



Evolution and significance of the *Lon* gene family in *Arabidopsis* organelle biogenesis and energy metabolism

Stamatis Rigas, Gerasimos Daras, Dikran Tsitsekian, Anastasios Alatzas and Polydefkis Hatzopoulos*

Department of Biotechnology, Agricultural University of Athens, Athens, Greece

Edited by:

Daniel H. Gonzalez, Universidad Nacional del Litoral, Argentina

Reviewed by:

Diego Fabian Gomez-Casati, National University of Rosario – National Scientific and Technical Research Council, National Research Council of Argentina, Argentina
Chris Carrie, Ludwig-Maximilians-Universität, Germany

*Correspondence:

Polydefkis Hatzopoulos, Department of Biotechnology, Agricultural University of Athens, Iera Odos 75, Athens 118 55, Greece
e-mail: phat@aua.gr

Lon is the first identified ATP-dependent protease highly conserved across all kingdoms. Model plant species *Arabidopsis thaliana* has a small *Lon* gene family of four members. Although these genes share common structural features, they have distinct properties in terms of gene expression profile, subcellular targeting and substrate recognition motifs. This supports the notion that their functions under different environmental conditions are not necessarily redundant. This article intends to unravel the biological role of Lon proteases in energy metabolism and plant growth through an evolutionary perspective. Given that plants are sessile organisms exposed to diverse environmental conditions and plant organelles are semi-autonomous, it is tempting to suggest that *Lon* genes in *Arabidopsis* are paralogs. Adaptive evolution through repetitive gene duplication events of a single archaic gene led to *Lon* genes with complementing sets of subfunctions providing to the organism rapid adaptability for canonical development under different environmental conditions. Lon1 function is adequately characterized being involved in mitochondrial biogenesis, modulating carbon metabolism, oxidative phosphorylation and energy supply, all prerequisites for seed germination and seedling establishment. Lon is not a stand-alone proteolytic machine in plant organelles. Lon in association with other nuclear-encoded ATP-dependent proteases builds up an elegant nevertheless, tight interconnected circuit. This circuitry channels properly and accurately, proteostasis and protein quality control among the distinct subcellular compartments namely mitochondria, chloroplasts, and peroxisomes.

Keywords: Lon, mitochondria, chloroplasts, protein dual-targeting, gene expression, gene evolution, molecular modeling, energy metabolism

NO GIFT WITHOUT A PRICE: LIFE IN AN AEROBIC WORLD IS NOT NECESSARILY PERFECT

The paradox of aerobic life or the “Oxygen Paradox” argues that organisms do not survive in oxygen depleted environments, yet oxygen is inherently dangerous to their existence. This “dark side” of oxygen is attributed to the damage of biomolecules (Davies, 1995). Life in an oxygenated environment contributed to the evolution of aerobic metabolic processes such as respiration and photosynthesis that unavoidably result in the production of molecular oxygen metabolites known as reactive oxygen species (ROS). Although increasing evidence indicates that ROS in plants could function as signaling molecules in regulating development and pathogen defense response, ROS have the capacity to stochastically cause oxidative damage to proteins, DNA, and lipids (Apel and Hirt, 2004; Møller et al., 2007).

Mitochondria, chloroplasts, and peroxisomes represent subcellular sources for ROS production and the principle targets for oxidative macromolecular damage. In particular, the electron transport chain of mitochondria transfers high energy electrons to oxygen through a series of inner membrane protein complexes. This process of electron transfer from NADH or FADH₂ to O₂ by the electron carriers, known as oxidative phosphorylation (OXPHOS), is leading to energy production in the form of ATP. However, through this process leakage of

electron occurs, ultimately generating highly reactive species, causing severe cell damage. This side-effect in mitochondria led to the “free-radical theory,” conceived in 1956, speculating that endogenous oxygen radicals were generated within cells and resulted in a pattern of cumulative damage (Harman, 1956). Nowadays, this theory is widely appreciated by an increasing number of scientists from an expanding circle of fields, including plant biologists, supporting the role of oxidants in cellular damage (Beckman and Ames, 1998; Finkel and Holbrook, 2000).

To cope with the hostile oxygenated environment, organisms have evolved sophisticated networks of defense (Finkel and Holbrook, 2000; Apel and Hirt, 2004; Friguet et al., 2008). The first line of defense against oxidative injury is composed of a complex array of ROS detoxifying enzymes and non-enzymatic antioxidants that counteract and regulate the overall ROS levels, maintaining physiological homeostasis. During physiological steady state conditions the cellular oxidants are efficiently scavenged by these antioxidative defense components that are often confined to particular compartments. However, under persisting adverse oxidative conditions the equilibrium between ROS production and scavenging is perturbed resulting in rapid intracellular accumulation of oxidants. These disturbances characterized as oxidative stress, induce modifications to

both the polypeptide backbone and amino acid side chains of proteins.

As plants are sessile organisms exposed to harsh environmental conditions, numerous abiotic conditions result in protein misfolding usually caused by ROS-mediated chemical modifications. These conditions include exposure to high light intensity, drought stress, low or high temperature and mechanical stress (Apel and Hirt, 2004; Møller et al., 2007). The misfolded proteins are particularly prone to oxidation (Dukan et al., 2000) leading to the formation of adducts that often bring in carbonyl groups and cross-links (Friguet et al., 2008). The carbonylated proteins are functionally impaired or completely inactive, creating toxic protein aggregates and cross-linked inclusion bodies that interfere with normal cellular function (Petropoulos and Friguet, 2006). Hence, the second line of defense against oxidative injury is composed of the protein quality control mechanisms that essentially ensure the proper level of functional proteins within the cell and eliminate non-functional proteins.

The ATP-dependent Lon protease is a key component of protein quality control highly conserved across the kingdoms of living organisms. This article presents important findings and the progress recently made in plants, whereas special emphasis is simultaneously given on major scientific breakthroughs regarding the Lon function in non-plant organisms. This comparative approach will contribute toward better understanding of Lon in organellar proteostasis and cellular homeostasis.

THE AAA⁺ Lon PROTEASE IS A MAJOR COMPONENT OF PROTEIN QUALITY CONTROL MECHANISM

Protein fate depends on an elegant protein quality control system that precisely orchestrates protein complex assembly and degradation, thereby safeguarding cellular homeostasis especially under stress conditions. The role of protein quality control is biphasic, as it is composed of energy-dependent repair molecular chaperones and degradation machines. Chaperones and proteases represent two sides of the same coin, acting in opposing pathways to clear unfolded proteins from the cell (Voos, 2013). The molecular chaperones within the cell facilitate the folding of newly synthesized proteins into their native conformations, prevent aggregation and assist in the assembly of multiprotein complexes. Conversely, ATP-dependent proteases degrade irreparably damaged or improperly synthesized proteins. In the cytosol and nucleus of higher eukaryotes, the proteins to be removed are ubiquitinated and delivered to the 26S proteasome for degradation (Hershko and Ciechanover, 1998). The 26S proteasome is the most elaborate AAA⁺ protease (ATPases associated with diverse cellular activities), consisting of a 20S protease core particle and two 19S regulatory caps modulating several aspects of plant development (Coux et al., 1996; Voges et al., 1999; Smalle and Vierstra, 2004). In contrast to these cellular compartments and as a legacy of their endosymbiotic heritage, eukaryotic organelles maintain independent AAA⁺ protein degradation machineries categorized into the soluble Lon and Clp (caseinolytic protease) families and the membrane-integrated FtsH-class (filament-forming temperature-sensitive) proteases (also called as AAA-proteases; Adam et al.,

2001; Sinvany-Villalobo et al., 2004; Sakamoto, 2006; Rigas et al., 2012; Janska et al., 2013). In the case of FtsH and Lon, the ATPase and proteolytic domains are formed by a single polypeptide, whereas in Clp these domains are separate proteolytic (ClpP) and chaperone-like (ClpX) subunits.

