	0.04ns	0.06ns	0.04ns	0.05ns	0.06ns	-0.01ns	0.13ns	0.10ns	
PDCAAS	-0.09ns	0.39***	0.23*	-0.21ns	-0.36***	-0.09ns	-0.33***	-0.13ns	-0.26**	-0.29**	-0.29**	-0.33**	-0.23*	-0.30**	-0.28**	-0.28**	-0.38***	0.00ns	-0.19*	0.88***

0 1

Significance levels for the correlation coefficient (r) is denoted by the symbols *, **, ***, for $P < 0.05$, $P < 0.01$, $P < 0.001$ or not significant (ns), respectively.

10.3389/fpls.2023.1083086

10.3389/fpls.2023.1083086

TABLE 4 ANOVA analysis on methionine, cysteine, tryptophan and lysine.

Methionine						
Source of Variation	DF	SS	MS	F	P-value	H^2 (%)
Block	13	0.0055	0.0004	6.2120	<0.01***	25.5%
Genotype	110	0.0858	0.0008	11.5416	<0.01***	
Station-year	6	1.4502	0.2417	3575.5540	<0.01***	
GE_Interaction	651	0.1154	0.0002	2.6215	<0.01***	
Model	780	1.6568	0.0021	31.4234	<0.01***	
Error	1385	0.0936	0.0001			
Total	2165	1.7504				
Cysteine						
Source of Variation	DF	SS	MS	F	P-value	H^2 (%)
Block	13	0.17	0.01	40.95	<0.01***	16.6%
Genotype	110	0.21	0.00	6.00	<0.01***	
Station-year	6	1.26	0.21	648.36	<0.01***	
GE_Interaction	651	0.38	0.00	1.80	<0.01***	
Model	780	2.03	0.00	8.02	<0.01***	
Error	1385	0.45	0.00			
Total	2165	2.47				
Tryptophan						
Source of Variation	DF	SS	MS	F	P-value	H^2 (%)
Block	13	0.0047	0.0004	7.3664	<0.01***	26.6%
Genotype	110	0.0507	0.0005	9.4113	<0.01***	
Station-year	6	1.9084	0.3181	6497.8110	<0.01***	
GE_Interaction	651	0.0488	0.0001	1.5312	<0.01***	
Model	780	2.0126	0.0026	52.7111	<0.01***	
Error	1385	0.0678	0.0000			
Total	2165	2.0804				
Lysine						
Source of Variation	DF	SS	MS	F	P-value	H^2 (%)
Block	13	0.58	0.04	11.13	<0.01***	15.8%
Genotype	110	4.00	0.04	9.13	<0.01***	
Station-year	6	12.41	2.07	519.56	<0.01***	
GE_Interaction	651	11.45	0.02	4.42	<0.01***	
Model	780	28.44	0.04	9.16	<0.01***	
Error	1385	5.52	0.00			
Total	2165	33.96				

Significance levels is denoted by the symbols ***, for P< 0.001. GE_Interaction was referred to gene-environment interactions. H^2 is the heritability of the assessed traits.

effects of genotype, station-year, and the interaction of genotype and station-year (Table 4). Frequency distribution of methionine + cysteine concentration, tryptophan concentration, lysine concentration, protein digestibility and PDCAAS presented the ranges, the distribution of phenotypic scores of PR-25 in each station-year (Figures 3–7).

As the primary interest of this research was related to both the abundant and limiting amino acids, hence, lysine, as the abundant

10.3389/fpls.2023.1083086

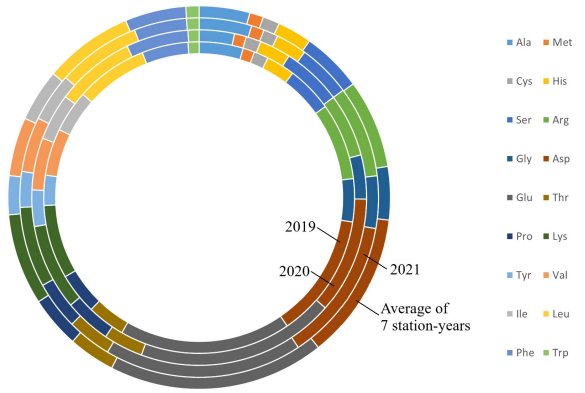


FIGURE 1
Amino acids profile of PR-25 from field trials conducted in Sutherland nursery in 2019, Sutherland, Rosthern, Lucky Lake nurseries in 2020 and Floral, Rosthern and Lucky Lake nurseries in 2021, and the average amino acid profiles of PR-25 across the seven station-years.

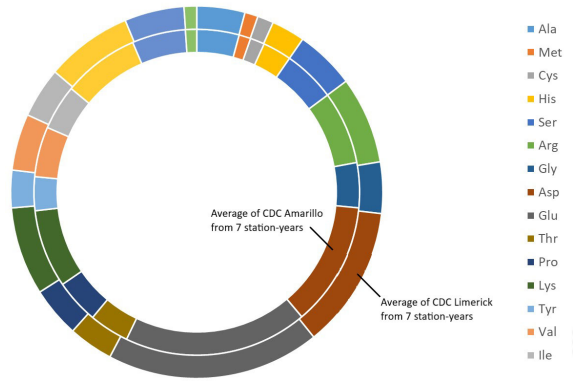


FIGURE 2
Average amino acids profile of CDC Amarillo and CDC Limerick from 7 station-years.

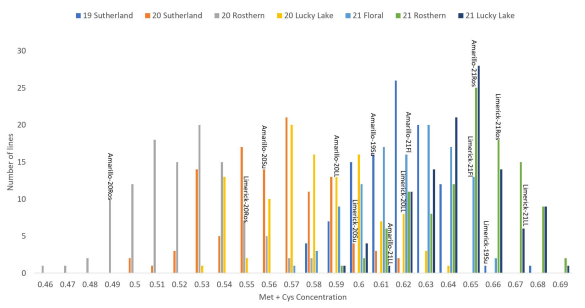


FIGURE 3
Frequency distribution of methionine + cysteine concentration of PR-25 across seven station-years based on the average of biological replicates for each line in each station-year (two biological replicates in 2019 Sutherland and three biological replicates for the rest of the station-years), and the averaged methionine + cysteine concentration for the parents of PR-25 in each station-year.

amino acid, and tryptophan, cysteine and methionine, as the limiting amino acids, were selected for QTL analysis (Table 4). Cysteine and methionine were combined in QTL analysis since they belong to same metabolic pathway (Ravanel et al, 1998). ANOVA showed that environment and genotype × environment interactions had significant contributions to the variation in amino acid concentrations. Therefore, QTL analysis was based on averaged data from selected station-years from which their individual analysis showed significant QTLs. A total of 14 QTLs were identified associated with protein quality traits of interest and they were significant in at least 4 of 7 station-years when compared individually. Four QTLs were identified associated with the methionine + cysteine concentration. Met+Cys-QTL-1 was found on chromosome 2 and it explained 15% of the phenotypic variation, its flanking markers were Chr2LG1_244771437/Chr2LG1_287501555 and there were 22 markers in between; Met+Cys-QTL-2 was found on chromosome 5 and it explained 11% of the phenotypic variation, its flanking markers were Chr5LG3_4173823/Chr5LG3_15801800 and there were 10 markers in between; Met+Cys-QTL-3 was also found on chromosome 5 and it explained 16% of the phenotypic variation, its flanking markers were Chr5LG3_101924498/Chr5LG3_137457380 and there were 10 markers in between; Met+Cys-QTL-4 was found on chromosome 3 and it explained 10% of the phenotypic variation, its flanking markers were Chr3LG5_120117355/Chr3LG5_408080154 and there were 8 markers in between (Figure 8). Chr2LG1_259006623, Chr5LG3_5113345, Chr5LG3_137457380 were the loci within the peak region of each of these three methionine + cysteine associated QTLs. Five QTLs were identified associated with tryptophan concentration, Trp-QTL-1 was found on chromosome 5 and it explained 8% of the phenotypic variation, its flanking markers were Chr5LG3_5127342/Chr5LG3_7509381 and there were 2 markers in between; Trp-QTL-2 was also found on chromosome 5 and it explained 13% of the phenotypic variation, its flanking markers were Chr5LG3_67663653/Chr5LG3_112710798 and there were 10 markers in between; Trp-QTL-3 was found on chromosome 3 and it explained 9% of the phenotypic variation, its flanking markers were Chr3LG5_185794949/Chr3LG5_198663551 and there were 4 markers in between; Trp-QTL-4 was also found on chromosome 3 and it explained 8% of the phenotypic variation, its flanking markers were Chr3LG5_424086163/Chr3LG5_455814220 and there were 2 markers in between; Trp-QTL-5 was found on chromosome 1 and it explained 9% of the phenotypic variation, its flanking markers were Chr1LG6_132953926/Chr1LG6_233117537 and there were 2 markers in between. Chr5LG3_5381756, Chr5LG3_84309239, Chr3LG5_197202364, Chr1LG6_233117537 were the loci within the peak region of each of these four tryptophan associated QTLs. Three QTLs were identified associated with lysine concentration. Lys-QTL-1 was found on chromosome 4 and it explained 21% of the phenotypic variation, its flanking markers were Chr4LG4_185310109/Chr4LG4_218381712 and there were 23 markers in between; Lys-QTL-2 was also found on chromosome 4 and it explained 15% of the phenotypic variation, its flanking markers were Sc02659_148875/Chr4LG4_417303831 and there were 2 markers

10.3389/fpls.2023.1083086

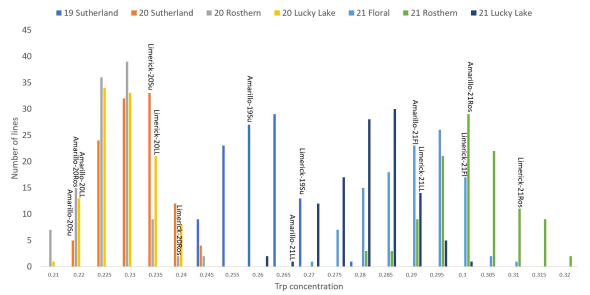


FIGURE 4
Frequency distribution of tryptophan concentration of PR-25 across seven station-years based on the average of biological replicates for each line in each station-year (two biological replicates in 2019 Sutherland and three biological replicates for the rest of the station-years), and the averaged tryptophan concentration for the parents of PR-25 in each station-year.

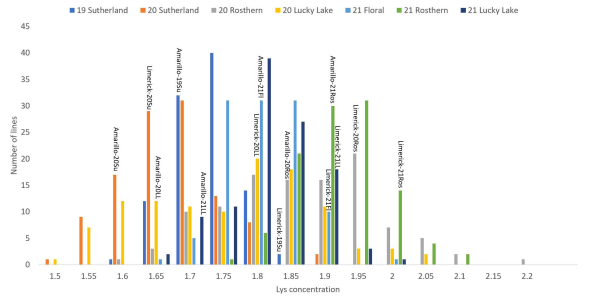


FIGURE 5
Frequency distribution of lysine concentration of PR-25 across seven station-years based on the average of biological replicates for each line in each station-year (two biological replicates in 2019 Sutherland and three biological replicates for the rest of the station-years), and the averaged lysine concentration for the parents of PR-25 in each station-year.

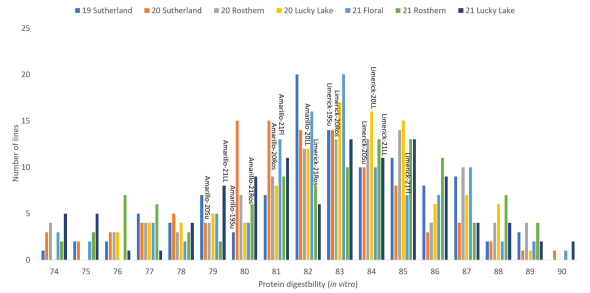


FIGURE 6
Frequency distribution of protein digestibility (*in vitro*) of PR-25 across seven station-years based on the average of biological replicates for each line in each station-year (two biological replicates in 2019 Sutherland and three biological replicates for the rest of the station-years), and the averaged protein digestibility (*in vitro*) for the parents of PR-25 in each station-year.

in between; Lys-QTL-3 was found on chromosome 1 and it explained 10% of the phenotypic variation, its flanking markers were Chr3LG5_424086163/Chr3LG5_437233435 and there were 2 markers in between. Chr4LG4_206035753, Chr4LG4_326486541, Chr3LG5_437233435 were the loci within the peak region of each of these three lysine associated QTLs. Two QTLs were identified associated with IVPD. IVPD-QTL-1 was found on chromosome 2 and it explained 10% of the phenotypic variation, its flanking markers were Chr2LG1_285985643/Chr2LG1_290867919 and there were 6 markers in between; IVPD-QTL-2 was found on chromosome 4 and it explained 11% of the phenotypic variation, its flanking markers were Chr1LG6_36689547/Chr1LG6_71617678 and there were 12 markers in between. Chr2LG1_291265214, Chr1LG6_41413580 were the loci within the peak region of each of these two *in vitro* protein digestibility associated QTLs. In addition, a set of QTL analyses were conducted in each station-year for methionine + cysteine concentration, tryptophan concentration, lysine concentration and protein digestibility, as shown in Table 5. The positions of identified QTLs, based on the individual QTL analysis from each station-year, were presented in Supplementary Figures 1–4.

Discussion

Development of pea varieties with high seed protein concentration and quality is necessary to fulfill the growing plant-based protein demand. To accomplish this goal, a better understanding of the underlying genetic control of the protein related traits is required.

Pea recombinant inbred line population PR-25, derived from the cross of CDC Amarillo and CDC Limerick, was developed specifically for the study of protein related traits. The variation in concentration of individual amino acids between CDC Amarillo and CDC Limerick ranged from 10 to 27%. They were significantly different in the concentration of almost all amino acids, except for alanine. These ensured sufficient diversity within PR-25 population to identify amino acid related QTLs. CDC Amarillo and CDC Limerick are widely grown in western Canada for their good yield and protein concentration, therefore, PR-25 is an ideal population for the research of protein-quality traits and could potentially avoid tradeoff between favorable traits.

HPLC analysis of individual amino acids is a destructive method which requires protein digestion steps. However, the unique attributes of some amino acids increase the complexity of the digestion steps. In comparison to HPLC analysis, NIR analysis is a non-destructive, high throughput and cost-efficient method to assess amino acid profile for pea protein. Amino acid profile of a thousand samples could be assessed in a week by using NIR analysis, while it may take months if using HPLC analysis. In the current study, the calibration formula developed based on HPLC quantification of amino acids in pea genotypes was used for NIR based prediction of amino acids. For most amino acids, their correlation coefficient values (r) were above 0.9. For the limiting amino acids, such as methionine, cystine and tryptophan, their r values were also acceptable, which were 0.733, 0.833 and 0.855, respectively.

10.3389/fpls.2023.1083086

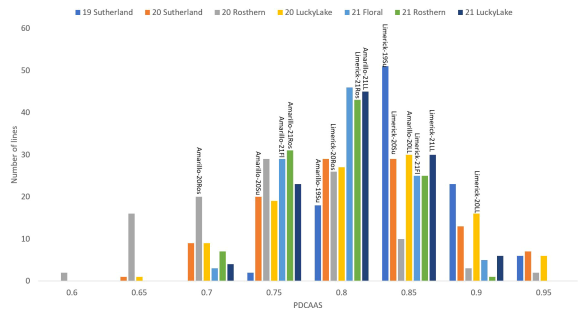


FIGURE 7
Frequency distribution of PDCAAS of PR-25 across seven station-years based on the average of biological replicates for each line in each station-year (two biological replicates in 2019 Sutherland and three biological replicates for the rest of the station-years), and the averaged PDCAAS for the parents of PR-25 in each station-year.

