



OPEN ACCESS

EDITED BY

Xiaosheng Tan,
Rutgers, The State University of New Jersey,
United States

REVIEWED BY

Zunsong Hu,
City of Hope National Medical Center,
United States
Piao Zhao,
Sichuan Agricultural University, China
Yifei Yang,
Cornell University, United States

*CORRESPONDENCE

Shigang Yin

✉ sgyin@swmu.edu.cn

Huangfan Xie

✉ xie1991@swmu.edu.cn;

Xin Tian

✉ tianxin202407@163.com

[†]These authors share first authorship

RECEIVED 12 June 2025

ACCEPTED 26 August 2025

PUBLISHED 12 November 2025

CITATION

Liu L, Mao X, Yang C, Li N, Zhou Y, Zhang L,
Jiang N, Huang Y, Yin S, Xie H and Tian X
(2025) Single-cell multi-omics analysis reveals
cellular subpopulations associated with
relapse in high-risk B-ALL following
intensified chemotherapy.
Front. Immunol. 16:1645546.
doi: 10.3389/fimmu.2025.1645546

COPYRIGHT

© 2025 Liu, Mao, Yang, Li, Zhou, Zhang, Jiang,
Huang, Yin, Xie and Tian. 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.

Single-cell multi-omics analysis reveals cellular subpopulations associated with relapse in high-risk B-ALL following intensified chemotherapy

Li Liu^{1,2†}, Xiaoyan Mao^{3†}, Chunhui Yang⁴, Na Li⁴, Yan Zhou⁴,
Li Zhang⁵, Nanjing Jiang⁶, Yu Huang⁶, Shigang Yin^{7*},
Huangfan Xie^{7*} and Xin Tian^{4*}

¹Department of Pediatrics, Qujing Medical College, Qujing, China, ²Institute of Advanced Biotechnology, Institute of Homeostatic Medicine, and School of Medicine, Southern University of Science and Technology, Shenzhen, China, ³Department of Pediatrics, Children Hematological Oncology and Birth Defects Laboratory, The Affiliated Hospital of Southwest Medical University, Sichuan Clinical Research Center for Birth Defects, Luzhou, Sichuan, China, ⁴Department of Hematology, Kunming Children's Hospital, Kunming, China, ⁵Department of Pediatrics, Da Li University, Da Li, China, ⁶Department of Hematology, Sichuan Provincial Women's and Children's Hospital/The Affiliated Women's and Children's Hospital of Chengdu Medical College, Chengdu, China, ⁷Luzhou Key Laboratory of Nervous System Disease and Brain Function, Southwest Medical University, Luzhou, Sichuan, China

Introduction: Acute lymphoblastic leukemia (ALL) is the most prevalent malignant tumor in children, with B-cell ALL (B-ALL) accounting for 85% of cases. Despite advancements in chemotherapy and supportive care, a subset of high-risk B-ALL patients still experience relapse post-treatment. The molecular mechanisms underlying the relapses after intensified chemotherapy remain poorly understood.

Methods: We performed an integrated single-cell multi-omics analysis combining single-cell RNA sequencing (scRNA-seq) and single-cell ATAC sequencing (scATAC-seq) on peripheral blood mononuclear cells (PBMCs) from pediatric high-risk B-ALL patients following early intensified chemotherapy, as well as from healthy controls. Bioinformatic pipelines were applied to assess cellular composition, chromatin accessibility, gene ontology enrichment, spectral clustering, and copy number variation.

Results: Significant differences in cellular composition were observed between the remission and non-remission groups, with the non-remission group exhibiting a notable increase in HSC/MPP and Pro-B cells. Copy number variation (CNV) analysis also revealed that the CNV levels in HSC/MPP and Pro-B cells were higher in the non-remission group compared to other cell types. We subsequently identified a subcluster associated with resistance to intensified therapy within both the HSC/MPP and Pro-B cell groups. The drug-resistant subcluster of HSC/MPP cells was characterized by high expression of TCF4, EBF1, ERG, AL589693.1, and CRIM1, as well as enrichment of the allograft rejection pathway and the Notch signaling pathway. The drug-resistant subcluster of Pro-B cells was characterized by high expression of RPS29, B2M, RPL41, RPS21, NEIL1, AC007384.1, and CRIM1, as well as enrichment of the B cell receptor signaling pathway.

Discussion: Our study identified distinct cellular subpopulations associated with treatment failure, provide insights into the molecular mechanisms underlying treatment resistance in B-ALL and may inform the development of targeted therapies for high-risk patients.

KEYWORDS

high-risk B-ALL, scRNA-seq, scATAC-seq, PBMCs, CNV

Introduction

Acute lymphoblastic leukemia (ALL) is the most common malignant tumor in children (1). B-cell acute lymphoblastic leukemia (B-ALL), which originates from B-lineage lymphoid progenitor cells, accounts for 85% of it and is characterized by the clonal expansion of immature B lymphocytes and suppression of normal hematopoiesis. With the continuous development of combined chemotherapy, risk stratification, the combination of new drugs, and advances in supportive treatment, the prognosis for children with B-ALL is now very high, with a five-year survival rate exceeding 90% (1, 2). However, for those patients categorized as high-risk, the prognosis remains less favorable, prompting the use of early intensified chemotherapy regimens such as CAML (Cyclophosphamide, Cytosine arabinoside, 6-Mercaptopurine, and Pegaspargase) to enhance treatment outcomes. Despite these efforts, a considerable subset of patients exhibits resistance to intensified chemotherapy, resulting in subsequent relapses, with the molecular underpinnings of this resistance remaining largely enigmatic.

Single-cell RNA sequencing (scRNA-seq) and single-cell Assay for Transposase Accessible Chromatin sequencing (scATAC-seq) are powerful tools that enable a comprehensive dissection of cellular heterogeneity and chromatin accessibility at the individual cell level (3). Wang et al. have uncovered the remarkable heterogeneity of exhausted T cells in B-ALL, characterizing these subsets by their unique expression signatures by scRNA-seq (4). Granja et al., utilizing an integrative approach of scRNA-seq and scATAC-seq in the study of mixed phenotype acute leukemia, have pinpointed potential regulatory elements and transcription factors linked to leukemia-specific genes, with a particular emphasis on the crucial role of the transcription factor RUNX1 (5). These technologies provide an unprecedented opportunity to identify novel biomarkers associated with treatment response and relapse in high-risk B-ALL patients.

Although single-cell RNA sequencing has been increasing applied to study leukemia heterogeneity, the integration of transcriptomic and epigenetic single-cell data remains rare in B-ALL research. Our study uniquely combines scRNA-seq and scATAC-seq to dissect the interplay between transcriptional states and chromatin accessibility at relapse, offering a more comprehensive

understanding of treatment resistance mechanisms. This dual-modality design allows us to pinpoint not only the cellular composition but also the regulatory logic underlying relapse, which has not been systematically investigated in high-risk pediatric B-ALL.

This study aims to leverage the power of scRNA-seq and scATAC-seq to delineate the molecular landscape of peripheral blood mononuclear cells (PBMCs) in high-risk B-ALL patients post early intensification treatment. We hypothesized that distinct immune subpopulations, particularly HSC/MPP and Pro-B cells, carry transcriptional and chromatin accessibility signatures that are associated with resistance to intensified chemotherapy and may serve as predictors of relapse. The identification of such biomarkers may pave the way for the development of targeted therapies and personalized treatment strategies, ultimately improving the prognosis for high-risk B-ALL patients.

Methods

Study design and patient cohort

Our study includes a cohort of six pediatric patients diagnosed with high-risk B-ALL who underwent early intensification treatment using the CAML protocol (Supplementary Table S1). High-risk B-ALL was diagnosed based on current Chinese pediatric ALL risk stratification guidelines. Peripheral blood samples were collected post-treatment, with a focus on the second round of treatment where three patients achieved remission and three experienced relapses. Additionally, peripheral blood samples from three age-matched healthy controls were obtained for comparative analysis.

Sample collection and processing

Peripheral blood mononuclear cells (PBMCs) were isolated from the collected samples using density gradient centrifugation at 600×g for 25 minutes at 20°C. Cell number and viability were assessed by fluorescence using Reward C100 cell counter. Informed consent was obtained from all participants or their guardians, and the study was approved by the Institutional Review Board.