Protease La encoded by the *Lon* gene homolog in *Escherichia coli*, was the first discovered AAA⁺ protease (Chung and Goldberg, 1981). As “La” is the sixth musical note of the solfège syllable, the nomenclature describes the order of Lon elution from the chromatographic analysis of *E. coli* soluble proteolytic enzymes (Swamy and Goldberg, 1981). Lon is an ubiquitous proteolytic machine present in unicellular and multicellular organisms. The Lon protease consists of three functional domains (Rotanova et al., 2006; Rigas et al., 2012). The long N-terminal domain that possibly together with the central AAA⁺ module selectively interact with target proteins and the C-terminal proteolytic domain (P-domain) with a typical Serine–Lysine catalytic dyad at the active center (Botos et al., 2004). In plants, the N-domain and the P-domain exhibit the highest evolutionary conservation. On the contrary, the AAA⁺ module that includes the Walker Box A and B motifs shows the highest degree of divergence in both amino acid composition and length, and is responsible for binding and hydrolysis of ATP (Rigas et al., 2009b). The orthologs of Lon are divided into two subgroups (Rotanova et al., 2006; Rigas et al., 2012): A type (A-Lons), which have a large multi-lobed N-terminal domain together with the ATPase and protease domains, and B type (B-Lons), which instead of the N domain have a membrane-anchoring region emerging from the ATPase domain. B-Lons are exclusively present in Archaea that lack FtsH and the Clp proteases and thereby B-Lons are the major ATP-dependent proteolytic machines in those cells. The soluble A-Lons are found in all bacteria and in eukaryotic cell organelles, such as mitochondria, chloroplasts, and peroxisomes (Lingard and Bartel, 2009; Rigas et al., 2009a,b, 2012). In the yeast *Saccharomyces cerevisiae*, Pim1 (proteolysis in mitochondria) the homologous Lon protease has a major role in mitochondrial proteostasis as this organism lacks Clp (Venkatesh et al., 2012).

As a chambered protease, the 26S proteasome degrades protein substrates that carry multiple ubiquitin moieties (Hershko and Ciechanover, 1998). Given that mitochondria do not exhibit a covalent tagging system for damaged proteins like the ubiquitin tag in the cytosol or nucleus, substrate selectivity of Lon ATP-dependent protease most likely depends on the intrinsic characteristics of the polypeptide to be degraded. Lon preferentially degrades to completion damaged or misfolded polypeptides having a 50–60 amino acid long unstructured and exposed protein segment with compromised conformational state (von Janowsky et al., 2005). Upon protein misfolding, specific sequences rich in aromatic and hydrophobic residues become accessible to be recognized by Lon (Gur and Sauer, 2008). Moreover, Lon can also degrade folded unassembled polypeptides that can be accommodated into the proteolytic central channel with surfaced-exposed hydrophobic residues located within a highly charged environment (Ondrovicová et al., 2005). Consequently, Lon selectively degrades untagged non-natively folded substrates or folded but unassembled subunits, ultimately protecting the functional integrity of the organellar proteome.

EXPRESSION AND PROTEIN TARGETING OF *Arabidopsis* Lon GENES

The protein isoforms of Lon are encoded by small nuclear gene families and predicted to be targeted to distinct subcellular organelles. In *Arabidopsis*, four nuclear genes have been identified that potentially encode for members of the Lon family (Sinvány-Villalobo et al., 2004; Janska et al., 2010; Rigas et al., 2012). On the basis of publicly available microarray data in the Genevestigator database and scientific reports (Rigas et al., 2009a) Lon genes in *Arabidopsis* are expressed in various cell types and tissues. Nevertheless, the Lon genes show distinct expression profiles (Figure 1). The expression of *Lon1* (At5g26860) is high in rapidly growing organs of embryonic origin and in high-energy dependent tissues, which have increased mitochondria population per cell to sustain increased energy requirements. *Lon1* is predominantly expressed in germinating seeds, embryonic organs, including cotyledons and primary roots, and in organs with high growth rates such as developing inflorescences, while it was hardly detected in mature roots or stems of adult plants (Figure 1; Rigas et al., 2009a). In comparison to *Lon1*, *Lon4* (At3g05790) shows the lowest level of expression, albeit *Lon4* gene response is qualitatively similar to *Lon1*. Among the members of the Lon gene family, *Lon2* (At5g47040) is highly expressed, while gene expression generally remains constant among the examined cell types and tissues without significant fluctuations. Due to the lack of experimental evidence to report the presence of gene transcripts, *Lon3* (At3g05780) is presumed to be a pseudogene (Ostersetzer et al., 2007; Rigas et al., 2009a). However, microarray data strongly support that *Lon3* expression dominates in sperm cells. This specific

Lon3 expression profile implies a potential role in plant reproduction and particularly in male gametes maturation and double fertilization. Apart from the sperm cells, the marginal level of *Lon3* expression detected in other tissues most likely represents experimental noise impossible to be filtered as *Lon3* and *Lon4* are almost identical.

Most of the nuclear-encoded proteins are specifically targeted to a single organelle. However, dual-targeting of proteins to mitochondria and chloroplasts has been surprisingly frequent due to their post-endosymbiotic evolution (Millar et al., 2006; Baudisch et al., 2014). Two types of dual-targeting configurations have been reported in plants: twin and ambiguous presequences (Peeters and Small, 2001; Silva-Filho, 2003; Mackenzie, 2005; Baudisch et al., 2014). The ambiguous presequence generates a single protein isoform with a targeting peptide recognized by the import apparatus of both mitochondria and chloroplasts. Although this configuration can be organized in domains determining targeting specificity to an individual organelle, the signals responsible for organelle targeting most frequently overlap (Berglund et al., 2009a). Hence, ambiguous presequences cannot be completely distinguished from organelle-specific targeting peptides and they have an intermediate amino acid composition using the same organelle import pathways as the organelle-specific proteins. Despite the fact that the determinants for dual-targeting are not fully understood, the physicochemical properties within the N-terminal of the ambiguous presequences including hydrophobicity, the charge of amino acids and secondary structure, modulate the double localization (Berglund et al., 2009b; Ge et al., 2014). The twin presequences include two distinct targeting domains arranged in tandem at