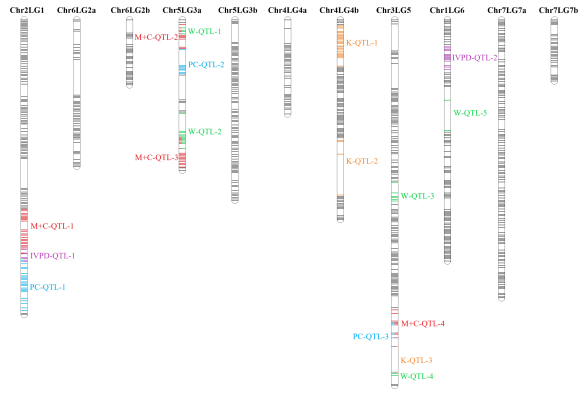


FIGURE 8
Combined QTL analysis in PR-25 reveals the QTLs associated for protein concentration (PC) (Zhou et al., 2022), *in vitro* protein digestibility (IVPD), lysine concentration (K), tryptophan concentration (W) and methionine + cysteine concentration (M+C). Combined analysis was based on averaged phenotypic data from station-years in which the QTL was significant (minimum of 3 of 7 station-years).

Pea protein is limited in sulfur amino acids and is abundant in lysine, thus complements cereal protein to provide complete plant protein. There is a debate on whether the improvements of limiting amino acids in certain crops are necessary since the other option would always be paired up with a different crop to provide complete nutrition. Though it seems less cost-effective to enhance the nutritional attributes of a single crop, this approach has several benefits. When introducing more ingredients into a food product, there is an increased risk of increasing the allergenicity of the product. Secondly, having more ingredients in a food product often means more food processing steps, including masking of unpleasant flavors, or more addition of food coloring which increases the cost. Meanwhile, over-processing is an issue that causes a decline in consumer acceptance. There is increased willingness among consumers to opt for products with fewer ingredients, driven by their consciousness for more healthy diets. Furthermore, climate change has different impacts on crops, and some are more severely influenced (Raza et al., 2019). The improvement of protein quality in single crop would contribute to food security and provide strength for agricultural businesses.

In PR-25, the averaged *in vitro* PDCAAS of 7 station-years ranged from 0.73 to 0.94. The PDCAAS of pea protein isolates was reported as 0.86 for children and 0.93 for adults by the FDA (2019). Some lines in PR-25 had lower PDCAAS than what was reported by the FDA. However, the measurements were conducted on whole pea seeds in the current study, while FDA measured PDCAAS on protein isolates. Since the fractionation process had positive effects on PDCAAS *via* improving the protein digestibility, whole seed samples were expected to have a lower PDCAAS than protein concentrates or isolates (Rivera Del Rio et al., 2022).

Some variation in amino acid profile was detected among station-years. These variations can be attributed to the effects of environments, and the interaction between genotype and environment. These abiotic factors impact individual amino acids differently and some amino acids, for instance, arginine and phenylalanine, had larger variations across station-years compared to others, which would lead to a change of their proportion in pea protein. However, these proportional changes

TABLE 5 Detected quantitative trait loci (QTL) for methionine + cysteine, tryptophan, lysine concentration, and *in vitro* protein digestibility in PR-25 using composite interval mapping (CIM).

Methionine + Cysteine - QTLs							
Name of QTL	Chromosome/ Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
Met+Cys-QTL-1	Chr2LG1	Combined	79.3-97.3/86.6	Chr2LG1_244771437/ Chr2LG1_287501555	6.4	15	0.006
		20Rosthern	87.7-101.2/89.0	Chr2LG1_257899664/ Chr2LG1_457180999	3.6	9	0.008
		20LuckyLake	88.3-94.8/89.6	Chr2LG1_259390540/ Chr2LG1_285985643	3.5	11	0.008
		21Floral	84.4-96.1/87.0	Chr2LG1_259006623/ Chr2LG1_287178718	3.8	10	0.006
		21LuckyLake	82.5-103.1/94.0		7.5	18	0.009

(Continued)

10.3389/fpls.2023.1083086

TABLE 5 Continued

Methionine + Cysteine - QTLs							
Name of QTL	Chromosome/ Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
				Chr2LG1_252806647/ Chr2LG1_293686819			
Met+Cys-QTL-2	Chr5LG3a	Combined	1.8-11.6/4.5	Chr5LG3_4173823/ Chr5LG3_15801800	4.1	11	-0.005
		20Sutherland	16.8-31.6/26.4	Chr5LG3_23306250/ Chr5LG3_47058351	4.2	14	-0.01
		21Floral	1.0-12.2/5.3	Chr5LG3_1308458/ Chr5LG3_15801800	3.8	11	-0.007
		21LuckyLake	1.5-5.9/3.3	Chr5LG3_4173823/ Chr5LG3_7509381	3.2	7	-0.005
Met+Cys-QTL-3	Chr5LG3a	Combined	48.4-60.9/59.1	Chr5LG3_101924498/ Chr5LG3_137457380	7.3	16	-0.006
		20Rosthern	46.9-61.1/56.2	Chr5LG3_83086224/ Chr5LG3_146553297	4.4	12	-0.008
		20LuckyLake	57.2-61.3/57.6	Chr5LG3_137457380/ Chr5LG3_146553297	3.5	11	-0.009
		21LuckyLake	48.4-61.1/61.0	Chr5LG3_101924498/ Chr5LG3_146553297	6.9	17	-0.008
		21Floral	56.7-57.6/57.2	Chr5LG3_131198978/ Chr5LG3_137724142	3.1	8	-0.005
Met+Cys-QTL-4	Chr3LG5	Combined	118.3-142.1/130.8	Chr3LG5_120117355/ Chr3LG5_408080154	4.4	10	-0.005
		20Sutherland	113.7-129.4/124.3	Chr3LG5_236004028/ Chr3LG5_422388773	3.5	9	-0.006
		20Rosthern	114.9-143.1/120.4	Chr3LG5_236004028/ Chr3LG5_437233435	7.1	20	-0.01
		21LuckyLake	135.3-149.5/141.5	Chr3LG5_424086163/ Chr3LG5_459895536	4.8	12	-0.007
Tryptophan – QTLs							
Name of QTL	Chromosome/Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
Trp-QTL-1	Chr5LG3a	Combined	2.0-10.9/4.5	Chr5LG3_5127342/ Chr5LG3_7509381	4.9	8	-0.001
		20Sutherland	5.2-11.6/9.9	Chr5LG3_7509381/ Chr5LG3_17377451	4.6	13	-0.002
		20Rosthern	3.2-11.0/4.3	Chr5LG3_5127342/ Chr5LG3_17355584	4.3	14	-0.002
		21LuckyLake	1.9-19.3/4.3	Chr5LG3_4173823/ Chr5LG3_23359112	5.5	12	-0.003
Trp-QTL-2	Chr5LG3a	Combined	56.5-60.9/59.1	Chr5LG3_67663653/ Chr5LG3_112710798	4.3	13	-0.001
		20Rosthern	47.7-60.5/57.4	Chr5LG3_459165288/ Chr5LG3_146553297	4.7	11	-0.002
		20LuckyLake	56.7-61.0/59.3	Chr5LG3_135166280/ Chr5LG3_146553297	3.4	8	-0.001
		21Floral	43.8-57.4/47.1	Chr5LG3_84309239/ Chr5LG3_137724142	4.5	11	-0.003

(Continued)

10.3389/fpls.2023.1083086

TABLE 5 Continued

Methionine + Cysteine - QTLs							
Name of QTL	Chromosome/ Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
		21LuckyLake	52.2-61.0/57.5	Chr5LG3_112710798/ Chr5LG3_146553297	5.4	12	-0.003
Trp-QTL-3	Chr3LG5	Combined	62.5-75.4/65.1	Chr3LG5_185794949/ Chr3LG5_198663551	3.8	9	-0.001
		20Sutherland	56.1-72.8/64.5	Chr3LG5_152019637/ Chr3LG5_197202364	4.8	11	-0.002
		20LuckyLake	62.5-74.8/66.9	Chr3LG5_163108809/ Chr3LG5_198663551	4.4	12	-0.002
		21LuckyLake	62.5-68.3/63.8	Chr3LG5_163108809/ Chr3LG5_185794949	3.8	8	-0.002
Trp-QTL-4	Chr3LG5	Combined	135.4-143.2/139.7	Chr3LG5_424086163/ Chr3LG5_455814220	3.4	8	-0.015
		20Sutherland	126.3-145.7/138.5	Chr3LG5_13646657/ Chr3LG5_455814220	4.6	13	-0.002
		20Rosthern	136.0-149.5/148.4	Chr3LG5_424086163/ Chr3LG5_459895536	3.7	9	-0.002
		20LuckyLake	118.4-124.6/120.5	Chr3LG5_120117355/ Chr3LG5_24108451	3.1	9	-0.002
		21LuckyLake	136.0-148.5/140.5	Chr3LG5_424086163/ Chr3LG5_455088720	4.3	11	-0.003
Trp-QTL-5	Chr1LG6	Combined	34.6-50.4/42.1	Chr1LG6_132953926/ Chr1LG6_233117537	4.6	9	0.001
		20Rosthern	27.0-47.2/43.5	Chr1LG6_97211857/ Chr1LG6_240693356	4.1	10	0.003
		20LuckyLake	45.1-49.6/47.1	Chr1LG6_233117537/ Chr1LG6_261967742	3.4	8	0.002
		21Floral	41.3-61.9/48.6	Chr1LG6_233117537/ Chr1LG6_305454704	4.2	10	0.003
Lysine – QTLs							
Name of QTL	Chromosome/Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
Lys-QTL-1	Chr4LG4b	Combined	3.6-18.9/8.3	Chr4LG4_185310109/ Chr4LG4_218381712	6.7	21	-0.02
		20Sutherland	7.8-13.2/10.8	Chr4LG4_194957049/ Chr4LG4_207119930	5.3	17	-0.01
		20LuckyLake	11.8-21.4/13.1	Chr4LG4_203961784/ Chr4LG4_226603806	4.9	14	-0.01
		21Floral	2.4-19.0/8.3	Chr4LG4_185310109/ Chr1LG6_334873830	7.7	20	-0.03
Lys-QTL-2	Chr4LG4b	Combined	52.3-60.8/55.5	Sc02659_148875/ Chr4LG4_417303831	3.5	15	0.02
		20Rosthern	45.15-55.3/50.4	Chr4LG4_316816169/ Chr4LG4_326486541	3.2	13	0.01
		21Floral	50.0-73.7/58.6	Chr4LG4_438891008/ Chr4LG4_418348946	3.7	16	0.03
		21LuckyLake	55.3-76.2/60.2	Chr4LG4_326486541/ Chr4LG4_438079536	3.1	12	0.01

(Continued)

10.3389/fpls.2023.1083086

TABLE 5 Continued

Methionine + Cysteine - QTLs							
Name of QTL	Chromosome/ Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
Lys-QTL-3	Chr3LG5	Combined	134.2-144.8/141.6	Chr3LG5_424086163/ Chr3LG5_437233435	4.3	10	-0.01
		20LuckyLake	107.3-118.4/113.2	Chr3LG5_257364623/ Chr3LG5_120117355	3.8	8	-0.01
		21Floral	125.6-150.1/140.2	Chr3LG5_13646657/ Chr3LG5_459895536	4.1	9	-0.02
		21LuckyLake	124.1-146.5/138.9	Chr1LG6_347327176/ Chr3LG5_455088720	4.1	10	-0.02
Protein Digestibility (*in vitro*)- QTLs							
Name of QTL	Chromosome/Linkage group	Station-year	Position/Peak (cM)	Flanking markers	LOD score	R^2 (%)	Additive effect
IVPD-QTL-1	Chr2LG1	Combined	95.5-100.5/98.4	Chr2LG1_285985643/ Chr2LG1_290867919	3.3	10	1.27
		21Floral	93.5-103.8/98.6	Chr2LG1_271201976/ Chr2LG1_295431002	4.1	13	1.55
		21Rosthern	92.0-101.7/94.5	Chr2LG1_267076305/ Chr2LG1_457180999	3.3	10	1.26
		21LuckyLake	94.0-98.6/94.5	Chr2LG1_278779827/ Chr2LG1_291265214	3.1	10	1.196
IVPD-QTL-2	Chr1LG6	Combined	10.9-20.5/14.4	Chr1LG6_36689547/ Chr1LG6_71617678	3.5	11	-1.12
		19Sutherland	0.2-1.9/0.9	Chr1LG6_24608817/ Chr1LG6_27085223	3.1	9	-1.03
		20Sutherland	9.7-23.8/15.5	Chr1LG6_33266075/ Chr1LG6_84517828	4.3	14	-1.34
		20Rosthern	9.0-24.0/14.4	Chr1LG6_30581935/ Chr1LG6_87699075	3.3	10	-1.00
		21Rosthern	11.0-21.3/14.6	Chr1LG6_34300922/ Chr1LG6_81358908	3.4	10	-1.27
		21LuckyLake	9.7-21.9/10.9	Chr1LG6_33266075/ Chr1LG6_81358908	3.8	12	-1.33

Combined analysis was based on averaged phenotypic data from station-years in which the QTL was significant. Individual QTL analysis was based on averaged phenotypic data of biological replicates in each station-year. Additive effects were calculated as the average performance of lines carrying A allele from CDC Amarillo minus the average performance of lines carrying B allele from CDC Limerick.

in some amino acids only had limited impact on the overall amino acid profile as the profile didn't change significantly across station-years.

All 18 amino acids assessed in this study were found positively correlated with total protein concentration, in most cases with correlation coefficients above 0.8. Meanwhile, close proximities or overlaps were found on several protein-related traits. Met+Cys-QTL-1 was found adjacent to PC-QTL-1 on chromosome 2, where IVPD-QTL-1 was located in between. Trp-QTL-1 was found within Met+Cys-QTL-2 on one end of chromosome 5 linkage group 3a and Trp-QTL-2 overlapped with Met+Cys-QTL-3 on another end. Met+Cys-QTL-4, PC-QTL-3, Lys-QTL-3 and Trp-QTL-4 were found overlapped on one end of Chromosome 3. Though overlaps were found among these protein-related traits, it didn't necessarily lead to strong correlations. For instance, the correlation between protein concentration and methionine + cysteine concentration was 0.55, between IVPD and protein concentrations was 0.10, between IVPD and methionine + cysteine concentration was 0.17, between tryptophan concentration and methionine + cysteine concentration was 0.69. The main reason was that all protein associated traits assessed in this study are quantitative traits that are regulated by multiple loci. The identified QTL regions only explained small portions of the phenotypic variations and hence, though some of the QTL regions of these traits were in close proximity, their correlations were not very strong. Yet, within these overlapped QTL regions, there are potentials to find pleiotropic QTLs that control multiple protein quality traits of pea, as the research by Li et al. (2018) had identified

several pleiotropic QTLs that associated with multiple amino acids concentrations in soybeans.

QTL analyses on pea seed traits including protein and mineral concentration, seed yield, thousand seed weight, seed number per plant, have been conducted by several research programs in the past two decades and numerous QTL had been identified associated with these traits (Tar'an et al., 2004; Ma et al., 2017; Dissanayaka et al., 2020; Klein et al., 2020; Zhou et al., 2022). Protein digestibility and amino acid profile have also been studied for deeper understanding of the functional attributes of pea protein (Nosworthy & House, 2017; Cabuk et al., 2018). However, none of the previous publications was related to QTL analysis for pea protein digestibility and amino acids concentration, and the QTLs identified in this study provide valuable information of the underlying genetic control of these traits. Twelve loci were found within the peak regions of the QTLs identified in this study, three were associated with methionine + cysteine concentration, four were associated with tryptophan concentration, three were associated with lysine concentration, and two were associated with *in vitro* protein digestibility. The information of these loci will be beneficial for developing markers to facilitate the selection of high protein-quality varieties in pea breeding. Several lines in PR-25, including PR-25-2, PR-25-46-PR-25-53, PR-25-86 and PR-25-96, had high concentrations of methionine, cysteine and tryptophan, as well as good protein concentration and grain yield. These lines had consistently good performance in protein quality traits across 7 station-years without compromise on other valued traits. These lines could be the potential high protein quality varieties or could be used as parental materials to develop varieties with better performance.

Conclusion

The effect of genotype x environment interaction on the amino acid concentrations of pea cultivars is significant. The segregation pattern of amino acid concentrations in PR-25 population combined with the NIR-based predictions offers a possibility for high throughput selection of breeding lines for amino acid concentrations. Three QTLs were found associated with methionine + cysteine concentration, four QTLs were found associated with the tryptophan concentration, three QTLs were found associated with lysine concentration and two QTLs were found associated with *in vitro* protein digestibility. Overlaps were found among protein-related traits on chromosome 2 and chromosome 5. These identified QTL regions have a potential for use in marker-assisted selection of protein quality traits.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

TW, BT and KG conceptualized the study and secured funding for this study through a research grant from the Saskatchewan Ministry of Agriculture. KG developed the RIL population and conducted genotyping and linkage map construction. JZ conducted field trials, collected phenotypic data, conducted IVPD experiments, data curation, QTL analysis and wrote the manuscript. MN provided expert advice on IVPD experiments. JH designed methodology for amino acid profiling *via* NIR analysis. ZW conducted amino acid profiling. All authors contributed to the article and approved the submitted version.