Single-cell RNA sequencing

PBMCs were subjected to scRNA-seq to capture the transcriptomic landscape of individual cells. Single-cell RNA sequencing libraries were prepared using the Chromium Single Cell 3'V2 Reagent Kit, and generated using the 10x Genomic platform. All libraries were sequenced on Illumina NovaSeq 6000 platform. Each sample generated approximately 110 – 140 million raw bases, with Q30 scores above 91%. Quality control, data normalization, and dimensionality reduction were performed using established computational methods.

Upon receiving the FASTQ files from the 10x Genomics single-cell RNA sequencing results, we utilized Cell Ranger software (v7.1.0, 10x Genomics, USA) to align barcodes and UMIs to the reference genome (GRCh38), followed by filtering and quantifying gene expression at the single-cell level, resulting in a cell-gene expression matrix. Putative doublets were identified and removed using Scrublet (Scanpy pipeline) and DoubletFinder (Seurat workflow), based on the expected doublet rate from 10x Genomics specifications. This matrix was then processed using Seurat (v4.3.0, Satija Lab, USA) and Scanpy (v1.9.3, Theis Lab, Germany) (6). Cells were retained if they exhibited between 200 and 5,000 detected genes, 500 to 25,000 unique molecular identifiers (UMIs), and less than 10% mitochondrial gene expression. To further ensure data quality, doublets were identified using DoubletFinder (v2.0.3), with an expected doublet rate of 8% based on cell loading density and a pK value of 0.09. Cells with a doublet prediction score greater than 0.25 were excluded prior to downstream normalization. Cells with fewer than 200 detected genes or with mitochondrial gene read counts exceeding 10% were excluded. We integrated the data across different samples using the IntegrateLayers function through Canonical Correlation Analysis (CCA). Dimensionality reduction was performed via Principal Component Analysis (PCA), and the resulting data were clustered using UMAP. The clusters were annotated using ScType (v1.1.0, Ianevski Lab, Finland) (7), and the expression of marker genes for various immune cell types was examined. Following annotation, the R software (v4.3.1, R Foundation for Statistical Computing, Austria) was employed to calculate the proportion of each cell type across sample groups.

Single-cell ATAC sequencing

For the single-cell ATAC-seq data, Cell Ranger ATAC (v2.1.0, 10x Genomics, USA) was used to align the reads to the reference genome and calculate the binding activity across genomic regions. The binding sites and abundance information were read using Signac (v1.9.0, Stuart Lab, USA) (8), and the genomic regions were annotated accordingly. Nuclei with 3,000 – 100,000 unique fragments, transcription start site (TSS) enrichment score ≥ 8 , and fraction of reads in peaks (FRiP) ≥ 0.20 were retained. Potential doublets were identified using the ArchR addDoubletScores() function with an expected rate of 5%, and nuclei with scores > 1.0 were excluded prior to downstream analysis. The distribution of

binding peaks across various genomic elements was computed and statistically analyzed. Data normalization was performed, and sample integration was achieved using the Harmony algorithm (v0.1.0, Korsunsky Lab, USA), followed by UMAP clustering. To integrate single-cell transcriptomic and ATAC-seq data, the FindTransferAnchors function was used to calculate transfer anchors, and single-cell ATAC-seq data were annotated based on the single-cell RNA-seq data. The R software was then used to calculate the distribution proportions of each cell type within the single-cell ATAC-seq data.

Using the Seurat package's FindAllMarkers function, cell-type-specific regulatory peaks were identified (9). Overrepresented motifs were found by FindMotifs function. Subsequent functional enrichment analysis of these marker genes was conducted using the ClusterProfiler (v4.6.0, Yu Lab, China) and GSVA (v1.46.0 Hanzelmann Lab, Germany) packages to identify cell types closely related to the study's focus (10). The distribution of these motifs across cell types and genes was analyzed using single-cell ATAC-seq data.

Data integration and analysis

The scRNA-seq and scATAC-seq data were integrated using advanced computational approaches to dissect the cellular heterogeneity and regulatory programs in high-risk B-ALL. Differential gene expressions and chromatin accessibility analysis were performed to identify markers associated with treatment response and relapse. We constructed regulatory networks by integrating scATAC-seq data with scRNA-seq data, using SCENIC to infer potential regulatory interactions between transcription factors and their target genes (11).

For subgroup analysis of HSC/MPP cells, the relevant clusters were extracted from the integrated dataset. Cells were re-normalized using SCTransform, incorporating donor ID and percentage of mitochondrial gene expression as covariates. We selected the top 1,500 highly variable genes and conducted PCA, using the first 15 principal components for UMAP visualization and Louvain clustering. Batch effect correction was applied using the Harmony algorithm, with donor ID specified as the batch variable, and performed prior to dimensionality reduction. This approach ensured that downstream clustering reflected true biological heterogeneity within the HSC/MPP compartment.

Bioinformatic pipeline

We employed a robust bioinformatic pipeline that includes quality control, doublet detection, data normalization, dimensionality reduction using tools like PCA and t-SNE, and clustering algorithms to classify cells into distinct subtypes. Pathway enrichment analysis and gene set variation analysis (GSVA) were conducted to interpret the biological significance of the observed changes. Using the inferCNV algorithm (v1.0, Broad Institute, USA) for CNV analysis, with normal pro-B cells as references.

“High CNV” refers to greater variability in chromosomal expression profiles as estimated by inferCNV, relative to reference Pro-B cells. Both gains and losses contribute to the CNV score.

All the visualizations, including UMAP plots, violin plots, heatmaps, dot plots, and bar charts, were generated using Seurat (v4.3.0), ggplot2 (v3.4.0), and ComplexHeatmap (v2.14.0) packages in R. Motif logos and accessibility tracks were plotted using ArchR and Signac.

Statistical analyses were performed in R unless otherwise stated. Comparisons between two groups were conducted using the Wilcoxon rank-sum test. For multiple comparisons, adjusted p -values were calculated using the Benjamini-Hochberg method to control the false discovery rate (FDR), with significance thresholds set at adjusted $p < 0.05$.

Results

Identification of cellular composition in B-ALL and healthy PBMC samples

To construct the landscape of cell type-specific open chromatin features and gene expression profile at the single cell level, we performed sc-RNA-seq and scATAC-seq analysis on PBMCs samples from six high risk pediatric B-ALL patients with and three healthy controls (CK1, CK2, CK3) to explore their cellular composition (Supplementary Figure S1). PBMCs were chosen due to their accessibility and reduced invasiveness, especially in pediatric patients undergoing frequent procedures. These high-risk patients had all undergone the standard CAML

(Cyclophosphamide, Cytosine arabinoside, 6-Mercaptopurine, and Pegaspargase) chemotherapy regimen. Specifically, we selected three patients who achieved complete remission (CR1, CR2, CR3) and three who experienced non-remission (NCR1, NCR2, NCR3) following two courses of intensified treatment.

After an initial quality control assessment and doublet removal, we obtained a total of 76458 cells, with 38773 cells from healthy donors and 37685 cells from B-ALL patients, median number of estimated cells were 6861, and the ATAC Median high-quality fragments per cell were 5071. After performing normalization and log1p transformation, highly-variable gene selection, dimensionality reduction, batch correction, and Leiden clustering, cells originating from samples were separately annotated into distinct broad cell types and visualized via Uniform Manifold Approximation and Projection (UMAP). We identified 9 main clusters in parallel according to the gene profile. These clusters were annotated as CD8+ NKT-like cells, classical monocytes, HSC/MPP cells, memory CD4+ T cells, naïve B cells, naïve CD4+ T cells, naïve CD8+ T cells, pre-B cells, pro-B cells, and progenitor cells based on the expression of known markers for each cell type (Figures 1A, B) (12–15).

We then analyzed the relative proportions of cell clusters (Figures 1C, D). The cellular composition of the ALL remission group (CR group) is similar to that of the normal group (CK), with naïve CD4 T cells and CD8+ NKT-like cells being predominant. However, the cellular composition of the non-remission group (NCR group) is significantly different from the CR group, characterized by the highest proportion of Pro-B cells ($p < 0.0001$), a marked increase in HSC/MPP cells ($p = 0.0001$), and a significant decrease in naïve CD4 T cells ($p = 0.0002$) and naïve B cells (Figure 1E).