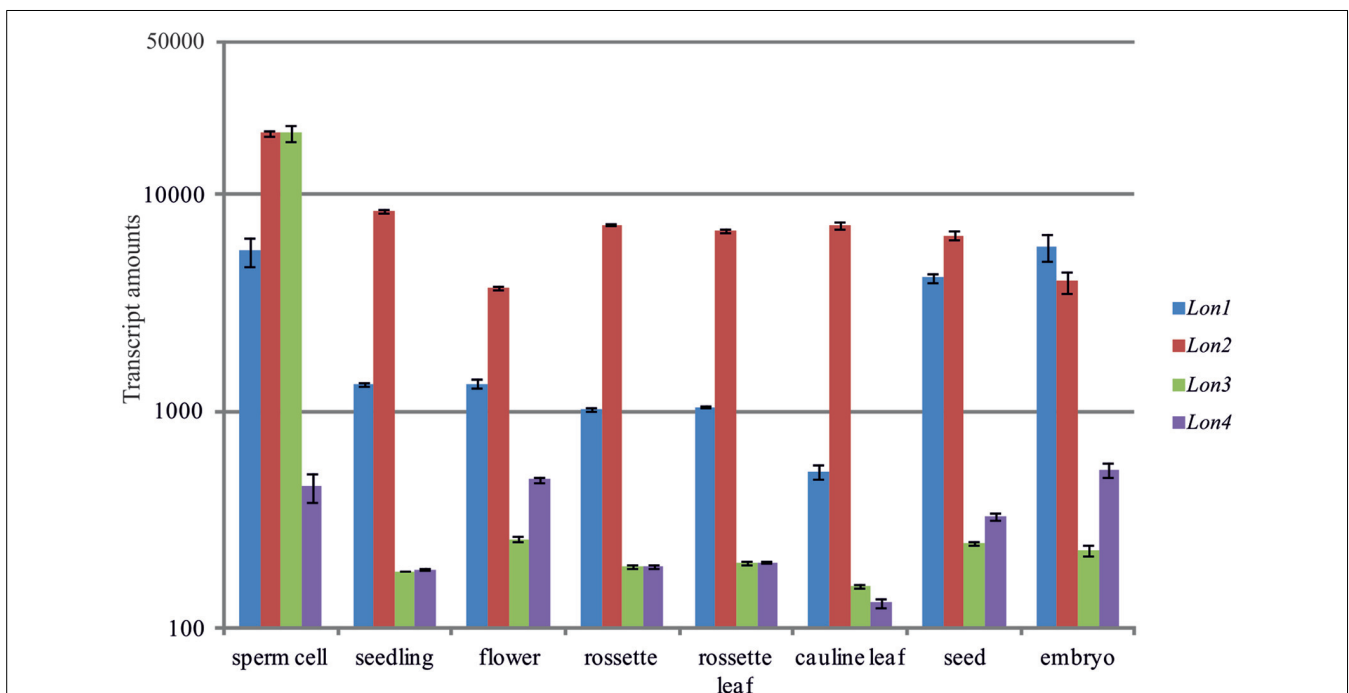


FIGURE 1 | Comparative analysis of the *Arabidopsis* Lon gene expression profiles in various cell types and tissues. Gene expression data were obtained from the Genevestigator server (Hruz et al., 2008). The vertical axis uses a base 10 logarithmic scale.

the N-terminus. In eukaryotes, twin presequences can confer dual-targeting to distinct subcellular compartments by employing two alternative in-frame translational initiation codons (Danpure, 1995; Silva-Filho, 2003; Carrie et al., 2009; Carrie and Small, 2013). Both ambiguous and twin presequences amplify the number of protein isoforms in subcellular compartments without affecting genome size. The majority of dual-targeted proteins in plants contain an ambiguous presequence showing an overall prevalence over twin presequences (Carrie et al., 2009).

The protein isoforms encoded by the nuclear *Lon* genes in *Arabidopsis* are scattered to plant cell organelles mainly involved in energy metabolism by utilizing different mechanisms of protein translocation (Figure 2). Lon4 is dual-targeted to mitochondria and chloroplasts displaying an ambiguous presequence (Sakamoto, 2006; Ostersetzer et al., 2007). The C-terminus of Lon2 bears a type 1 peroxisome-targeting signal (PTS1) conferring protein localization in peroxisomes (Lingard and Bartel, 2009). Computational analysis of Lon3 N-terminal domain identified a potential ambiguous presequence for dual-organelle localization to chloroplasts and mitochondria. The Lon1 dual-targeting is regulated both at the transcriptional and translational level (Daras et al., 2014).

THE SUBSTRATE RECOGNITION MOTIFS ARE HIGHLY VARIABLE AMONG THE *Arabidopsis* LON PROTEIN ISOFORMS

The AAA⁺ Lon protease is a soluble single-ringed multimeric holoenzyme. Adjacent to the AAA⁺ module is the sensor- and

substrate-discrimination (SSD) domain mainly involved in modulating selective substrate recognition by Lon so as the target protein to be degraded. In line with its highly selective mode of action, the SSD domain exhibits substantial interspecies and within the same species diversity (Rigas et al., 2009b). The yeast *S. cerevisiae* Lon purified from mitochondria is a ring-shaped complex with seven flexible subunits as determined by analytical ultracentrifugation and cryoelectron microscopy (Stahlberg et al., 1999). Subunits of *E. coli* Lon are known to assemble into ring-shaped homohexamers that enclose an internal degradation chamber. These hexamers may also interact to form a dodecamer at physiological protein concentrations (Vieux et al., 2013). Insights may also be gained from the structure of an intact, assembled Lon protease from the hyperthermophilic archaea *Thermococcus onnurineus* (*TonLon*) that is ubiquitously present in various deep-sea hydrothermal vent systems. *TonLon* is a 635-residue protein belonging to the B-Lon family having the protease domains arranged with a near perfect sixfold symmetry relative to the axial pore (Cha et al., 2010). This crystal structure suggests that the P-domains of each subunit form a bowl-like chamber with a lid formed by the AAA⁺ domains, such that substrates and degradation products may enter and exit the proteolytic chamber via opposing axial pores. Likewise, the homology model of human Lon suggests an hexameric complex formation that has an asymmetric, open-ring arrangement reminiscent of yeast Lon, albeit the yeast Lon fails to be modeled as a hexamer (Venkatesh et al., 2012). The structural features of Lon proteases in line with homology modeling provide conclusive

