frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 12 September 2024
DOI 10.3389/fimmu.2024.1408109

Check for updates

OPEN ACCESS

EDITED BY
Erika Adriana Eksioglu,
Moffitt Cancer Center, United States

REVIEWED BY
Yun Tan,
Shanghai Jiao Tong University, China
Joshua Zeidner,
University of North Carolina at Chapel Hill, United States
Lucia Silla,
Federal University of Rio Grande do Sul, Brazil

*CORRESPONDENCE
Junmin Li
drlijunmin@126.com
Zhen Jin
jane-king@163.com
Han Liu
liuhan68@sjtu.edu.cn
Huijin Zhao
celine_zh@outlook.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 March 2024
ACCEPTED 26 August 2024
PUBLISHED 12 September 2024

CITATION
Li Z, Jin P, Xiang R, Li X, Shen J, He M, Liu X, Zhu H, Wu S, Dong F, Zhao H, Liu H, Jin Z and Li J (2024) A CD8+ T cell related immune score predicts survival and refines the risk assessment in acute myeloid leukemia. *Front. Immunol.* 15:1408109. doi: 10.3389/fimmu.2024.1408109

COPYRIGHT
© 2024 Li, Jin, Xiang, Li, Shen, He, Liu, Zhu, Wu, Dong, Zhao, Liu, Jin and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A CD8$^+$ T cell related immune score predicts survival and refines the risk assessment in acute myeloid leukemia

Zeyi Li[1†], Peng Jin[1†], Rufang Xiang[2†], Xiaoyang Li[1†], Jie Shen[1], Mengke He[1], Xiaxin Liu[1], Hongming Zhu[1], Shishuang Wu[1], Fangyi Dong[1], Huijin Zhao[1*], Han Liu[1*], Zhen Jin[1*] and Junmin Li[1,3*]

[1]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Department of General Practice, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]Wuxi Branch of Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

Although advancements in genomic and epigenetic research have deepened our understanding of acute myeloid leukemia (AML), only one-third of patients can achieve durable remission. Growing evidence suggests that the immune microenvironment in bone marrow influences prognosis and survival in AML. There is a specific association between CD8$^+$ T cells and the prognosis of AML patients. To develop a CD8$^+$ T cell-related immune risk score for AML, we first evaluated the accuracy of CIBERSORTx in predicting the abundance of CD8$^+$ T cells in bulk RNA-seq and found it significantly correlated with observed single-cell RNA sequencing data and the proportions of CD8$^+$ T cells derived from flow cytometry. Next, we constructed the CTCG15, a 15-gene prognostic signature, using univariate and LASSO regression on the differentially expressed genes between CD8$^+$ T^{High} and CD8$^+$ T^{Low} groups. The CTCG15 was further validated across six datasets in different platforms. The CTCG15 has been shown to be independent of established prognostic markers, and can distill transcriptomic consequences of several genetic abnormalities closely related to prognosis in AML patients. Finally, integrating this model into the 2022 European LeukemiaNet contributed to a higher predictive power for prognosis prediction. Collectively, our study demonstrates that CD8$^+$ T cell-related signature could improve the comprehensive risk stratification and prognosis prediction in AML.

KEYWORDS

acute myeloid leukemia, CD8$^+$ T cell, the European LeukemiaNet, prognosis, CIBERSORTx

Introduction

Acute myeloid leukemia (AML) is a heterogeneous hematopoietic malignancy with diverse genetic abnormalities (1). New standard therapies such as Azacitidine and Venetoclax have enhanced survival for older, unfit AML patients at initial diagnosis. Despite these improvements, most AML patients eventually relapse, with survival rates of 32% at two years and 24% at five years (2). Novel therapeutic strategies, such as checkpoint blockade (3) and chimeric antigen receptor T-cell therapies (4), have achieved promising impacts on the outcomes in hematologic malignancies. Similarly, numerous ongoing clinical trials are exploring the effectiveness of stimulating the immune system in the treatment of AML (5–7).

To date, our understanding of the classification of AML has been based on the somatic mutations and cytogenetic abnormalities, which is also the cornerstone of the European LeukemiaNet (ELN) genetic risk classification (8, 9). Nevertheless, leukemia is not just a genetic disorder; it represents a complex microenvironment within the bone marrow, consisting of tumor cells, various immune cells, and other cell types (10). Recently, studies have addressed the significance of immune microenvironment in AML, disclosing that the immune-related genes may predict the therapeutic response and patients' prognosis (11–13). T cells in the bone marrow (BM) are the prerequisite for anti-leukemia response. It has been reported that the percentage and the dysfunctional state of T cells in BM were correlated with the response to treatment and the survival rates (14–17). Given the substantial influence of $CD8^+$ T cells on both the efficiency of immunotherapy and prognosis in AML, we intend to explore whether a prognostic model built on CD8 T-related genes can accurately forecast the survival rates of AML patients.

In order to accurately evaluate $CD8^+$ T cells in AML BM, we chose CIBERSORTx as the machine learning method that enables the estimation of cell type abundances from bulk transcriptome data (18, 19). CIBERSORTx, a computational framework for digitally isolating individual cell type transcriptomes from mixed bulk RNA data without physical separation, now includes functions for cross-platform data normalization and virtual cell purification (18). CIBERSORTx, widely employed as an effective tool in AML research for quantifying immune cell abundance (20–22), has been confirmed for its accuracy compared to other methods (21, 22). Utilizing this method, we calculated $CD8^+$ T cell abundance in AML BM specimens.

In our study, we aim to build a CD8 tumor infiltrating lymphocytes (TILs) related prognostic score which is strongly correlated with overall survival (OS) in AML patients. Firstly, we used four distinct methods to validate the accuracy of CIBERSORTx's prediction of CD8 TILs abundance from bulk RNA-seq. We then analyzed RNA-sequencing data from AML patients in the RJAML cohort (RNA-sequencing data of bone marrow samples from *de novo* AML patients collected at Ruijin Hospital) and generated fifteen CD8 TILs-related genes from the differentially expressed genes (DEGs) to construct the CTCG15 prognostic model. We demonstrated the effectiveness of this model across seven distinct cohorts from RJAML, HOVON, TCGA-LAML, BeatAML, and GEO. Further, we explored the relationship between CTCG15 and the high-frequency mutations in AML. Finally, we integrated CTCG15 into the ELN2022 framework, this resulted in a notable enhancement in the predictive accuracy of ELN. Through our findings, we aim to provide CD8 TILs insights that could refine risk stratification in AML.

Methods

Analysis of scRNA-seq

Before conducting single-cell data analysis, we extracted the immune cells types and expression counts based on the annotations provided in the original text (23). The single-cell RNA sequencing (scRNA-seq) data were normalized using the 'Seurat' R package (24), followed by log-transformation with an offset of 1 and subsequent scaling. We identified genes with substantial expression variation by employing the 'FindVariableFeatures' function in Seurat, specifying 'vst' as the value for the 'vst.method' parameter. Then, we utilized 'ScaleData' function to normalize the gene expression values. We performed linear dimensionality reduction on the high-dimensional dataset through Principal Component Analysis (PCA) on the highly variable genes. The 'IntegrateLayers' was applied to mitigate batch effects across different experimental batches. We then applied the Uniform Manifold Approximation and Projection (UMAP) algorithm for nonlinear dimensionality reduction and visualized the distribution of the data in the reduced-dimensional space. Utilizing 'FindMarkers', we identified genes with statistically significant differences in expression across various groups.

Collection and RNA-sequencing of AML patient samples

We collected bone marrow samples from 157 patients newly diagnosed with AML at Ruijin Hospital, Shanghai, China, during the period from June 2019 to September 2020 (Supplementary Table S6). The collection of the specimens was approved by the Institutional Review Boards of Ruijin Hospital, and the written informed consent for specimen collection and research was obtained following the Declaration of Helsinki (RJ-AML2014-65 & RJ-AML2016). RNA was extracted by AllPrep PowerFecal RNA Kit, and RNA-seq libraries were prepared with TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA, USA). The quality of the RNA was assessed using an RNA 6000 Bioanalyzer Nano Kit (Agilent, Palo Alto, CA, USA). Sequencing was conducted on Illumina's NovaSeq 6000, with adapter-trimmed reads aligned using STAR (v2.7.8) and quantified with HT-Seq (v0.13.5) against GRCh38. Expression levels were normalized to TPM with a custom script. In this manuscript, the cohort is designated as 'RJAML.' The details of the examined loci are available for review in Supplementary Table S7.

Patient cohorts

The scRNA-seq data of AML BM cells from 16 patients were acquired from GSE116256 (23) along with the bulk RNA-seq data of these 16 patients. The HOVON cohort [n = 618; ref. (25)] was obtained from Array Express (Dataset ID: E-MTAB-3444) and used as the training set. Gene expression profiles and survival information were extracted from the GSE146173 [n = 246; ref. (26)], GSE37642 [n = 553; ref. (27)], GSE12417 [n = 240; ref. (28)] cohorts from GEO. TCGA-LAML [n = 179; ref. (29)] and BeatAML [n = 244; ref. (30)] cohorts were accessed through the GDC data portal for the RNA-sequencing (RNA-seq) data, clinical information, and processed mutational variants. It is noteworthy that, only samples from newly diagnosed adult AML patients were retained in all cohorts, and patients with incomplete survival information in these cohorts were omitted.