Funding

This research was part of the research project Pea Protein 'Omics Determination (P-POD) supported by the Saskatchewan Ministry of Agriculture, Strategic Research Initiative (project # 20180436).

Acknowledgments

We thank Brent Barlow and the rest of the pulse crop breeding staff at the Crop Development Centre for planting, maintaining, and harvesting PR-25 from 2019 to 2021. We thank Perkin Elmer for providing access to the DA 7250 NIR spectrophotometer used in the current study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1083086/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Individual QTL analysis from each station-year in PR-25 reveals the QTLs associated for methionine + cysteine concentration. Individual QTL analysis was based on average phenotypic data of biological replicates in each station-year.

SUPPLEMENTARY FIGURE 2
Individual QTL analysis from each station-year in PR-25 reveals the QTLs associated for tryptophan concentration. Individual QTL analysis was based on average phenotypic data of biological replicates in each station-year.

SUPPLEMENTARY FIGURE 3
Individual QTL analysis from each station-year in PR-25 reveals the QTLs associated for lysine concentration. Individual QTL analysis was based on average phenotypic data of biological replicates in each station-year.

SUPPLEMENTARY FIGURE 4
Individual QTL analysis from each station-year in PR-25 reveals the QTLs associated for *in vitro* protein digestibility. Individual QTL analysis was based on average phenotypic data of biological replicates in each station-year.

References

Alcorta, A., Porta, A., Tárrega, A., Alvarez, M. D., and Vaquero, M. P. (2021). Foods for plant-based diets: Challenges and innovations. *Foods* 10, 293. doi: 10.3390/foods10020293

Aschemann-Witzel, J., Gantriis, R. F., Fraga, P., and Perez-Cueto, F. J. A. (2021). Plant-based food and protein trend from a business perspective: Markets, consumers, and the challenges and opportunities in the future. *Crit. Rev. Food Sci. Nutr.* 61 (18), 3119–3128. doi: 10.1080/10408398.2020.1793730

Bessada, S. M. F., Barreira, J. C. M., and Oliveira, B. P. P. (2019). Pulses and food security: Dietary protein, digestibility, bioactive and functional properties. *Trends Food Sci. Technol.* 93, 53–68. doi: 10.1016/j.tifs.2019.08.022

Cabuk, B., Nosworthy, M. G., Stone, A. K., Korber, D. R., Tanaka, T., House, J. D., et al. (2018). Effect of fermentation on the protein digestibility and levels of non-nutritive compounds of pea protein concentrate. *Food Technol. Biotechnol.* 56 (2), 257–264. doi: 10.17113/ftb.56.02.18.5450

Dissanayaka, D. N., Gali, K. K., Jha, A. B., Lachagari, V. B. R., and Warkentin, T. D. (2020). Genome-wide association study to identify single nucleotide polymorphisms associated with fe, zn, and Se concentration in field pea. *Crop Sci.* 60, 2070–2084. doi: 10.1002/csc2.20161

Dołowy, M., and Pyka, A. (2014). Application of TLC, HPLC and GC methods to the study of amino acid and peptide enantiomers: A review. *Biomed. Chromatogr.* 28 (1), 84–101. doi: 10.1002/bmc.3016

Duodu, K. G., Nunes, A., Delgadillo, I., Parker, M. L., Mills, E. N. C., Belton, P. S., et al. (2002). Effect of grain structure and cooking on sorghum and maize *in vitro* protein digestibility. *J. Cereal Sci.* 35 (2), 161–174. doi: 10.1006/jcrs.2001.0411

Estell, M., Hughes, J., and Grafenauer, S. (2021). Plant protein and plant-based meat alternatives: Consumer and nutrition professional attitudes and perceptions. *Sustainability.* 13, 1478. doi: 10.3390/su13031478

FDA (2019) *GRAS notification for roquette freres pea protein isolate.* Available at: https://www.fda.gov/media/134207/download#:~:text=Pea%20Protein%20Isolate%20is%20a,and%2093%25%20PDCAAS%20for%20adults (Accessed 24 October 2022).

Gali, K. K., Sackville, A., Tafesse, E. G., Lachagari, V. B. R., McPhee, K., Hybl, M., et al. (2019). Genome-wide association mapping for agronomic and seed quality traits of field pea (*Pisum sativum* l.). *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01538

Hang, J., Shi, D., Neufeld, J., Bett, K. E., and House, J. D. (2022). Prediction of protein and amino acid contents in whole and ground lentils using near-infrared reflectance spectroscopy. *LWT.* 165, 113669. doi: 10.1016/j.lwt.2022.113669

Hoffmann, B., Munch, S., Schwagele, F., Neusu, C., and Jira, W. (2017). A sensitive HPLC-MS/MS screening method for the simultaneous detection of lupine, pea, and soy proteins in meat products. *Food Control* 71, 200–209. doi: 10.1016/j.foodcont.2016.06.021

Jiang, G. (2020). Comparison and application of non-destructive NIR evaluations of seed protein and oil content in soybean breeding. *Agron.* 10 (1), 77. doi: 10.3390/agronomy10010077

Kambhampati, S., Li, J., Evans, B. S., and Allen, D. K. (2019). Accurate and efficient amino acid analysis for protein quantification using hydrophilic interaction chromatography coupled tandem mass spectrometry. *Plant Methods* 15, 46. doi: 10.1186/s13007-019-0430-z

Ketnawa, S., and Ogawa, Y. (2021). *In vitro* protein digestibility and biochemical characteristics of soaked, boiled and fermented soybeans. *Sci. Rep.* 11, 14257. doi: 10.1038/s41598-021-93451-x

Klein, A., Houtin, H., Rond-Coissieux, C., Naudet-Huart, M., Touratier, M., Marget, P., et al. (2020). Meta-analysis of QTL reveals the genetic control of yield-related traits and seed protein content in pea. *Sci. Rep.* 10, 15925. doi: 10.1038/s41598-020-72548-9

Leterme, P., Monmart, T., and Baudart, E. (1990). Amino acid composition of pea (Pisum sativum) proteins and protein profile of pea flour. *J. Sci. Food. Agric.* 53, 107–110. doi: 10.1002/jsfa.2740530112

Li, X., Tian, R., Kamala, S., Du, H., Li, W., Kong, Y., et al. (2018). Identification and verification of pleiotropic QTL controlling multiple amino acid contents in soybean seed. *Euphytica.* 214, 93. doi: 10.1007/s10681-018-2170-y

Ma, Y., Coyne, C. J., Grusak, M. A., Mazourek, M., Cheng, P., Main, D., et al. (2017). Genome-wide SNP identification, linkage map construction and QTL mapping for seed mineral concentrations and contents in pea (Pisum sativum l.). *BMC Plant Biol.* 17, 43. doi: 10.1186/s12870-016-0956-4

Nosworthy, M. G., Franczyk, A., Zimoch-Korzycka, A., Appah, P., Utioh, A., Neufeld, J., et al. (2017). Impact of processing on the protein quality of Pinto bean (Phaseolus vulgaris) and buckwheat (Fagopyrum esculentum moench) flours and blends, as determined by in vitro and in vivo methodologies. *J. Agric. Food Chem.* 65 (19), 3919–3925. doi: 10.1021/acs.jafc.7b00697

Nosworthy, M. G., and House, J. D. (2017). Factors influencing the quality of dietary proteins: Implications for pulses. *Cereal Chem.* 94 (1), 49–57. doi: 10.1094/CCHEM-04-16-0104-FI

Pilorge, E., Kezeya, B., Stauss, W., Muel, F., and Mergenthaler, M. (2021). Pea and rapeseed acreage and land use for plant-based meat alternatives in the EU. *OCL.* 28, 54. doi: 10.1051/ocl/2021037

Ravanel, S., Gakiere, B., Job, D., and Douce, R. (1998). The specific features of methionine biosynthesis and metabolism in plants. *PNAS.* 95 (13), 7805–7812. doi: 10.1073/pnas.95.13.7805

Raza, A., Razzaq, A., Mehmood, S. S., Zou, X., Zhang, X., Lv, Y., et al. (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. *Plants* 8, 34. doi: 10.3390/plants8020034

Rivera Del Rio, A., Boom, R. M., and Janssen, A. E. M. (2022). Effect of fractionation and processing conditions on the digestibility of plant proteins as food ingredients. *Foods* 11 (6), 870. doi: 10.3390/foods11060870

Sabate, J., and Soret, S. (2014). Sustainability of plant-based diets: Back to the future. *Am. J. Clin. Nutr.* 100, 476–482. doi: 10.3945/ajcn.113.071522

Schaafsma, G. (2005). The protein digestibility-corrected amino acid score (PDCAAS)–a concept for describing protein quality in foods and food ingredients: A critical review. *J. AOAC Int.* 88 (3), 988–994. doi: 10.1093/jaoac/88.3.988

Sharma, G., Attri, S. V., Behra, B., Bhisikar, S., Kumar, P., Tageja, M., et al. (2014). Analysis of 26 amino acids in human plasma by HPLC using AQC as derivatizing agent and its application in metabolic laboratory. *Amino Acids* 46, 1253–1263. doi: 10.1007/s00726-014-1682-6

Shi, D. (2021) *Effects of genotype, environment and genotype by environment interaction on soybean protein and amino acid content using near-infrared reflectance spectroscopy* (Winnipeg: University of Manitoba). Available at: https://mspace.lib.umanitoba.ca/bitstream/handle/1993/35965/Shi_Da.pdf?sequence=1&isAllowed=y (Accessed July 28th, 2022).

Shi, D., Hang, J., Neufeld, J., Zhao, S., and House, J. D. (2022). Estimation of crude protein and amino acid contents in whole, ground and defatted ground soybeans by different types of near-infrared (NIR) reflectance spectroscopy. *J. Food Compos. Anal.* 111, 104601. doi: 10.1016/j.jfca.2022.104601

Srivastava, S., Neerubala,, Singh, S., and Shamim, M. Z. (2015). Nutritional composition of weaning food using malted cereal and pulses flour for infants. *Int. J. Pure App. Biosci.* 3 (1), 171–185.

Tar'an, B., Warkentin, T., Somers, D., Miranda, D., Vandenberg, A., Blade, S., et al. (2004). Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (Pisum sativum l.). *Euphytica.* 136, 297–306. doi: 10.1023/B:EUPH.0000032721.03075.a0

Tavano, O. L., Neves, V. A., and Junior, S. I. S. (2016). *In vitro* versus *in vivo* protein digestibility techniques for calculating PDCAAS (protein digestibility-corrected amino acid score) applied to chickpea fractions. *Int. Food Res. J.* 89 (1), 756–763. doi: 10.1016/j.foodres.2016.10.005

Taylor, S. L., Marsh, J. T., Koppelman, S. J., Kabourek, J. L., Johnson, P. E., and Baumert, J. L. (2021). A perspective on pea allergy and pea allergens. *Trends Food Sci. Technol.* 116, 186–198. doi: 10.1016/j.tifs.2021.07.017

Van Loo, E. J., Caputo, V., Jayson, L., and Lusk, J. L. (2020). Consumer preferences for farm-raised meat, lab-grown meat, and plant-based meat alternatives: Does information or brand matter? *Food Policy* 95) 101931. doi: 10.1016/j.foodpol.2020.101931

Wang, S., Basten, C. J., and Zeng, Z. B. (2012). *Windows QTL cartographer 2.5. department of statistics* (Raleigh, NC: North Carolina State University).

Zhou, J., Gali, K. K., Jha, A. B., Tar'an, B., and Warkentin, T. D. (2022). Identification of quantitative trait loci associated with seed protein concentration in a pea recombinant inbred line population. *Genes* 13, 1531. doi: 10.3390/genes13091531

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 17 March 2023
DOI 10.3389/fpls.2023.1092493

Check for updates

OPEN ACCESS

EDITED BY
Eric Von Wettberg,
University of Vermont, United States

REVIEWED BY
Gyanendra Kumar Rai,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India
Thi My Linh Hoang,
Queensland University of Technology, Australia

*CORRESPONDENCE
Bunyamin Tar'an
bunyamin.taran@usask.ca

SPECIALTY SECTION
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

RECEIVED 08 November 2022
ACCEPTED 03 March 2023
PUBLISHED 17 March 2023

CITATION
Jahan TA, Kalve S, Belak Z, Eskiw C and Tar'an B (2023) Iron accumulation and partitioning in hydroponically grown wild and cultivated chickpea (*Cicer arietinum* L). *Front. Plant Sci.* 14:1092493. doi: 10.3389/fpls.2023.1092493

COPYRIGHT
© 2023 Jahan, Kalve, Belak, Eskiw and Tar'an. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Iron accumulation and partitioning in hydroponically grown wild and cultivated chickpea (*Cicer arietinum* L)

Tamanna A. Jahan[1], Shweta Kalve[1], Zachery Belak[2], Christopher Eskiw[2] and Bunyamin Tar'an[1]*

[1]Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Saskatoon, SK, Canada, [2]Department of Food and Bioproduct Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Saskatoon, SK, Canada

Chickpea (*Cicer arietinum* L.) is a staple food in many developing countries where iron (Fe) deficiency often occurs in their population. The crop is a good source of protein, vitamins, and micronutrients. Fe biofortification in chickpea can be part of long-term strategy to enhance Fe intake in human diet to help to alleviate Fe deficiency. To develop cultivars with high Fe concentration in seeds, understanding the mechanisms of absorption and translocation of Fe into the seeds is critical. An experiment was conducted using a hydroponic system to evaluate Fe accumulation in seeds and other organs at different growth stages of selected genotypes of cultivated and wild relatives of chickpea. Plants were grown in media with Fe zero and Fe added conditions. Six chickpea genotypes were grown and harvested at six different growth stages: V3, V10, R2, R5, R6, and RH for analysis of Fe concentration in roots, stems, leaves, and seeds. The relative expression of genes related to Fe-metabolism including *FRO2*, *IRT1*, *NRAMP3*, *V1T1*, *YSL1*, *FER3*, *GCN2*, and *WEE1* was analyzed. The results showed that the highest and lowest accumulation of Fe throughout the plant growth stages were found in the roots and stems, respectively. Results of gene expression analysis confirmed that the *FRO2* and *IRT1* were involved in Fe uptake in chickpeas and expressed more in roots under Fe added condition. All transporter genes: *NRAMP3*, *V1T1*, *YSL1* along with storage gene *FER3* showed higher expression in leaves. In contrast, candidate gene *WEE1* for Fe metabolism expressed more in roots under Fe affluent condition; however, *GCN2* showed over-expression in roots under Fe zero condition. Current finding will contribute to better understanding of Fe translocation and metabolism in chickpea. This knowledge can further be used to develop chickpea varieties with high Fe in seeds.

KEYWORDS

Fe accumulation, Fe translocation, gene expression, chickpea, biofortification

1 Introduction

Globally, the population is increasing and is projected to reach 9.1 billion by 2050. As a result, around 70% higher demand for food is needed during the same period (WSFS, 2009). This situation challenges support the rapid growth of the global economy. Although added food production has been generated through high yielding cultivars of staple crops, hundreds of millions of people still suffer from micronutrient deficiency (Satterthwaite et al., 2010). Among a variety of micronutrient deficiencies, Fe deficiency is one of the most common and extensive malnutrition world-wide (Baltussen et al., 2004). Globally, Fe deficiency is considered the 6th highest cause of mortality and top 10 health challenges in present day (Briat, 2011; Campion et al., 2013). Fe deficiency may occur throughout a lifetime if the diets are mainly based on cereals and legumes (WHO, 2002). Anemia, which is mainly due to Fe deficiency, affects around 2 billion people in the world (Briat, 2011). By increasing the amount of Fe in the diet, Fe deficiency can be overcome (WHO, 2005). However, solving Fe deficiency in developing countries is difficult as the population relies mostly on staple food crops as sources of micronutrients which are inherently low in Fe (Gómez-Galera et al., 2010). To address this problem, micronutrient dense cultivars with high yielding capacity are needed (Zhu et al., 2007). Chickpea is an important legume crop annually grown over 14 million hectares in 59 countries around the globe. It is an inexpensive, high-quality source of protein along with vitamins and minerals including Fe. In many countries such as India and the Middle East, chickpea is a staple food crop and a major component of the diets (Zhu et al., 2005; Millán et al., 2015). Chickpea consumption has been rising in many countries around the world (Tan et al., 2017). As such, chickpea has a potential to serve as a vehicle to address Fe deficiency in human. Chickpea cultivars with high Fe concentration in seeds are needed.