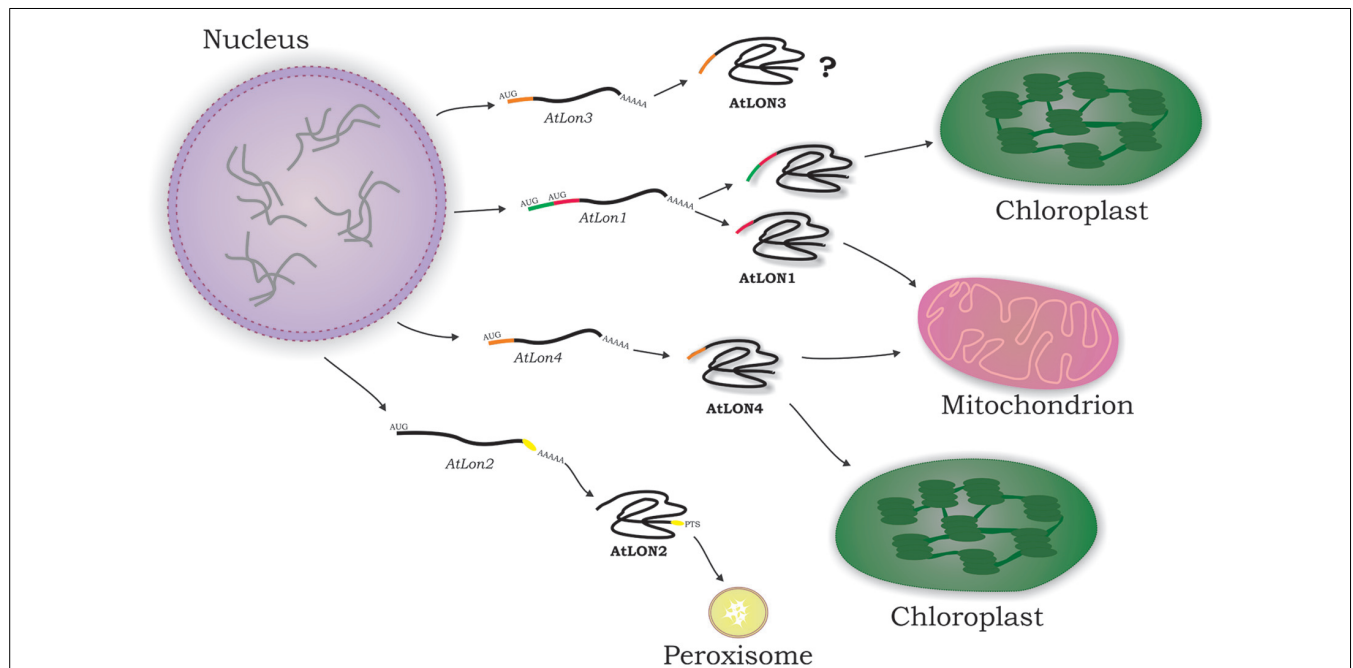


FIGURE 2 | *Arabidopsis* Lon protein isoforms are present in the subcellular compartments involved in energy metabolism. The majority of nuclear-encoded organellar proteins are translated on cytosolic ribosomes. In a distinct post-translational process, they have to be transported to their final destination in the organelle directed by specific targeting signals. Lon1 and

Lon4 are dual-targeted to mitochondria and chloroplasts due to twin and ambiguous presequences, respectively, at the N-terminus of the encoded protein isoforms. Lon2 is imported into peroxisomes by a canonical C-terminal type 1 peroxisome-targeting signal (PTS1). Lon3 subcellular localization remains elusive as yet.

evidence that distinguishes bacterial and human Lon proteases as hexameric complexes from yeast Lon, which is uniquely heptameric.

As the SSD domain is the most variable domain among the *Arabidopsis* Lon proteases, the architectural features of protein monomers were analyzed by molecular modeling. These ribbon models were in turn compared with the hexameric complexes of bacterial (*EcLon*) and human (*HsLon*) Lon proteases and with the heptameric complex of yeast Pim1. Homology modeling confirmed that *EcLon* and *HsLon* share the same structural features but differ from the heptameric Pim1 complex (**Figure 3A**). As reported by Venkatesh et al. (2012), this is most likely explained by the primary amino acid sequences of *EcLon* and *HsLon*, which are significantly shorter than Pim1 and *Arabidopsis* Lon sequences (Rigas et al., 2009b). Interestingly, the analysis revealed that structurally the *Arabidopsis* Lon proteases deviate from the hexameric *EcLon* and *HsLon* proteases fitting best with the heptameric yeast structure and due to distinct structural features they are classified into two groups (**Figure 3A**). The first group includes *AtLon1* and *AtLon3* that share similar structural properties with Pim1, all preserving a single pair of parallel β -sheets (depicted in red color) that is typical of *EcLon* and *HsLon*. *AtLon2* and *AtLon4* belong in the second group bearing between the α -helices (depicted in gray color) an internal loop polypeptide configuration (depicted in yellow color) with different secondary structure (depicted in purple color) compared to the members of the first group. Additionally, a surprising core structure is discovered in Pim1 and *Arabidopsis* Lon proteases likely originating from the hexameric *EcLon* and *HsLon* complexes (**Figure 3B**). The alignment between the models of

AtLon1 or *AtLon4* with the core structure of *HsLon* show that *AtLon1* internal geometry differs from *AtLon4*. The internal loop domain of *AtLon4* shows a right-handed extension, whereas in *AtLon1* is left-handed. This structural difference between the two major representative proteases of *Arabidopsis* suggests that *Lon1* and *Lon4* are gene paralogs performing specialized functions without being necessarily redundant. However, the possibility of recognizing similar protein targets under adverse environmental conditions that considerably modify the internal milieu of mitochondria and chloroplasts cannot be excluded. Future studies are required to assess the homology models of *Arabidopsis* Lon proteases and to solve the crystal structure of the holoenzyme, providing insights on Lon structural dynamics and functional versatility.

Lon1 AND Lon4 PARALOGS EVOLVED DISTINCT STRUCTURAL AND FUNCTIONAL FEATURES

Ancient invasions by eubacteria gave rise through symbiosis to mitochondria and chloroplasts that have enormous impact on bioenergetic and metabolic homeostasis of plants (Dyall et al., 2004). Mitochondria originated first from an endosymbiotic event of α -proteobacterial fusion. A second cyanobacterial invasion supplied the plant cell with the present-day chloroplast capable for photosynthesis. During the endosymbiotic process, the symbionts lost their autonomy by massive transfer of their genetic information to the host nucleus resulting in genetic redundancy. The evolution and establishment of the protein translocation machinery caused bulk gene loss leading to organellar genome reduction. The translocation process involved N-terminal extensions of the nuclear-encoded precursor protein