Estimation of immune cell proportion

The CIBERSORTx (18) algorithm was employed to calculate the relative abundance scores for non-leukemic immune cell types based on the bulk RNA-seq data of AML BM. Specifically, gene expression data from these cells was processed with CIBERSORTx to create a signature matrix, with the minimum expression parameter being tuned to 0.25 from the default settings. Deconvolution was applied to TPM-normalized RNA-seq data using S-mode batch correction in absolute mode and the relative abundance scores were subsequently normalized to indicate the proportion of each cell type.

Flow cytometry analysis

Twenty-four BM specimens from AML patients were randomly selected from the aforementioned 157 RJAML samples. Specimens were thawed and then washed with Dulbecco's phosphate-buffered saline (DPBS) containing 2% fetal bovine serum. Cells were stained with multiple monoclonal antibodies including anti-human CD45-PE, CD3-BV510, and CD8-APC-CY7 at 4°C for 30 minutes. Before FCM analysis, samples were further washed to remove antibodies resuspended in staining buffer. In the flow cytometry analysis, a minimum of 10,000 cells were collected for each sample.

Differential gene expression analysis and enrichment analysis

Differential expression analysis, comparing $CD8^+$ T^{High} and $CD8^+$ T^{Low} patients, were conducted using the raw count data with the DESeq2 R package. We identified DEGs using a cutoff of $|logFC| > 1$ and FDR<0.05, finding 2925 DEGs. GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) analyses of these DEGs were conducted using 'ClusterProfiler'. Enrichment pathways from both GO and KEGG with p-values and q-values below 0.05 were considered significant.

Construction of a prognostic signature based on $CD8^+$ T cells related genes

The normalized gene expression profiles of $CD8^+$ T cells from AML patients in the HOVON cohort were used to serve as the training set. Univariate cox analysis of OS was performed to screen for $CD8^+$ T cells genes with potential prognostic value. Significant candidate genes were refined using the least absolute shrinkage and selection operator (LASSO) model, which was implemented by the glmnet package, to identify a subset of pivotal genes for constructing a predictive model. The LASSO model selected 15 key genes and their regression coefficients were utilized to calculate the CTCG15 score for each sample. Patients were stratified into high-risk and low-risk groups based on the median threshold derived from the CTCG15 scores.

Statistical analyses

Statistical analyses were performed by R software (version 4.0.3). The Wilcoxon rank-sum test assessed differences between two groups. The package "survival" and "survminer" were used to determine the significance of survival analysis. We utilized Kaplan-Meier plots and log-rank tests to evaluate the impact of the signature on OS and event-free survival (EFS). The univariate regression was utilized to select genes of prognostic value with the "survival" package. A two-tailed P-value < 0.05 indicated statistical significance.

Data availability statement

The original data collected in Ruijin Hospital is accessible in the GEO database under accession number GSE201492.

Results

Validation of CIBERSORTx's ability of immune cell deconvolution by comparing with scRNA-seq results

For better understanding, the workflow of this study was illustrated in Figure 1A. Firstly, our goal was to verify the accuracy of CIBERSORTx in measuring $CD8^+$ T cell levels in RNA-seq data. We found that CIBERSORTx's results correlate positively with scRNA-seq data, two RNA-seq scoring methods (Cytolytic Score and Activated $CD8^+$ T Score), and flow cytometry. Secondly, we developed a prognostic model with 15 $CD8^+$ T cell-related genes (CTCG15). After identifying 2925 DEGs in $CD8^+$ T

10.3389/fimmu.2024.1408109

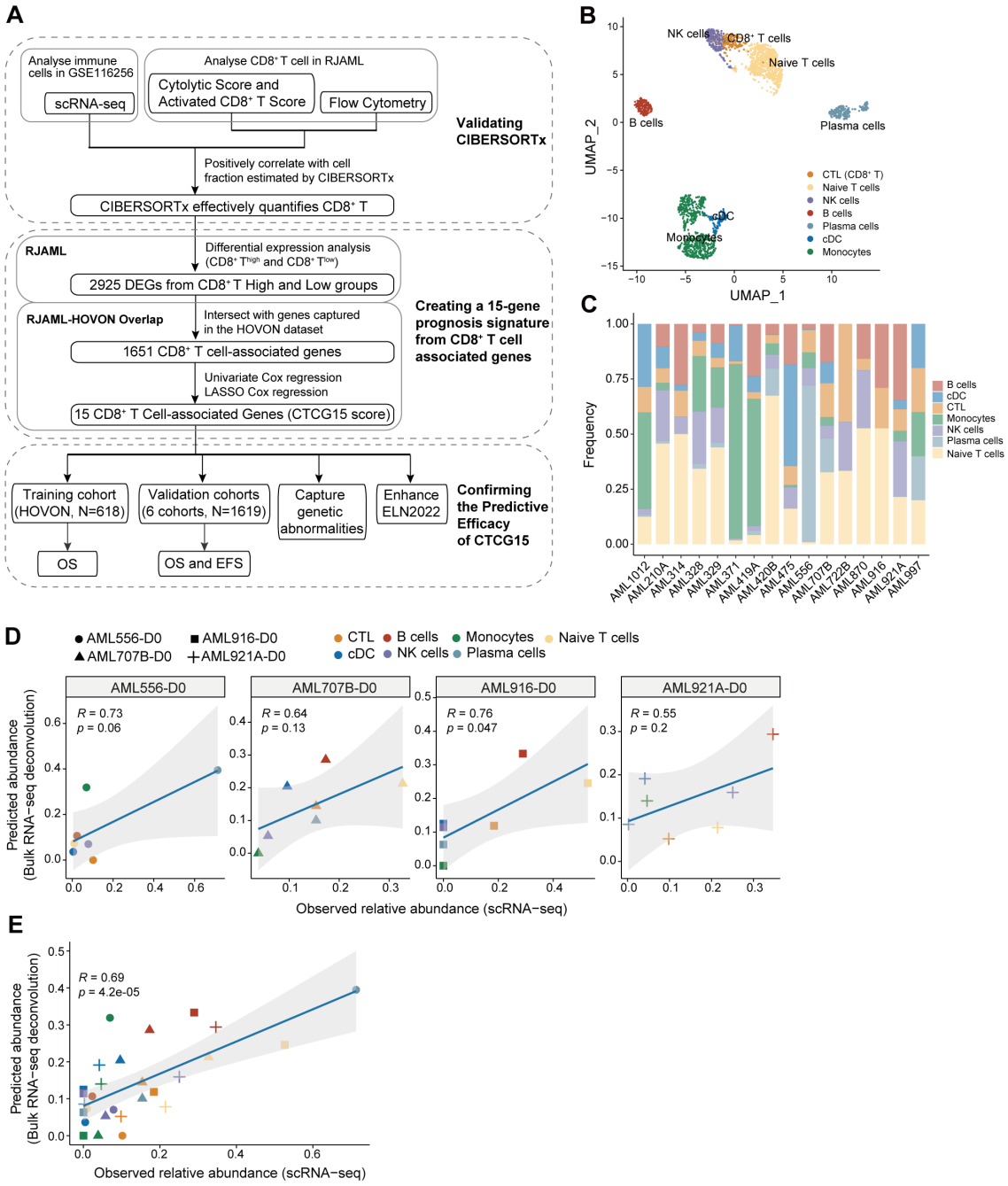

FIGURE 1
Validating the accuracy of CIBERSORTx to deconvolute the immune cells in acute myeloid leukemia (AML) bulk RNA-seq. **(A)** A flowchart for study design. **(B)** Uniform Manifold Approximation and Projection (UMAP) clustering map of bone marrow (BM) immune cells from 16 AML patients in GSE116256 shows the distribution of 7 cellular clusters, representing 7 types of immune cells. **(C)** Bar plot displaying the frequencies of seven types of immune cells from **(B)** separately in 16 AML patients in GSE116256. **(D)** Pearson correlation between observed abundance from scRNA-seq and predicted abundance from bulk RNA-seq for each of the four patients separately in GSE116256. The bulk RNA-seq profile utilized patient-specific reference signatures derived from the scRNA-seq data of each respective patient. **(E)** The Pearson correlation was performed between observed and predicted abundance for four patients combined from **(D)**.

cell expression groups, we used the HOVON cohort to refine these genes. The CTCG15 score was then created using univariate and Lasso Cox regression analyses. Finally, we validated the predictive power of CTCG15 score across multiple cohorts, confirming its ability to forecast patient survival and identify prognostic genetic markers. It also improves the predictive accuracy of the ELN2022 scoring system.