Fe is commonly found in an abundant amount in the earth's crust; however, its limited solubility results in limited uptake by plants. Consequently, only low amount of Fe is accumulated in the edible parts of the plants (Zuo and Zhang, 2011). Besides the uptake of Fe from soil to the roots, Fe accumulation in the seeds depends on the translocation of Fe to the vegetative tissues and loading it into the seeds. Fe concentration in plant organs also varies based on the species and cultivars. Plants' ability to acquire and accumulate Fe in different tissues is under genetic Fe zero (Sankaran and Grusak, 2014). Fe deficiency in plants will result in low levels of Fe in the seeds that ultimately affect human nutrition (Zuo and Zhang, 2011).

One strategy to mitigate Fe deficiency in human population is to increase Fe concentration and its bioavailability in the edible part of the plants (Briat, 2011). Fe biofortification is a long-term approach to improve Fe nutrition. To reach this goal, genetic modifications and plant breeding offer a possibility for improving Fe amount and bioavailability of crops (Briat, 2011).

In the efforts toward generating Fe biofortified crops, several genes associated with Fe metabolism in plants have been identified in different species. In nongraminaceous plant species, the dominant genes responsible for Fe uptake are *FRO2 (Ferric-chelate reductase oxidase gene)* and *IRT1 (Fe-regulated transporter gene)*. These two-uptake genes strongly responded under Fe zero environment (Eide et al., 1996; Robinson et al., 1999). Genes responsible for Fe translocation such as *YELLOW STRIPE 1-like (YSL)* gene family (Koike et al., 2004) has also been identified. *Natural resistance-associated macrophage protein 3 (NRAMP3)* and *vacuolar Fe transporter 1 (V1T1)* are two core genes that are involved in subcellular transportation of Fe (Thomine et al., 2000; Lanquar et al., 2005; Kim et al., 2006). *Ferritin (FER)* gene is responsible for high-capacity Fe storage and sequestration (Petit et al., 2001).

One obstacle to Fe biofortification is the lack of knowledge of how Fe is accumulated into the seeds (Sankaran and Grusak, 2014). As a result, uncertainty occurs when it comes to select the appropriate pathways or genes to target for selection or modifications in genetic improvement program. This problem supports the urgency to understand the mechanism of Fe partitioning from the roots to the seeds for breeding of Fe rich crops.

Using genetic transformation technique, Lee and An (2009) reported that overexpression of Fe acquisition gene *IRT1* increased the Fe concentration by 1.1-folds in brown rice grain. In rice and wheat, the overexpression of nicotinamine synthase (*NAS)* improved Fe concentration by up to 2-folds (Johnson et al., 2011; Beasley et al., 2019). Furthermore, the expression of *GmFERH1* gene increased Fe concentration up to 3-folds in both brown and polished rice. The overexpression of *TaVIT2* in wheat and *OsVIT1* or *OsVIT2* in rice increased Fe amount by 2 and 1.3-folds, respectively (Zhang et al., 2012; Connorton et al., 2017). Successful results also found by the introduction of transgene combination, such as the combined overexpression of *IRT1* and *PvFER1* increased up to 4-folds Fe in the endosperm of polished rice (Boonyaves et al., 2017). When the transgene combination of *PvFERRITIN*, *AtNAS1* and *Afphytase* expressed together resulting in a 6-folds increase in rice seed Fe concentration (Wirth et al., 2009). In field grown polished rice, another different multiple transgene combination (*HvNAS1*, *OsYSL2*, and *GmFERRITIN*) was reported to increase Fe amount by 4.4-folds (Masuda et al., 2013). Overexpression of nicotinamine synthase (*NAS*) gene resulted in increase nicotinamine (NA) amount in rice and wheat that ultimately increased Fe bioavailability in mice or Caco-2 cell culture model (Zheng et al., 2010; Beasley et al., 2019).

Information of Fe translocation, partitioning and accumulation to the seeds at different growth stages as well as the related genes associated with the process is needed to develop cultivars with high Fe concentration. To date, information on the dynamics of Fe accumulation in chickpea seeds is lacking. Therefore, the first step to increase Fe amount in chickpea seeds is to examine the mobilization of Fe accumulation in plant organs of diverse chickpea genotypes. The output of the study will help in the development of new breeding strategies to improve Fe concentration in chickpea seeds. The main objectives of this study were: 1) to evaluate Fe accumulation in organs at different growth stages of diverse chickpea genotypes, and 2) to evaluate the

expression levels of genes associated with Fe uptake, transportation, and accumulation into the seeds of chickpea.

2 Materials and methods

The research was conducted in a Fe zeroled chamber at the College of Agriculture and Bioresources, University of Saskatchewan. The whole experiment was repeated twice with four replications each time. The chamber was adjusted with 16 h, 22°C, and 8 h, 15°C Day-night regime. The light intensity of PAR was 220 µmol m^{-2} s^{-1} of photon flux density. Light was provided with a combination of florescent tubes and incandescent bulbs.

2.1 Plant materials

Six chickpea genotypes, namely CDC-551-1, CDC Verano, FLIP97-677C, Kalka 064, Sarik 067, and Cermi 075 were used in this study (Figure 1).

These genotypes were collected from the chickpea breeding program at the Crop Development Centre, University of Saskatchewan. Three genotypes (Figure 1) belong to the cultivated species (*Cicer arietinum*: CDC Verano, FLIP97-677C, and CDC-551-1), and the other three are wild species accession *(C. reticulatum*: Sarik 067 and Kalka 064 and *C. echinospermum*: Cermi 075). Prior analysis from field grown plants showed that these six genotypes had varying Fe concentrations in seeds (Table 1).

Plants were grown using hydroponic system in the phytotron chamber. Seeds were pre-germinated before transferring them to polyethylene containers with nutrient solution using hydroponic system (Grusak et al., 1990). Only plants from V3(3^{rd} multifoliate leaf has unfolded from the stem) and V10 (10^{th} multifoliate leaf has unfolded from the stem) growth stage were grown in both Fe added (Fe^{+}, 5 µM Fe (III)-EDDHA [ethylenediamine- *N*, N bis(*o*-hydroxyphenyl) acetic acid] and Fe zero [Fe^{-}, 0 µM Fe (III)-EDDHA] conditions. The other stages were only grown under Fe added condition to skip the chlorosis. The volume of the nutrient mixture was adjusted according to the number of plants grown in each container. For the vegetative stages (V3 and V10), twelve plants of each genotype were grown in each container. However, for the reproductive stages R2 (full bloom stage), R5 (early seed stage) and R6 (full seed stage) and full maturity stage (RH), four plants were grown at equal space on the container to minimize the tangle of roots. Plant samples were collected from the 2^{nd} week at the 3^{rd} node stage (first multifoliate leaf stage) to physiological maturity.

2.2 Media preparation

The hydroponically grown plants were provided with nutrient solution containing the following macronutrients: 3.6mM KNO_3, 2.4mM $Ca(NO_3)_2$, 0.3mM $NH_4H_2PO_4$, 0.6mM $MgSO_4$, 75µM $CaCl_2$, 75µM H_3BO_3, 6µM $MnSO_4$, 6µM ZnSO4, 1.5µM $CuSO_4$, 1.5µM H_2MoO_4, and 0.3µM $NiSO_4$, and was buffered with 3mM MES[2-(N-morpholino) ethane sulfonic acid] to maintain the pH between 5.5-6.0. Another buffer, pH down©, was used to lower the

FIGURE 1
Six diverse chickpea genotypes used for the study of Fe absorption and accumulation in the chickpea.

TABLE 1 Description of six chickpea genotypes with the seed size (g/1000seeds) and mean Fe concentration in seeds (μg g^{-1} ± SE) evaluated in the study of Fe absorption and accumulation in different organs.

No.	Name	Species	Type	Status	Seed size (g/1000 seeds)	Average Fe conc. in seeds (μg g^{-1} ± SE)
1.	CDC-551-1	*Cicer arietinum*	Desi	Breeding line	297	41.6 ± 0.8
2.	CDC Verano	*Cicer arietinum*	Kabuli	Cultivar	165	60.1 ± 0.4
3.	FLIP97-677C	*Cicer arietinum*	Kabuli	Breeding line	396	52.4 ± 0.4
4.	Kalka 064	*Cicer reticulatum*	Wild species	Germplasm	143	49.3 ± 0.3
5.	Sarik 067	*Cicer reticulatum*	Wild species	Germplasm	164	95.2 ± 0.8
6.	Cermi 075	*Cicer echinospermum*	Wild species	Germplasm	149	40.3 ± 0.2

Source of information and initial Fe concentration: Diapari et al., 2014; Mengistu (2016).

pH if needed. Fe solution was prepared according to Chaney and Bell (1987). Fe was added as 5μM Fe (III)-EDDHA per four plants according to Grusak (1994).

2.3 Experimental design

The Fe concentration, Fe amount and gene expression analysis in different organs of six different chickpea plants was done with four replications with two repeats in both Fe added and Fe zero conditions. In this study, seven growth stages: two vegetative stages (V3 = 3rd multifoliate leaf has unfolded, and V10 = full vegetative stage), three reproductive stages (R2 = full bloom, R5 = early seed, and R6 = fully developed seed) and one physiological maturity stage (RH = 90% of pods are golden brown) were used for measuring the Fe concentration (μg g^{-1}). Fe analysis was done for each organ separately including roots, stems, leaves, and mature seeds. Roots, stems, and leaves at the vegetative stage, and seeds at the reproductive and maturity stage were digested and prepared for Fe analysis using inductively couple plasma (ICP) –atomic emission spectrometry (iCAP 6500 series: Thermo Jarrell Ash Corp., Franklin, MA, USA) at plant biochemistry and molecular physiology lab, University of Saskatchewan. However, for gene expression analyses, only two vegetative stages (V3 = 3rd multifoliate leaf has unfolded, and V10 = full vegetative stage) and two reproductive growth stages (R2 = full bloom, and R5 = early seed) were used. The gene expression analysis was done in root and leaf using qPCR. For growth stages R2, R5, R6, and RH, one plant was harvested for each replication. However, at V3 growth stage, to get 0.5g tissue sample, six plants were harvested for each replication for measuring Fe concentration (μg g^{-1}) and gene expression analyses. For V10, three plants were harvested for each replication. For dry weight in roots, shoots, and roots to shoots (dry weight/dry weight) ratio, six and three plants were harvested for each replication for measuring root and shoot dry weight at V3 and V10 growth stages, respectively.

2.4 Data collection

2.4.1 Fe concentrations

Samples were prepared from roots, stems, leaves, and seeds. From V3 and V10 vegetative growth stages, three samples: roots, stems, and leaves were collected. Seeds were collected from the reproductive growth stages R5 and R6 as well as the physiological maturity growth stage (RH) along with roots, stems, and leaves. Pods were tagged by using different colored threads to ensure appropriate growth stage for collecting seeds. For the R5 growth stage, pods were harvested at 16 days after anthesis. However, for R6 and RH growth stages, pods were harvested at 24 and 32 days after anthesis. For each stage around 30 pods were harvested to get enough tissue samples. Roots were collected at each growth stage. Roots were rinsed (2.5 min each rinse) twice with aerated deionized water. The root samples were blotted dry and were placed in paper bags for oven drying and subsequent dry weight determination. Besides V3 and V10, per time point, a total of 4 plants were analyzed. All tissue samples were dried at 60°C followed by weighing. After weighing, each tissue samples were transferred to polycarbonate tubes and homogenized with geno grinder (SPEX^{TM} SamplePrep, 65 Liberty Street, Metuchen, NJ) to get the powdered samples (Waters and Grusak, 2008). Each sample was digested by taking 0.5g of dried tissue sample using 4mL of concentrated nitric acid and 2mL of perchloric acid at 200°C temperature followed by drying. Digests were suspended again in 1mL of 2M HNO_3 and after 1h was brought to 10mL with deionized water. The acids used were as trace metal grade (Fisher Scientific, Pittsburgh, Pennsylvania, USA) and the water was deionized *via* a MilliQ system (Millipore, Billerica, Massachusetts, USA). Samples were then analyzed for Fe concentration by using inductively couple plasma –atomic emission spectrometry. Fe amount from each tissue was calculated by multiplying the average Fe concentration from each tissue by the average tissue weight per plant at a given time point. Fe amount, root and shoot dry weight were measured from the average tissue weight of six and three plants at V3 and V10 vegetative growth stages, respectively.

2.4.2 Gene expression analysis

Young root and leaf samples were harvested and immediately put in liquid nitrogen and stored at -80 ° C until RNA extraction. Tissue samples were ground using sterilized mortar and pestle before RNA extraction. RNA was extracted and treated with DNase I using Qiagen RNeasy plant mini kit (Qiagen, Valencia, CA, USA). Thereafter, RNA quantity and purity was measured using an optical density reading at 260nm and the 260/280 and the 260/230 absorption ratios using NanoDrop 800 UV-vis spectrophotometer (Thermo Fisher Scientific Inc, USA). RNA integrity was measured on 1% agarose gel using MOPS buffer (3-morpholinopropane-1-sulfonic acid). 1 µg of total RNA was reverse transcribed to cDNA using SensiFast cDNA synthesis kit (Bioline, Inc.). Before running qPCR, cDNA was diluted 5X with DNase/RNase free water to get the specific amount of cDNA according to the manufacturer instructions (applied biosystem qPCR protocol). Primers were designed for each of the selected genes and the reference gene *GAPDH* by using IDT primer quest tool (IDT, 2018). *GAPDH* was selected and used as internal Fe zero to normalize the relative quantities of the target genes due to its consistency across the different growth stages and genotypes. The reverse and forward primer sequences are presented in Table 2.

Each primer pair was designed to span exon-exon junction with PCR product size 89 to 115 bp, length of primer sequence 20 to 22 nucleotides, temperature 61 to 62°C, and the GC content 47.6 to 50%. Before checking primer efficiencies of each selected and reference gene, cDNA was diluted 5X. The target gene expression analysis was done using SensiFAST SYBR No-ROX kit using optical 96 well plate on QuantStudio™ 3 Real-Time PCR System. The genomic DNA contamination or primer dimers was checked on PCR by using Fe zeros [negative reverse transcription Fe zero (-RTC)] and no template Fe zero (NTC)] for each genotype at each time.

After completing 40 amplification cycles, specificity of PCR product per gene was observed by analysing melting curve. All samples for each amplicon had a single sharp peak at the amplicon melting temperature. Genes involved in Fe metabolism, were selected for the analysis. Candidate genes: *FRO2* (García et al., 2012), *IRT1* (García et al., 2012), *NRAMP3* (Lanquar et al., 2010), *VIT1*(Brear et al., 2013), *YSL1*(Kim et al., 2006), *FER3* (Briat et al., 2010), *WEE1* (Mendes, 2014) and *GCN2* (Mendes, 2014) were selected based on their crucial role in Fe metabolisms.

Gene expression analyses were completed using three biological replications along with two technical replications. The gene expression of each tissue sample at different growth stages was averaged over four replications with two repeats. The standard curve method was used for the absolute quantification of the expression of selected genes and the expression were calculated by $2^{(-\Delta CT)}$ method (Schmittgen and Livak, 2008).

2.5 Statistical analysis

The Fe concentrations from different organs of the plants grown under Fe added and Fe zero conditions were averaged over four replications and two repeats. Fe amount from each tissue sample was calculated by the following formula:

Fe amount (g) = Mean Fe concentration of organ (µg g^{-1}) x Mean organ dry weight (g)/1000

Descriptive statistics were used to calculate mean ± standard error (SE). The PROC GLM of SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) was used to compare the means of roots and shoots dry weight and roots to shoots (dry weight/dry weight) ratio.