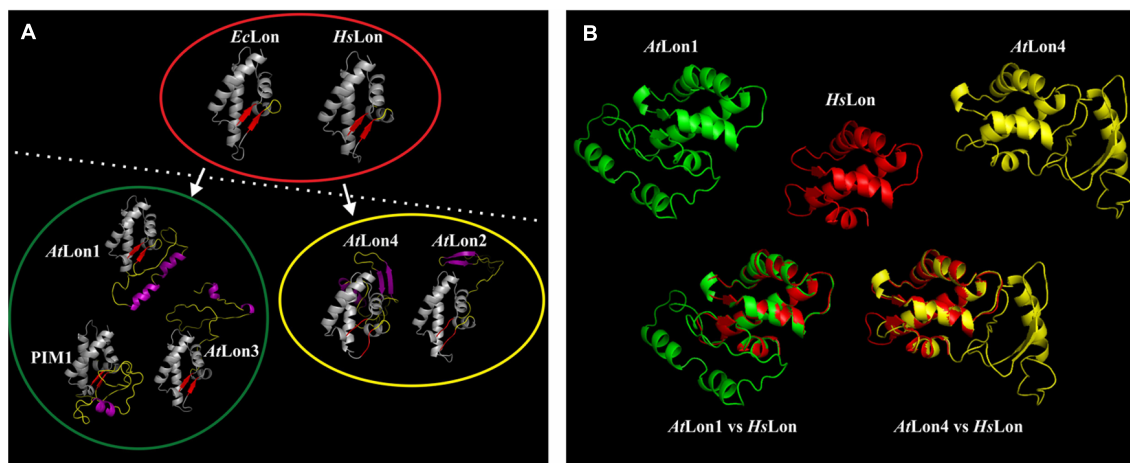


FIGURE 3 | Molecular modeling provides insights into the structural features of *Arabidopsis* Lon proteases. (A) Monomer structure comparison of the ribbon model of the sensor- and substrate-discrimination (SSD) domain discriminates the hexameric bacterial and human Lon complexes encompassed by red eclipse from possibly heptameric complexes of Pim1 and *Arabidopsis* Lon homologs. On the basis of discrete structural features the *Arabidopsis* proteases could be further categorized into two groups encompassed by green and yellow eclipses. The Lon protein accessions and the coordinates of the SSD domains given in parentheses are *AtLon1*: NP_568490 (603–739), *AtLon2*: NP_568675 (547–784), *AtLon3*: NP_566258

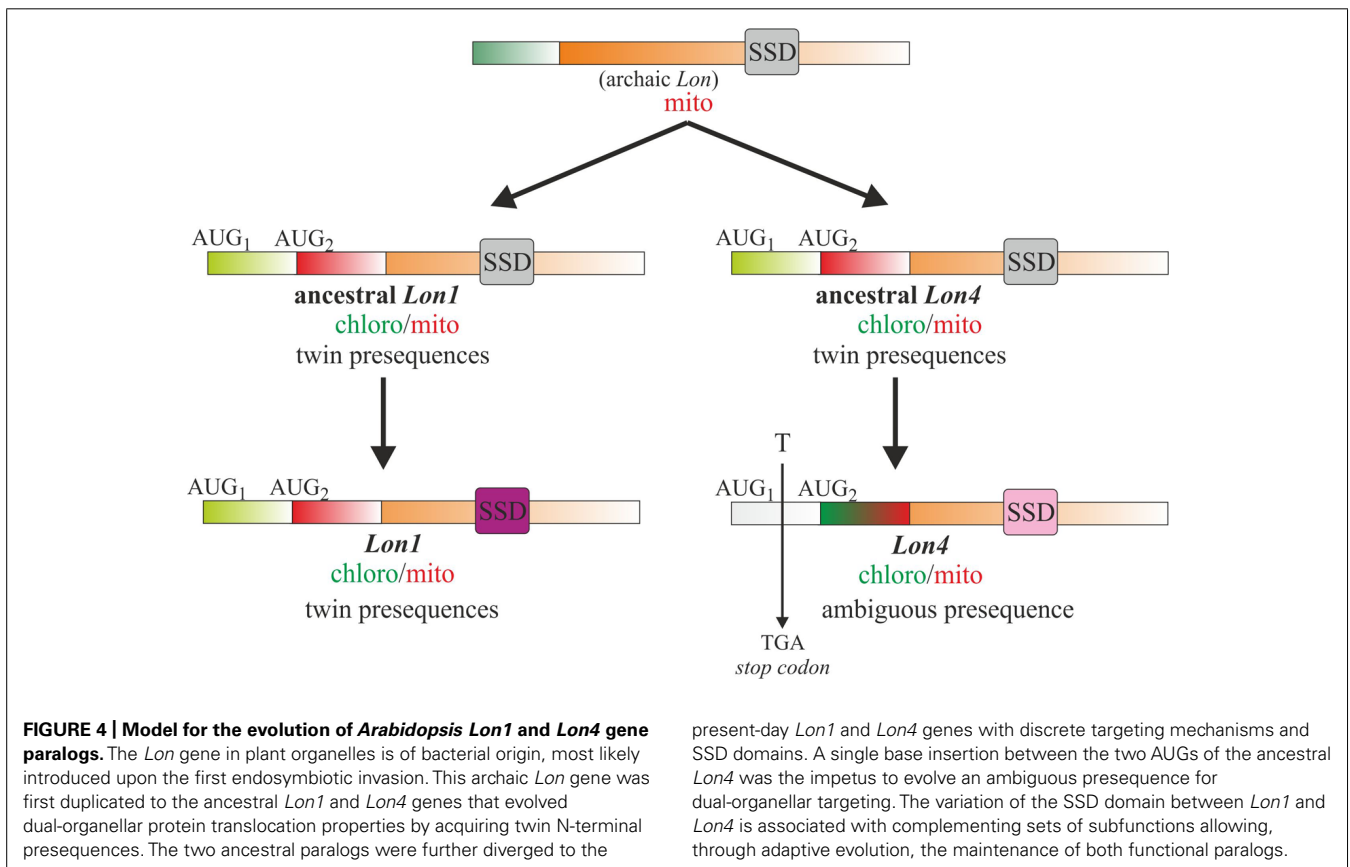
(586–726), and *AtLon4*: NP_566259 (585–733) from *Arabidopsis thaliana*, the *Homo sapiens* Lon: NP_004784 (662–747), Pim1: P36775 (772–911) from *Saccharomyces cerevisiae* and *EcLon*: AAC36871 (494–580) from *Escherichia coli*. Modeling of the SSD domain was performed on the basis of known crystallographic data mainly available from AAA⁺ proteins and bacterial Lon proteases, which were automatically detected by the Phyre2 Protein Fold Recognition Server (www.sbg.bio.ic.ac.uk/phyre2). The ribbon model was generated in PyMol (www.pymol.org). (B) Homology modeling distinguishes *AtLon1* (green) protease from *AtLon4* (yellow), albeit both preserve the core structure of the hexameric human Lon (red) homolog.

synthesized on cytoplasmic ribosomes. Coordinated evolution of protein import machineries from ancient symbionts led to dual-targeting of nuclear-encoded proteins to both mitochondria and chloroplasts.

Contrary to *Lon1* dual-targeting that is attributed to twin presequences, an ambiguous presequence confers *Lon4* dual-targeting specificity (Sakamoto, 2006; Ostersetzer et al., 2007). Besides the annotated initiation codon, surprisingly an additional upstream AUG is present in *Lon4* at the same place as the first initiation codon of *Lon1*. However, the reading frame between the first and second initiation codons of *Lon4* is interrupted by a single thymine base insertion that results in a TGA stop codon (Figure 4). Upon removal of this base the reading frame becomes open encoding an N-terminal extension conferring *Lon4* targeting specificity to chloroplasts similarly to the N-terminal extension of *Lon1*. Moreover, *Lon1* and *Lon4* are remarkably similar in terms of amino acid identity and similarity of the structural domains besides the SSD (Rigas et al., 2009b). As the structure of the SSD domain is tightly associated with Lon proteolytic activity (Figure 3), *Lon1* and *Lon4* gene duplicates were at the molecular level preserved through adaptive evolution with complementing sets of subfunctions (Lynch and Conery, 2000). The process of subfunctionalization provides an adaptive advantage by permitting a dynamic model of gene regulation so that each daughter protein performs a specialized function with greater precision than the ancestor. Taken together, these observations

support the notion that *Lon1* and *Lon4* are gene paralogs that evolved distinct mechanisms for dual-targeting and subsets of function.