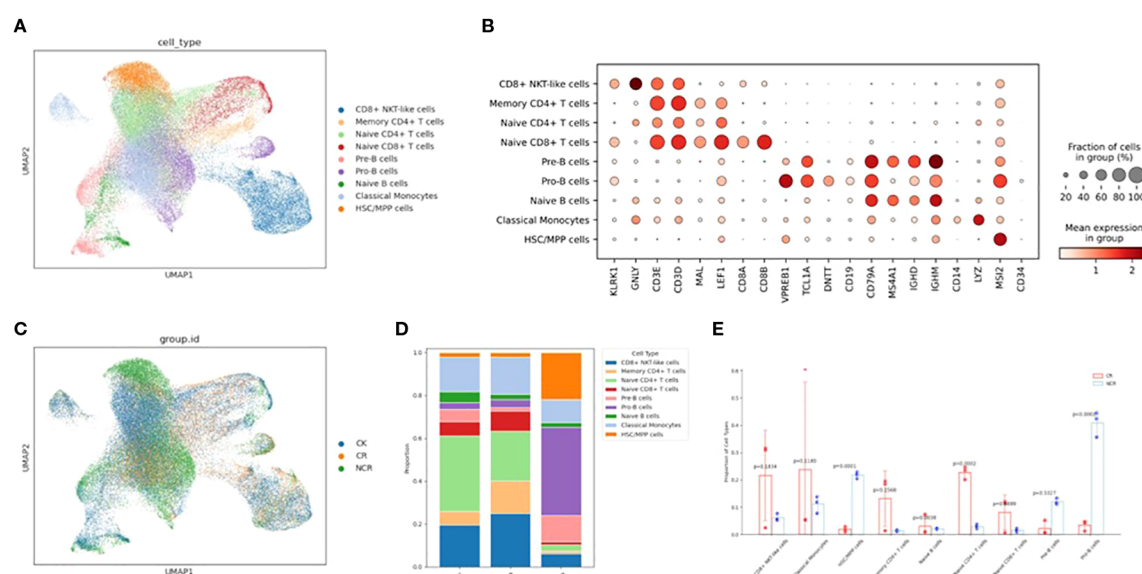


FIGURE 1

Single-cell transcriptomic analysis in ALL patient PBMC cells and healthy donor PBMC cells. (A) UMAP plots of nine color-coded cell clusters of all the samples. (B) Dot plots show the signature gene expression levels across the 9 distinct cell types. (C) UMAP plots of the CK, CR and NCR groups. (D) Relative proportion of cell clusters in each group across the 9 cellular clusters. (E) Histogram of the difference in cell proportion between the CR group and the NCR group. PBMC, Peripheral blood mononuclear cell; UMAP, uniform manifold approximation and projection for dimension reduction; CK, control group; CR, remission group; NCR, non-remission group.

Chromatin accessibility different in each cell type

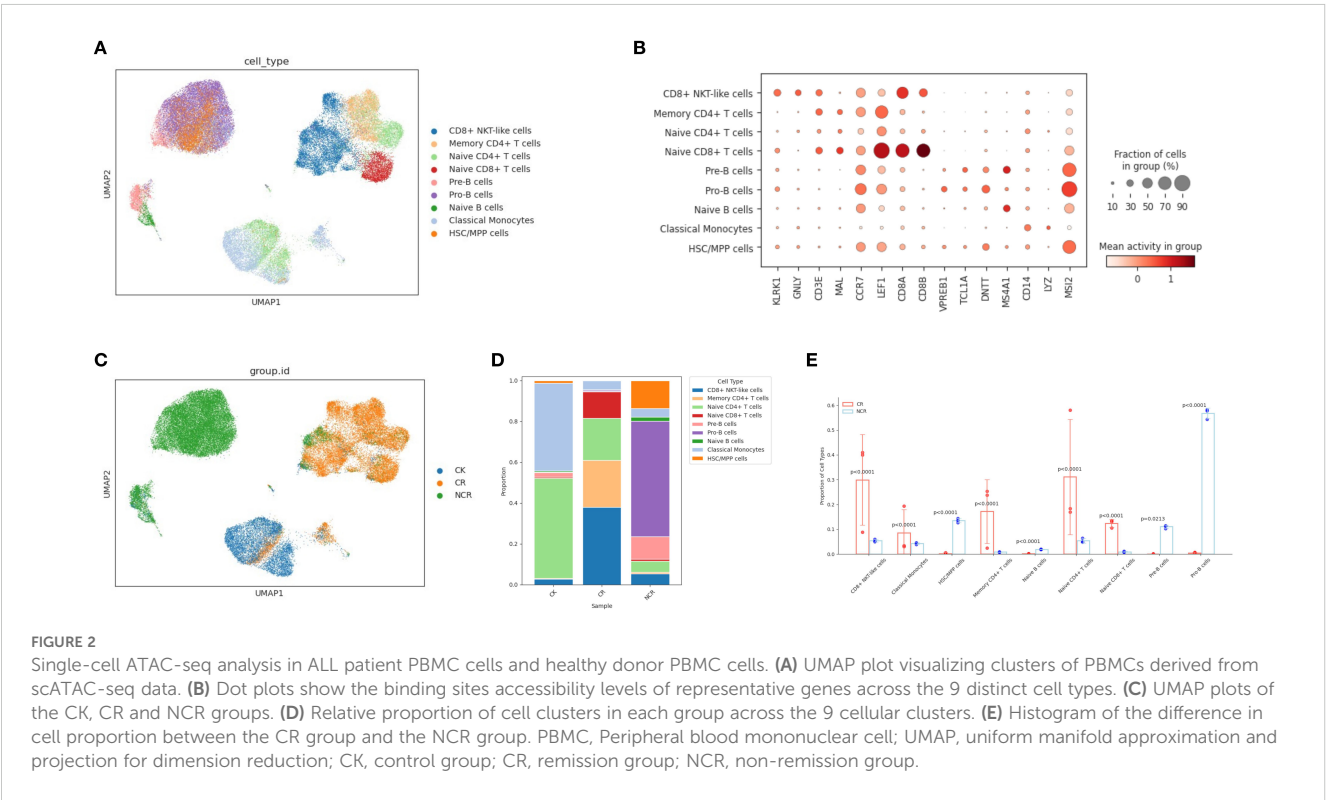
The chromatin accessibility profile of individual cells is captured by scATAC-seq. Cell-type-specific chromatin accessibility profiles are relatively unknown. So, we used Seurat's label transfer method to predict scATAC-seq cell types based on annotated scRNA-seq data. The detailed analysis methods were executed as previously described. Comparison between scATAC-seq cell-type predictions obtained by label transfer and curated annotations of unsupervised clusters indicates that all major cell types were present in both datasets and that scATAC-seq is comparable to scRNA-seq in the detection and assignment of cell identities (Figures 2A, B). Subsequently, we investigated the differences in chromatin accessibility among CK, CR and NCR groups (Figure 2C). In the CK group, naïve CD4 T cells were predominant, followed by classical monocytes and HSC/MPP cells (Figure 2D). In the CR group, CD8+ NKT-like cells were dominant, which is in line with the findings from scRNA-seq. Relative to the CK group, the CR group exhibited a notable increase in the proportion of memory CD4+ T cells and CD8+ NKT-like cells, alongside a significant reduction in the proportion of naïve CD4 T cells. The cellular composition of the NCR group diverged markedly from the other groups, with Pro-B cells predominating and accounting for more than half of the cells. In contrast to the CR group, the NCR group showcased a significant upsurge in the proportion of HSC/MPP cells ($p<0.0001$), naïve B cells($p<0.0001$) and Pro-B cells ($p<0.0001$), coupled with a substantial decrease in the proportion of CD8+ NKT-like cells ($p<0.0001$), classical monocytes ($p<0.0001$), memory CD4+ T cells ($p<0.0001$), naïve CD4 T cells ($p<0.0001$), and naïve CD8 T cells ($p<0.0001$) (Figure 2E).