TABLE 2 Name and general function of selected genes involved in Fe metabolism along with the forward and reverse primer sequences used in qPCR analysis.

Gene	General Function	Primer sequence
GAPDH	Reference Genes	F 5'- CCAAGGTCAAGATCGGAATCA -3' R 5'- CAAAGCCACTCTAGCAACCAAA -3'
FRO2	Fe Root Uptake	F 5'- CTGCAGAGGATGGCGATAAA -3' R 5'- GAACCACGAGTCACTGGAAA -3'
IRT1	Fe Root uptake	F 5'- GCTTTCGCTTCTGGTGTTATAC -3' R 5'- CCAAGGACGCTGAGGTAAA -3'
NRAMP3	Fe Transport	F 5'- CACGGCTATGGGACTTCTTATT-3' R 5'- TCCTAGCCCAACTAGGATACTC-3'
V1T1	Fe Transport	F 5'- GAGAAACCAGATCCAAGGAGAG -3' R 5'- GGAATGAACGCGTAAGGAATG -3'
YSL1	Fe Transport	F 5'- GTGTGGTAGCAGGACTTGTAG -3' R 5'- CGGAGAGGTACGTGTGTAATG -3'
FER3	Fe Storage	F 5'- CCTATGTGTACCATTCCATGTTTG -3' R 5'- ACTCTTCCACCACGATTGTTC -3'
WEE1	Fe Metabolism	F 5'- GCAAGTTGCCACTACTACCT -3' R 5'- CTCTCTAGCCGAAGGTCTCTTA -3'
GCN2	Fe Metabolism	F 5'- ACGACAGTGAAGGTGAGAAAG -3' R 5'- GATGAGGCAAGGAGACAGAAG -3'

10.3389/fpls.2023.1092493

3 Results

3.1 Fe concentrations

Fe concentrations in roots, stems, and leaves of six chickpea genotypes at V3 and V10 growth stages are presented in Figure 2.

The Fe concentration in roots of six genotypes at V3 and V10 growth stages are presented in Figures 2A, D. Results showed that Fe concentration in roots from plants grown under Fe zero (no Fe added) condition decreased from V3 to V10 growth stage in all six genotypes. The reduction rate of Fe concentration from V3 to V10 under Fe zero condition was higher in cultivated species (33%-62%) compared to wild species (6%-50%). In contrast, under Fe added condition, Fe concentration level increased From V3 to V10 growth stages (Figures 2A, D). Like in roots, Fe concentration levels in stem and leaf tissues in plants grown under Fe zero condition decreased from V3 to V10 growth stage in the selected genotypes, except the stems of CDC-551-1 (Figures 2B, C, E, F).

The present study also showed that at V3 growth stage, the highest Fe concentration level under Fe zero condition was observed in leaves followed by roots and stems. However, under Fe added condition, the highest Fe concentration level was observed in roots over six different growth stages (Figures 2A, D). As in the Fe zero environment, under Fe added condition, Fe in leaves decreased dramatically from V3 to V10 growth stages (Figures 2C, F). During the regenerative stage (R2 to RH), Fe concentration in roots increased with slight fluctuation at R6 and RH stages over six genotypes (Supplementary Table S1). However, Fe concentration in stems and leaves were relatively stable with slight variations across the genotypes (Supplementary Tables S2 and S3).

3.2 Fe amount

Fe amount in leaves, stems, and roots of six chickpea genotypes at V3 and V10 growth stages are presented in Figure 3.

The Fe amount in roots of six genotypes at V3 and V10 growth stages are presented in Figures 3A, D. Results showed that Fe amount in roots from plants grown under Fe zero condition slightly increased from V3 to V10 growth stage in all six genotypes except CDC Verano. Likewise, Fe absorption and accumulation in roots increased from V3 to V10 growth stage under Fe added condition (Figures 3A, D). Similar observation was found in reproductive growth stage R2 to physiological maturity stage R6 (Supplementary Table S4).

The Fe amount in stems of six genotypes at V3 and V10 growth stages are presented in Figures 3B, E. Fe amount under Fe zero condition increased more from V3 to V10 growth stage compared to Fe added condition. Like in the roots, Fe amount in the stems showed higher level at the physiological maturity stage (Supplementary Table S5).

The Fe amount in leaves of six genotypes at V3 and V10 growth stages are presented in Figures 3C, F. As in stems, Fe amount under Fe zero condition increased more from V3 to V10 growth stage compared to Fe added condition. Similarly, the lowest Fe amount in leaves was found in the V3 growth stage compared to V10 under both conditions (Figures 3C, F). In contrast to the roots and stems, the Fe amount in leaves decreased sharply from the physiological maturity stage R6 to RH (Supplementary Table S6).

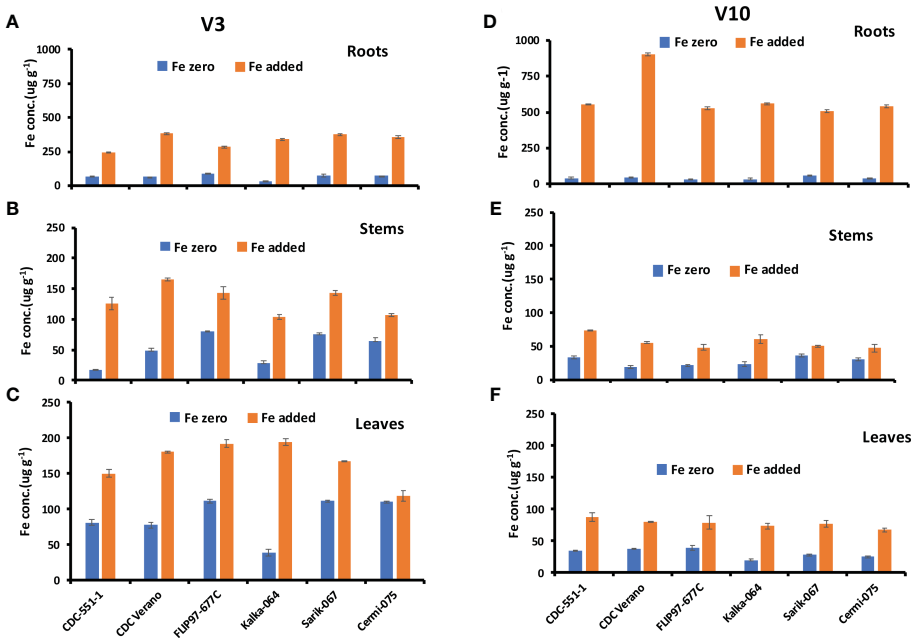


FIGURE 2
Mean Fe concentration (μg g^{-1}, ± SE, n = 48) at V3 and mean Fe concentration (μg g^{-1}, ± SE, n = 24) at V10 growth stages in roots **(A, D)**, stems **(B, E)**, and leaves **(C, F)** of six different genotypes of chickpea under Fe zero and Fe added conditions.

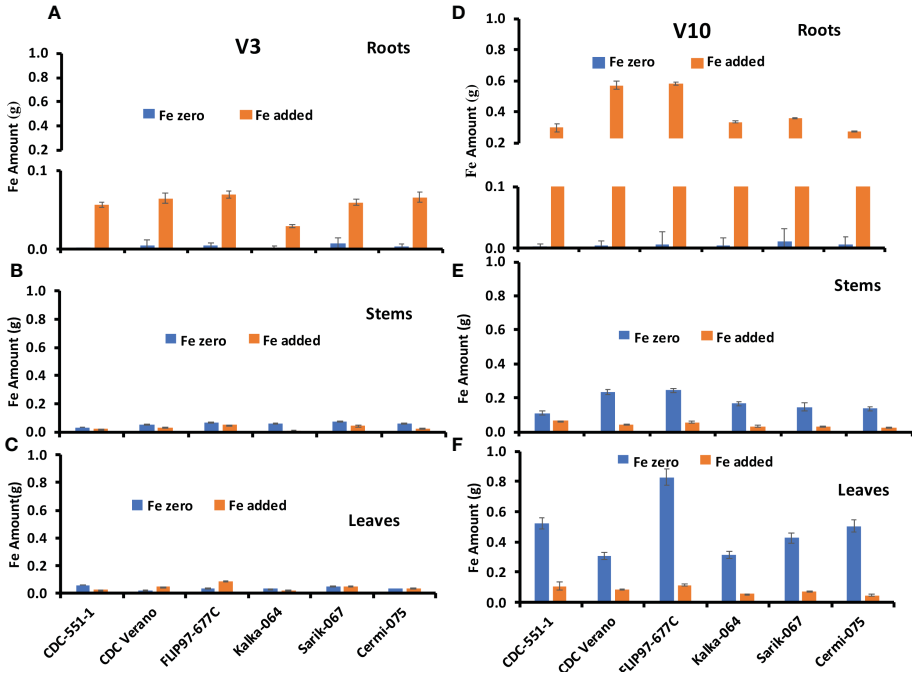


FIGURE 3
Mean Fe amount (g, ± SE, n = 48) at V3 and mean Fe amount (g, ± SE, n = 24) at V10 growth stages in roots **(A, D)**, stems **(B, E)**, and leaves **(C, F)**, of six different genotypes of chickpea under Fe zero and Fe added conditions.

3.3 Fe concentrations vs Fe amount in seeds

Fe partitioning in seeds showed that Fe concentration levels decreased gradually over three growth stages: R5, R6, and RH among the six genotypes (Figure 4A).

In contrast, data from three reproductive growth stages R5, R6, and RH showed that Fe amount in seeds increased gradually from R5 to RH. Genotypes Sarik 067 (1.7 g) and CDC Verano (0.5 g) showed the highest and the lowest Fe amount at RH stage, respectively (Figure 4B). Similarly, at RH stage, genotypes Sarik 067 had the highest Fe concentration (76 μg g^{-1}) in seeds followed by FLIP97-677C (68 μg g^{-1}), whereas Kalka 064 had the lowest (38 μg g^{-1}) (Figure 4A). Our findings showed that Fe concentration level in seeds was higher in reproductive stages (R5 and R6) compared to full maturity stage (RH) (Figure 4A). However, Fe amount level in seeds was higher in full maturity stage (RH) compared to reproductive stages (R5 and R6) (Figure 4B).

3.4 Dry weight

Dry weight in roots, shoots, and roots to shoots ratio of six chickpea genotypes in Fe zero and Fe added conditions at V3 and V10 growth stages are presented in Table 3.

Results showed that the mean values for roots dry weight in Fe added conditions significantly higher ($p < 0.05$) than for the Fe zero conditions at both V3 and V10 stages in all six chickpea

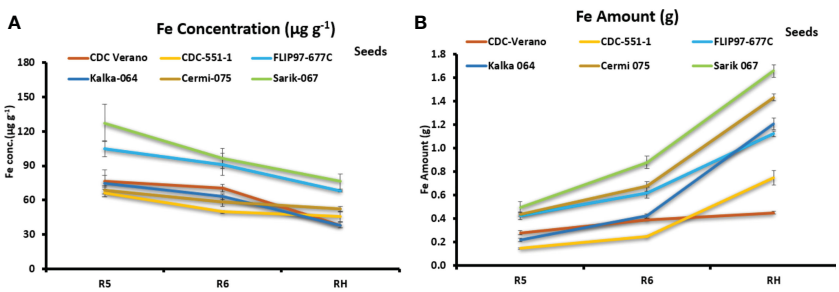


FIGURE 4
Mean Fe concentration (μg g^{-1}, ± SE, n = 8) **(A)** and mean Fe amount (g total $seeds^{-1}$ $plant^{-1}$, ± SE, n = 8) **(B)** in seeds of six chickpea genotypes at R5, R6, and RH growth stages.

TABLE 3 Root to shoot (dry weight/dry weight) ratio of six chickpea genotypes under Fe zero and Fe added conditions at V3 and V10 growth stages.

Genotypes	Growth Stages	Growth Conditions	Roots DW (g ± SE)	Shoots DW (g ± SE)	Roots/shoots (DW/DW ± SE) ratio
CDC-551-1	V3	Fe zero	0.03 ± 0.00^{a}	0.09 ± 0.00^{a}	0.40 ± 0.03^{def}
		Fe added	0.23 ± 0.01^{f}	0.32 ± 0.01^{b}	0.73 ± 0.02^{h}
	V10	Fe zero	0.07 ± 0.00^{ab}	0.63 ± 0.00^{e}	0.11 ± 0.01^{a}
		Fe added	0.83 ± 0.01^{j}	2.06 ± 0.04^{k}	0.41 ± 0.02^{def}
CDC Verano	V3	Fe zero	0.08 ± 0.00^{abd}	0.07 ± 0.00^{a}	1.06 ± 0.08^{g}
		Fe added	0.16 ± 0.01^{bde}	0.44 ± 0.01^{bc}	0.47 ± 0.05^{eg}
	V10	Fe zero	0.10 ± 0.00^{abcd}	0.54 ± 0.00^{de}	0.19 ± 0.02^{ab}
		Fe added	0.69 ± 0.01^{i}	1.91 ± 0.00^{j}	0.36 ± 0.01^{d}
FLIP97-677C	V3	Fe zero	0.05 ± 0.01^{a}	0.10 ± 0.00^{a}	0.53 ± 0.01^{g}
		Fe added	0.32 ± 0.00^{l}	0.77 ± 0.01^{f}	0.40 ± 0.02^{def}
	V10	Fe zero	0.22 ± 0.00^{ef}	1.07 ± 0.00^{g}	0.21 ± 0.02^{b}
		Fe added	1.04 ± 0.00^{k}	2.55 ± 0.02^{l}	0.41 ± 0.01^{df}
Kalka-064	V3	Fe zero	0.06 ± 0.00^{ab}	0.09 ± 0.04^{a}	0.64 ± 0.01^{h}
		Fe added	0.12 ± 0.01^{cd}	0.19 ± 0.10^{a}	0.65 ± 0.04^{h}
	V10	Fe zero	0.15 ± 0.01^{cde}	0.48 ± 0.03^{cd}	0.32 ± 0.02^{cd}
		Fe added	0.57 ± 0.02^{h}	1.28 ± 0.02^{h}	0.47 ± 0.02^{eg}
Sarik-067	V3	Fe zero	0.11 ± 0.02^{bcd}	0.12 ± 0.06^{a}	0.89 ± 0.03^{i}
		Fe added	0.12 ± 0.01^{cd}	0.60 ± 0.05^{de}	0.21 ± 0.01^{b}
	V10	Fe zero	0.21 ± 0.01^{bef}	0.57 ± 0.03^{de}	0.66 ± 0.03^{df}
		Fe added	0.71 ± 0.03^{i}	1.55 ± 0.01^{i}	0.46 ± 0.02^{efg}
Cermi-075	V3	Fe zero	0.06 ± 0.02^{ab}	0.09 ± 0.05^{a}	0.66 ± 0.02^{h}
		Fe added	0.17 ± 0.02^{be}	0.52 ± 0.02^{de}	0.34 ± 0.01^{cd}
	V10	Fe zero	0.16 ± 0.01^{cde}	0.64 ± 0.05^{e}	0.24 ± 0.01^{bc}
		Fe added	0.45 ± 0.02^{g}	1.12 ± 0.03^{g}	0.40 ± 0.02^{def}

Data are means ± standard error (SE), n = 48 at V3, n = 24 at V10. Means in the same column with the different letters are significantly different based on LSD tests ($p < 0.05$). DW, Dry weight.

genotypes. The maximum and minimum mean value was observed in FLIP97-677C (1.04g) and CDC-551-1(0.03g) at V10 and V3 growth stages under Fe added and Fe zero conditions, respectively (Table 3). Like in roots dry weight, the mean values of shoots dry weight in Fe added condition also showed significantly higher ($p < 0.05$) than for the Fe zero conditions at both V3 and V10 growth stages except Kalka-064 at V3 stage. The highest and lowest mean score was observed in FLIP97-677C (2.55g) and CDC Verano (0.07g) at V10 and V3 growth stages under Fe added and Fe zero conditions, respectively (Table 3). Similarly, for roots to shoots (dry weight/dry weight) ratio, the mean values of roots to shoots ratio in Fe added condition also showed significantly higher (($p < 0.05$) than for the Fe zero conditions at both V3 and V10 stages except Kalka-064 at V3 stage. The maximum average score for roots to shoots ratio was observed in CDC Verano (1.06) at V3 stage under Fe zero condition, whereas the lowest was observed in CDC-551-1(0.11) at V10 under Fe zero condition (Table 3).