We analyzed the scRNA-seq data from GSE116256 (23) and classified the immune cells into 7 types: CD8$^+$ T cells, natural killer (NK) cells, naive T cells, B cells, plasma cells, monocytes, and

conventional dendritic cells (cDCs) (Figure 1B). The proportion of immune cells for each AML patient was calculated and shown in Figure 1C. We then employed CIBERSORTx (18) to deconvolute the bulk RNA-seq data of four patients in GSE116256 using matched patients-specific reference signatures derived from scRNA-seq data (31). For each patient, we performed Pearson correlation analysis between the predicted abundances from bulk RNA-seq data with CIBERSORTx and the observed abundances from scRNA-seq. Positive correlations were found between CIBERSORTx and scRNA-seq results in both individual (Figure 1D) and combined data of four AML patients (Figure 1E), which demonstrated the reliable performance of CIBERSORTx to deconvolute immune cells.

Further validation of CIBERSORTx's accuracy for CD8$^+$ T cell analysis via additional algorithms and flow cytometry results

We employed other methods at both transcriptomic and proteomic levels to validate CIBERSORTx. To begin with, the predicted abundance of CD8$^+$ T cells from the bulk RNA-seq data in the RJAML cohort was analyzed and quantified by the CIBERSORTx (18) (Figure 2A). It was stratified into three categories based on quartiles for subsequent analysis: High, Intermediate, and Low. In the analysis results from CIBERSORTx, the abundance of CD8$^+$ T cells in the High group was significantly greater than that in the Low group (Figure 2B). This observation was further validated by two independent methods: the Activated CD8$^+$ T Score (31) (Figure 2C) and the Cytolytic Score (32) (Figure 2D), which consistently demonstrated a significantly higher proportion of CD8$^+$ T cells in the High group compared to the Low group. Next, twelve RJAML specimens were randomly selected from both the CTCG^{High} group and the CTCG^{Low} group for flow cytometry analysis, aiming to calculate the percentages of CD8$^+$ T cells among non-leukemic immune cells. The FCM results indicate that the proportion of CD8$^+$ T cells in the CD8$^+$ T High group defined by CIBERSORTx is significantly higher than that in the Low group (Figure 2E), which is consistent with the results of CIBERSORTx in Figure 2B. Pearson correlation revealed a significant positive correlation between the predicted CD8$^+$ T cell abundance from bulk RNA-seq deconvolution and the CD8$^+$ T cell percentages from flow cytometry (Figure 2F). Figure 2G displays five representative flow cytometry plots for each group, and their tendencies align with the results obtained from CIBERSORTx. Overall, the capability of CIBERSORTx in evaluating CD8$^+$ TILs has been further validated with the methods above.

Development of a 15-gene prognostic signature derived from CD8$^+$ T cell related genes in AML

To identify genes associated with CD8$^+$ T cell, we performed differential analysis between the CD8 TILs High and Low groups with the DESeq2 package (32) and found a total of 2925 DEGs (with |logFC| >1 and FDR<0.05) (Supplementary Table S1, Figure 3A). We performed GO and KEGG enrichment analysis (Supplementary Tables S2, S3) to further elucidate the biological functions and to identify signaling pathways linked to these DEGs. The top 15 GO (Figure 3B) and KEGG (Figure 3C) terms associated with the DEGs cover various immune responses and immune cell activation pathways, which underline that DEGs are associated with immune reactions.

1651 CD8$^+$ T cell genes were retained after we intersected 2925 DEGs with genes in the training cohort HOVON (25) (Figure 3D). Using univariate cox regression and LASSO regression, we yielded an optimal 15-gene CD8$^+$ T cell signature (the CTCG15 score) as shown in Figure 3E. These fifteen genes involved are *PLIN2*, *MSLN*, *MYH10*, *RXFP1*, *OLFML2A*, *ST6GALNAC4*, *WASIR2*, *MRC1*, *CA3*, *SHANK3*, *C8orf88*, *IL2RA*, *ITGA2B*, *PRUNE2* and *SEMA4F*. Kaplan–Meier analysis revealed a significant association between high CTCG15 risk scores and reduced OS in the training cohort HOVON, suggesting its ability to effectively predict prognosis for AML patients (Figure 3F).

Extensive assessment of the predictive ability of CTCG15 in multiple external cohorts

After the development of CTCG15, we explored its correlation with OS and EFS in 1619 AML patients across six cohorts and two technology platforms. We first validated the CTCG15 score with the RNA-seq data of 157 *de novo* AML patients from the RJAML cohort. Consistent with our observations from the HOVON dataset, patients exhibiting high CTCG15 scores (CTCG15^{High}) showed notably worse OS (Figure 4A) and EFS (Figure 4B) than those with low CTCG15 scores (CTCG15^{Low}). When applied to other RNA-seq datasets including TCGA (n = 179) (29), BeatAML (n = 244) (30), and GSE146173 (n = 246) (26), the CTCG15 all remained strongly associated with clinical outcomes (Figures 4C–E). Similarly, the significant difference of OS between CTCG15^{High} and CTCG15^{Low} groups was also observed in the GSE37642 (n = 553) (27) (Figures 4F, G) and GSE12417 (n = 240) (28) (Figures 4H, I), with both datasets being analyzed using the GPL96 and GPL570 platforms. In multivariate survival analysis with Cox proportional hazards (CPH) models, the CTCG15 score demonstrated ability to constitute a novel independent prognostic indicator in the TCGA, BeatAML, and RJAML cohorts, apart from the established outcome markers like patient age, white blood cell (WBC) count, ELN2022, and the presence of *NPM1* and *FLT3-ITD* mutations (Supplementary Table S4).

CTCG15 captured specific genetic abnormalities related to AML prognosis

For a more thorough understanding of the mutational landscape linked to the CTCG15 score, we investigated the recurrently mutated somatic driver genes within the combined dataset comprising the BeatAML, RJAML and TCGA cohorts