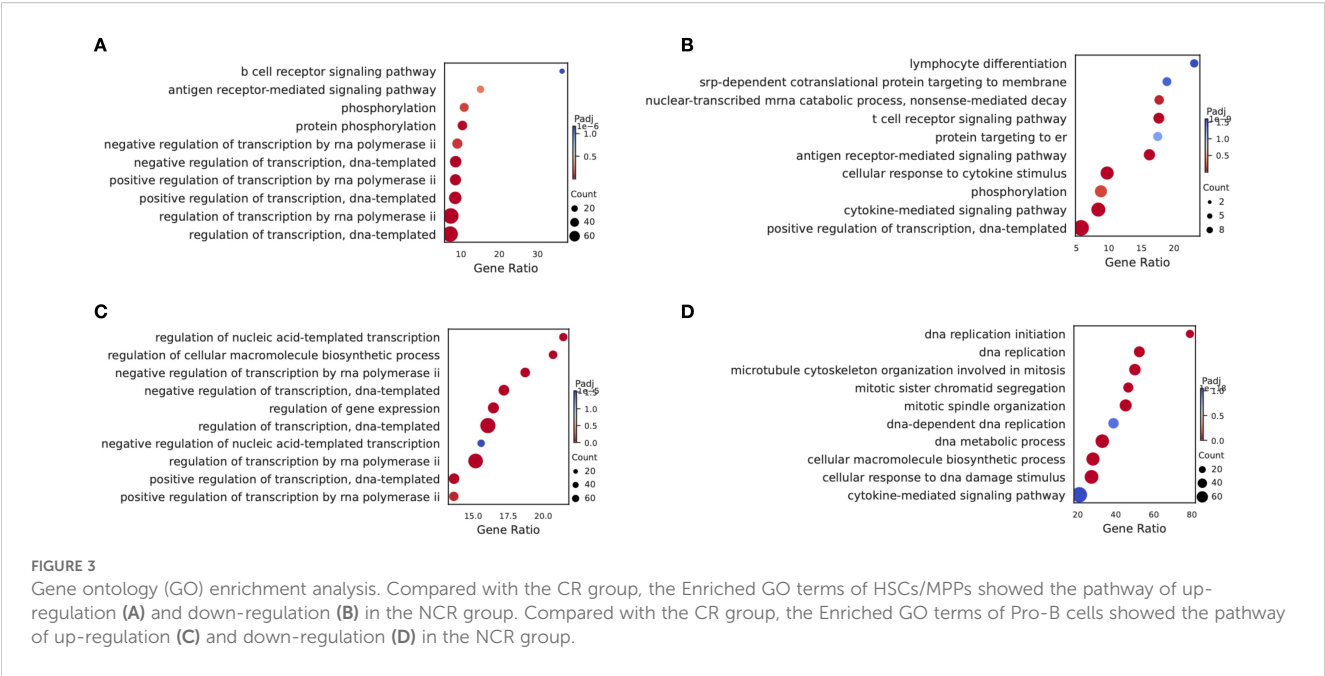
Gene ontology enrichment in HSC/MPP cells and Pro-B cells

Given the most pronounced differences in the relative proportions of HSC/MPP cells and Pro-B cells between the CR group and the NCR group, and the concordance of findings from scRNA-seq and scATAC-seq multi-omics analyses, we concentrated our investigation on these two specific subgroups. We analyzed the differentially expressed genes (DEGs) in HSC/MPP cells and Pro-B cells between the NCR and CR groups. The gene ontology (GO) analysis of DEGs in HSC/MPP revealed that the B cell receptor signaling pathway and antigen receptor-mediated signaling pathway were upregulated in NCR (Figure 3A), while lymphocyte differentiation and T cell receptor signaling pathway were downregulated in NCR (Figure 3B). GO analysis of DEGs in Pro-B revealed that pathways governing the regulation of transcription, gene expression, and biosynthetic processes were enhanced in the NCR group (Figure 3C), contrasted with a suppression of signaling pathways involved in DNA replication within the NCR group (Figure 3D).

Spectral clustering of specific subtype cells (HSC/MPP and Pro-B)

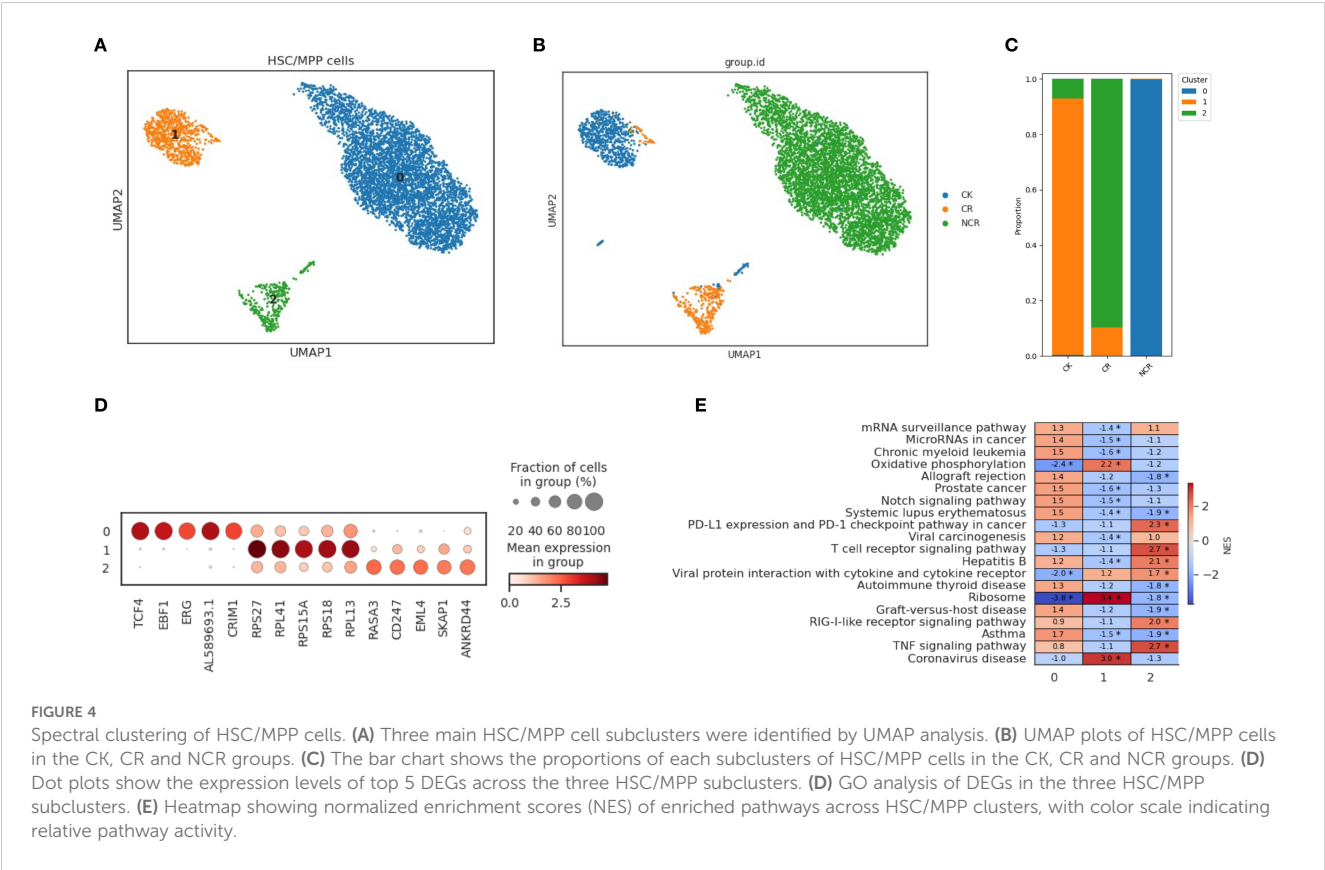
In order to further characterize the intrinsic structure and potential functional subtypes, we performed spectral clustering on HSCs cells using UMAP. As a result, we identified three stable subclusters in HSCs cells (Figure 4A). The CK group was mainly composed of subcluster 1, the CR group was predominantly made





up of cells from subcluster 2, and the Pro-B cells of the NCR group were found in subcluster 0 (Figure 4B). Subcluster 0 included cells expressing TCF4, EBF1, ERG, AL589693.1, and CRIM1 genes (Figure 4C). Subcluster 1 included cells expressing genes like RPS27, RPL41, RPS15A, RPS18, and RPL13. Subcluster 2 included cells expressing genes such as RASA3, CD247, EML4, SKAP1, and ANKRD44. The TNF signaling pathway and T cell receptor signaling pathway were enriched in subcluster 2, while allograft rejection pathway and Notch signaling pathway were enriched in subcluster 0 (Figure 4D).

For the spectral clustering of Pro-B cells, five subclusters were identified (Figure 5A). The CK group was predominantly composed of



subcluster 1, along with some cells from subcluster 4 (Figure 5B). The CR group was mainly made up of cells from subcluster 2, while the Pro-B cells from the NCR group belonged to subclusters 0 and 3. Each subcluster expressed its unique set of characteristic genes. The relative expression levels of cluster-specific markers in the five cellular subclusters of Pro-B have been presented in Figure 5C. RPS29, B2M, RPL41, and RPS21 were expressed across subclusters 0, 1, 2, and 3, but exhibited the highest expression in subcluster 1. In addition to RPS29, B2M, RPL41, and RPS21, subcluster 0 also included cells expressing the genes NEIL1, AC007384.1, and CRIM1. Subcluster 2 further included cells expressing genes such as HMGB2, PTPRC, SKAP1, and TUBA1B. Subcluster 4 was relatively distinct from the other subclusters, primarily consisting of cells expressing genes like GP1BB, NRGN, PF4, OAZ1, and PPBP. GSEA enrichment analysis showed that the B cell receptor signaling pathway was enriched in subcluster 0, while the cell cycle, DNA replication, and mismatch repair pathways were enriched in subclusters 3 and 2 (Figure 5D). Gene ontology enrichment analysis further revealed distinct functional programs among the HSC/MPP and Pro-B clusters (Figures 4E, 5E).

Copy number variation analysis to distinguish malignant cells

Copy number variation (CNV) analysis has been widely used in scRNA-seq to investigate disease evolution and development. To identify the cellular origin of malignant cells, we utilized the inferCNV algorithm to analyze the chromosomal CNV levels between different cell types (Figure 6A). As expected, the CNV levels in the NCR group were significantly higher than those in

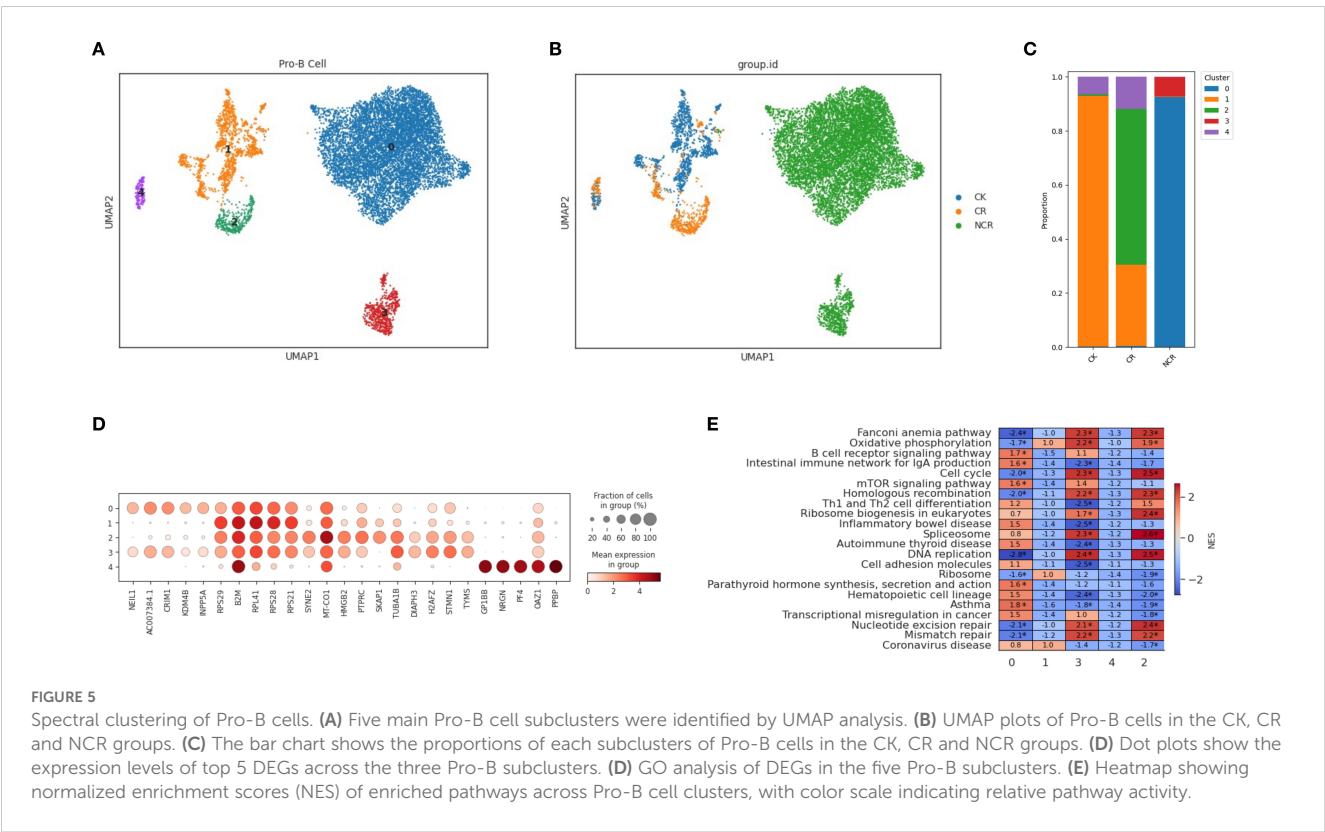
the CR group (Figure 6A). Within the NCR group, HSS/MPP cells and pro-B cells exhibited higher CNV levels than other cell types.

We further analyzed the CNV levels in the three subclusters of HSC/MPP cells and the five subclusters of Pro-B cells. The results indicated that subcluster 0 of HSC/MPP cells had the highest CNV levels (Figure 6B), while in Pro-B cells, subgroup 0 exhibited the highest CNV levels, followed by subgroup 3 (Figure 6C).

Correlation between cell-type-specific chromatin accessible degree and gene expression level in HSC/MPP cells and Pro-B cells

Compared to other cell types, HSC/MPP cells in scATAC-seq had a large number of differential peaks of chromatin accessibility (Figure 7A). With transcription factor motif analysis, we found that the binding sites of FOSL2, FOS, FOSL1, BATF, JUN, JUNB had higher accessibility in the HSC/MPP than in other cell types (Figures 7B, C). However, only FOSL2 had a significant level of expression in the HSC/MPP group, while the expression levels of other genes were relatively low. When comparing HSC/MPP cells between NCR and CR groups, no differential peaks in binding site accessibility were observed (Figure 7B). Furthermore, expression patterns of representative immune-related genes across cell types were also assessed (Figure 7D).

The analysis of the Pro-B subgroup in scATAC-seq showed that the binding sites of EBF3, EBF1, MAZ, ZNF148, KLF15, NHLH1, and KLF5 had higher accessibility in Pro-B cells than in other cell types (Figures 8A-C). We further analyzed the expression differences of these genes in scRNA-seq and found that EBF1 and



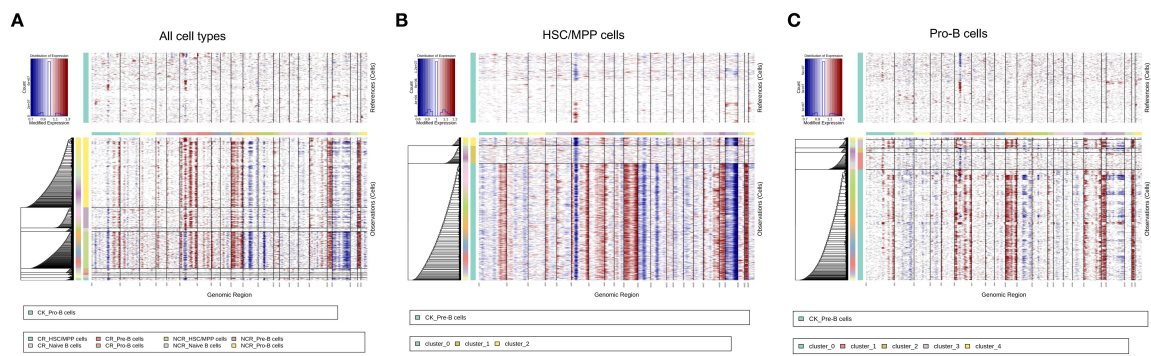


FIGURE 6
Copy number variation (CNV) analysis of all cell types (A), HSC/MPP cells (B) and Pro-B cells (C).

ZNF148 were highly expressed in Pro-B cells (Figure 8D). When comparing Pro-B cells between NCR and CR groups, the differential peaks in NCR were mainly downregulated, which means they were predominantly enriched in the CR group (Figure 8E). We found that the accessibility of KLF15, SP1, SP3, NRF1, ZBTB14, EGR1, EGR3 binding sites in Pro-B was higher in the CR group than in the NCR groups (Figure 8F). Among them, SP1, SP3, NRF1 also showed higher expression levels in scRNA-seq in the CR group compared to the NCR group, which is consistent with scATAC-seq (Figure 8G).

Discussion

The standard intensive therapy for high-risk B-ALL patients currently involves two cycles of the CAML protocol. Despite this, some patients exhibit resistance to early intensified treatment, resulting in therapeutic failure. However, our understanding of the molecular mechanisms underlying chemotherapy sensitivity and the identification of patients who are likely to benefit from such treatments remains limited. Literature has indicated that infants with CD10-positive blasts, in the absence of 11q23/MLL

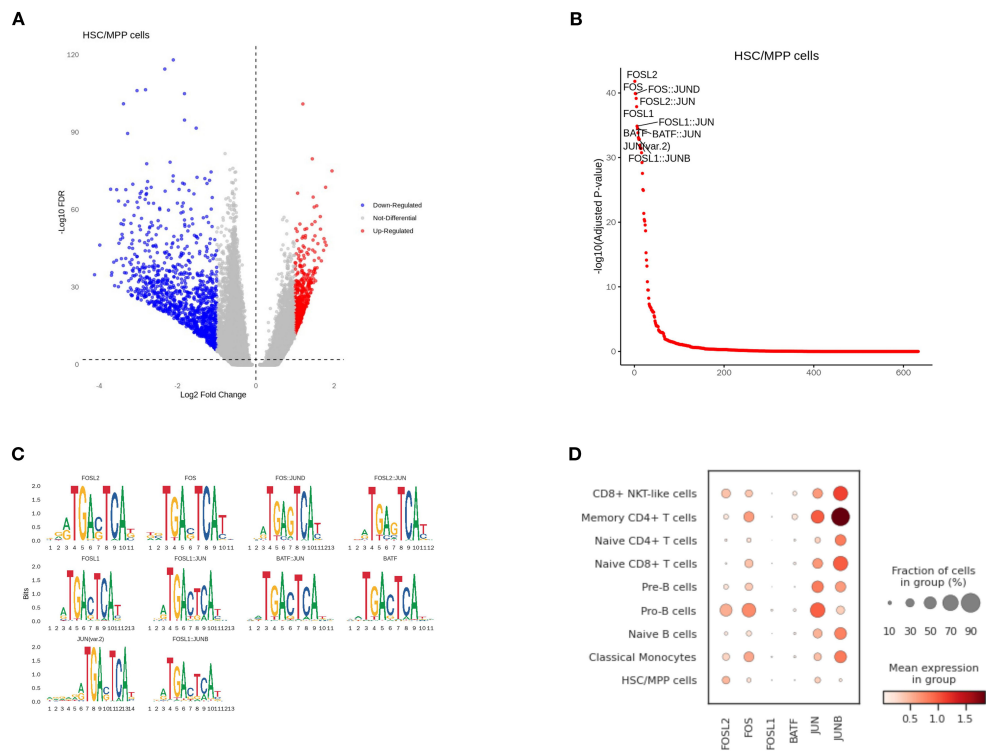
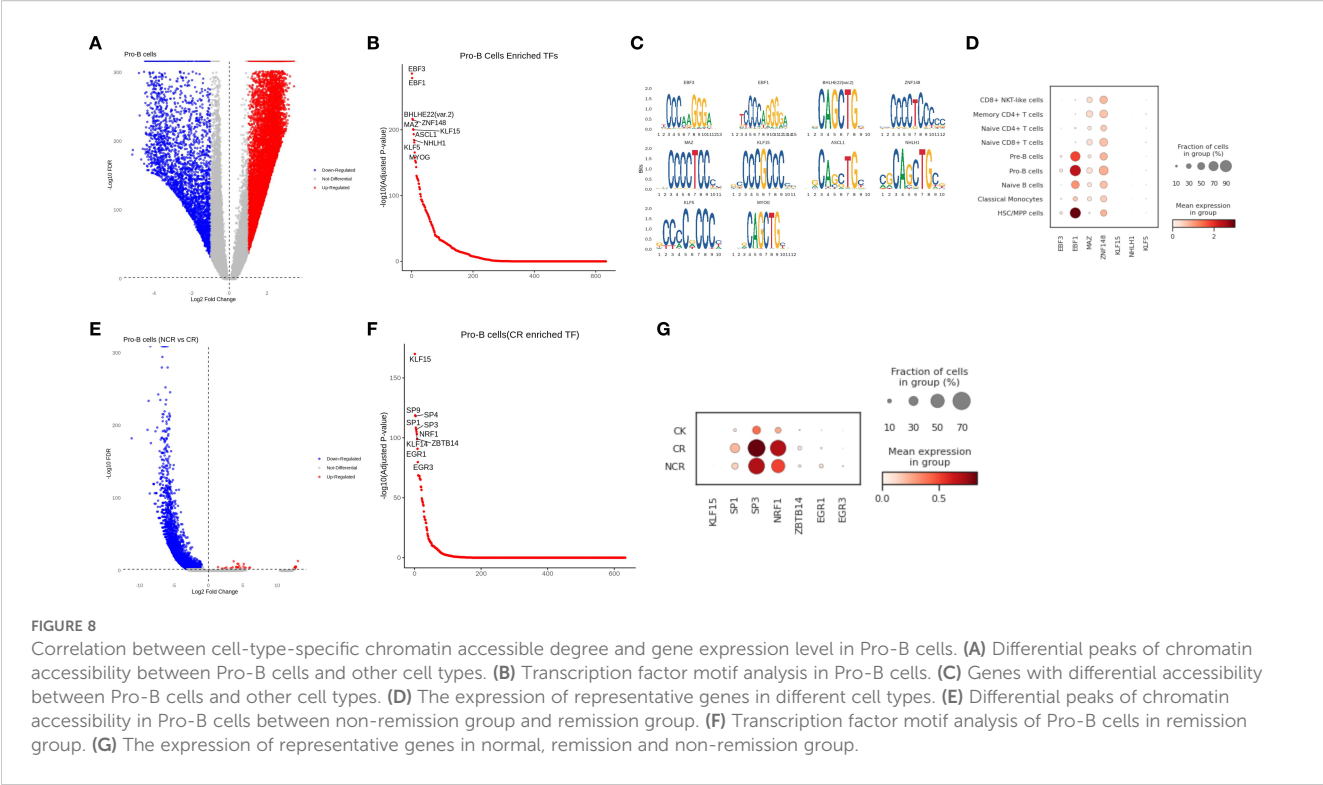


FIGURE 7
Correlation between cell-type-specific chromatin accessible degree and gene expression level in HSC/MPP cells. (A) Differential peaks of chromatin accessibility between HSC/MPP cells and other cell types. (B) Transcription factor motif analysis. (C) Genes with differential accessibility between HSC/MPP cells and other cell types. (D) The expression of representative genes in different cell types.



rearrangements, tend to have improved outcomes and respond better to intensified therapy (16). Infants with B-ALL harboring germline MLL genes show superior responses to chemotherapy compared to those with rearranged MLL genes, achieving a 5-year event-free survival rate of 95.5% with manageable treatment toxicity (17). Traditional diagnostic methods, which assess the average mutations and expressions across a population of cells, can obscure the signals of cellular heterogeneity, particularly overlooking the signals from rare cell types. To address this, we employed single-cell RNA sequencing (scRNA-seq) combined with single-cell ATAC sequencing (scATAC-seq) to analyze differential gene expressions in specific cellular subsets within peripheral blood mononuclear cells (PBMCs) from B-ALL patients who experienced remission or relapse following intensive therapy, as well as from healthy controls. By clustering these expression profiles, we inferred genes associated with relapse within particular cell types.

The scRNA-seq results elucidate that the cellular composition of the remission group is similar to that of the normal group, which also demonstrates the effectiveness of the treatment at the cellular level. However, the cellular composition of the non-remission group is significantly different from the other groups, with the most pronounced changes observed in the HSC/MPP and Pro-B cell populations. In both the remission and normal groups, the proportions of these two types of cells are relatively low, whereas in the non-remission group, the proportions of these cells have increased dramatically, becoming the predominant types. This may be due to the failure of intensified therapy, where normal hematopoiesis in leukemia continues to be suppressed, leading to anemia, thrombocytopenia, and neutropenia. The body compensates by increasing the number of HSC/MPP cells, which

possess multilineage differentiation and self-renewal capabilities (18). Simultaneously, there is an arrest of B cell development at the Pro-B stage, impeding cellular maturation and precipitating the accumulation of these immature cell forms.

Compared to the scRNA-seq analysis, the scATAC-seq analysis reveals even greater differences in cellular proportions among the normal, remission, and non-remission groups, possibly because the epigenome is more prone to change than the transcriptome. In the non-remission group, the cell proportions based on chromatin accessibility also show that HSC/MPP and Pro-B cells are the most predominant, which is consistent with the scRNA-seq findings. CNV analysis also demonstrates that the CNV levels in HSC/MPP and Pro-B cells are higher in the NCR group than in other cell types, indicating that malignant cells in PBMCs are primarily concentrated in these two cell populations. Therefore, we focus our analysis on these two subgroups. CNV differences observed between CR and NCR groups may reflect individual genomic background rather than therapy-specific changes. We interpret CNV variation as a correlate of clonal genomic instability, but not direct proof of malignancy.

Further stratification of HSC/MPP identified three distinct subclusters, with the normal, remission, and non-remission groups each corresponding to different subclusters. The HSC/MPP cells from the NCR group are predominantly found in subcluster 0, which shows high expression of TCF4, EBF1, ERG, AL589693.1, and CRIM1, along with enrichment of the allograft rejection pathway and the Notch signaling pathway. These associations suggest that this subpopulation may be involved in resistance to intensified therapy. However, these findings are observational and require functional validation to determine any causal role. TCF4 is a transcription factor belonging to the basic helix-loop-helix (bHLH) family, which is involved in

lymphoid cell differentiation (19). In B-ALL, TCF4 may be linked in the regulation of B cell proliferation and differentiation, as well as the self-renewal of leukemic stem cells. EBF1 is a key transcription factor in B cell development, essential for the proliferation and maturation of early B cells (20). Its aberrant expression may be associated with the proliferation and survival of leukemia cells (21). ERG is a member of the ETS family, involved in multiple cellular processes, including cell proliferation, differentiation, and migration (22). In B-ALL, ERG may be involved in the migration and invasion of leukemic cells, as well as resistance to therapy. CRIM1 is a transmembrane protein that modulates the bone morphogenetic protein (BMP) signaling pathway, which plays a role in various cellular processes, including cell differentiation and proliferation (23, 24). It may be involved in regulating the differentiation of leukemia cells and their response to treatment. Current research on the biology of ALL has identified several recurrent ALLs with targetable pathways, including the Notch signaling pathway (25). The upregulation of the Notch signaling pathway in subcluster 0 of HSC/MPP cells suggests a possible role in relapse biology, but functional studies are necessary to determine whether this pathway is directly involved in therapy resistance.

In addition to its role in hematopoietic development and leukemogenesis, Notch signaling has been implicated in immune modulation and resistance to immunotherapy (26). Studies in T-cell acute lymphoblastic leukemia (T-ALL) and other malignancies have shown that aberrant Notch activation may suppress T cell activation, promote an immunosuppressive tumor microenvironment, or alter immune checkpoint expression (27, 28). These mechanisms may contribute to poor responses to immunotherapies such as CAR-T cells or immune checkpoint inhibitors. Although our study did not directly examine immunotherapy response, the enrichment of Notch signaling in the NCR-associated HSC/MPP subpopulation raises the possibility that this pathway may influence immune escape and warrants further investigation in the context of relapsed B-ALL.

Pro-B cells were stratified into 5 subclusters through spectral clustering, with the NCR group corresponding to subclusters 0 and 3, among which the majority of the pro-B cells in NCR were in subcluster 0. This subcluster is characterized by high expression of RPS29, B2M, RPL41, RPS21, NEIL1, AC007384.1, and CRIM1, as well as enrichment of the B cell receptor signaling pathway. Previous studies have shown that treatment with glucocorticoids leads to enrichment of the B cell receptor signaling pathway in B-ALL cells, and that B cell development is crucial for glucocorticoid resistance in B-ALL cells (29). Therefore, the upregulation of the B cell receptor signaling pathway in subcluster 0 of Pro-B cells may also contribute to the patients' resistance to intensified chemotherapy.

Due to the clinical constraints and ethical consideration, the sample size in our study is limited and not longitudinally paired. As a result, caution should be exercised in extrapolating our findings, and larger, paired cohort studies will be necessary to validate the identified subpopulations and mechanisms. Future validations are needed. First, the potential pathogenic genes identified have not yet undergone functional validation and require further functional experiments for confirmation in the future. Secondly, although these genes may all be related to the relapse of intensified therapy in B-ALL, our sample size is small, and a larger sample is needed to verify the universality and

specificity of these relapse genes in specific cell types. Thirdly, our study is limited by the absence of full molecular subtype classification for each B-ALL patient, which poses a challenge given the extensive heterogeneity of the disease. Over 20 subtypes have been recognized, many of which are associated with distinct prognoses and treatment responses. While our inferCNV-based approach provides insights into global genomic instability, it may not detect subtype defined by gene fusion or single-nucleotide variants. Fourthly, all single-cell data in this study were derived from peripheral blood mononuclear cells (PBMCs), rather than from the bone marrow microenvironment, which is the primary site of leukemic development in B-ALL. While PBMCs provide a minimally invasive snapshot of systemic immune and leukemic cell states, they may not fully capture the complexity of interactions occurring in the bone marrow niche. Future studies involving bone marrow samples will be essential to validate and extend these findings.

Although in the absence of *in vitro* validation, our conclusions are supported by consistency with previous studies. The relapse-associated HSC/MPP subcluster enriched for *TCF4*, *EBF1*, and *ERG* expression, together with Notch signaling, aligns with reports implicating these factors in B-cell development, leukemic stemness, and therapy resistance (20–22). Similarly, the Pro-B subcluster characterized by *RPS29*, *B2M*, and enrichment of the BCR pathway corresponds with published evidence linking BCR signaling to glucocorticoid resistance in B-ALL (26, 28). Prior studies have also emphasized the role of CRIM1 and BMP signaling in leukemic differentiation and drug response (23, 24, 29). Taken together, these consistencies with the literature reinforce the robustness of our multi-omics findings, while also highlighting Notch and BCR signaling as promising axes for future mechanistic and therapeutic exploration.

In conclusion, our study has identified genes within specific cellular subsets that are associated with the failure of intensified therapy in B-ALL, thereby enriching the existing knowledge base for the investigation of drug resistance mechanisms in B-ALL. Additionally, we posit that the disruption of signaling pathways due to aberrant expression of these genes may be a significant contributor to therapeutic resistance. Consequently, the development of targeted inhibitors against these genes and their associated pathways could potentially overcome drug resistance in B-ALL, thereby enhancing therapeutic outcomes and providing increased benefits to patients.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://www.ncbi.nlm.nih.gov/> accession number PRJNA1333439 and PRJNA1336723.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Sichuan Provincial Woman's and Children's Hospital under a project license (No. 2024-D.-22). The studies were

conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

LL: Methodology, Data curation, Investigation, Conceptualization, Writing – review & editing, Software, Writing – original draft, Formal Analysis. XM: Methodology, Data curation, Conceptualization, Writing – original draft, Formal Analysis, Investigation. CY: Investigation, Writing – original draft, Data curation, Methodology, Formal Analysis. NL: Methodology, Data curation, Writing – original draft, Formal Analysis. YZ: Data curation, Methodology, Writing – original draft, Investigation. LZ: Investigation, Methodology, Writing – original draft, Data curation. NJ: Methodology, Investigation, Formal Analysis, Writing – original draft. YH: Methodology, Investigation, Writing – original draft, Resources. SY: Investigation, Methodology, Supervision, Software, Validation, Writing – original draft. HX: Validation, Methodology, Supervision, Investigation, Project administration, Writing – original draft, Resources, Visualization, Writing – review & editing, Software. XT: Writing – original draft, Software, Visualization, Formal Analysis, Funding acquisition, Resources, Conceptualization, Project administration, Methodology, Data curation, Writing – review & editing, Validation, Supervision, Investigation.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the Joint Funds of Southwest Medical University and Luzhou Government (2021LZXNYD-Z05), the Sichuan Science and Technology Program, China (2022YFS0615), Southwest Medical University (2022YFS0615-A2), and the National Natural Science Foundation of China (82301658).

Acknowledgments

We thank Chongqing Knorigene Technologies for bioinformatics analysis support and Kunming KingMed Center for Clinical Laboratory, Kunming, China, for sequencing support.

References

- Hunger SP, Mullighan CG. Acute lymphoblastic leukemia in children. *New Engl J Med*. (2015) 373:1541–52. doi: 10.1056/NEJMra1400972
- Inaba H, Mullighan CG. Pediatric acute lymphoblastic leukemia. *Haematologica*. (2020) 105:2524–39. doi: 10.3324/haematol.2020.247031
- Ranzoni AM, Tangherloni A, Berest I, Riva SG, Myers B, Strzelecka PM, et al. Integrative single-cell RNA-seq and ATAC-seq analysis of human developmental hematopoiesis. *Cell Stem Cell*. (2021) 28:472–487.e7. doi: 10.1016/j.stem.2020.11.015
- Wang X, Chen Y, Li Z, Huang B, Xu L, Lai J, et al. Single-cell RNA-seq of T cells in B-ALL patients reveals an exhausted subset with remarkable heterogeneity. *Adv Sci*. (2021) 8:e2101447. doi: 10.1002/advs.202101447
- Granja JM, Klemm S, McGinnis LM, Kathiria AS, Mezger A, Corces MR, et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol*. (2019) 37:1458–65. doi: 10.1038/s41587-019-0332-7
- Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. (2018) 19:15. doi: 10.1186/s13059-017-1382-0

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fimmu.2025.1645546/full#supplementary-material>

SUPPLEMENTARY TABLE S1

Clinical and demographic information of all patient samples used in this study.

SUPPLEMENTARY FIGURE S1

Schematic overview of the study workflow. Peripheral blood samples were collected from three CR (complete remission), three NCR (non-complete remission), and three healthy control children. After mixing with PBS/EDTA, samples underwent density gradient centrifugation to separate plasma, red blood cells, and PBMCs. PBMCs were subjected to 10x Genomics-based 3' single-cell RNA sequencing and ATAC sequencing. Downstream analyses focused on identifying and comparing HSC/MPP and Pro-B cell populations across the three groups to explore immune remodeling and recurrence mechanisms in high-risk B-ALL.

7. Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. *Nat Commun.* (2022) 13:1246. doi: 10.1038/s41467-022-28803-w
8. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods.* (2021) 18:1333–41. doi: 10.1038/s41592-021-01282-5
9. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. *Cell.* (2019) 177:1888–1902.e21. doi: 10.1016/j.cell.2019.05.031
10. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf.* (2013) 14:7. doi: 10.1186/1471-2105-14-7
11. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* (2017) 14:1083–6. doi: 10.1038/nmeth.4463
12. Zhang C, Lei L, Yang X, Ma K, Zheng H, Su Y, et al. Single-cell sequencing reveals antitumor characteristics of intratumoral immune cells in old mice. *J Immunother Cancer.* (2021) 9:e002809. doi: 10.1136/jitc-2021-002809
13. Stewart A, Ng JC-F, Wallis G, Tsioligka V, Fraternali F, Dunn-Walters DK. Single-cell transcriptomic analyses define distinct peripheral B cell subsets and discrete development pathways. *Front Immunol.* (2021) 12:602539. doi: 10.3389/fimmu.2021.602539
14. Andreatta M, Corria-Osorio J, Müller S, Cubas R, Coukos G, Carmona SJ. Interpretation of T cell states from single-cell transcriptomics data using reference atlases. *Nat Commun.* (2021) 12:2965. doi: 10.1038/s41467-021-23324-4
15. Zhang Y, Wang S, Zhang J, Liu C, Li X, Guo W, et al. Elucidating minimal residual disease of paediatric B-cell acute lymphoblastic leukaemia by single-cell analysis. *Nat Cell Biol.* (2022) 24:242–52. doi: 10.1038/s41556-021-00814-7
16. Nishimura S, Kobayashi M, Ueda K, Ishii E, Okamura J, Kawaha K, et al. Treatment of infant acute lymphoblastic leukemia in Japan. Childhood Leukemia Study Group of the Ministry of Health and Welfare (Kouseisho). *Int J Hematol.* (1999) 69:244–52.
17. Nagayama J. Infants with acute lymphoblastic leukemia and a germline MLL gene are highly curable with use of chemotherapy alone: results from the Japan Infant Leukemia Study Group. *Blood.* (2006) 107:4663–5. doi: 10.1182/blood-2005-11-4728
18. Yamamoto R, Morita Y, Oechara J, Hamanaka S, Onodera M, Rudolph KL, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell.* (2013) 154:1112–26. doi: 10.1016/j.cell.2013.08.007
19. in 't Hout FEM, van Duren J, Monteferrario D, Brinkhuis E, Mariani N, Westers TM, et al. TCF4 promotes erythroid development. *Exp Hematol.* (2019) 69:17–21.e1. doi: 10.1016/j.exphem.2018.10.002
20. Sigvardsson M. Early B-cell factor 1: an archetype for a lineage-restricted transcription factor linking development to disease. (2024) 1459:143–56. doi: 10.1007/978-3-031-62731-6_7
21. Li L, Zhang D, Cao X. EBF1, PAX5, and MYC: regulation on B cell development and association with hematologic neoplasms. *Front Immunol.* (2024) 15:1320689. doi: 10.3389/fimmu.2024.1320689
22. Mounir Z, Lin F, Lin VG, Korn JM, Yu Y, Valdez R, et al. TMPRSS2:ERG blocks neuroendocrine and luminal cell differentiation to maintain prostate cancer proliferation. *Oncogene.* (2015) 34:3815–25. doi: 10.1038/onc.2014.308
23. Iyer S, Chou FY, Wang R, Chiu HS, Raju VKS, Little MH, et al. Crim1 has cell-autonomous and paracrine roles during embryonic heart development. *Sci Rep.* (2016) 6:19832. doi: 10.1038/srep19832
24. Nakashima Y, Morimoto M, Toda K, Shinya T, Sato K, Takahashi S. Inhibition of the proliferation and acceleration of migration of vascular endothelial cells by increased cysteine-rich motor neuron 1. *Biochem Biophys Res Commun.* (2015) 462:215–20. doi: 10.1016/j.bbrc.2015.04.118
25. Raetz EA, Teachey DT. T-cell acute lymphoblastic leukemia. *Hematology.* (2016) 2016:580–8. doi: 10.1182/asheducation-2016.1.580
26. Zhou B, Lin W, Long Y, Yang Y, Zhang H, Wu K, et al. Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduct Target Ther.* (2022) 7:95. doi: 10.1038/s41392-022-00934-y
27. Lu J, Xue X, Wang H, Hao Y, Yang Q. Notch1 activation and inhibition in T-cell acute lymphoblastic leukemia subtypes. *Exp Hematol.* (2025) 148:104771. doi: 10.1016/j.exphem.2025.104771
28. Aster JC, Blacklow SC, Pear WS. Notch signalling in T-cell lymphoblastic leukaemia/lymphoma and other haematological Malignancies. *J Pathol.* (2011) 223:263–74. doi: 10.1002/path.2789
29. Sarno J, Domizi P, Liu Y, Merchant M, Pedersen CB, Jedoui D, et al. Dasatinib overcomes glucocorticoid resistance in B-cell acute lymphoblastic leukemia. *Nat Commun.* (2023) 14:2935. doi: 10.1038/s41467-023-38456-y