3.5 Gene expression analysis

3.5.1 V3 growth stage

Genes involved in Fe metabolism of six chickpea genotypes grown under Fe zero (Fe^{-}) and Fe added (Fe^{+}) conditions at V3 growth stage are presented in Figure 5.

Two Fe uptake genes *FRO2* and *IRT1* expressed differently across the six chickpea genotypes (Figure 5). At V3 growth stage, under Fe added (Fe^{+}) condition, *FRO2* was expressed more in root tissues than in leaves compared to Fe zero (Fe^{-}) (Figures 5A, E). Among the cultivated genotypes, genotype CDC 551-1 showed the highest level of expression of *FRO2* under Fe added condition in roots, which is 14, 12, and 10-folds higher than the *FRO2* expression in wild genotypes Kalka 064, Sarik 067, and Cermi 075, respectively. Like CDC 551-1, the cultivated genotype FLIP97-677C also showed high expression of *FRO2* under added Fe condition in roots, which is 9, 8 and 6-folds higher than the *FRO2* expression in wild genotypes Kalka 064,

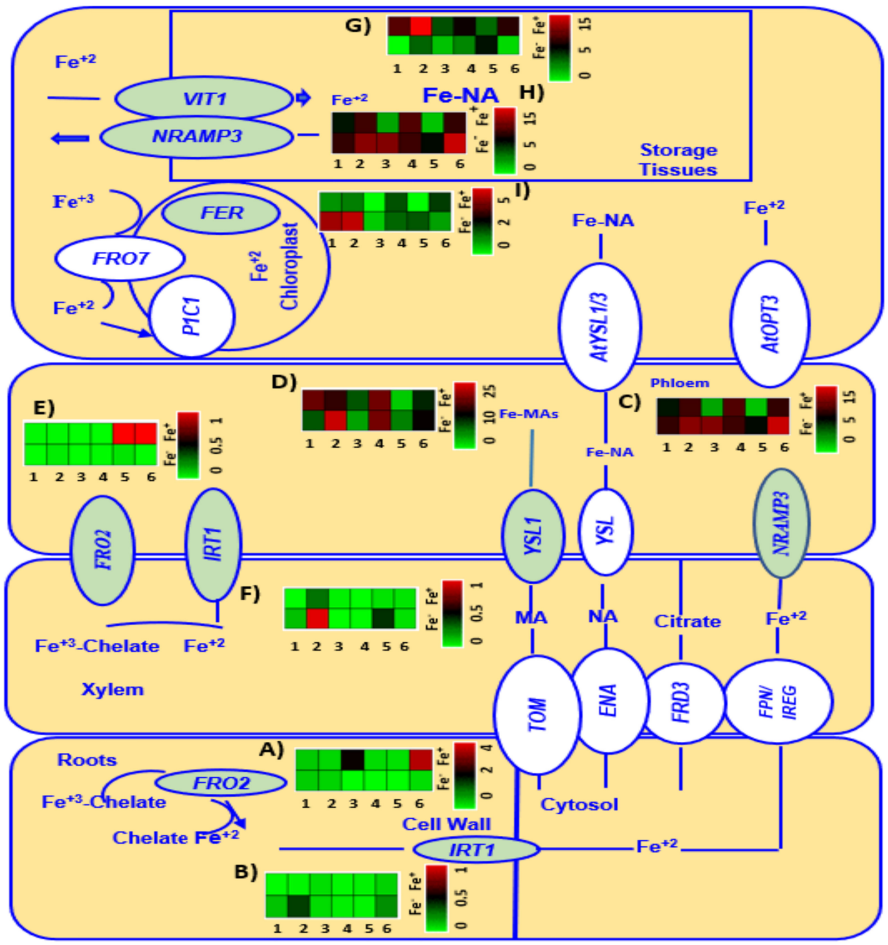


FIGURE 5
A heatmap analysis showing the gene expression patterns of Fe metabolism related genes *FRO2* **(A, E)**, *IRT1* **(B, F)**, *NRAMP3* **(C, H)**, *YSL1* **(D)**, *V1T1* **(G)**, and *FER3* **(I)** in roots and leaves of six genotypes (1 = CDC Verano, 2 = Cermi 075, 3 = FLIP97-677C, 4 = Sarik 067, 5 = Kalka 064, and 6 = CDC 551-1). The data at V3 growth stage was taken from both Fe zero (Fe^{-}) and Fe added (Fe^{+}) conditions. Green and red color represents down-regulation, and up-regulation in the color scale, respectively.

Sarik 067, and Cermi 075, respectively (Figure 5A). However, in leaves, most of the genotypes showed lower expression of *FRO2* under both Fe added and Fe zero conditions (Figure 5E). Another Fe regulated gene in roots, *IRT1*, showed higher expression in leaves (Figure 5F) rather than in roots (Figure 5B). The highest level of *IRT1* expression was observed in Cermi 075 under Fe zero condition in leaves, whereas the lowest *IRT1* expression was shown by FLIP97-677C and Sarik 067 (Figure 5F). Our findings also showed that under Fe zero condition, the expression of *IRT1* was higher in most of the genotypes in both leaves (Figure 5F) and roots (Figure 5B) compared to Fe added condition.

Three selected Fe transporter genes: *NRAMP3* (Figures 5C, H), *V1T1* (Figure 5G), *and YSL1* (Figure 5D) were mainly expressed in leaves compared to roots (Supplementary Figure S1). Under Fe zero condition, an increase of *NRAMP3* expression was observed in leaves of the six genotypes compared to Fe added condition (Figures 5C and 6H). In roots, the *NRAMP3* showed lower expression in all six genotypes under both conditions compared to leaves. However, the expression levels of *NRAMP3* were relatively high in most genotypes except CDC Verano and Cermi 075 under Fe zero condition compared to Fe added condition (Supplementary Figure S1). Gene *V1T1* expressed more in leaves under Fe added condition in all six genotypes (Figure 5G). Under Fe added condition the highest level of *V1T1* expression (5 folds) was obtained in the leaves of Cermi 075 compared to Fe zero (Figure 5G). We also found that under Fe deficient condition higher level of *YSL1* expression compared to Fe added conditions in all genotypes except CDC Verano and FLIP97-677C (Figure 5D). However, in roots, we found that the expression of *YSL1* was higher under Fe added than Fe zero conditions in all genotypes except FLIP97-677C and Kalka 064 (Supplementary Figure S1). Moreover, leaf tissues have higher expression of *FER3* under Fe zero condition in comparison to Fe added condition, except Sarik 067 and CDC-551-1 which showed relatively high expression of *FER3* under Fe added condition (Figure 5I).

At V3 stage, we found that the *WEE1* gene was up regulated in both tissues when the plants were subjected to Fe zero condition (Supplementary Figure S1). In the current study at V3 stage, both root and leaf tissues showed elevated expression of *WEE1* and

GCN2 genes when the plants were subjected to Fe zero environment (Supplementary Figure S1).

3.5.2 V10, R2, and R5 growth stage

Expression of genes (highlighted in green) involved in Fe metabolism (*FRO2*, *IRT1*, *NRAMP3*, *V1T1*, *YSL1*, *FER3*, *WEE1*, and *GCN2)* of six chickpea genotypes grown under Fe added condition at V10, R2, and R5 growth stages are presented in Figure 6.

In the current study expression of Fe uptake genes, *FRO2* (Figures 6A, E) and *IRT1*(Figures 6B, F), was higher in roots (Figures 6A, B) compared to leaves (Figures 6E, F) across all six genotypes (Figure 6) at V10, R2, and R5 stages. The highest level of expression of *FRO2* was observed in V10 roots compared to R2 and R5 stages across all six genotypes (Figure 6A). Unlike *FRO2*, expression of *IRT1* was highest in V10 leaves among all other stages in root and leaf tissues (Figures 6B, F).

This study also found the highest expression of *NRAMP3* (Figures 6C, H), and *VIT1*(Figure 6G) in R5 leaves compared to V10 and R2 leaves except Kalka 064 (Figures 6C, H, G). Like in *NRAMP3* and *VIT1*, the highest expression of *YSL1* was also found in R5 leaves except in leaves of Kalka 064 and FLIP97-677C (Figure 6D). However, at V10 stage, we found that the expression of both *NRAMP3* (Figures 6C, H), and *YSL1*(Figures 6B, F) was higher in leaves (Figure 6) compared to roots (Supplementary Figure S2) except CDC Verano and CDC 551-1. However, at R2, roots showed higher expression of *NRAMP3* than leaves except Cermi -075. Similarly, at R2, the higher expression of *YSL1* was also found in roots compared to leaves except Cermi -075, FLIP97-677C, and CDC-551-1. At R5, leaves showed higher expression of

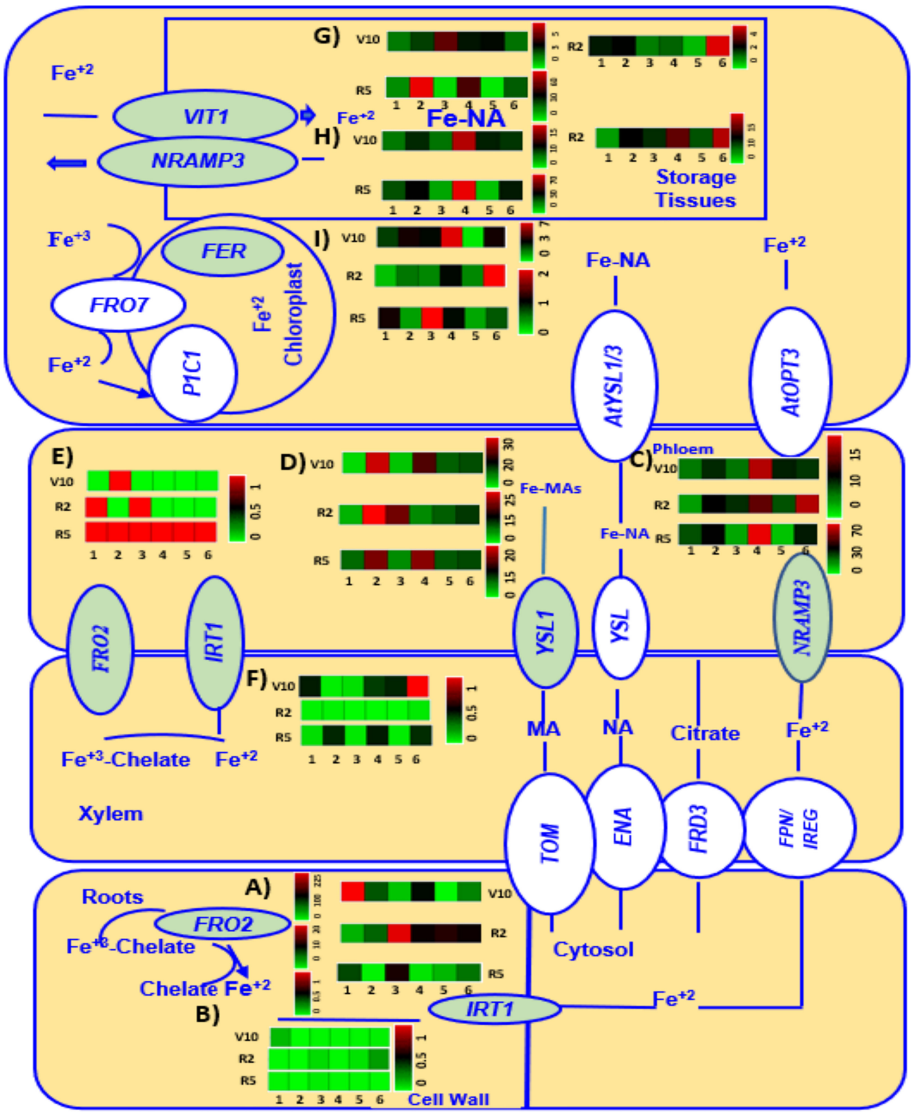


FIGURE 6
A heatmap analysis showing the gene expression patterns of Fe metabolism related genes *FRO2* **(A, E)**, *IRT1* **(B, F)**, *NRAMP3* **(C, H)**, *YSL1* **(D)**, *V1T1* **(G)**, and *FER3* **(I)** in roots and leaves of six different genotypes (1 = CDC Verano, 2 = Cermi 075, 3 = FLIP97-677C, 4 = Sarik 067, 5 = Kalka 064, and 6 = CDC 551-1). The data at V10, R2 and R5 growth stages taken only from Fe added (Fe^{+}) conditions. Green and red color represents down-regulation, and up-regulation in the color scale, respectively.

NRAMP3 and *YSL1* than roots except FLIP97-677C and Kalka 064 of *NRAMP3* and FLIP97-677C of *YSL1* (Figure 6 and Supplementary Figure S2). However, the higher expression of *V1T1* was found in leaves compared to roots across all genotypes at V10 to R5 stages except Kalka 064 at R2 (Figure 6G and Supplementary Figure S2).

In this study, we also found that roots showed higher expression of *FER3* compared to leaves across all genotypes and stages except Kalka 064 at V10 leaf tissues (Figure 6I and Supplementary Figure S2). Like in *FER3*, the expression of WEE1 was higher in roots compared to leaves across all genotypes and stages except FLIP97-677C and Kalka 064 at R5 roots (Supplementary Figure S2). Another gene involved in Fe metabolism, *GCN2*, expressed more in roots compared to leaves across all genotypes and growth stages. In addition, at V10 and R2 stages both in root and shoot tissues compared to R5 (Supplementary Figure S2).

3.6 Correlation study of the gene expressions and Fe concentrations

The correlation between the Fe concentration and the gene expression in roots and leaves at V3 to R5 growth stages of the six chickpea genotypes grown under Fe zero and Fe added conditions are shown in Table 4.

Among the selected growth stages, under Fe added condition in roots, the relationship between the Fe concentration and *FRO2* showed a strong positive correlation (r = 0.95; $p < 0.01$) at V3, and (r = 0.86; $p < 0.05$) at V10, respectively (Table 4). Similarly, under Fe added condition in leaves, *FRO2* showed a strong positive correlation (r = 0.82; $p < 0.05$) at V10 in leaves (Table 4). Like *FRO2*, under Fe added medium, *IRT1* also showed a strong positive correlation (r = 0.99; $p < 0.01$) at V3, (r = 0.99; $p < 0.01$) at V10, and (r = 0.87; $p < 0.05$) at R2 growth stages, respectively (Table 4). Likewise, under Fe added environment in roots, *NRAMP3, YSL1, FER3*, and *GCN2* showed a strong positive correlation (r = 0.95; $p < 0.01$) at V10, (r = 0.87; $p < 0.05$) at V3, (r = 0.88; $p < 0.01$) at V10, and (r = 0.81; $p < 0.05$) at V3, respectively (Table 4). Moreover, under Fe zero condition in leaves, *FRO2, V1T1*, and *GCN2* showed a strong positive correlation (r = 0.82; $p < 0.05$) at V10, (r = 0.88; $p < 0.05$) at V3, and (r = 0.86; $p < 0.05$) at V3, respectively (Table 4).

4 Discussions

A better understanding of the mechanism of Fe accumulation in seeds is a prerequisite for developing chickpea cultivars with high Fe amount in seeds. Although major advances have been made in generating engineered biofortified crops, the Fe loading pathway into seeds is still unclear (Grillet et al., 2014). To enhance our knowledge in this regard, a study was conducted to evaluate Fe accumulation in organs at different growth stages of chickpea. Our results showed that Fe was remobilized from roots to the sink tissues under Fe zero environment are consistent with the previous studies done in wheat, *Arabidopsis thaliana*, and chickpea (Mahmoudi et al., 2005; Peng and Li, 2005; Waters and Grusak, 2008; Waters et al., 2009; Sankaran and Grusak, 2014). Our Fe accumulation study also showed more Fe remobilization under Fe zero condition compared to Fe added condition. Also, our findings showed that Fe concentration level in seeds was higher in reproductive stages (R5 and R6) compared to full maturity stage (RH). This finding suggests that at the reproductive stage Fe remobilize more from the other tissues to the seeds than at physiological maturity stage, which is consistent with previous studies of wheat, and *Arabidopsis thaliana* (Himelblau and Amasino, 2001; Garnett and Graham, 2005).