The features of the two paralogs support the existence of an archaic *Lon* gene originated from the first bacterial endosymbiotic event that duplicated leading to the ancestral *Lon1* and *Lon4* genes with twin-presequence structures for dual-organellar targeting (Figure 4). This duplication event most likely occurred upon land plant evolution over 400 million years ago. Both ancestral paralogs evolved through adaptive evolution discrete SSD domains and thereby specialized roles in plant development and stress-response. Moreover, the ancestral *Lon4* gene evolved the mitochondrial presequence into an ambiguous one. This evolutionary step was characterized by a single thymine insertion interrupting the reading frame of *Lon4* to prevent the synthesis of the chloroplast transit peptide. This evolutionary process drifts protein dual-targeting from twin presequences to the ambiguous one. Considering that the cases of protein dual-targeting by ambiguous presequence prevail over twin presequences (Carrie et al., 2009), the dual-targeting driven by twin presequences plausibly represents an evolutionary fossil. In line with this model, additional duplication events of the archaic or the ancestral *Lon* genes likely occurred generating the present-day *Lon3* gene that is in close proximity to *Lon4* in a head-to-tail orientation and *Lon2*. While *Lon2* acquired a peroxisome-targeting signal, the mitochondrial presequence deteriorated. This series of duplication events does not exclude other



evolutionary pathways resulting in quadruple Lons in *Arabidopsis* genome. Nevertheless, the features of *Arabidopsis* Lon genes that determine protein isoform translocation in plant organelles together with the structure of the SSD domains argue in favor of the proposed model.

Lon1-DEPENDENT MITOCHONDRIAL BIOGENESIS IS ASSOCIATED WITH OXPHOS CAPACITY

Seed germination and seedling establishment depend on the assembly or biogenesis of mitochondria and the mobilization of storage reserves. In oilseed species like *Arabidopsis*, seedling establishment is supported by soluble sugars that are generated by storage oil mobilization. The mobilization of storage oil to sucrose involves main biochemical pathways compartmentalized into distinct organelles. The triacylglycerols contained in oil bodies are hydrolyzed to free fatty acids (FFAs). The FFAs are imported into the peroxisome entering the reactions of β -oxidation and the glyoxylate cycle. Seedling establishment additionally depends on the mitochondrial tricarboxylic acid (TCA) cycle and on gluconeogenesis that operates in the cytosol.

Molecular genetics revealed that *Lon1* is involved in the biogenesis and maintenance of mitochondrial function to ensure the proper operation of such biochemical network. Transmission electron microscopy studies of *lon1* mutants revealed the presence of mitochondria with abnormal morphology. The *lon1* mitochondria are swollen, having a poorly developed internal membrane system composed of few discernible cristae (Rigas et al., 2009a). These ultrastructural features of *lon1* mitochondria are reminiscent of the pro-mitochondrial morphology of dry seeds, supporting the role of *Arabidopsis* Lon1 protease in mitochondrial biogenesis during germination. Likewise, electron microscopy performed on Lon-deficient mitochondria of yeast (Suzuki et al., 1994) and human (Bota et al., 2005) cells revealed aberrant mitochondrial morphology with electron-dense inclusion bodies in the mitochondrial matrix most likely representing oxidatively modified and aggregated proteins. These severe phenotypes of Lon deficiency across eukaryotes demonstrate the importance of this proteolytic machine to maintain proper mitochondrial function. As the main mitochondrial electron transport chain consists of coupled respiratory chain complexes found in the inner mitochondrial membrane, the OXPHOS capacity of *lon1* mitochondria is expected to be impaired due to the abnormal mitochondrial morphology. Mitochondria isolated from *lon1* mutants showed reduced respiratory capacity when oxidizing succinate and cytochrome *c* via decreased activity of complexes II and IV, respectively (Rigas et al., 2009a). Additionally, in the absence of Lon1 the activities of at least five TCA cycle enzymes were significantly decreased. Analysis of the mitochondrial proteome revealed that complex I was additionally affected in *lon1* mutants (Solheim et al., 2012). Taken together, these results support the notion that Lon protease sustains the activity of major OXPHOS complexes during germination in *Arabidopsis*.

Despite primary metabolism and energy supply through OXPHOS, mitochondria also play a crucial role in cell signaling and communication. In mammalian cells, Lon protease under

hypoxic conditions optimizes the activity of the electron transport chain by modulating the equilibrium between cytochrome *c* oxidase (COX; complex IV) subunits COX4-1 and COX4-2 (Fukuda et al., 2007). Under reduced O₂ availability, the hypoxia-inducible transcription factor HIF-1 α binds to hypoxia response elements (HRE) of *Lon* gene promoter leading to the induction of *Lon* expression for the degradation of COX4-1. At the same time, HIF-1 α up-regulates an alternate isoform-COX4-2, which is assembled into the COX complex replacing COX4-1. In hypoxic cells, the COX4-2 containing complexes are better optimized for transporting electrons and increasing the efficiency of respiration. Additionally, *Lon* expression is enhanced *in vitro* by hypoxia or under endoplasmic reticulum (ER) stress and *in vivo* by brain ischemia (Hori et al., 2002). Under hypoxia or ER stress, a novel signaling pathway from ER to mitochondria disturbs the expression and assembly of COX, whereas the expression of *Lon* protects the mitochondria from unassembled complexes. Intriguingly, Lon was recently reported to be implicated into the cellular homeostasis of the *bZip* transcription factor *ATFS-1* (activating transcription factor associated with stress-1) that is required for the mitochondrial unfolded protein response (UPR^{mt}) cascade (Nargund et al., 2012). During mitochondrial stress, *ATFS-1* accumulates in the nucleus and activates the UPR^{mt} as *ATFS-1* import in mitochondria is inhibited due to reduced mitochondrial import efficiency by the localized in the inner mitochondrial membrane ATP-binding cassette transporter HAF-1 [half transporter (P-glycoprotein related)]. In healthy cells, the UPR^{mt} is not activated as *ATFS-1* is compartmentalized away from the nucleus efficiently imported in mitochondria matrix, where is rapidly degraded by the Lon protease. Consequently, mitochondrial homeostasis is maintained by the conditional-dependent translocation of a transcription activator between the nucleus, where it activates the stress response cascade, and mitochondria where it is removed by Lon protease.

CONCLUSIONS AND FUTURE PERSPECTIVES

Protein misfolding and degradation, especially in mitochondria which are the main source for oxidants in the cell, are processes that determine protein fate causing mitochondrial dysfunction. Mitochondrial dysfunction has now been implicated in aging, cancer and in a variety of age-related degenerative diseases. Lon, in association with other AAA⁺ proteases, modulate protein quality control, constitutive metabolism and adaptive responses to cellular or environmental stress. Our understanding of the physiological role of Lon proteases in plants is still evolving, although great advancement is made in non-plant species. However, contrary to the bacterial, yeast and mammalian counterparts, *Arabidopsis* has a genetic pluralism in terms of *Lon* gene copies within the nuclear genome. This could be attributed to the presence of an additional organelle in plants, the chloroplast, and to the fact that plants are sessile organisms exposed to extreme environmental conditions. The *Arabidopsis* *Lon* genes could be considered paralogs that evolved distinct structural and functional features including gene regulation and expression, subcellular targeting localization and substrate recognition mechanisms. Moreover, *Arabidopsis* has the genetic and molecular tools to contribute toward better understanding of the functional role of Lon as key

controller of proteostasis in organelles and in response to intrinsic or environmental cues. These *Arabidopsis* paralogs could be proven valuable assets to unravel the substrate recognition mechanisms and organelle-to-nucleus communication circuits. This knowledge might be of use to precisely comprehend the role of Lon in non-plant species including humans and thereby to improve life quality and expectancy.

ACKNOWLEDGMENTS

We apologize to colleagues whose relevant work could not be included due to space restrictions. Polydefkis Hatzopoulos acknowledges support by GSRT ARISTEIA/1200 2012–2015 grant. Gerasimos Daras and Dikran Tsitsekian are indebted for funding to IKY Fellowships of Excellence for Postgraduate Studies in Greece-Siemens Program.

REFERENCES

- Adam, Z., Adamska, I., Nakabayashi, K., Osterstetzer, O., Haussuhl, K., Manuell, A., et al. (2001). Chloroplast and mitochondrial proteases in *Arabidopsis*. A proposed nomenclature. *Plant Physiol.* 125, 1912–1918. doi: 10.1104/pp.125.4.1912
- Apel, K., and Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress and signal transduction. *Annu. Rev. Plant Biol.* 55, 373–399. doi: 10.1146/annurev.arplant.55.031903.141701
- Baudisch, B., Langner, U., Garz, I., and Klösgen, R. B. (2014). The exception proves the rule? Dual targeting of nuclear-encoded proteins into endosymbiotic organelles. *New Phytol.* 201, 80–90. doi: 10.1111/nph.12482
- Beckman, K. B., and Ames, B. N. (1998). The free radical theory of aging matures. *Physiol. Rev.* 78, 547–581.
- Berglund, A. K., Pujol, C., Duchene, A. M., and Glaser, E. (2009a). Defining the determinants for dual targeting of amino acyl-tRNA synthetases to mitochondria and chloroplasts. *J. Mol. Biol.* 393, 803–814. doi: 10.1016/j.jmb.2009.08.072
- Berglund, A. K., Spänning, E., Biverstahl, H., Maddalo, G., Tellgren-Roth, C., Mäler, L., et al. (2009b). Dual targeting to mitochondria and chloroplasts: characterization of Thr-tRNA synthetase targeting peptide. *Mol. Plant* 2, 1298–1309. doi: 10.1093/mp/ssp048
- Bota, D. A., Ngo, J. K., and Davies, K. J. (2005). Downregulation of the human Lon protease impairs mitochondrial structure and function and causes cell death. *Free Radic. Biol. Med.* 38, 665–677. doi: 10.1016/j.freeradbiomed.2004.11.017
- Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulo, F., et al. (2004). The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J. Biol. Chem.* 279, 8140–8148. doi: 10.1074/jbc.M312243200
- Carrie, C., Giraud, E., and Whelan, J. (2009). Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. *FEBS J.* 276, 1187–1195. doi: 10.1111/j.1742-4658.2009.06876.x
- Carrie, C., and Small, I. (2013). A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 1833, 253–259. doi: 10.1016/j.bbamcr.2012.05.029
- Cha, S. S., An, Y. J., Lee, C. R., Lee, H. S., Kim, Y. G., Kim, S. J., et al. (2010). Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J.* 29, 3520–3530. doi: 10.1038/emboj.2010.226
- Chung, C. H., and Goldberg, A. L. (1981). The product of the lon (capR) gene in *Escherichia coli* is the ATP-dependent protease, protease La. *Proc. Natl. Acad. Sci. U.S.A.* 78, 4931–4935. doi: 10.1073/pnas.78.8.4931
- Coux, O., Tanaka, K., and Goldberg, A. L. (1996). Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.* 65, 801–847. doi: 10.1146/annurev.bi.65.070196.004101
- Danpure, C. J. (1995). How can the products of a single gene be localized to more than one intracellular compartment? *Trends Cell Biol.* 5, 230–238. doi: 10.1016/S0962-8924(00)89016-9
- Daras, G., Rigas, S., Tsitsekian, D., Zur, H., Tuller, T., and Hatzopoulos, P. (2014). Alternative transcription initiation and the AUG context configuration control dual organellar targeting and functional competence of *Arabidopsis* Lon1 protease. *Mol. Plant* doi: 10.1093/mp/ssu030 [Epub ahead of print].
- Davies, K. J. (1995). Oxidative stress: the paradox of aerobic life. *Biochem. Soc. Symp.* 61, 1–31.
- Dukan, S., Farewell, A., Ballesteros, M., Taddei, F., Radman, M., and Nyström, T. (2000). Protein oxidation in response to increased transcriptional or translational errors. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5746–5749. doi: 10.1073/pnas.100422497
- Dyall, S. D., Brown, M. T., and Johnson, P. J. (2004). Ancient invasions: from endosymbionts to organelles. *Science* 304, 253–257. doi: 10.1126/science.1094884
- Finkel, T., and Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature* 408, 239–247. doi: 10.1038/35041687
- Friguet, B., Bulteau, A. L., and Petropoulos, I. (2008). Mitochondrial protein quality control: implications in ageing. *Biotechnol. J.* 3, 757–764. doi: 10.1002/biot.200800041
- Fukuda, R., Zhang, H., Kim, J. W., Shimoda, L., Dang, C. V., and Semenza, G. L. (2007). HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. *Cell* 129, 111–122. doi: 10.1016/j.cell.2007.01.047
- Ge, C., Spänning, E., Glaser, E., and Wieslander, A. (2014). Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in *Arabidopsis thaliana*. *Mol. Plant* 7, 121–136. doi: 10.1093/mp/sst148
- Gur, E., and Sauer, R. T. (2008). Recognition of misfolded proteins by Lon, a AAA⁺ protease. *Genes Dev.* 22, 2267–2277. doi: 10.1101/gad.1670908
- Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* 11, 298–300. doi: 10.1093/geronj/11.3.298
- Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479. doi: 10.1146/annurev.biochem.67.1.425
- Hori, O., Ichinoda, E., Tamatani, T., Yamaguchi, A., Sato, N., Ozawa, K., et al. (2002). Transmission of cell stress from endoplasmic reticulum to mitochondria: enhanced expression of Lon protease. *J. Cell Biol.* 157, 1151–1160. doi: 10.1083/jcb.200108103
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics* 2008:420747. doi: 10.1155/2008/420747
- Janska, H., Kwasniak, M., and Szczepanowska, J. (2013). Protein quality control in organelles – AAA/FtsH story. *Biochim. Biophys. Acta* 1833, 381–387. doi: 10.1016/j.bbamcr.2012.03.016
- Janska, H., Piechota, J., and Kwasniak, M. (2010). ATP-dependent proteases in biogenesis and maintenance of plant mitochondria. *Biochim. Biophys. Acta* 1797, 1071–1075. doi: 10.1016/j.bbmbio.2010.02.027
- Lingard, M. J., and Bartel, B. (2009). *Arabidopsis* LON2 is necessary for peroxisomal function and sustained matrix protein import. *Plant Physiol.* 151, 1354–1365. doi: 10.1104/pp.109.142505
- Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151–1155. doi: 10.1126/science.290.5494.1151
- Mackenzie, S. A. (2005). Plant organellar protein targeting: a traffic plan still under construction. *Trends Cell Biol.* 15, 548–554. doi: 10.1016/j.tcb.2005.08.007
- Millar, A. H., Whelan, J., and Small, I. (2006). Recent surprises in protein targeting to mitochondria and plastids. *Curr. Opin. Plant Biol.* 9, 610–615. doi: 10.1016/j.pbi.2006.09.002
- Møller, I. M., Jensen, P. E., and Hansson, A. (2007). Oxidative modifications to cellular components in plants. *Annu. Rev. Plant Biol.* 58, 459–481. doi: 10.1146/annurev.arplant.58.032806.103946
- Nargund, A. M., Pellegrino, M. W., Fiorese, C. J., Baker, B. M., and Haynes, C. M. (2012). Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation. *Science* 337, 587–590. doi: 10.1126/science.1223560
- Ondrovicová, G., Liu, T., Singh, K., Tian, B., Li, H., Gakh, O., et al. (2005). Cleavage site selection within a folded substrate by the ATP-dependent Lon protease. *J. Biol. Chem.* 280, 25103–25110. doi: 10.1074/jbc.M502796200
- Osterstetzer, O., Kato, Y., Adam, Z., and Sakamoto, W. (2007). Multiple intracellular locations of Lon protease in *Arabidopsis*: evidence for the localization of AtLon4 to chloroplasts. *Plant Cell Physiol.* 48, 881–885. doi: 10.1093/pcp/pcm052
- Peeters, N., and Small, I. (2001). Dual targeting to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 1541, 54–63. doi: 10.1016/S0167-4889(01)00146-X
- Petropoulos, I., and Friguet, B. (2006). Maintenance of proteins and aging: the role of oxidized protein repair. *Free Radic. Res.* 40, 1269–1276. doi: 10.1080/10715760600917144
- Rigas, S., Daras, G., Laxa, M., Marathias, N., Fasseas, C., Sweetlove, L. J., et al. (2009a). The role of Lon1 protease in post-germinative growth and maintenance of mitochondrial function in *Arabidopsis thaliana*. *New Phytol.* 181, 588–600. doi: 10.1111/j.1469-8137.2008.02701.x

- Rigas, S., Daras, G., Sweetlove, L. J., and Hatzopoulos, P. (2009b). Mitochondria biogenesis via Lon1 selective proteolysis: who dares to live for ever? *Plant Signal. Behav.* 4, 221–224. doi: 10.4161/psb.4.3.7863
- Rigas, S., Daras, G., Tsitsekian, D., and Hatzopoulos, P. (2012). The multifaceted role of Lon proteolysis in seedling establishment and maintenance of plant organelle function: living from protein destruction. *Physiol. Plant.* 145, 215–223. doi: 10.1111/j.1399-3054.2011.01537.x
- Rotanova, T. V., Botos, I., Melnikov, E. E., Rasulova, F., Gustchina, A., Maurizi, M. R., et al. (2006). Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. *Protein Sci.* 15, 1815–1828. doi: 10.1110/ps.052069306
- Sakamoto, W. (2006). Protein degradation machineries in plastids. *Annu. Rev. Plant Biol.* 57, 599–621. doi: 10.1146/annurev.arplant.57.032905.105401
- Silva-Filho, M. C. (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Curr. Opin. Plant Biol.* 6, 589–595. doi: 10.1016/j.pbi.2003.09.008
- Sinvany-Villalobo, G., Davydov, O., Ben-Ari, G., Zaltsman, A., Raskind, A., and Adam, Z. (2004). Expression in multigene families. Analysis of chloroplast and mitochondrial proteases. *Plant Physiol.* 135, 1336–1345. doi: 10.1104/pp.104.043299
- Smalle, J., and Vierstra, R. D. (2004). The ubiquitin 26S proteasome proteolytic pathway. *Annu. Rev. Plant Biol.* 55, 555–590. doi: 10.1146/annurev.arplant.55.031903.141801
- Solheim, C., Li, L., Hatzopoulos, P., and Millar, A. H. (2012). Loss of Lon1 in *Arabidopsis* changes the mitochondrial proteome leading to altered metabolite profiles and growth retardation without an accumulation of oxidative damage. *Plant Physiol.* 160, 1187–1203. doi: 10.1104/pp.112.203711
- Stahlberg, H., Kutejová, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G., et al. (1999). Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6787–6790. doi: 10.1073/pnas.96.12.6787
- Suzuki, C. K., Suda, K., Wang, N., and Schatz, G. (1994). Requirement for the yeast gene LON in intramitochondrial proteolysis and maintenance of respiration. *Science* 264, 273–276. doi: 10.1126/science.8146662
- Swamy, K. H., and Goldberg, A. L. (1981). *E. coli* contains eight soluble proteolytic activities, one being ATP dependent. *Nature* 292, 652–654. doi: 10.1038/292652a0
- Venkatesh, S., Lee, J., Singh, K., Lee, I., and Suzuki, C. K. (2012). Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim. Biophys. Acta* 1823, 56–66. doi: 10.1016/j.bbamcr.2011.11.003
- Vieux, E. F., Wohlever, M. L., Chen, J. Z., Sauer, R. T., and Baker, T. A. (2013). Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2002–E2008. doi: 10.1073/pnas.1307066110
- Voges, D., Zwickl, P., and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* 68, 1015–1068. doi: 10.1146/annurev.biochem.68.1.1015
- von Janowsky, B., Knapp, K., Major, T., Krayl, M., Guiard, B., and Voos, W. (2005). Structural properties of substrate proteins determine their proteolysis by the mitochondrial AAA+ protease Pim1. *Biol. Chem.* 386, 1307–1317. doi: 10.1515/BC.2005.149
- Voos, W. (2013). Chaperone-protease networks in mitochondrial protein homeostasis. *Biochim. Biophys. Acta* 1833, 388–399. doi: 10.1016/j.bbamcr.2012.06.005

Conflict of Interest Statement: 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.

Received: 26 November 2013; accepted: 26 March 2014; published online: 11 April 2014.

Citation: Rigas S, Daras G, Tsitsekian D, Alatzas A and Hatzopoulos P (2014) Evolution and significance of the Lon gene family in *Arabidopsis* organelle biogenesis and energy metabolism. *Front. Plant Sci.* 5:145. doi: 10.3389/fpls.2014.00145

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

Copyright © 2014 Rigas, Daras, Tsitsekian, Alatzas and Hatzopoulos. 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) or licensor 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.