10.3389/fimmu.2024.1408109

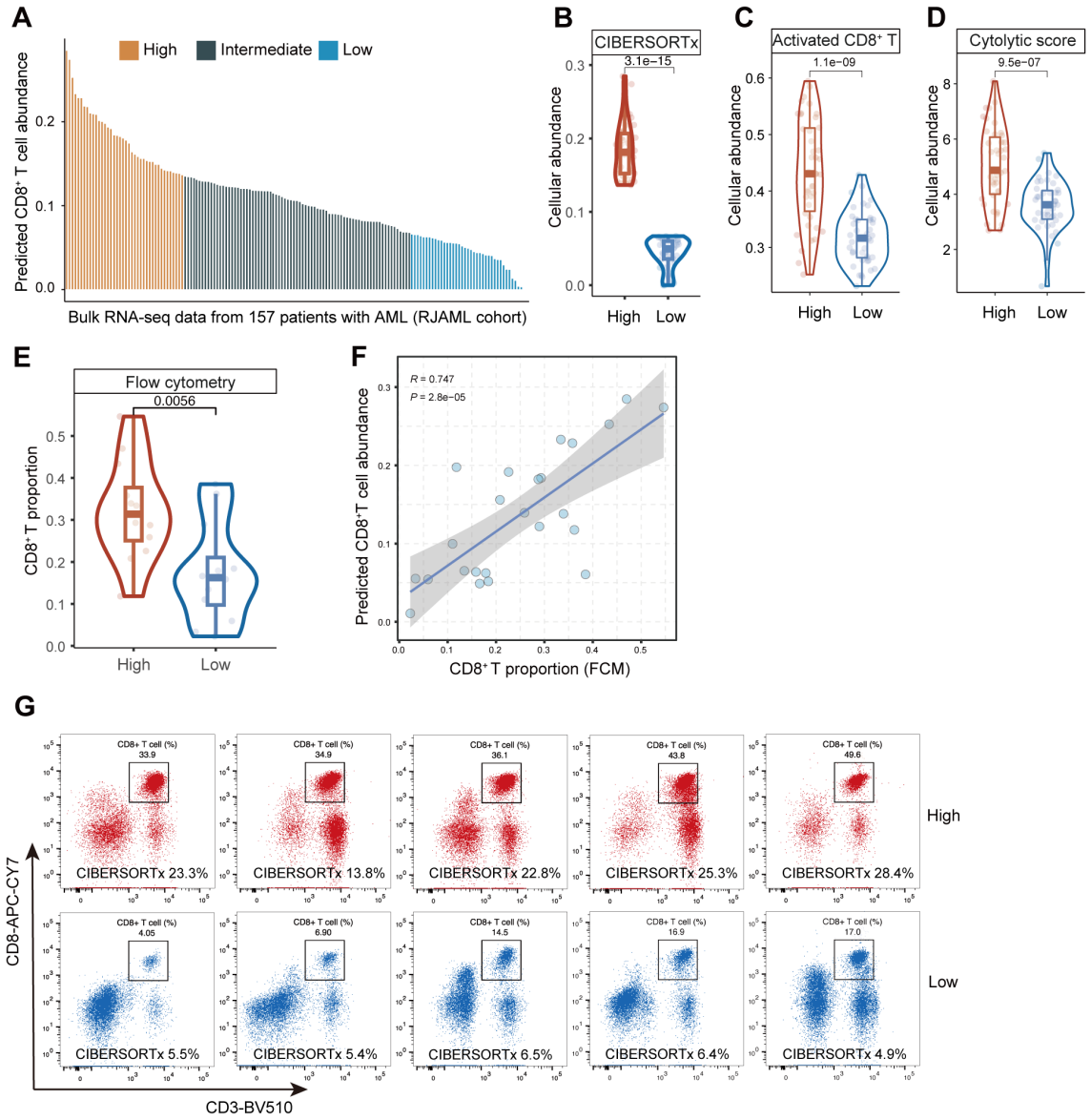

FIGURE 2
Assessment of predictive performance of CIBERSORTx in CD8 tumor infiltrating lymphocytes (TILs) from bulk RNA-seq data in RJAML cohort. **(A)** The predicted abundance of $CD8^+$ T cells in RJAML with CIBERSORTx, and the three groups (High, Intermediate and Low) categorized based on quartiles. **(B–D)** The abundance of $CD8^+$ T cells between $CD8^+$ T^{High} and $CD8^+$ T^{Low} groups calculated by CIBERSORTx, the Activated $CD8^+$ T Score and the Cytolytic Score. **(E)** The comparison of $CD8^+$ T cell proportions detected by flow cytometry (FCM) of 24 patients randomly selected from $CD8^+$ T^{High} and $CD8^+$ T^{Low} groups in RJAML (12 from the $CD8^+$ T^{High} group and 12 from $CD8^+$ T^{Low} group). **(F)** Pearson correlation between the $CD8^+$ T proportion from CIBERSORTx and FCM of the 24 patients. **(G)** Ten representative $CD8^+$ T cells FCM plots are shown in $CD8^+$ T^{High} and $CD8^+$ T^{Low} groups.

(Figure 5A). Four molecular markers showed significant frequency variations between the $CTCG15^{High}$ group and $CTCG15^{Low}$ group (Figure 5B). The $CTCG15^{High}$ group exhibited higher frequencies of *SRSF2* and *RUNX1* mutations, which are also markers of the ELN2022 (8) Adverse Risk group. Conversely, patients with low CTCG15 score more commonly presented mutations in the *CEBPA* (including *CEBPA* bZIP mutation and other types) and *SMC1A* genes. In ELN2022 risk classification, bZIP in-frame mutated *CEBPA* is presented in the Favorable Risk group. Notably, the presence of the *CEBPA* mutation indicates a relatively favorable clinical outcome, while *SRSF2* mutation and *RUNX1* mutation an adverse outcome, as demonstrated by univariate analysis (Figures 5C–E). In a combined cohort included TCGA cohort, BeatAML cohort and RJAML cohort, a multivariate CPH regression analysis revealed that the CTCG15 score was independent of well-known clinical parameters, ELN2022 classification and the above mentioned three genetic abnormalities (Supplementary Table S5).

10.3389/fimmu.2024.1408109

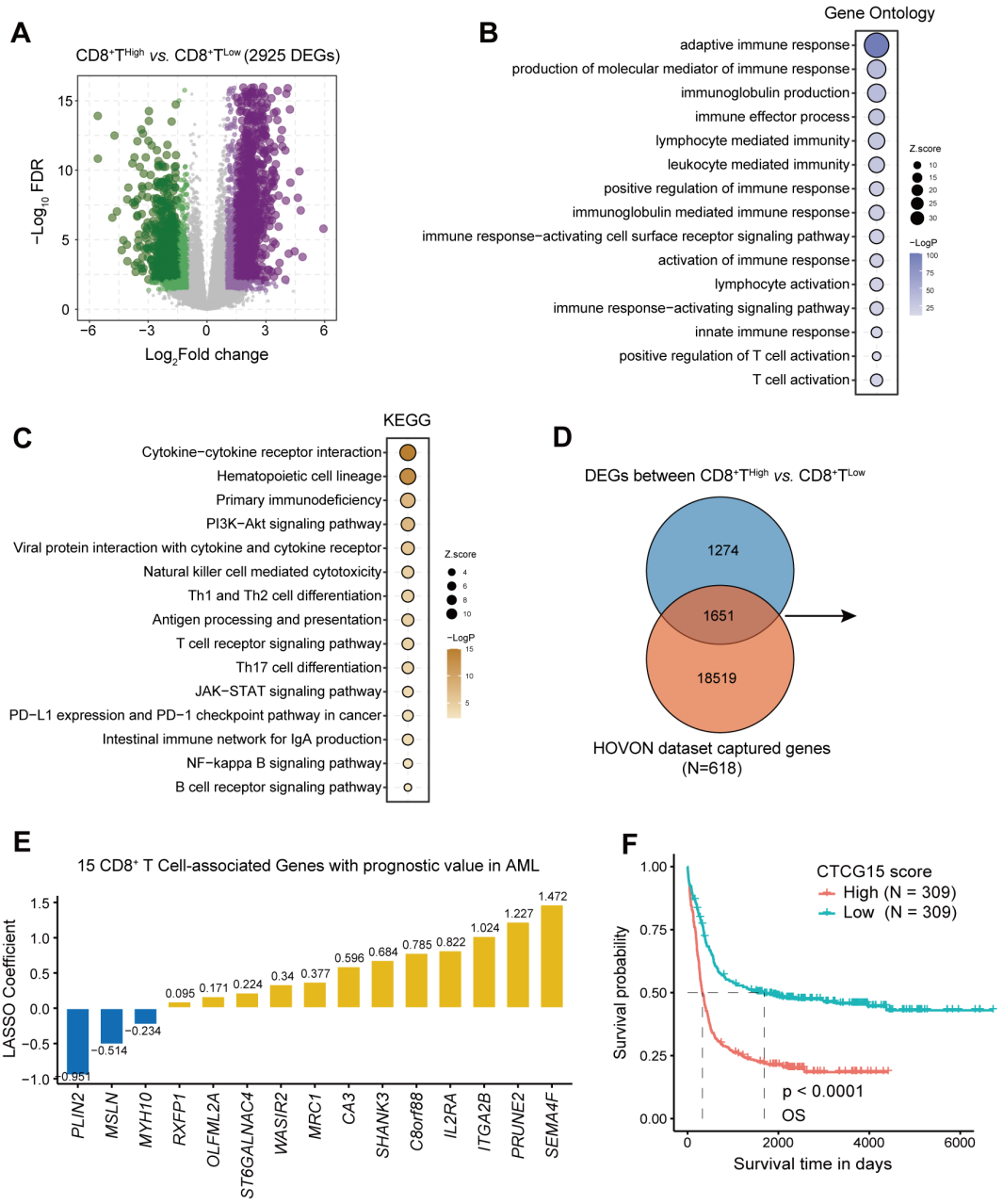

FIGURE 3
Identification of CD8$^+$ T cell-related genes in AML and construction of 15-gene prognostic signature. **(A)** Volcano plot showing differentially expressed genes (DEGs) between groups with high and low CD8$^+$ T cell levels, determined using the threshold of |log2(fold change)| ≥ 1 and FDR < 0.05. **(B, C)** Bubble chart displaying GO and KEGG analysis results for CD8$^+$ T cell-related genes. **(D)** A Venn diagram identifying the overlap between DEGs and HOVON dataset. **(E)** Distribution of LASSO coefficients for prognosis-related genes. **(F)** Kaplan-Meier curves depicting the difference of overall survival (OS) between CTCG15^{High} and CTCG15^{Low} groups in the HOVON cohort.

The CTCG15 can serve as a valuable complement to the ELN2022 risk classification

The 2022 ELN risk classification, which is predicated on genetic abnormalities, is widely accepted for the risk assessment of pediatric and adult patients with AML (8). Given the strong association between CTCG15 score and the prognosis of AML patients, we integrated the CTCG15 signature into the ELN2022 scheme with the intention to refine the AML patients risk classification from the immunological perspective. Among all patients categorized within the ELN2022 framework, the incorporation of the CTCG15 score led to a reclassification for approximately half of the patients. Within the ELN2022-Favorable group, 33.3% of patients were reassigned to the Intermediate category of the ELN2022 +CTCG15. In the ELN2022-Intermediate group, 45.5% of patients were shifted to the Adverse category of the ELN2022+CTCG15. Furthermore, in the ELN2022-Adverse group, 42.4% of patients

10.3389/fimmu.2024.1408109

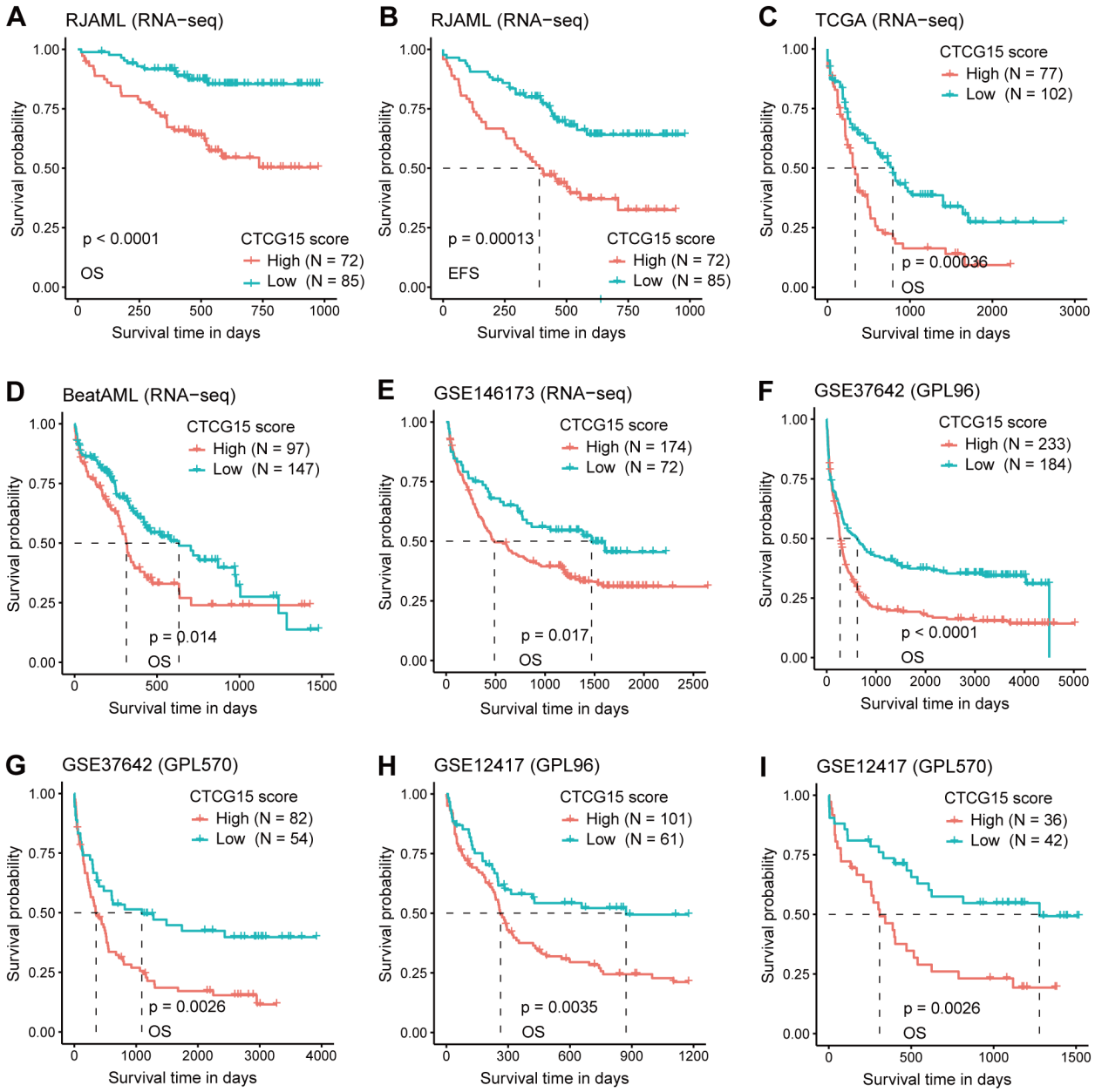

FIGURE 4
The CTCG15 score is strongly linked with OS and event free survival (EFS) across various independent AML cohorts in different analysis platforms. Kaplan–Meier curves depicting the difference of OS **(A)** and EFS **(B)** between CTCG15^{High} and CTCG15^{Low} groups in the RJAML cohort. Kaplan-Meier estimates of OS according to the CTCG15 score in the RNA-seq–based cohorts: TCGA **(C)**, BeatAML **(D)**, and GSE146173 **(E)**. Kaplan–Meier estimates of OS based on the CTCG15 score in the GSE37642 **(F, G)** and GSE12417 **(H, I)** cohorts, quantified on GPL96 and GPL570 microarray platforms.

were reclassified to the Intermediate category of the ELN2022 +CTCG15 (Figure 6A). The restructured risk stratification method improves the classification of patients with AML into distinct risk groups from the perspective of immunological factors (Figures 6B, C). The revised risk scheme exhibits an elevated Harrell C-index in the combined cohorts, signifying an enhancement in the model's predictive accuracy and reliability for assessing patient prognosis (Figure 6D). This refinement has the potential to make an effective tool for tailoring individualized treatment plans and predicting patient outcomes.

Discussion

To our understanding, leukemogenesis is driven by the genomic and epigenetic abnormalities. However, the tumor microenvironment, especially the CD8$^+$ T in the BM, plays a critical role in adaptive immune reactions, and further affect the efficacy of immunotherapy (3, 33, 34). In this investigation, we tried to elucidate the impacts of BM CD8$^+$ T cells in AML.

To quantify the abundance of CD8 TILs in AML, we chose CIBERSORTx (18) among various computational methods (35).

10.3389/fimmu.2024.1408109

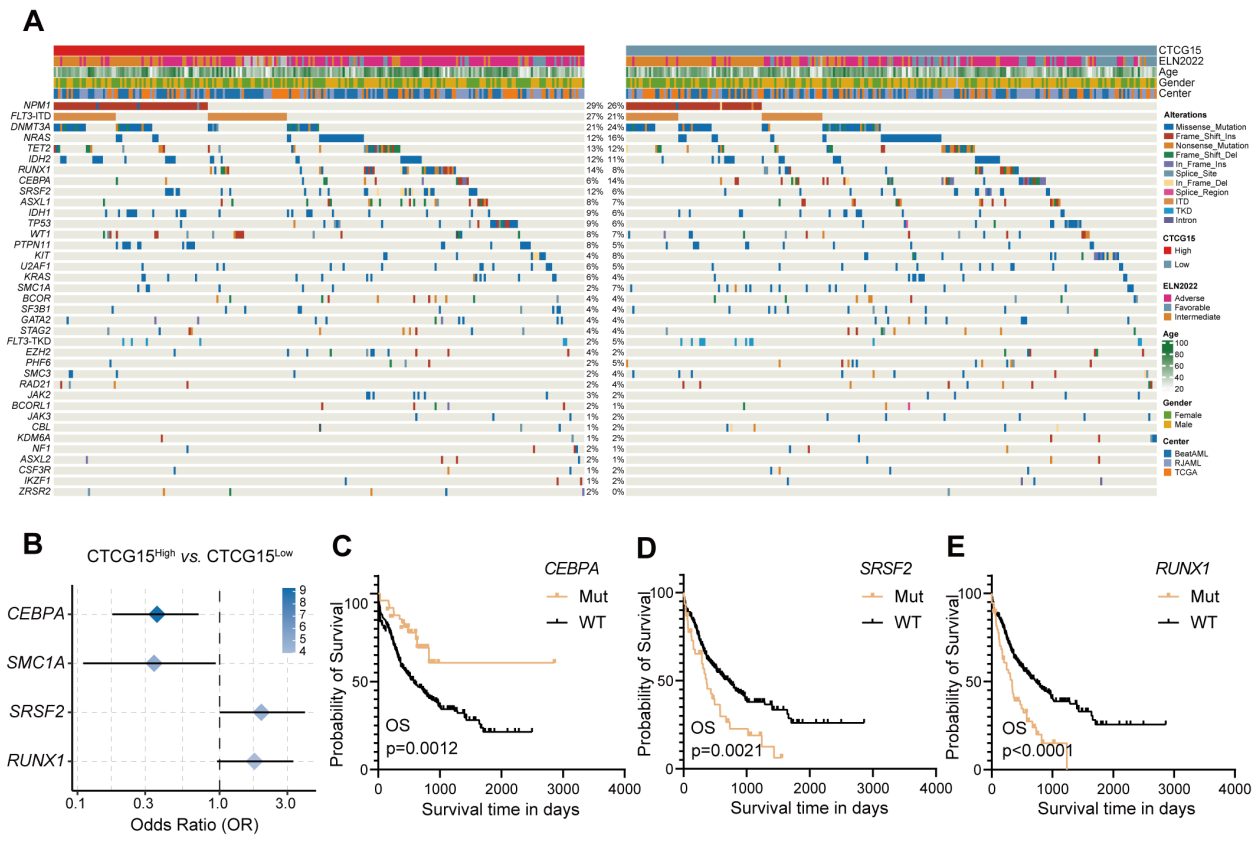

FIGURE 5
CTCG15 identified distinct genetic abnormalities associated with AML prognosis. **(A)** Somatic mutations and clinical information are presented in a heatmap featuring CTCG15^{High} and CTCG15^{Low} patient groups. **(B)** The forest plot displays gene mutations with significantly different frequencies between CTCG15^{High} and CTCG15^{Low} groups. Kaplan–Meier estimates OS based on the status of *CEBPA* **(C)**, *SRSF2* **(D)** and *RUNX1* **(E)**.

Several studies have reported its accuracy in calculating the levels of immune cells in AML (21, 31). Zeng et al. reported in their study that 73% of the cell type estimates generated by CIBERSORTx with S-mode batch correction did not deviate by more than 5% from the scRNA-seq results (31). To validate the accuracy of the CIBERSORTx algorithm, we firstly investigated the discrepancy between the deconvoluted cell type abundance and true cell type abundance from scRNA-seq data. A significant positive association between predicted and observed relative abundance is demonstrated by the Pearson correlation. We further employed two supplementary algorithms (36, 37) to certify the stratification effectiveness. The results aligned coherently with the outcomes produced by the CIBERSORTx evaluation. At the protein level, we adopted the flow cytometry technique (21) to corroborate the proportions of CD8$^+$ T cells in BM across both CD8$^+$ T^{High} and CD8$^+$ T^{Low} groups and the results were consistent with our findings at RNA-level. Concurrently, we employed Pearson correlation to validate their association. The results above strongly supported the reliability of CIBERSORTx as a computational tool to estimate CD8$^+$ T cell abundance in bulk RNA-seq.

After calculating the abundance of CD8 TILs using CIBERSORTx in the RJAML cohort, we identified DEGs associated with CD8$^+$ T cells for the subsequent analysis. A total of 2925 DEGs that showed significant enrichment in immune activation-related responses and pathways were identified by GO and KEGG enrichment analysis. Using univariate CPH regression and LASSO algorithms, we derived an optimal 15-gene signature (the CTCG15 score) based on CD8 TILs-associated genes. The CTCG15 score was validated using RNA-seq and microarray data from six datasets in total. These datasets encompass a diverse range of ethnicities, including Asians, Africans, and Caucasians, with RJAML dataset specifically being sourced from Ruijin Hospital in Shanghai, China. With multiple validation across different datasets and various populations, the CTCG15 score has demonstrated excellent predictive ability and can be applied to extensive cohorts.

CTCG15 consists of fifteen CD8$^+$ T cell marker genes, including *PLIN2*, *MSLN*, *MYH10*, *RXFP1*, *OLFML2A*, *ST6GALNAC4*, *WASIR2*, *MRC1*, *CA3*, *SHANK3*, *C8orf88*, *IL2RA*, *ITGA2B*, *PRUNE2* and *SEMA4F*. Most of these genes have shown associations with the prognosis of AML or the activity of CD8$^+$ T cells, which is in accordance with our results. Among the 15 genes used to construct the prognostic model, higher expression of PLIN2, MSLN and MYH10 are associated with favorable prognosis. Research has revealed that the upregulation of lipid droplets associated genes (*PLIN2*) is associated with the enhanced cytotoxic T lymphocytes activity (38). Mesothelin (MSLN) is a new cell surface indicator of the disease and a prospective therapeutic target for AML (39). Recent studies have shown that mesothelin is a key therapeutic target in pediatric AML, and two

10.3389/fimmu.2024.1408109

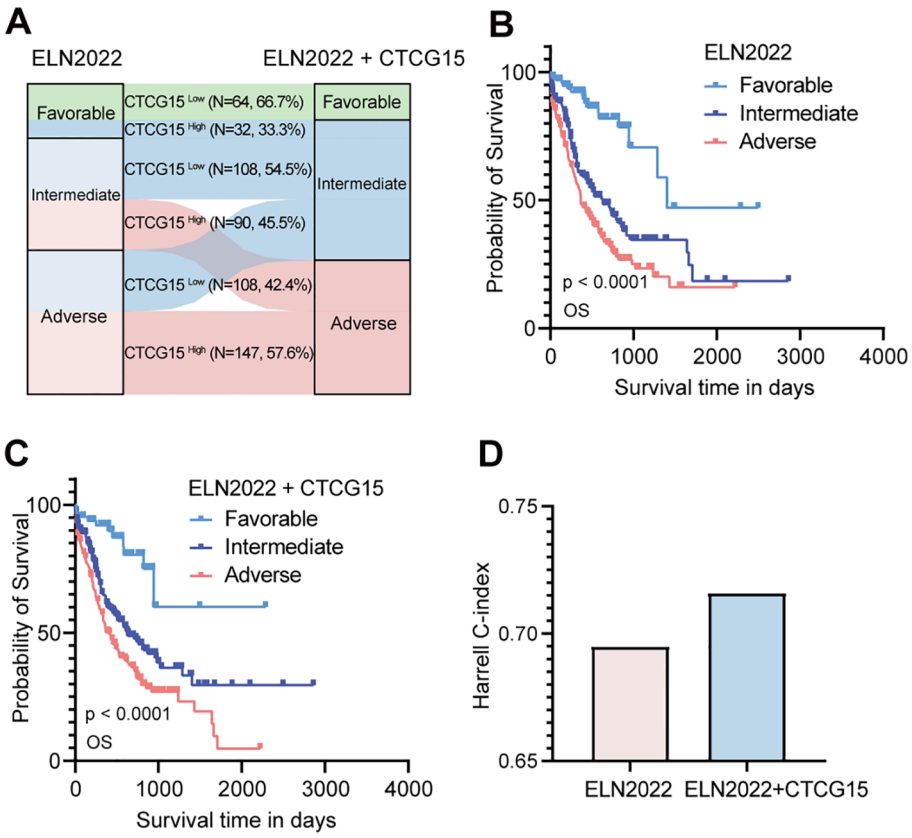

FIGURE 6
Refine the 2022 European LeukemiaNet (ELN2022) classification by integrating CTCG15. **(A)** Patients in different risk status are recategorized from the original ELN2022 schema categories (Favorable, Intermediate, and Adverse) to the ELN2022 integrated with CTCG15. **(B, C)** Kaplan–Meier estimates of OS based on the risk categories of AML patients in the ELN2022 **(B)** and ELN2022+CTCG15 **(C)** within the combined cohort (RJAML, TCGA-LAML and BeatAML). **(D)** The Harrell C-index assesses performance of risk classification of ELN2022 alone and the ELN2022 integrated CTCG15.

MSLN/CD3-targeting bispecific antibodies have achieved complete remission in mouse models (40). Genes negatively correlated with CTCG15 are associated with T cell activation.

As for the remaining 12 genes whose expression is negatively associated with prognostic, IL2RA and MRC1 have been extensively studied and are closely related to immune cells. In AML, IL2RA's overexpression correlates with poor treatment response and adverse outcomes (41, 42), and it ranks as a crucial gene in survival prediction analyses (43). Recently, studies have revealed CD206, encoded by the *MRC1* gene, is an independent adverse prognostic indicator for AML patients (44). Among the other 10 genes positively correlated with CTCG15, each exhibits a profound association with diverse facets of neoplastic initiation and subsequent progression. *SEMA4F* (45) and *C8orf88* (46) are significant contributor to prostate cancer progression. Research has indicated that carbonic anhydrases(CA) play an indispensable role in ensuring leukemic cell viability within an oxygen-deprived environment (47). *ST6GALNAC4* (48), *OLFML2A* (49) and *RXFP1* (50) exhibit a notable association with the initiation and progression of various neoplastic entities. The influential role of most genes that positively correlated with CTCG15 in tumorigenesis and tumor development may contribute to the shorter OS observed in $CTCG15^{High}$.

The goal of establishing an AML prognostic signature is to enhance the effectiveness of both diagnosis and treatment, as well as to provide a more accurate prognosis. Somatic mutations and chromosomal abnormalities drive the onset and development of AML (51, 52). Therefore, we investigated the gene abnormalities related to AML between $CTCG15^{High}$ and $CTCG15^{Low}$. We found four mutations with significantly different frequencies between the $CTCG15^{High}$ and $CTCG15^{Low}$ subgroups. As mentioned, the $CTCG15^{High}$ group is associated with poor prognosis in AML across multiple validation datasets. The occurrence of *SRSF2* and *RUNX1* alterations is higher in this group, aligning with the conclusions of 2022 ELN risk classification (8) and other studies (53–55). Conversely, the $CTCG15^{Low}$ group exhibits a higher incidence of *CEBPA* and *SMC1A* alterations. We have demonstrated that patients with *CEBPA* mutation (including *CEBPA* bZIP mutation and other types) have a prolonged survival. Studies have confirmed that mutations in the bZIP domain of *CEBPA* are associated with a favorable prognosis for patients (8, 56). In the TCGA and BeatAML datasets, the mutation details for *CEBPA* were not provided. However, bZIP mutations account for 90.9% of all *CEBPA* mutations in RJAML. In results, we observed that patients with *CEBPA* mutations have a better prognosis than those without mutations, which may be due to the high proportion of bZIP mutations in the combined cohort. However, the situation is somewhat complicated for *SMC1A*. Structural

maintenance of chromosomes protein 1A (SMC1A) is a core unit of the cohesin complex regulating chromosome segregation during meiosis and mitosis (57), which has not been reported with association of prognosis in AML (58, 59). Our results show that alterations in *SMC1A* are more prevalent in the CTCG15^{Low} group, and the underlying mechanism of this observation needs to be further investigated.

To further validate the clinical applicability of CTCG15, we integrated it with the ELN2022 risk classification, leading to more effective patient stratification and increased precision in prognostic accuracy. Several prognostic scores such as LSC17 (60), GENE4 (61), and AFG16 (62) have estimated the outcomes for AML patients recently. Compared to these, our model aims to improve ELN2022's stratification and prediction ability by incorporating insights from the CD8 TILs in tumor microenvironment. Patients who were reclassified in the three groups were closely associated with CD8 TIL. It has been reported that the function of CD8$^+$ T cells is associated to the gene expression in leukemia cells (14, 15). In patients with favorable-risk AML, the proliferation and stemness of leukemic stem and progenitor cells are driven by a limited number of intrinsic molecular abnormalities. Additionally, bone marrow-infiltrating CD8 T cells play a key role in regulating these leukemic stem and progenitor cells (14). However, more aggressive AML is propagated mainly by cell-intrinsic mechanisms and develops independent of immune cells (14). These factors may contribute to the reallocations in ELN2022+CTCG15. Multiple studies have demonstrated that the extent of CD8$^+$ T cell infiltration in tumors is closely correlated with patients prognosis and the efficacy of chemotherapy (63–65). These findings coincide with our discovery that ELN2022+CTCG15 improved predictive performance.

To conclude, our study validated the accuracy of CIBERSORTx to estimate the CD8 TILs abundance from the bulk RNA-seq data. We defined fifteen DEGs associated with CD8 TIL abundance and constructed the CTCG15 prognostic model. The CTCG15 score can predict the survival of AML patients, capture gene abnormalities in AML, and significantly enhance the predictive precision of ELN2022. Our research highlights the capability of CTCG15 score to serve as a valuable tool to refine risk stratification in AML and to indicate patient selection for the potential immunotherapy.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GEO database under accession number GSE201492.

Ethics statement

The studies involving humans were approved by Institutional Review Boards from Ruijin Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZYL: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. PJ: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. RFX: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. XYL: Writing – review & editing, Conceptualization. JS: Writing – review & editing. MKH: Writing – review & editing. XXL: Writing – review & editing. HMZ: Writing – review & editing. SSW: Writing – original draft, Resources. FYD: Writing – review & editing. HJZ: Writing – review & editing, Funding acquisition. HL: Resources, Writing – review & editing, Funding acquisition. ZJ: Writing – review & editing, Writing – original draft, Funding acquisition. JML: Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Key Research and Development Program of China (2019YFA0905900), National Natural Sciences Foundation of China (82000196, 82370157, 81870110, 82200158) and Shanghai Rising-Star Program (21QA1405700).

Acknowledgments

The computational analysis was executed on the π 2.0 cluster, which is maintained by the Center for High Performance Computing at Shanghai Jiao Tong University.

Conflict of interest

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

The reviewer YT declared a shared affiliation, though no other collaboration, with one of the authors RX at the time of the review.

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.2024.1408109/full#supplementary-material

References

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. (2015) 373:1136–52. doi: 10.1056/NEJMra1406184

2. Shallis RM, Wang R, Davidoff A, Ma X, Zeidan AM. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev*. (2019) 36:70–87. doi: 10.1016/j.blre.2019.04.005

3. Xu-Monette ZY, Zhou J, Young KH. PD-1 expression and clinical PD-1 blockade in B-cell lymphomas. *Blood*. (2018) 131:68–83. doi: 10.1182/blood-2017-07-740993

4. June CH, Sadelain M. Chimeric antigen receptor therapy. *N Engl J Med*. (2018) 379:64–73. doi: 10.1056/NEJMra1706169

5. Zeidner JF, Vincent BG, Ivanova A, Moore D, McKinnon KP, Wilkinson AD, et al. Phase II trial of pembrolizumab after high-dose cytarabine in relapsed/refractory acute myeloid leukemia. *Blood Cancer Discov*. (2021) 2:616–29. doi: 10.1158/2643-3230.BCD-21-0070

6. Jin X, Zhang M, Sun R, Lyu H, Xiao X, Zhang X, et al. First-in-human phase I study of CLL-1 CAR-T cells in adults with relapsed/refractory acute myeloid leukemia. *J Hematol Oncol*. (2022) 15:88. doi: 10.1186/s13045-022-01308-1

7. Baumeister SH, Murad J, Werner L, Daley H, Trebeden-Negre H, Gicobi JK, et al. Phase I trial of autologous CAR T cells targeting NKG2D ligands in patients with AML/MDS and multiple myeloma. *Cancer Immunol Res*. (2019) 7:100–12. doi: 10.1158/2326-6066.CIR-18-0307

8. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. (2022) 140:1345–77. doi: 10.1182/blood.2022016867

9. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. (2017) 129:424–47. doi: 10.1182/blood-2016-08-733196

10. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell*. (2023) 41:374–403. doi: 10.1016/j.ccell.2023.02.016

11. Fu D, Zhang B, Wu S, Zhang Y, Xie J, Ning W, et al. Prognosis and characterization of immune microenvironment in acute myeloid leukemia through identification of an autophagy-related signature. *Front Immunol*. (2021) 12:695865. doi: 10.3389/fimmu.2021.695865

12. Huang S, Zhang B, Fan W, Zhao Q, Yang L, Xin W, et al. Identification of prognostic genes in the acute myeloid leukemia microenvironment. *Aging (Albany NY)*. (2019) 11:10557–80. doi: 10.18632/aging.102477

13. Zeng T, Cui L, Huang W, Liu Y, Si C, Qian T, et al. The establishment of a prognostic scoring model based on the new tumor immune microenvironment classification in acute myeloid leukemia. *BMC Med*. (2021) 19:176. doi: 10.1186/s12916-021-02047-9

14. Radpour R, Riether C, Simillion C, Höpner S, Bruggmann R, Ochsenbein AF. CD8+ T cells expand stem and progenitor cells in favorable but not adverse risk acute myeloid leukemia. *Leukemia*. (2019) 33:2379–92. doi: 10.1038/s41375-019-0441-9

15. Knaus HA, Berglund S, Hackl H, Blackford AL, Zeidner JF, Montiel-Esparza R, et al. Signatures of CD8+ T cell dysfunction in AML patients and their reversibility with response to chemotherapy. *JCI Insight*. (2018) 3:e120974. doi: 10.1172/jci.insight.120974

16. Le Dieu R, Taussig DC, Ramsay AG, Mitter R, Miraki-Moud F, Fatah R, et al. Peripheral blood T cells in acute myeloid leukemia (AML) patients at diagnosis have abnormal phenotype and genotype and form defective immune synapses with AML blasts. *Blood*. (2009) 114:3909–16. doi: 10.1182/blood-2009-02-206946

17. Lamble AJ, Kosaka Y, Laderas T, Maffit A, Kaempf A, Brady LK, et al. Reversible suppression of T cell function in the bone marrow microenvironment of acute myeloid leukemia. *Proc Natl Acad Sci USA*. (2020) 117:14331–41. doi: 10.1073/pnas.1916206117

18. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*. (2019) 37:773–82. doi: 10.1038/s41587-019-0114-2

19. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. (2015) 12:453–7. doi: 10.1038/nmeth.3337

20. Jiang G, Jin P, Xiao X, Shen J, Li R, Zhang Y, et al. Identification and validation of a novel CD8+ T cell-associated prognostic model based on ferroptosis in acute myeloid leukemia. *Front Immunol*. (2023) 14:1149513. doi: 10.3389/fimmu.2023.1149513

21. Wang Y, Cai Y-Y, Herold T, Nie R-C, Zhang Y, Gale RP, et al. An immune risk score predicts survival of patients with acute myeloid leukemia receiving chemotherapy. *Clin Cancer Res*. (2021) 27:255–66. doi: 10.1158/1078-0432.CCR-20-3417

22. Chen H, Wu M, Xia H, Du S, Zhou G, Long G, et al. FLT3LG and IFITM3P6 consolidate T cell activity in the bone marrow microenvironment and are prognostic factors in acute myelocytic leukemia. *Front Immunol*. (2022) 13:980911. doi: 10.3389/fimmu.2022.980911

23. van Galen P, Hovestadt V, Ii MHW, Hughes TK, Griffin GK, Battaglia S, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. (2019) 176:1265–1281.e24. doi: 10.1016/j.cell.2019.01.031

24. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. (2021) 184:3573–3587.e29. doi: 10.1016/j.cell.2021.04.048

25. Valk PJM, Verhaak RGW, Beijen MA, Erpelinck CAJ, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med*. (2004) 350:1617–28. doi: 10.1056/NEJMoa040465

26. Bamopoulos SA, Batcha AMN, Jurinovic V, Rothenberg-Thurley M, Janke H, Ksienzyk B, et al. Clinical presentation and differential splicing of SRSF2, U2AF1 and SF3B1 mutations in patients with acute myeloid leukemia. *Leukemia*. (2020) 34:2621–34. doi: 10.1038/s41375-020-0839-4

27. Li Z, Herold T, He C, Valk PJM, Chen P, Jurinovic V, et al. Identification of a 24-gene prognostic signature that improves the european leukemiaNet risk classification of acute myeloid leukemia: an international collaborative study. *JCO*. (2013) 31:1172–81. doi: 10.1200/JCO.2012.44.3184

28. Metzeler KH, Hummel M, Bloomfield CD, Spiekermann K, Braess J, Sauerland M-C, et al. An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood*. (2008) 112:4193–201. doi: 10.1182/blood-2008-02-134411

29. Cancer Genome Atlas Research Network, Timothy JL, Christopher M, Ding L, Benjamin JR, Andrew JM, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *Engl J Med*. (2013) 368:2059–74. doi: 10.1056/NEJMoa1301689

30. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, et al. Functional genomic landscape of acute myeloid leukemia. *Nature*. (2018) 562:526–31. doi: 10.1038/s41586-018-0623-z

31. Zeng AGX, Bansal S, Jin L, Mitchell A, Chen WC, Abbas HA, et al. A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. *Nat Med*. (2022) 28:1212–23. doi: 10.1038/s41591-022-01819-x

32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. (2014) 15:550. doi: 10.1186/s13059-014-0550-8

33. Sweeney C, Vyas P. The graft-versus-leukemia effect in AML. *Front Oncol*. (2019) 9:1217. doi: 10.3389/fonc.2019.01217

34. Brudno JN, Kochenderfer JN. Chimeric antigen receptor T-cell therapies for lymphoma. *Nat Rev Clin Oncol*. (2018) 15:31–46. doi: 10.1038/nrclinonc.2017.128

35. Bortolomeazzi M, Keddar MR, Ciccarelli FD, Benedetti L. Identification of non-cancer cells from cancer transcriptomic data. *Biochim Biophys Acta (BBA) - Gene Regul Mech*. (2020) 1863:194445. doi: 10.1016/j.bbagrm.2019.194445

36. Dufva O, Pölönen P, Brück O, Keränen MAI, Klievink J, Mehtonen J, et al. Immunogenomic landscape of hematological Malignancies. *Cancer Cell*. (2020) 38:380–399.e13. doi: 10.1016/j.ccell.2020.06.002

37. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019

38. Nava Lauson CB, Tiberti S, Corsetto PA, Conte F, Tyagi P, Machwirth M, et al. Linoleic acid potentiates CD8+ T cell metabolic fitness and antitumor immunity. *Cell Metab*. (2023) 35:633–650.e9. doi: 10.1016/j.cmet.2023.02.013

39. Kaeding AJ, Barwe SP, Gopalakrishnapillai A, Ries RE, Alonzo TA, Gerbing RB, et al. Mesothelin is a novel cell surface disease marker and potential therapeutic target in acute myeloid leukemia. *Blood Adv*. (2021) 5:2350–61. doi: 10.1182/bloodadvances.2021004424

40. Gopalakrishnapillai A, Correnti CE, Pilat K, Lin I, Chan MK, Bandaranayake AD, et al. Immunotherapeutic targeting of mesothelin positive pediatric AML using bispecific T cell engaging antibodies. *Cancers (Basel)*. (2021) 13:5964. doi: 10.3390/cancers13235964

41. Du W, He J, Zhou W, Shu S, Li J, Liu W, et al. High IL2RA mRNA expression is an independent adverse prognostic biomarker in core binding factor and intermediate-risk acute myeloid leukemia. *J Transl Med*. (2019) 17:191. doi: 10.1186/s12967-019-1926-z

42. Fujiwara S-I, Muroi K, Yamamoto C, Hatano K, Okazuka K, Sato K, et al. CD25 as an adverse prognostic factor in elderly patients with acute myeloid leukemia. *Hematology*. (2017) 22:347–53. doi: 10.1080/10245332.2016.1276240

43. Nguyen CH, Glüxam T, Schlerka A, Bauer K, Grandits AM, Hackl H, et al. SOCS2 is part of a highly prognostic 4-gene signature in AML and promotes disease aggressiveness. *Sci Rep*. (2019) 9:9139. doi: 10.1038/s41598-019-45579-0

44. Xu Z-J, Gu Y, Wang C-Z, Jin Y, Wen X-M, Ma J-C, et al. The M2 macrophage marker CD206: a novel prognostic indicator for acute myeloid leukemia. *OncoImmunology*. (2020) 9:1683347. doi: 10.1080/2162402X.2019.1683347

45. Ding Y, He D, Florentin D, Frolov A, Hilsenbeck S, Ittmann M, et al. Semaphorin 4F as a critical regulator of neuro-epithelial interactions and a biomarker of aggressive prostate cancer. *Clin Cancer Res*. (2013) 19:6101–11. doi: 10.1158/1078-0432.CCR-12-3669

46. Peng Y, Song Y, Wang H. Systematic elucidation of the aneuploidy landscape and identification of aneuploidy driver genes in prostate cancer. *Front Cell Dev Biol*. (2021) 9:723466. doi: 10.3389/fcell.2021.723466

47. Chen F, Licarete E, Wu X, Petrusca D, Maguire C, Jacobsen M, et al. Pharmacological inhibition of Carbonic Anhydrase IX and XII to enhance targeting of acute myeloid leukaemia cells under hypoxic conditions. *J Cell Mol Med*. (2021) 25:11039–52. doi: 10.1111/jcmm.17027

48. Smith BAH, Deutzmann A, Correa KM, Delaveris CS, Dhanasekaran R, Dove CG, et al. MYC-driven synthesis of Siglec ligands is a glycoimmune checkpoint. *Proc Natl Acad Sci USA*. (2023) 120:e2215376120. doi: 10.1073/pnas.2215376120

49. Zhao Q, Zhang K, Li Y, Ren Y, Shi J, Gu Y, et al. OLFML2A is necessary for anti-triple negative breast cancer effect of selective activator protein-1 inhibitor T-5224. *Transl Oncol*. (2021) 14:101100. doi: 10.1016/j.tranon.2021.101100

50. Feng S, Agoulnik IU, Bogatcheva NV, Kamat AA, Kwabi-Addo B, Li R, et al. Relaxin promotes prostate cancer progression. *Clin Cancer Res*. (2007) 13:1695–702. doi: 10.1158/1078-0432.CCR-06-2492

51. Mrózek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. *Blood Rev*. (2004) 18:115–36. doi: 10.1016/S0268-960X(03)00040-7

52. Rowley JD. Chromosomal translocations: revisited yet again. *Blood*. (2008) 112:2183–9. doi: 10.1182/blood-2008-04-097931

53. Venugopal S, DiNardo CD, Loghavi S, Qiao W, Ravandi F, Konopleva M, et al. Differential prognostic impact of RUNX1 mutations according to frontline therapy in patients with acute myeloid leukemia. *Am J Hematol*. (2022) 97:1560–7. doi: 10.1002/ajh.26724

54. Tefferi A, Barbui T. Polycythemia vera: 2024 update on diagnosis, risk-stratification, and management. *Am J Hematol*. (2023) 98:1465–87. doi: 10.1002/ajh.27002

55. Zhang S-J, Rampal R, Manshouri T, Patel J, Mensah N, Kayserian A, et al. Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. *Blood*. (2012) 119:4480–5. doi: 10.1182/blood-2011-11-390252

56. Stahl M, Derkach A, Farnoud N, Bewersdorf JP, Robinson T, Famulare C, et al. Molecular predictors of immunophenotypic measurable residual disease clearance in acute myeloid leukemia. *Am J Hematol*. (2023) 98:79–89. doi: 10.1002/ajh.26757

57. Waldman T. Emerging themes in cohesin cancer biology. *Nat Rev Cancer*. (2020) 20:504–15. doi: 10.1038/s41568-020-0270-1

58. Eckardt J-N, Stasik S, Röllig C, Sauer T, Scholl S, Hochhaus A, et al. Alterations of cohesin complex genes in acute myeloid leukemia: differential co-mutations, clinical presentation and impact on outcome. *Blood Cancer J*. (2023) 13:18. doi: 10.1038/s41408-023-00790-1

59. Thol F, Bollin R, Gehlhaar M, Walter C, Dugas M, Suchanek KJ, et al. Mutations in the cohesin complex in acute myeloid leukemia: clinical and prognostic implications. *Blood*. (2014) 123:914–20. doi: 10.1182/blood-2013-07-518746

60. Ng SWK, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. (2016) 540:433–7. doi: 10.1038/nature20598

61. Chen Z, Song J, Wang W, Bai J, Zhang Y, Shi J, et al. A novel 4-mRNA signature predicts the overall survival in acute myeloid leukemia. *Am J Hematol*. (2021) 96:1385–95. doi: 10.1002/ajh.26309

62. Jin P, Jin Q, Wang X, Zhao M, Dong F, Jiang G, et al. Large-scale *in vitro* and *in vivo* CRISPR-cas9 knockout screens identify a 16-gene fitness score for improved risk assessment in acute myeloid leukemia. *Clin Cancer Res*. (2022) 28:4033–44. doi: 10.1158/1078-0432.CCR-22-1618

63. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med*. (2003) 348:203–13. doi: 10.1056/NEJMoa020177

64. Pagès F, Kirilovsky A, Mlecnik B, Asslaber M, Tosolini M, Bindea G, et al. *In situ* cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *J Clin Oncol*. (2009) 27:5944–51. doi: 10.1200/JCO.2008.19.6147

65. Halama N, Michel S, Kloor M, Zoernig I, Benner A, Spille A, et al. Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer Res*. (2011) 71:5670–7. doi: 10.1158/0008-5472.CAN-11-0268