In this study, we also measured Fe amount which is a better measurement option for selecting cultivars for Fe improvement than is Fe concentration. This is because Fe amount allow to assess total Fe accumulation as well as Fe remobilization in different organs. Our findings showed that the Fe amount in root, stem, and leaf increased gradually over the six time points suggests that

TABLE 4 Correlation between Fe concentration and gene expression in roots and leaves at V3 to R5 growth stage of six genotypes grown under Fe zero and Fe added conditions.

	Roots					Leaves				
	Fe zero	Fe added				Fe zero	Fe added			
Genes	V3	V3	V10	R2	R5	V3	V3	V10	R2	R5
FRO2	ns	0.95**	0.87*	ns	ns	ns	ns	0.82*	ns	ns
IRT1	ns	0.99**	0.99**	0.87*	ns	ns	ns	ns	ns	ns
NRAMP3	ns	ns	0.95**	ns	ns	ns	ns	ns	ns	ns
VIT1	ns	ns	ns	ns	ns	ns	0.88*	ns	ns	ns
YSL1	ns	0.87*	ns	ns	ns	ns	ns	ns	ns	ns
FER3	ns	ns	0.88*	ns	ns	ns	ns	ns	ns	ns
GCN2	0.81*	ns	ns	ns	ns	ns	0.86*	ns	ns	ns
WEE1	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

ns, non-significant; * = significant at $p < 0.05$; and ** = significant at $p < 0.01$.

little or no Fe remobilization occurred at the later growth stages, which is consistent with the results in the previous study in pea (Sankaran and Grusak, 2014). We also found that Fe amount in seeds was higher at physiological maturity stage (RH) compared to reproductive (R5 and R6) stages suggested that at reproductive stages, total incoming Fe either from continuous uptake or remobilization from other tissues might be distributed to new leaves, roots, stems, flowers, and seeds. However, at physiological maturity stage (RH), under Fe added condition, continuous uptake of Fe replaced remobilization that ultimately help to increase seed Fe accumulation. Similar results have been reported in wheat, *Arabidopsis thaliana*, and pea (Himelblau and Amasino, 2001; Garnett and Graham, 2005; Sankaran and Grusak, 2014).

This study also observed that under Fe added condition, the cultivated genotypes (such as CDC Verano, FLIP97-677C and CDC 551-1) contained higher Fe amount in roots than the wild genotypes (Supplementary Table S4). However, in seeds, the higher Fe amount was found in the three selected wild genotypes Sarik 067, Cermi 075, and Kalka 064 than the selected cultivated species FLIP97677C, CDC-551-1, and CDC Verano (Figures 4A, B). Therefore, our results indicated that Fe absorption and accumulation rate also depend on cultivars, which is consistent with previous studies in common bean and rice (Blair et al., 2010; Pereira et al., 2014).

In this study, we also measured root and shoot dry weight as well as root to shoot ratio under Fe zero and Fe added conditions at V3 and V10 growth stages (Table 3). Our findings showed that under Fe added conditions both root and shoot dry weight as well as root to shoot ratio was higher compared to Fe zero conditions at both V3 and V10 growth stages. This study also observed that shoot dry weight under both Fe zero and Fe added conditions was higher than root dry weight at both V3 and V10 growth stages. Our findings indicated that Fe zero condition reduced both root and shoot dry weight compared to Fe added conditions at both V3 and V10 growth stages. In addition, root dry weight was lower than shoot dry weight at both conditions and growth stages. Similar observations were reported in lentils, chickpeas, and soybean (Mahmoudi et al., 2005; Vasconcelos et al., 2006).

Another objective of this study was to determine the relative expression levels of Fe-related transporter and Fe metabolism genes in chickpea seeds. For this, a study was conducted to evaluate the expression levels of genes associated with Fe uptake, transportation, and accumulation into the seeds of chickpea. We analysed gene expression of selected genes for six chickpea genotypes grown under Fe zero and Fe added conditions at V3 (Figure 5), V10, R2 and R5 stages (Figure 6) under Fe added condition. Fe uptake gene *FRO2* generally expressed more under Fe zero environment. However, here, both at V3 and V10 to R5 growth stages, our study showed *FRO2* expressed more in roots than leaves under Fe added condition, which is consistent with the previous studies in soybean and common bean (Blair et al., 2010; Santos et al., 2013). In the present study, under Fe added condition at V3 in roots, we also observed genotype CDC 551-1 showed the highest level of expression among all selected genotypes (Figure 5A). Consistent with our results previous reports showed that the activity of *FRO2* depends on both species and cultivar (Blair et al., 2010; Santos et al., 2013; Pereira et al., 2014). However, under Fe zero condition at V3, another Fe regulates gene, *IRT1*, showed higher expression in leaves as well as root tissue (Figures 5B, D). However, under Fe added condition, we found higher expression of *IRT1* mostly in leaves of six genotypes compared to roots except R2 (Figures 6B, F). Although *IRT1* express more under Fe zero condition, a previous study in *Arabidopsis thaliana* showed that the regulation of *IRT1* depends on the root Fe status and shoot Fe demands (Vert et al., 2003). Moreover, as *IRT1* belongs to the ZIPs family that are not only involved in Fe uptake and transport Fe to the other plant organs, but also stores and detoxifies excess Fe (Grotz et al., 1998; Yang et al., 2009; Li et al., 2013).

In the present study, we also found three selected Fe transporter genes (*NRAMP3, V1T1, and YSL1)* were mainly expressed in leaves compared to roots (Figures 5 and 6). Our study showed under Fe zero condition, at V3 growth stage, an increase of *NRAMP3* expression was observed in leaves of the six genotypes compared to under Fe added condition (Figures 5C, D). In roots, the expression levels of *NRAMP3* were relatively low in all six genotypes under both Fe zero (Fe zero) and Fe added conditions (Supplementary Figure S1). Like in V3, at V10 -R5 stages, a higher expression of *NRAMP3* was found in leaves of most of the selected genotypes compared to roots under Fe added condition (Figures 6C, H). These results supported the fact that this gene is mainly predominant in leaves (Figures 5C, H and 6C, H). The conclusion that emerged from this study is similar to that in *Arabidopsis thaliana* reported by Lanquar et al. (2010). Gene *V1T1* is responsible for Fe loading into the vacuole, which is completely the opposite role of *NRAMP3*. Thus, it is expected that this gene would be expressed more under Fe added leaves as shown in all six genotypes at V3 and V10 - R5 as well. Our results also showed both at V3 and V10 - R5 stages, gene *V1T1* expressed more under Fe added shoot as shown in all six genotypes at V3 and V10 to R5 compared to roots (Figures 5G and 6G, Supplementary Figures S1, S2). These findings are consistent with previous studies in which the expression of *V1T1* increased in parallel with the accumulation of Fe^{2+} into the vacuoles, to lessen waste, and to cut down the oxidative damage (Bakhshi and Karimian, 2004).

Like the other two transporter genes *NRAMP3* and *V1T1*, *YSL1* also transports Fe in the form of Fe_2^+ NA complexes, which is the crucial transportable Fe form in the phloem (Kim et al., 2006). Thus, it is expected that this gene will be expressed more under Fe added condition. This is because under Fe added condition Fe transport activity is needed to fulfil plant requirement. Our findings also found that the root tissues at V3 stage under Fe-treated condition showed higher level of *YSL1* expression compared to Fe zero (Fe zero) conditions (Supplementary Figure S1). These findings support the previous reports that Fe was translocated to various plant organs (Kim et al., 2006). Similar observation was also found at V10 - R5 stages under Fe added condition (Figure 6D). However, other findings from this study showed that the *YSL1* expression in shoot tissues was higher in plants grown under Fe zero than Fe added condition at V3 stage (Figure 5D), which is in contrast with previous study in *Arabidopsis thaliana* (Kim et al., 2006).

Ferritin, which is the Fe storage protein, provides insurance for meeting Fe demands of cell without reacting with oxygen when Fe is

in affluent (Briat et al., 2010). In this study, at V3 stage, leaf tissues under Fe zero condition showed higher level of *FER3* expression compared to leaf under Fe-available condition, except for the CDC-551-1 (Figure 5I). The conclusion that emerged from this study is contrast to that in (Ravet et al., 2009). However, at V10 – R5 stages, our study found at V10 stage leaves showed an increase level of *FER3* relative expression compared to leaves at flowering (R2) and physiological maturity stage (R5). This process led the movement of the cellular Fe demand from source (leaf tissues) to other parts of plants (flowers and seeds) at R2 and R5 stages (Figure 6I) when the plants were under Fe added condition. The conclusion that emerged from this study is similar to that in (Ravet et al., 2009)

In this study, other two genes: *WEE1* and *GCN2*, which are responsible for Fe metabolism were also studied. *WEE1* is a protein that is associated with cell cycle regulation (de Schutter et al., 2007). Thus, the absence of *WEE1* function led to stunted plant growth (Jones et al., 2013). At V3 stage, we found higher expression of *WEE1* and *GCN2* in both root and leaf tissues of plants grown under Fe zero condition than under Fe added condition (Supplementary Figure S1). However, under Fe added condition both at V3 and V10 to R5 stages, the expression of *WEE1* and *GCN2* were observed in roots compared to leaves (Supplementary Figures S1 and S2). Our findings were previously reported in *G. max* and *M. truncatula* (Mendes, 2014).

In the current study, we also observed the correlations between the Fe concentrations and the gene expressions in roots and leaves at V3 to R5 growth stages of the six chickpea genotypes grown under Fe zero and Fe added media (Table 4). At V3 growth stage under Fe added medium, we found genes *FRO2*, showed a highly significant positive correlation at V3 and V10 in roots, and at V10 in leaves, respectively, with Fe concentration (Table 4). Although *FRO2* is needed to reduce ferric to ferrous Fe mainly in the roots under Fe zero condition, current findings observed *FRO2* expressed more under Fe added condition. This finding suggests that the *FRO2* requires more Fe to activate itself during the Fe zero condition. As a result, *FRO2* expression reduced during Fe zero condition. In addition, higher expression of *FRO2* helps to increase more available ferrous Fe. Thus, the large amount of Fe is stored in roots and leaves under Fe available environment. Previous studies conducted in common bean, soybean, and *Medicago truncatula* reported that an elevated *FRO2* expression increased Fe concentration in roots and leaves (Blair et al., 2010; Mendes, 2014). Current findings with chickpea provided further confirmation of these reports. Like *FRO2*, under Fe added medium, *IRT1* also showed a strong positive correlation at V3, V10, and R2 growth stages, respectfully (Table 4). A study in pea reported that *FRO2* is the rate limiting enzyme for Fe acquisition, Fe transporter gene *IRT1* cannot achieve saturation level of Fe concentration without the Fe reductase activity (Grusak et al., 1990). This suggests that *FRO2* and *IRT1* are coregulated. Previous studies in soybean and medicago truncatula, also reported that the level of *IRT1* expression is equivalent with *FRO2* expression in both species and tissues (Mendes, 2014). Similar observation was also found in *Arabidopsis thaliana* (Kim and Guerinot, 2007). Our findings were consistent with these reports.

Likewise, under Fe added environment in roots, *NRAMP3*, and *YSL1* showed a strong positive correlation. Although the role of *NRAMP3* is to remobilize Fe from the vacuole, *NRAMP3* express more in roots at V10 growth stage along with Fe concentration under Fe available media. The reason behind this finding is that under Fe added medium, plants do not require more Fe to be remobilized to the leaves, which is contrast to Fe zero condition. Thus, higher level of *NRAMP3* expression was found along with Fe concentration in roots under Fe available media. This finding is in contrast with reports on *Arabidopsis thaliana*, soybean and *Medicago truncatula* where *NRAMP3* express more in leaves under Fe zero media (Lanquar et al., 2010; Mendes, 2014). Since *YSL1* transport Fe^{2+}-NA complex, which is the main transportable Fe form in the phloem, it is expected to increase more transporter like *YSL1* to fulfil plant's demand under Fe available media. This finding agrees with reports on soybean and medicago where *YSL1* expression increase more in both tissues along with Fe concentration under Fe added condition (Kim et al., 2006; Mendes, 2014). Current work also found that a highly significant positive correlation between the Fe concentration and *V1T1* expression at V3 under Fe added environment in leaf (Table 4). Since gene *V1T1* plays the opposite role of *NRAMP3*, which is Fe loading into the vacuole, it is expected that this gene would be expressed more under Fe added leaves. The findings from this study is similar to that in (Bakhshi and Karimian, 2004).

The result from the present work also obtained that a highly significant positive correlation was detected between the Fe concentration and *FER3* expression in roots at V10 under Fe added condition (Table 4). This finding matches with reports on *Arabidopsis thaliana* where *FER3* express more with Fe concentration under Fe added condition (Ravet et al., 2009). We also found a strong positive correlation between the Fe concentration and *GCN2* expression in roots and leaves at V3 under Fe zero and Fe added conditions, respectively (Table 4). Our findings that *GCN2* plays a significant role in Fe metabolism were previously reported in *G. max* and *M. truncatula* (Mendes, 2014).

In conclusions, results from our present study indicates that the genes responsible for Fe uptake and translocation such as *FRO2*, *IRT1*, *NRAMP3*, *VITI*, and *YSL1*, as well as Fe storage (*FER3*) and Fe metabolism genes (*GCN2* and *WEE1*), can be targeted for selection and modification to increase Fe concentration in chickpea seeds. Moreover, in terms of Fe improvement in chickpea seeds, wild species could be a better option than cultivated species.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

TJ and BT designed the experiment. TJ conducted the experiments, analysed the data, and prepared the draft

manuscript. SK and BT assisted with data analysis and edited the manuscript. BT conceived and directed the research. All authors reviewed all documents critically and approved the final manuscript for submission in the journal. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Saskatchewan Ministry of Agriculture.

Acknowledgments

We are thankful for technical assistance provided by Brent Barlow and Kevin Mikituk of the Crop Development Centre, University of Saskatchewan. We acknowledge the Food and Bioproduct Science lab for providing us FeEDDHA. Thanks to Carmen Breitkreutz, for the technical support. Many thanks to Amit Deokar for sharing knowledge during this study. Thanks also goes to Jeremy Packet for his collaboration during this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1092493/full#supplementary-material

References

Magitsu, M. A. (2016). *Response of chickpea cultivars to dryland management in southern Ethiopia* (Saskatoon (SK: University of Saskatchewan).

Bakhshi, M., and Karimian, V. (2004). "Effect of fe application on soybean yield and chemical composition," in *Proceeding of 3rd National Congress in Efficient Application of Fertilizers in Agriculture*, Iran. (pp. 174–179).

Baltussen, R., Knai, C., and Sharan, M. (2004). Fe fortification and fe supplementation are cost-effective interventions to reduce fe deficiency in four subregions of the world. *J. Nutr.* 134 (10), 2678–2684. doi: 10.1093/jn/134.10.2678

Beasley, J., Bonneau, J., Sánchez-Palacios, J., Moreno-Moyano, L., Callahan, D., Tako, E., et al. (2019). Metabolic engineering of bread wheat improves grain fe concentration and bioavailability. *Plant Biotechnol. J.* 17 (8), 1514–1526. doi: 10.1111/pbi.13074

Blair, M. W., Knewtson, S. J. B., Astudillo, C., Li, C. M., Fernandez, A. C., and Grusak, M. A. (2010). Variation and inheritance of fe reductase activity in the roots of commonbean (*Phaseolusvulgaris* l.) and association with seed fe accumulation QTL. *BMC Plant Biol.* 10, 215. doi: 10.1186/1471-2229-10-215

Boonyaves, K., Wu, T., Gruissem, W., and Bhullar, N. (2017). Enhanced grain fe levels in rice expressing an FE-regulated metal transporter, NICOTIANAMINE synthase and ferritin gene cassette. *Front. Plant Sci.* 8, 130. doi: 10.3389/fpls.2017.00130

Brear, E. M., Day, D. A., and Smith, P. M. C. (2013). Fe: an essential micronutrient for the legume-rhizobium symbiosis. *Front. Plant Sci.* 4, 359. doi: 10.3389/fpls.2013.00359

Briat, J. F. (2011). "Fe nutrition and implications for biomass production and the nutritional quality of plant products," in *Molecular and physiological basis of nutrient use efficiency in crops*. Eds. M. J. Hawkesford and P. Barraclough (Oxford, UK: Wiley-Blackwell), 309–328.

Briat, J., Ravet, K., Arnaud, N., Duc, C., Boucherez, J., Touraine, B., et al. (2010).). new insights into ferritin synthesis and function highlight a link between fe homeostasis and oxidative stress in plants. *Ann. Bot.* 105 (5), 811. doi: 10.1093/aob/mcp128

Campion, B., Glahn, R. P., Tava, A., Perrone, D., Doria, E., Sparvoli, F., et al. (2013). Genetic reduction of antinutrients in common bean (Phaseolus vulgaris l.) seed, increases nutrients and *in vitro* fe bioavailability without depressing main agronomic traits. *Field Crops Res.* 141, 27–37. doi: 10.1016/j.fcr.2012.10.015

Chaney, R., and Bell, P. (1987). Complexity of fe nutrition: Lessons for plant-soil interaction research. *J. Plant Nutr.* 10 (9-16), 963–994. doi: 10.1080/01904168709363626

Connorton, J. M., Balk, J., and Rodrı´guez-Celma, J. (2017). Fe homeostasis in plants–a brief overview. *Metallomics* 9, 813–823. doi: 10.1039/C7MT00136C

de Schutter, K., Joubes, J., Cools, T., Verkest, A., Corellou, F., Babiychuk, E., et al. (2007). Arabidopsis WEE1 kinase fe zeros cell cycle arrest in response to activation of the DNA integrity checkpoint. *Plant Cell* 19 (1), 211–225. doi: 10.1105/tpc.106.045047

Diapari, M., Sindhu, A., Bett, K., Deokar, A., Warkentin, T. D., and Tar'an, B. (2014). Genetic diversity and association mapping of iron and zinc concentrations in chickpea (Cicer arietinum l.). *Genome* 57 (8), 459–468. doi: 10.1139/gen-2014-0108

Eide, D., Broderius, M., Fett, J., and Guerinot, M. (1996). A novel fe-regulated metal transporter from plants identified by functional expression in yeast. *Proc. Natl. Acad. Sci. United States America* 93 (11), 5624–5628. doi: 10.1073/pnas.93.11.5624

García, M. J., Romera, F. J., Stacey, M. G., Stacey, G., Villar, E., Alcántara, E., et al. (2012). Shoot to root communication is necessary to fe zero the expression. of fe-acquisition genes in strategy I plants. *Planta* 237 (1), 65–75. doi: 10.1007/s00425-012-1757-0

Garnett, T., and Graham, R. (2005). Distribution and remobilization of fe and copper in wheat. *Ann. Bot.* 95 (5), 817–826. doi: 10.1093/aob/mci085

Gómez-Galera, S., Rojas, E., Sudhakar, D., Zhu, C., Pelacho, A. M., Capell, T., et al. (2010). Critical evaluation of strategies for mineral fortification of staple food crops. *Transgenic Res.* 19 (2), 165–180. doi: 10.1007/s11248-009-9311-y

Grillet, L., Mari, S., and Schmidt, W. (2014). Fe in seeds – loading pathways and subcellular localization. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00535

Grotz, N., Fox, T., Connolly, E., Park, W., Guerinot, M., and Eide, D. (1998). Identification of a family of zinc transporter genes from arabidopsis that respond to zinc deficiency. *Proc. Natl. Acad. Sci. United States America* 95 (12), 7220–7224. doi: 10.1073/pnas.95.12.7220

Grusak, M. A. (1994). Fe transport to developing ovules of pisum sativum. i. seed import characteristics and phloem fe-loading capacity of source regions. *Plant Physiol.* 104 (2), 649–655. doi: 10.1104/pp.104.2.649

Grusak, M. A., Welch, R. M., and Kochian, L. V. (1990). Physiological characterization of a single-gene mutant of pisum sativum exhibiting excess fe accumulation. i. root fe reduction and fe uptake. *Plant Physiol.* 3), 976–981. doi: 10.1104/pp.93.3.976

Himelblau, E., and Amasino, R. (2001). Nutrients mobilized from leaves of arabidopsis thaliana during leaf senescence. *J. Plant Physiol.* 158 (10), 1317–1323. doi: 10.1078/0176-1617-00608

Integrated DNA Technologies IDT (2018) *Integrated DNA technologies*. Available at: https://www.idtdna.com/pages.

Johnson, A., Kyriacou, B., Callahan, D., Carruthers, L., Stangoulis, J., Lombi, E., et al. (2011). Constitutive overexpression of the OsNAS gene family reveals single-gene strategies for effective fe- and zinc-biofortification of rice endosperm (Fe and zinc

biofortification of rice endosperm). *PloS One* 6 (9), E24476. doi: 10.1371/journal.pone.0024476

Jones, R., Ougham, H., Thomas, H., and Waaland, S. (2013). The molecular life of plants. *Am. Soc. Plant Biologists.* 1 (1), 390–392. doi: 10.1098/rstb.2010.0136

Kim, S. A., and Guerinot, M. L. (2007). Mining fe: Fe uptake and transport in plants. *FEBS Lett.* 581, 2273–2280. doi: 10.1016/j.febslet.2007.04.043

Kim, S., Punshon, T., Lanzirotti, A., Li, L., Alonso, J. M., and Ecker, J. R. (2006). Localization of fe in arabidopsis seed requires the vacuolar membrane transporter VIT1. *Science* 314 (5803), 1295–1298. doi: 10.1126/science.1132563

Koike, S., Inoue, H., Mizuno, D., Takahashi, M., Nakanishi, H., Mori, S., et al. (2004). OsYSL2 is a rice metal-nicotianamine transporter that is regulated by fe and expressed in the phloem. *Plant J.: For Cell Mol. Biol.* 39 (3), 415–424. doi: 10.1111/j.1365-313X.2004.02146.x

Lanquar, V., Lelièvre, F., Bolte, S., Hamès, C., Alcon, C., Neumann, D., et al. (2005). Mobilization of vacuolar fe by AtNRAMP3 and AtNRAMP4 is essential for seed germination on low fe. *EMBO J.* 24 (23), 4041–4051. doi: 10.1038/sj.emboj.7600864

Lanquar, V., Ramos, M. S., Lelievre, F., Barbier-Brygoo, H., Krieger-Liszkay, A., and Kramer, U. (2010). Export of vacuolar manganese by AtNRAMP3 and AtNRAMP4 is required for optimal photosynthesis and growth under manganese deficiency. *Plant Physiol.* 152 (4), 1986–1999. doi: 10.1104/pp.109.150946

Lee, S., and An, G. (2009). Over-expression of OsIRT1 leads to increased fe and zinc accumulations in rice. *Plant Cell Environ.* 32 (4), 408–416. doi: 10.1111/j.1365-3040.2009.01935.x

Li, S., Zhou, X., Huang, Y., Zhu, L., Zhang, S., Zhao, Y., et al. (2013). Identification and characterization of the zinc-regulated transporters, fe-regulated transporter-like protein (ZIP) gene family in maize. *BMC Plant Biol.* 13 (1), 114. doi: 10.1186/1471-2229-13-114

Mahmoudi, H., Ksouri, R., Gharsalli, M., and Lachaâl, M. (2005). Differences in responses to fe deficiency between two legumes: lentil (Lens culinaris) and chickpea (Cicer arietinum). *J. Plant Physiol.* 162 (11), 1237–1245. doi: 10.1016/j.jplph.2004.12.009

Masuda, H., Aung, M., and Nishizawa, S. (2013). Fe biofortification of rice using different transgenic approaches. *Rice* 6 (1), 1–12. doi: 10.1186/1939-8433-6-40

Mendes, I. V. S. B. (2014). *Evaluation of the physiological and mocular impact of fe deficiency in two legume species* (Porto, Portugal: University of Católica). Available at: https://repositorio.ucp.pt/bitstream/10400.14/16216/1/tese.pdf.

Mengistu, M. A. (2016). Response of Chickpea Cultivars to Dryland Management in Southern Ethiopia. PhD Thesis. College of Graduate and Postdoctoral Studies, University of Saskatchewan, Saskatoon, Canada. 156.

Mengistu, M. A. (2016). Response of chickpea cultivars to dryland management in Southern Ethiopia. PhD thesis. College of graduate and postdoctoral studies, University of Saskatchewan, Saskatoon, Canada. 156

Millán, T., Madrid, E., Cubero, J. I., Amri, M., Patricia, C., and Rubio, J. (2015). "Chickpea," in *Handbook of plant breeding*. Ed. A. M. D. R.(Pontevedra, Spain: Springer Science+Business Media New York), (pp. 85–88. doi: 10.1007/978-1-4939-2797-5

Peng, Z. P., and Li, C. J. (2005). Transport and partitioning of phosphorus in wheat as affected by p withdrawal during flag-leaf expansion. *Plant Soil* 268, 1–11. doi: 10.1007/s11104-004-0297-1

Pereira, M., Santos, C., Gomes, A., and Vasconcelos, M. (2014). Cultivar variability of fe uptake mechanisms in rice (Oryza sativa l.). *Plant Physiol. Biochem.* 85, 21–30. doi: 10.1016/j.plaphy.2014.10.007

Petit, J., Briat, J., and Lobréaux, S. (2001). Structure and differential expression of the four members of the arabidopsis thaliana ferritin gene family. *Biochem. J.* 359 (3), 575–582. doi: 10.1042/bj3590575

Ravet, K., Touraine, B., Boucherez, J., Briat, J.-F., and Gaymard, F., C. (2009). Ferritins fe zero interaction between fe homeostasis and oxidative stress in arabidopsis. *Plant J.* 57 (3), 400–412. doi: 10.1111/j.1365-313X.2008.03698.x

Robinson, N. J., Procter, C. M., Connolly, E. L., and Guerinot, M. L. (1999). A ferric-chelate reductase for fe uptake from soils. *Nature* 397, 694–697. doi: 10.1038/17800

Sankaran, R. P., and Grusak, M. A. (2014). Whole shoot mineral partitioning and accumulation in pea (Pisum sativum). *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00149

Santos, C., Silva, A., Serrão, I., Carvalho, A., and Vasconcelos, M. (2013). Transcriptomic analysis of fe deficiency related genes in the legumes. *Food Res. Int.* 54 (1), 1162–1171. doi: 10.1016/j.foodres.2013.06.024

Satterthwaite, D., McGranahan, G., and Tacoli, C. (2010). Urbanization and its implications for food and farming. *Philos. Trans. R. Soc. B: Biol. Sci.* 365 (1554), 2809–2820. doi: 10.1098/rstb.2010.0136

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3 (6), 1101–1108. doi: 10.1038/nprot.2008.73

Tan, G. Z. H., Das Bhowmik, S. S., Hoang, T. M. L., Karbaschi, M. R., Johnson, A. A. T., Williams, B., et al. (2017). Finger on the pulse: Pumping fe into chickpea. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01755

Thomine, S., Wang, R., Ward, J., Crawford, N., and Schroeder, J. (2000). Cadmium and fe transport by members of a plant metal transporter family in arabidopsis with homology to nramp genes. *Proc. Natl. Acad. Sci. U.S.A.* 97 (9), 4991–4996. doi: 10.1073/pnas.97.9.4991

Vasconcelos, M., Eckert, H., Arahana, V., Graef, G., Grusak, M. A., and Clemente, T. (2006). Molecular and phenotypic characterization of transgenic soybean expressing the arabidopsis ferric chelate reductase gene, FRO2. *Planta* 224 (5), 1116–1128. doi: 10.1007/s00425-006-0293-1

Vert, G., Briat, J., and Curie, C. (2003). Dual regulation of the arabidopsis high-affinity root fe uptake system by local and long-distance signals1. *Plant Physiol.* 132 (2), 796–804. doi: 10.1104/pp.102.016089

Waters, B. M., and Grusak, M. A. (2008). Whole-plant mineral partitioning throughout the life cycle in arabidopsis thaliana ecotypes Columbia, landsberg erecta, cape Verde islands, and the mutant line ysl1ysl3. *New Phytol.* 177 (2), 389–405. doi: 10.1111/j.1469-8137.2007.02288.x

Waters, B., Uauy, C., Dubcovsky, J., and Grusak, M. (2009). Wheat (Triticum aestivum) NAM proteins regulate the translocation of fe, zinc, and nitrogen compounds from vegetative tissues to grain. *J. Exp. Bot.* 60 (15), 4263–4274. doi: 10.1093/jxb/erp257

WHO (2002) *World health organization*. Available at: https://www.who.int/.

Wirth, J., Poletti, S., Aeschlimann, B., Yakandawala, N., Drosse, B., Osorio, S., et al. (2009). Rice endosperm fe biofortification by targeted and synergistic action of nicotianamine synthase and ferritin. *Plant Biotechnol. J.* 7 (7), 631–644. doi: 10.1111/j.1467-7652.2009.00430.x

World Health Organization (2005). "Worldwide prevalence of anaemia: WHO global database on anaemia," in *WHO global database on anaemia*. doi: 10.1017/S1368980008002401

World Summit on Food Security (2009) *Declaration of the world summit on food security*. Available at: http://www.fao.org/fileadmin/templates/wsfs/summi.

Yang, X., Huang, J., Jiang, Y., and Zhang, H. (2009). Cloning and functional identification of two members of the ZIP (Zrt, irt-like protein) gene family in rice (Oryza sativa l.). *Mol. Biol. Rep.* 36 (2), 281–287. doi: 10.1007/s11033-007-9177-0

Zhang, Y., Xu, Y., Yi, H., and Gong, J. (2012). Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate fe translocation between flag leaves and seeds in rice. *Plant J.* 72 (3), 400–410. doi: 10.1111/j.1365-313X.2012.05088.x

Zheng, L., Ying, Y., Wang, L., Wang, F., Whelan, J., and Shou, H. (2010). Identification of a novel fe regulated basic helix-loop-helix protein involved in fe homeostasis in oryza sativa. *BMC Plant Biol.* 10 (1), 166. doi: 10.1186/1471-2229-10-166

Zhu, H., Choi, H., Cook, D. R., and Shoemaker, R. C. (2005). Bridging model and crop legumes through comparative genomics. *Plant Physiol.* 137 (4), 1189–1196. doi: 10.1104/pp.104.058891

Zhu, C., Naqvi, S., Gomez-Galera, S., Pelacho, A. M., Capell., T., and Christou, P. (2007). Transgenic strategies for the nutritional enhancement of plants. *Trends Plant Sci.* 12 (12), 1360–1385. doi: 10.1016/j.tplants.2007.09.007

Zuo, Y., and Zhang, F. (2011). Soil and crop management strategies to prevent fe deficiency in crops. *Plant Soil* 339 (1), 83–95. doi: 10.1007/s11104-010-0566-0

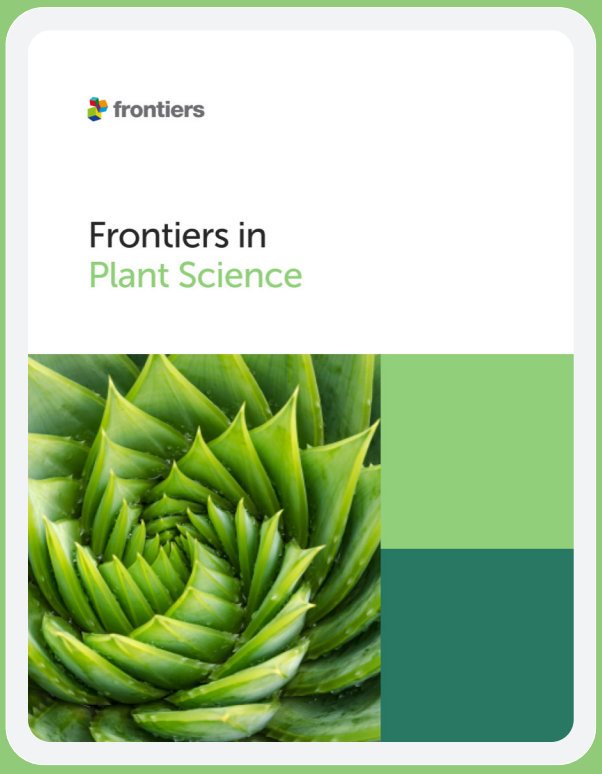
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact