RESEARCH ARTICLE

Blood Science

OPEN

Transcriptomics-based multi-omics approach for optimizing risk stratification in acute myeloid leukemia

Yang Yang[a,b,c], Lun Yan[a,b,c], Jianjun Fang[a,b,c], Hong Liu[a,b,c], Xu Tan[a,b,c], Yingying Ma[a,b,c], Xiao Han[a,b,c], Guo Chen[a,b,c], Ying Chen[a,b,c], Desheng Gong[d], Shichun Tu[d], Qin Wen[a,b,c], Jing Li[a,b,c,*], Xi Zhang[a,b,c,*], Cheng Zhang[a,b,c,*]

[a]Medical Center of Hematology, Xinqiao Hospital of Army Medical University, Chongqing, 400037 China; [b]State Key Laboratory of Trauma and Chemical Poisoning, Chongqing Key Laboratory of Hematology and Microenvironment, Chongqing, 400037 China; [c]Jinfeng Laboratory, Chongqing, 401329 China; [d]Clinical Laboratories, Shenyou Bio, Zhengzhou, China

Abstract

Acute myeloid leukemia (AML) is a subtype of hematopoietic neoplasm affecting myeloid cells in blood and bone marrow. Molecular subtyping of AML provides in-depth insights into disease pathogenesis, facilitating clinical diagnostics and therapeutic management. While current molecular screening is mainly established on DNA-based approaches, including targeted sequencing of gene panels and polymerase chain reaction (PCR)-based fusion gene detection, transcriptome-based RNA-seq has gained increasing attention for its potential in improving molecular subtyping as well as in facilitating prognostic prediction and therapeutic selection. In this study, we aimed to identify transcriptomic signatures in molecular genotyping and risk stratification in a cohort of 125 patients with AML excluding those harboring *PML::RARA* fusion genes. We subsequently established a 2-step diagnostic assay based on transcriptomic expression, which can be further explored for predicting the outcome of AML patients with increased confidence and accuracy. Together with the identification of genetic mutations and fusion genes, transcriptome-based stratification might significantly enhance the ability of outcome prediction. In summary, we developed a transcriptomics-based multi-omics prediction model for improved risk stratification of patients with AML.

Key Words: Acute myeloid leukemia; Olfactory receptors; RNA sequencing

1. INTRODUCTION

Acute myeloid leukemia (AML) is a type of hematopoietic neoplasm affecting myeloid cells and characterized by heterogeneity in both pathogenetic development and therapeutic responses.[1–4] The newest classification of AML (the 5th edition) was updated in 2022 by the World Health Organization (WHO) with a trend showing more emphasis on scalable genetic abnormalities than on traditional blast enumeration based on recent breakthroughs in how the disease is understood and managed.[5] The genetic risk classification by the European Leukemia Net (ELN) was also upgraded in the same year stratifying AML patients into favorable, intermediate, and adverse groups.[6,7] However, these classifications for some endpoints fail to distinguish patients in therapeutic outcomes between different genetic subgroups, although different genetic features, including tumor-associated single-nucleotide variants (SNVs), small insertion/deletions, and fusion genes, are clearly manifested. Therefore, alternative approaches are required to classify patients by prognostic and therapeutic value rather than just molecular characteristics, to guide treatment selection and outcome prediction.[8,9]

Current strategies for precision medicine are largely relied on genomic underpinnings for AML, including cytogenetic testing for structural genomic abnormalities and DNA-based targeted sequencing for prognostically and therapeutically applicable genetic variants.[10–13] However, the genome-phenome relationship can be complicated by biological regulation at the level of transcription and beyond, and subsequently compromise the process of mechanistic uncovering and therapeutic development.[14–18] While DNA-based next-generation sequencing (NGS) of small gene panels is typically used for screening disease management-relevant genetic abnormalities, offering great benefits for diagnostic and therapeutic purposes, RNA-based NGS assays may provide great diagnostic returns by providing transcriptome expression information and by screening for SNVs, short insertion/deletion variants, gene fusions, and more for disease management.[19,20] In a study, the transcriptome repertoire was evaluated in a large cohort of AML patients resulting in 8 molecular subgroups with both

*Address correspondence: Jing Li, Xi Zhang, Cheng Zhang, Medical Center of Hematology, Xinqiao Hospital of Army Medical University, Chongqing, 400037 China. E-mail address: 519054661@qq.com (J. Li), zhangxi@tmmu.edu.cn (X. Zhang), chzhang2014@tmmu.edu.cn (C. Zhang).

Y.Y., L.Y., and J.F. contributed equally to this work.

Noncommunicable Chronic Diseases-National Science and Technology Major Project (2025ZD0545804). This study was supported by grants from the National Natural Science Foundation (No. 82170212, 82370181). The Second Affiliated Hospital of the Army Medical University is a special project for clinical research (No. 2024F010).

No datasets were generated or analyzed during the current study.

Blood Science (2026) 8, 1–14:e00295.

Received January 27, 2026; Accepted April 7, 2026.

http://dx.doi.org/10.1097/BS9.0000000000000295

novel and redefined subtypes.[21] This stratification system exhibits robust clinical features, genetic lesions, and developmental hierarchies, especially the complexity of molecular interplay in this disease, and provides further prognostic and therapeutic value in complementary to current well-established genome-based classification frameworks. Interestingly, it has also been proposed that RNA-seq may be employed as a standalone assay for diagnostics in cancer, as shown in a recent work, which employed an innovative strategy to improve the performance of the 2022 ELN risk stratification in patients with favorable outcomes to overcome its limitations to further distinguish these patients with disparate outcomes over time.[22] Furthermore, it has also been demonstrated in a separated study that the authors, using RNA-seq via a customized algorithm, have identified 7 genes of that the expression levels were strongly associated with clinical outcomes and further developed a transcriptome-based score to distinguish patients with distinct gene expression patterns and prognosis.[23] Therefore, elucidating transcription-disease ontology may hold promise for better disease characterization and therapeutic development.

To further investigate the potential of RNA-seq assay for AML, we performed multi-omics tests, including DNA-sequencing of small gene panels, reverse transcription quantitative polymerase chain reaction (RT-qPCR)–based detection of fusion genes, and RNA-seq in a cohort of 125 patients with AML (excluding *PML::RARA* positive). While genomic alterations showed some prognostic association, transcriptome-based classification demonstrated a stronger correlation with treatment outcomes. We developed a 2-step assay based on identified transcriptomic signatures to facilitate risk stratification. This approach may provide a superior tool for prognostic prediction and therapeutic guidance, yet its clinical utility requires further validation in larger cohorts.

2. MATERIALS AND METHODS

2.1. Patients and samples

A total of 125 patients diagnosed with primary AML excluding those carrying *PML::RARA* fusion genes (M3 subtype) were enrolled in this study from February 2021 to October 2022. All patients were from the Xinqiao Hospital affiliated with the Army Medical University (China) and treated with the conventional chemotherapy regimen recommended by the 2017 Chinese Guidelines for the Diagnosis and Treatment of Acute Myeloid Leukemia. This study was approved by the Ethics Committee of the Xinqiao Hospital affiliated with the Army Medical University (China). Informed consent for both treatment and sample collection was given to all participating patients according to the Declaration of Helsinki.

Both bone marrows and blood samples were collected from all participating patients for molecular screening. Sample inclusion criteria include newly diagnosed AML (non-M3 type) patients confirmed by cell morphology and/or genetics; availability of blood samples from patients; informed consent was voluntarily signed by patients or their families. Sample exclusion criteria include non-newly diagnosed AML patients; M3 subtype of AML confirmed by cell morphology and/or genetics; patients diagnosed with other malignant tumors requiring corresponding treatments; and any other factors that researchers believe to be not suitable for participating in this study. Two to 3 mL of either blood or bone marrow samples were collected and mixed well in an ethylenediaminetetraacetic acid (EDTA) anticoagulant tubes, which are ready for transportation at 4°C. On arrival, samples underwent red blood cell lysis and leukocyte precipitation. Genomic DNA (gDNA) and total RNA were then co-extracted using the TIANGEN Blood Genomic DNA/RNA Extraction Kit (Tiangen, Beijing, China). The extracted DNA/RNA samples were immediately used for further analysis or stored at −20°C or −80°C for longer periods of time.

2.2. Targeted sequencing of genetic mutations

The above gDNA were subjected to targeted sequencing using a NGS panel for hematological malignancies containing either 38 genes for amplicon-based sequencing (Rightongene, Shanghai, China) or 129 genes hybridization-capture–based sequencing (Shenyou Bio, Zhengzhou, China).

The detailed procedures were described elsewhere.[24,25] In brief, for amplicon-based sequencing, the NGS libraries were constructed using multiplex amplification and paired-end sequenced (PE150) on a DNBSEQ-T7RS sequencer. For capture-based sequencing, purified DNA samples were fragmentated (150–200 bp) and hybridized using probes for the 129 gene panel to build a library for sequencing. The quality of raw sequencing data was evaluated with FastQC (v0.11.8). Adaptor sequences and low-quality bases (Phred <Q20) were filtered out using Trimomatic (v0.39). The remaining high-quality sequencing data were then aligned to the reference genome (Genome Reference Consortium Human Reference 37, GRCh37) using BWA (v0.7.17-r1188). Mutations were detected using VarDict (v1.7.0) focusing on SNVs and the detecting results were output in Variant Call Format (VCF) file format for subsequent analysis and annotation.

The genes included in the 38 gene panel are *ASXL1*, *BCOR*, *BCORL1*, *CBL*, *CEBPA*, *CSF3R*, *DNMT3A*, *ETV6*, *EV11*, *EZH2*, *FLT3*, *FLT3-ITD*, *FLT3-TKD*, *GATA2*, *HOX11*, *IDH1*, *IDH2*, *JAK2*, *KIT*, *KMT2A*, *KMT2A-PTD*, *KRAS*, *MECOM*, *MPL*, *MYC*, *NF1*, *NPM1*, *NRAS*, *NTRK3*, *RUNX1*, *SF3B1*, *SH2B3*, *SRSF2*, *TET2*, *TP53*, *TPMT*, *U2AF1*, and *ZRSR2*. And the genes included in the 129 gene panel are *ABCB1*, *ABL1*, *ANKRD26*, *ASXL1*, *ASXL2*, *ATG2B*, *ATRX*, *BCL2*, *BCOR*, *BCORL1*, *BIRC3*, *BRAF*, *BTK*, *CALR*, *CARD11*, *CBL*, *CCND1*, *CCND2*, *CD79B*, *CDKN2A*, *CDKN2B*, *CEBPA*, *CREBBP*, *CRLF2*, *CSF1R*, *CSF3R*, *CUX1*, *CXCR4*, *CYP2C19*, *CYP3A5*, *DDX41*, *DHX15*, *DKC1*, *DNM2*, *DNMT3A*, *EBF1*, *EGR2*, *ELANE*, *EP300*, *EPOR*, *ERG*, *ETNK1*, *ETV6*, *EZH2*, *FAT1*, *FAT3*, *FBXW7*, *FGFR1*, *FLT3*, *GATA1*, *GATA2*, *GATA3*, *GFI1*, *GSKIP*, *GSTP1*, *HAX1*, *IDH1*, *IDH2*, *IKZF1*, *IL7R*, *JAK1*, *JAK2*, *JAK3*, *KDM6A*, *KIT*, *KMT2A*, *KMT2C*, *KMT2D*, *KRAS*, *MBD4*, *MECOM*, *MEF2B*, *SAMD9L*, *MPL*, *MTHFR*, *MTOR*, *MYD88*, *NF1*, *NFKBIE*, *NOTCH1*, *NOTCH2*, *NPM1*, *NQO1*, *NRAS*, *NT5C2*, *PAX5*, *PDGFRA*, *PDGFRB*, *PHF6*, *PIGA*, *PIK3CA*, *PIK3CD*, *PLCG2*, *PML*, *PPM1D*, *PRPF8*, *PTEN*, *PTPN11*, *RAD21*, *RARA*, *RB1*, *RHOA*, *ROBO1*, *ROBO2*, *RUNX1*, *SAMD9*, *SETBP1*, *SETD2*, *SF1*, *SF3B1*, *SH2B3*, *SMC1A*, *SMC3*, *SRP72*, *SRSF2*, *STAG2*, *STAT3*, *STAT5B*, *TERC*, *TERT*, *TET2*, *TNFAIP3*, *TP53*, *TPMT*, *TTN*, *U2AF1L5*, *WT1*, *ZBTB7A*, and *ZRSR2*.

2.3. Detection of fusion genes by RT-qPCR

Detection of fusion genes was performed through RT-qPCR as described previously.[24,25] In brief, the prepared total RNA was reverse-transcribed, and the resulting cDNA was stored at 4°C or used immediately for RT-qPCR on a Fluorescence Quantitative Analyzer (ABI 7500). The experiment was performed according to the manufacturer's instruction using a Multiplex RT-PCR Fusion Gene Kits (Rightongene, Shanghai, China). The fusion gene panel contains 43 fusion genes: *BCR-ABL*, *AML1-ETO*, *PML-RARα*, *CBFβ-MYH11*, *MLL-AF4*, *MLL-AF9*, *E2A-PBX1*, *TEL-AML1*, *DEK-CAN*, *NPM-MLF1*, *E2A-HLF*, *SIL-TAL1*, *TEL-ABL*, *TEL-JAK2*, *AML1-MDS1/EVI1/MTG16*, *MLL-AF6/AF10/ELL/ENL*, *MLL-AF17/AF1q/AF1p/AFX/SEPT6*, *FIP1L1-PDGFRα*, *ETV6-PDGFRα*, *TEL-PDGFRβ*, *PLZF/STAT5b-RARα*, *F1P1L1/PRKR1A/NUMA/NPM-RARα*, *NUP98-HoxA13/HoxC11/HoxD13/HoxA9/HoxA11/PMX1*, *NPM-ALK*, *SET-CAN*, *TLS-ERG*.

2.4. Transcriptomic sequencing and data processing

High-quality total RNA extracted from blood or bone marrow as described in the previous section was used for the construction of

the transcriptome library. rRNA was removed from the total RNA using the Ribo off rRNA Depletion kit (Vazyme # N406, Nanjing, China) according to the manufacturer's instruction. Double-strand cDNA was synthesized using the Universal V8 RNA seq Library Prep kit (Vazyme # NRM605) and then connected with cDNA adapters using the Dual UMI ADB Adapter Set 1 for MGI (Vazyme # NM35101). After purification with VAHTS DNA Clean Beads (Vazyme # N411) with 0.45X magnetic beads, the cDNA library was amplified through PCR using the reaction reagents from the above library Prep kit. A pre-library was obtained after the PCR products were purified with 0.9X magnetic beads and subjected to the construction of a single-stranded circular DNA library, using the Universal Circulation kit (NM202), specifically designed for MGI high-throughput sequencing. After passing the quality inspection, high-throughput sequencing was performed using DNBSEQ-T7 sequencer (MGI Tech, Shenzhen, China) to generate raw offline data (fastq format file), resulting in a data volume of 12Gb per sample. We used FastQC (version 0.11.8) to evaluate the quality of the raw data, and then Trimomatic (version 0.39) for quality control processing (default parameters) to remove low-quality reads, adapter sequences, and reads containing unknown bases. Based on HISAT2 (version 2.2.1.0) software, the filtered data were aligned to the hg38 reference genome (default parameters), and the resulting file (BAM format) was processed using Samtools (version 1.9). The expression level of each transcript was calculated with StringTie (version 2.2.1) software. The differential gene expression was analyzed using the DESeq2 package (version 1.26.0) of the R language (version 3.6.2), with the following criteria considered as significant: p value <0.050; the expression difference is greater than 2 times. RNA-seq data of 30 normal human samples were downloaded from the public database GTEx (Prototype Issue Expression) (https://gtexportal.org/home/) and served as negative controls.

2.5. Unsupervised hierarchical clustering of gene expression

Unsupervised hierarchical clustering was conducted on 6303 genes that were differentially expressed ($p < 0.050$ and |log2FoldChange| >1) between AML patients and normal controls. This classification was developed based on early studies as described elsewhere with modification.[26,27] The cloud tool Hiplot (https://hiplot.com.cn/cloud-tool/drawing-tool/detail/106) was used to calculate the Canberra distance between samples based on gene expression levels, and hierarchical clustering was performed using ward.D2 algorithm with a tree-cut height of 0.6. Through this analysis, the samples were divided into different expression subtypes and heatmaps were drawn to visually display the clustering results.

2.6. Enrichment analysis of gene function and pathways

To further explore the biological functions of characteristic genes and reveal the biological pathways they participate in, we utilized the online tool DAVID (https://davidbioinformatics.nih.gov/) for gene function enrichment analysis.[28] The analysis includes Gene Ontology (GO) functional enrichment analysis, covering biological processes (BP) and molecular functions (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Subsequently, the analysis results were visualized using the ggplot2 package (version 3.5.1) in R language (version 3.6.2) to demonstrate the significant enrichment of gene functions.

2.7. Building a regression analysis model for risk stratification

This study used the R language (version 3.6.2) to construct a regression model for predicting treatment outcome as described previously.[29] First, the survival package (version 3.8.3) for Cox multivariate regression analysis was used to screen variables significantly correlated with prognosis: variables with a univariate Cox regression p value <0.100 were included in the multivariate Cox regression model, and a p value <0.050 in multivariate analysis was considered statistically significant for prognostic correlation, forming an initial prognostic prediction model. Second, the Lasso regression analysis of the glmnet package (version 4.1.8) was then utilized to reduce the dimensionality of variables and optimize the model. The optimal λ value was selected based on 10-fold cross-validation, with the λ.min (the λ value giving the minimum cross-validation error) chosen as the penalty parameter to eliminate genes with non-contributory prognostic value (coefficient = 0). Third, the pROC package (version 1.18.5) was subsequently used to calculate the area under the curve (AUC) value of the model to evaluate the performance of the regression model. Finally, an receiver operating characteristic (ROC) curve was drawn based on the plot function to visually demonstrate the predictive ability of the model.

2.8. Statistical analysis

Statistical analyses were conducted using the R language (version 3.6.2) unless otherwise indicated. $p < 0.050$ was considered as statistically significant unless otherwise indicated.

3. RESULTS

3.1. Genetic clustering based on transcriptomic sequencing in *PML-RARA*-neg AML patients

Main clinical characteristics of 125 participating patients are provided in Table 1. To better determine the biological effectors combining both genomic mutations and biological processes in

Table 1

Characteristics of participating patients with AML in this study.

Types	Cases (%)
Gender	
Female	53 (42.4)
Male	72 (57.6)
Age (y)	
≥60	25 (20)
<60	100 (80)
FAB Category	
M1	18 (14.4)
M2	41 (32.8)
M4	29 (23.2)
M5	23 (18.4)
M7	3 (2.4)
Not classified	11 (8.8)
Treatment	
Ara-C	64 (51.2)
Ara-C + AZA + Venclexta	6 (4.8)
AZA	13 (10.4)
AZA + IA	2 (1.6)
HHT	5 (4)
IA	12 (9.6)
AZA + HAG + Venclexta	2 (1.6)
Unknown	21 (16.8)
ELN classification	
Favorable	37 (29.6)
Intermediate	40 (32)
Adverse	48 (38.4)
Sample type	
Bone marrow	93 (74.4)
Peripheral blood	32 (25.6)

AML = acute myeloid leukemia, Ara-C = cytarabine, AZA = azacitidine, ELN = European Leukemia Net, FAB = French-American-British classification systems, HAG = homoharringtonine + cytarabine + granulocyte colony-stimulating factor, HHT = homoharringtonine, IA = idarubicin + cytarabine.

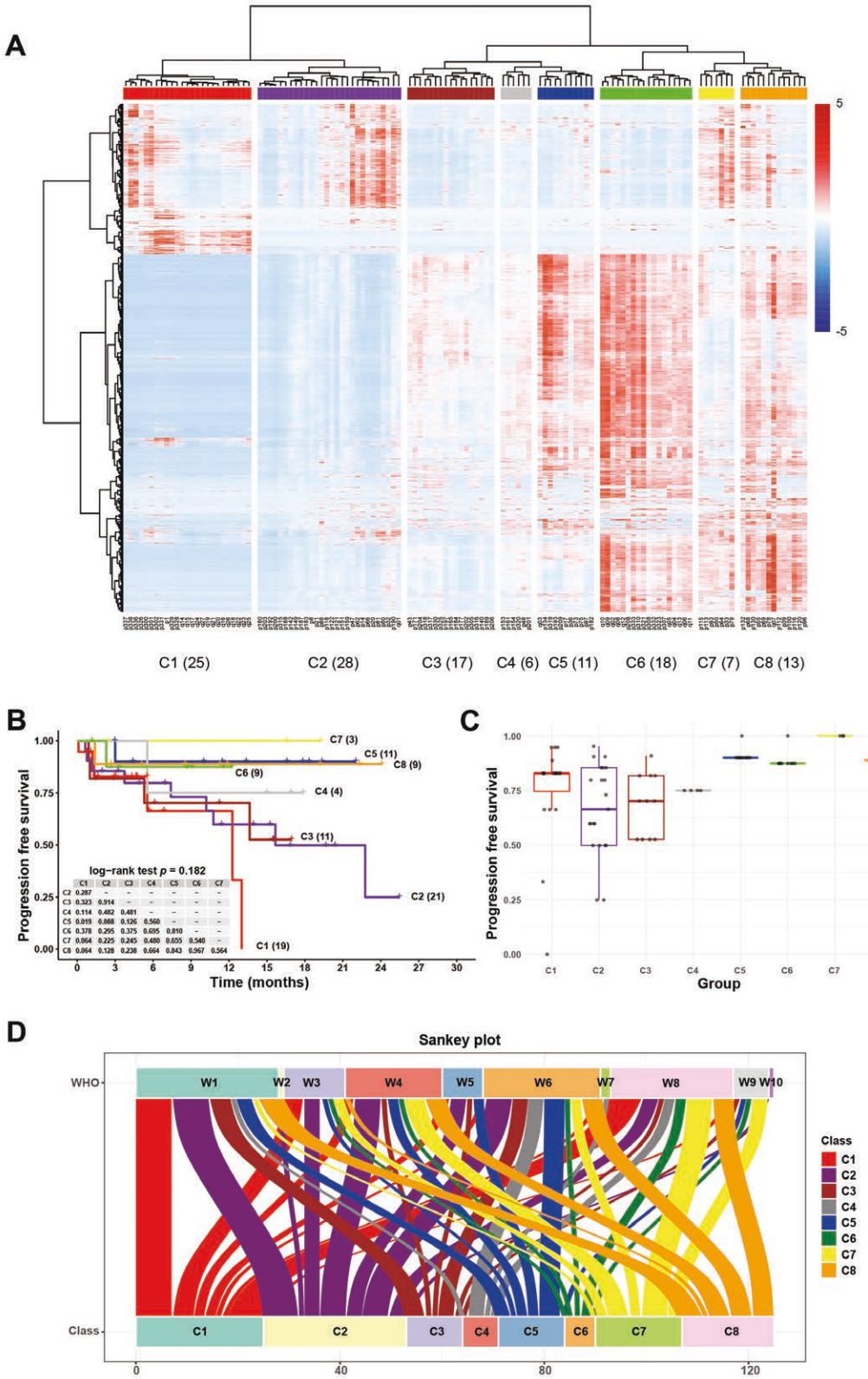

Figure 1. Molecular classification of AML based on unsupervised hierarchical clustering of gene expression. (A) Heatmap of gene expression classified by unsupervised hierarchical clustering on 6303 differentially expressed genes between AML patients and normal controls. These patients were classified into 8 clusters (C1–C8) labeled with different colors on the top of the map. Patients and patient clusters are shown at the bottom. (B) Survival curve of 87 AML patients classified into 8 clusters. Log-rank tests were performed by pairwise comparison and resulting *p* values are listed in the table at the bottom left corner. (C) Scatter plot of PFS in C1 to C8 clusters of patients. (D) Sankey plot showing the correlation between patients classified by our approach (C1–C8 clusters) and by the 2022 edition of WHO classification: W1 (AML-MR), W2 (AML with *BCR::ABL1* fusion), W3 (AML with *CBFB::MYH11* fusion), W4 (AML with *CEBPA* mutation), W5 (AML with *KMT2A* rearrangement), W6 (AML with *NPM1* mutation), W7 (AML with *NUP98* rearrangement), W8 (AML with defined by differentiation), W9 (AML with *RUNX1::RUNX1T1* fusion), W10 (AML with other defined genetic alterations). AML = acute myeloid leukemia, PFS = progression-free survival, WHO = World Health Organization.

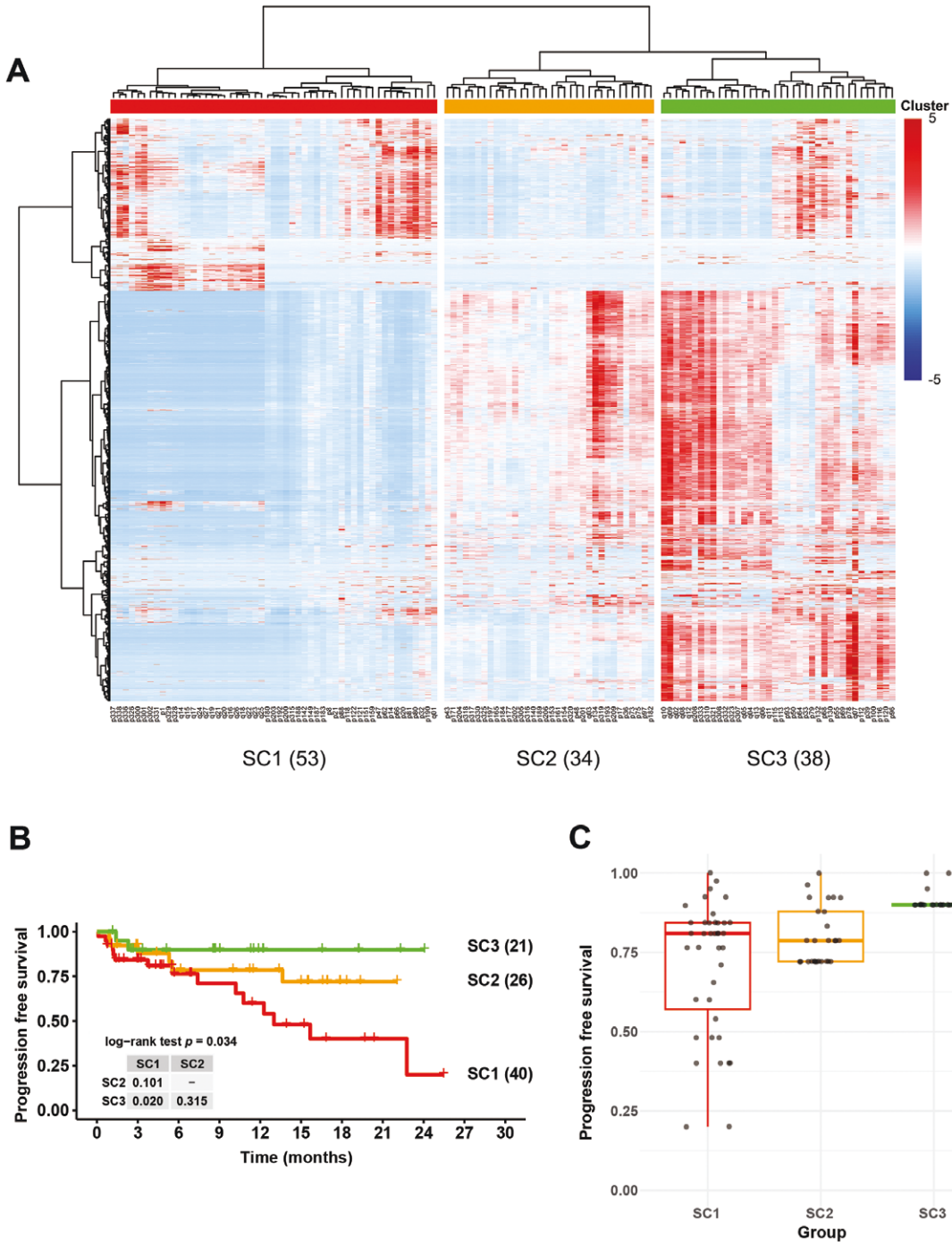

Figure 2. Eight clusters can be regrouped based on their position in dendrography showing similar prognostic characteristics. (A) Three superclusters (SC1–SC3) show distinctive patterns of gene expression and prognostic characteristics. SC1–SC3 were indicated by different colors on the top of the graph. (B) Survival curve of 3 superclusters of patients. Log-rank tests indicated significant differences observed in PFS within 3 superclusters ($p = 0.034$). (C) Scatter plot showing distinct patterns in prognosis among 3 superclusters. PFS = progression-free survival.

responding to both intrinsic and environmental alternations, we performed DNA-based NGS for genetic mutations, RT-PCR–based techniques for fusion genes, and RNA-seq for transcriptome expression in 125 *PML-RARA*-neg AML patients to identify potential risk factors associated with cancer. Based on RNA-seq data, we first identified 6303 genes that show a significant difference ($p < 0.050$ and ratio ≥ 2) in expression in 125 AML patients compared to normal controls. By hierarchical clustering of these 6303 genes, these patients can be classified into 8 clusters: C1–C8 (**Fig. 1A**). Among these patients, 87 of them have prognostic data, showing distinct prognostic characteristics in different clusters (**Fig. 1B**). Interestingly, patients in the clusters on the left half show significantly worse survival (progression-free survival [PFS]), while those on the right show better survival (**Fig. 1C**), suggesting that our RNA-based classification show better association with disease prognosis compared with classification by the newest WHO criteria (**Fig. 1D**).

3.2. Transcriptome-based risk stratification

For prognostic prediction, we grouped the C1 to C8 clusters into 3 superclusters (SC1–SC3): SC1 (C1–C2), SC2 (C3–C5),

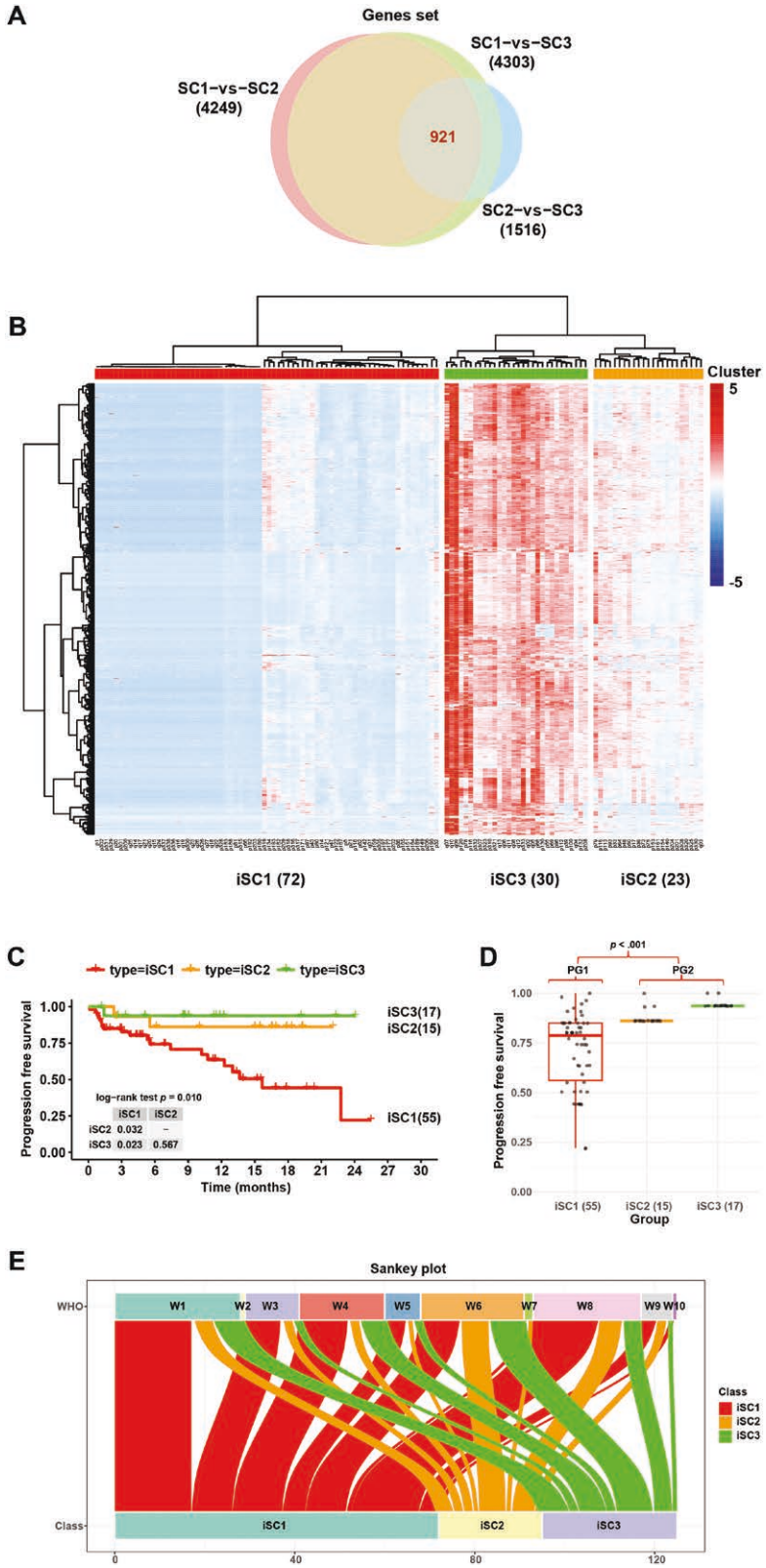

Figure 3. Improved prognostic classification on genes differentially expressed between any 2 superclusters as defined in Figure 2. (A) Pairwise comparison among 3 superclusters from Figure 2 resulted in 921 genes. (B) Three iSC (iSC1–iSC3) showing even more striking patterns in gene expression. (C) Survival curve of 3 iSCs of patients. Log-rank tests indicated significant differences in prognosis within these iSCs (p = 0.010). (D) Scatter plot showing prognostic patterns in 3 iSCs. Due to the similarity in prognosis, iSC2 and iSC3 can be combined as a new group (PG2), while iSC1 is renamed as PG1. The difference between PG1 and PG2 in PFS was highly significant (p < .001). (E) Sankey plot indicates the relationship between the defined 3 molecular subgroups (iSC1–iSC3) and disease entities defined by WHO classification. iSC = improved superclusters, PFS = progression-free survival, PG = prognostic group, WHO = World Health Organization.

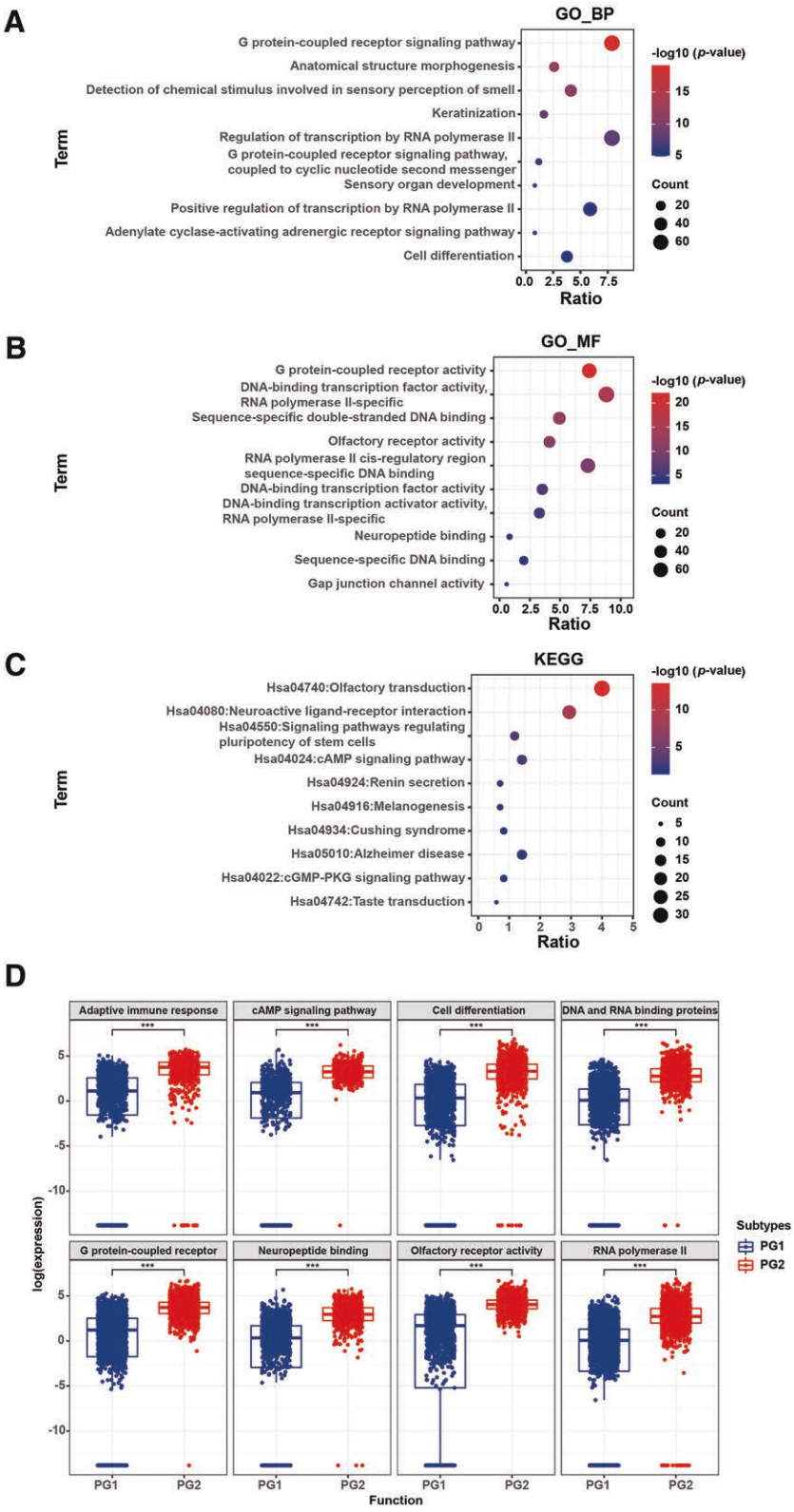

Figure 4. Functional enrichment analysis on 921 characteristic genes. (A–C) Eight top hits are shown by each of the following enrichment analysis: (A) Biological process functional enrichment analysis. (B) Molecular functional enrichment analysis. (C) KEGG pathway enrichment analysis. (D) Comparison of differential expressions of significantly enriched functional genes between the PG1 and PG2 groups. Eight top hit functional groups are presented in the box plots. Among them, blue: PG1 samples; red: PG2 samples. Statistical testing was conducted using *t* test, with the notation "***" indicating significant differences ($p < .001$). BP = biological processes, GO = Gene Ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, MF = molecular functions, PG = prognostic group.

SC3 (C6–C8) (**Fig. 2A**). Patients in these 3 SCs are distinctively different in prognosis (log-rank test p = 0.034, **Fig. 2B–C**). Specifically, the patients in SC1 show significantly worse PFS values than those in SC3 (p = 0.020), while a trend to be significant was observed in PFS between patients in SC1 and SC2.

To improve the accuracy of prognostic prediction of this approach, we extracted 921 genes out of the 6303 genes that were differentially expressed between any 2 SCs (ratio ≥2, p < .001) by performing pairwise comparison among these 3 SCs (**Fig. 3A**; Tables S1.1–1.3 and S2, https://links.lww.com/BS/A163). Based on the expression pattern of newly elected 921 genes, we reclassified these patients into 3 improved SCs (iSC1–iSC3) showing even more striking differences in the patterns of RNA expression (**Fig. 3B**) and prognosis (log-rank test p = 0.010, **Fig. 3C**). The PFS in iSC1 is significantly worse than those in both iSC2 and iSC3 (p = 0.032 and 0.023), respectively (**Fig. 3C**). As shown in scattergram of PFS (**Fig. 3D**), patients in cluster 2 and cluster 3 usually show higher values in PFS (both >75%), while patients in cluster 1 show mixed prognostic values. For the purpose of prognostics, we subjectively re-group these patients into 2 prognostic groups (PG1 = iSC1; PG2 = iSC2 + iSC3), and found that the PFS values between these 2 prognostic groups were significantly high (p < .001). The expression of 921 genes was also significantly different between PG1 and PG2 groups (Table S2, https://links.lww.com/BS/A163). By the Sankey plot, a similar comparison pattern was observed when compared to the WHO-defined classification in the aspect of PFS (**Fig. 3E**), consistent with the notion that patients in each WHO-classified group may have variable prognostic values.

3.3. Molecular pathways characterization by GO and KEGG enrichment analysis

We performed GO and KEGG enrichment analysis to uncover the molecular pathways associated with tumorigenesis and prognosis in this cohort of patients. Top hits include G protein–coupled receptor signaling pathway (GPCRs), olfactory receptor activity (ORs), and other sensory-related pathways, DNA binding, RNA polymerase II–associated transcription, cell differentiation, and among others (**Fig. 4A–C**; Tables S3.1–3.3, https://links.lww.com/BS/A163), which have been suggested to play significant roles or as potential biomarkers in cancer. For instance, GPCRSs have been shown to play a critical role in tumor invasion and metastasis,[30,31] while the expression specificity has been identified as potential biomarkers associated with cancer progression, including OR51E1 in small intestine neuroendocrine carcinomas, OR7C1 in colorectal cancer, PSGR in prostate cancer, and OR2B6 in breast carcinomas.[32] Meanwhile, pathways associated with DNA and RNA binding as well as RNA polymerase II can regulate numerous cellular processes contributing to cancer development and progression.[33,34]

Moreover, pathways regulating cell differentiation are likely involved in metabolism, angiogenesis, and immunity to influence the interaction between cancer cells and immune system and thus treatment outcomes.[35] It has also been demonstrated that cyclic adenosine monophosphate (cAMP) signaling pathways confer both tumor suppressive and tumor-promoting roles depending on tumor types and context,[36] while adaptive immune response can either promote or suppress cancer development and progression via regulating the immune reaction to cancer.[37] Finally, we demonstrated that the expression patterns of 8 significantly enriched functional genes show a significant difference between PG1 and PG2, indicating a likely association between these pathways and prognosis (**Fig. 4D**).

3.4. Characteristics of genetic mutations and fusion genes in *PML-RARA*-neg AML patients

Genetic mutations were determined via both DNA-based targeted sequencing and RNA-seq as described in Materials and Methods. Genetic mutations were detected in all participated patients (87 patients with follow-up information) (Table S4, https://links.lww.com/BS/A163). In these patients, there are 8.08 mutations (ranging 2–16) per patient on average. The top

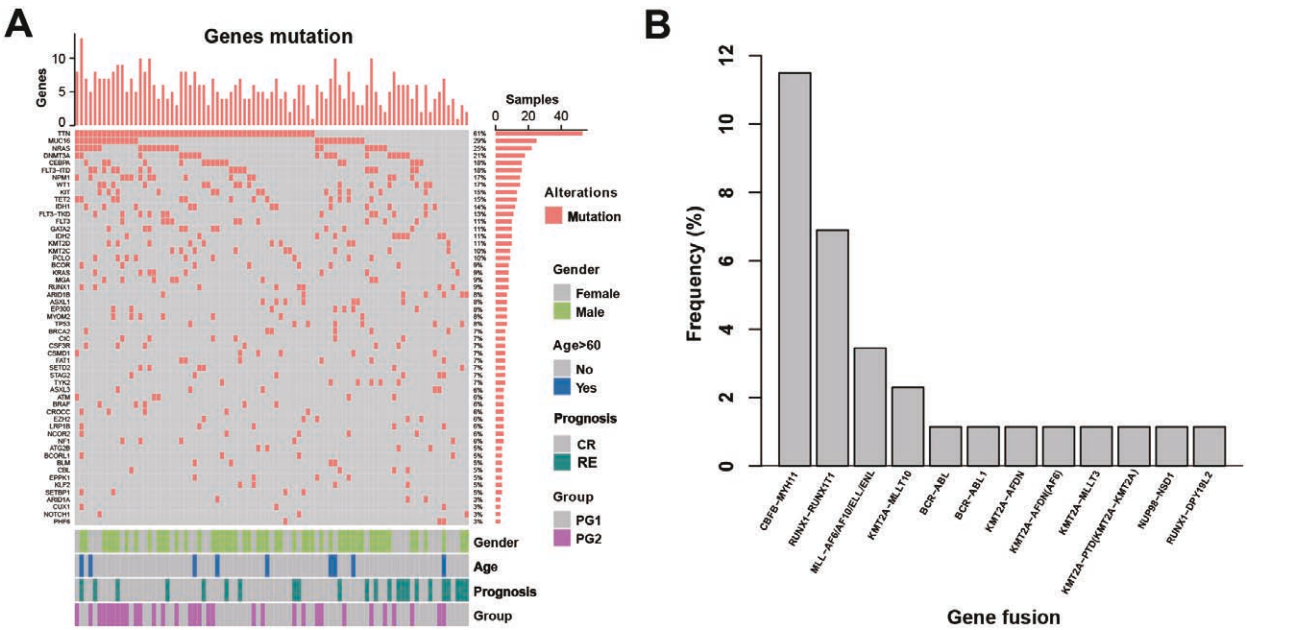

Figure 5. Genetic mutations and fusion genes detected in AML patients. (A) Heatmap showing mutations and clinical information in 87 AML patients with prognostic information. Only genes with mutation rate exceeding 5% are included in the map. Gene names are listed on the left. The upper bar displays the number of mutated genes in each sample. The number of patients and occurring frequency for each mutated gene are shown at the right. The information about the gender, age, prognosis (CR, RE), and group (PG1, PG2) of the patients are presented at the bottom. (B) Bar graph showing the frequency of fusion genes in AML patients. AML = acute myeloid leukemia, CR = complete remission, PG1 = prognostic group 1, PG2 = prognostic group 2, PFS = progression-free survival, RE = recurrence.

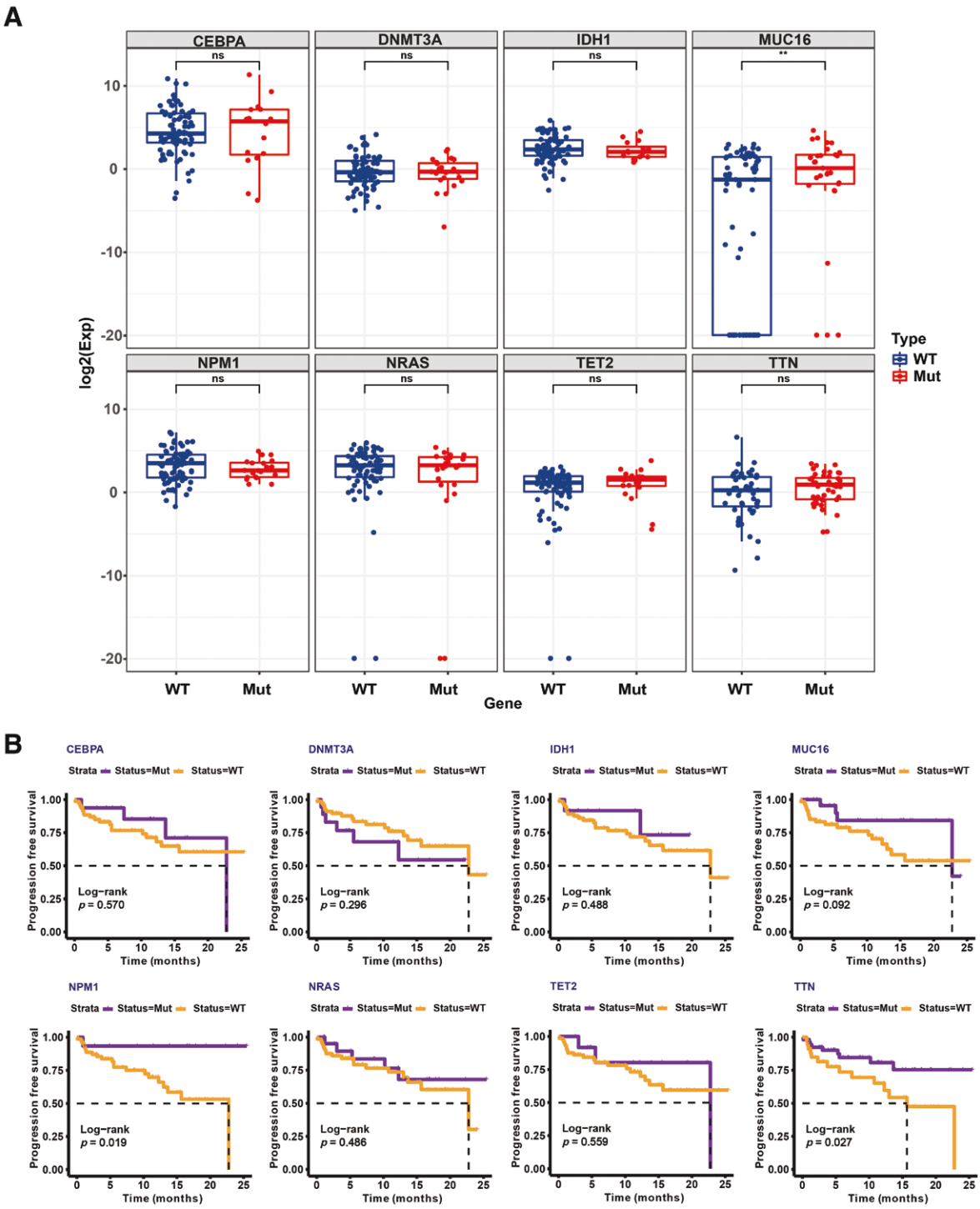

Figure 6. Clinical characteristics associated with the top mutated genes. (A) Expression of top mutated genes in patients with and without indicated mutations. $^{**}p$ <0.010. (B) Surviving curves of the top mutated genes in patients with and without indicated mutations. Among these genes, only *NPM1* and *TTN* genes show a significant difference in survival status between patients with and without mutations. ns = not significant.

mutated genes include *TTN* (61%), *MUC16* (29%), *NRAS* (25%), *DNMT3A* (21%), *CEBPA* (18%), *FLT3-ITD* (18%), *NPM1* (17%), *WT1* (17%), *KIT* (15%), *TET2* (15%), and *IDH1* (14%) (**Fig. 5A**).

We also identified 27.6% (24/87) of patients harbored at least 1 fusion gene, including well-characterized *CBFB*::*MYH11* at 11.5% and *RUNX1*::*RUNX1T1* at 6.9% (**Fig. 5B**; Table S5, https://links.lww.com/BS/A163). Among these patients, 20 of them carried only 1 fusion gene, while the remaining 4 patients carried 2 fusion genes. Interestingly, the fusion gene *MLL*::*AF6/AF10/ELL/ENL* was identified as one of the 2 fusion genes in 3 of the 4 patients carrying 2 fusion genes.

Our results indicated that genetic mutations do not usually lead to changes in the expression of affected genes between patients with (Mut) and without (WT) selected mutations

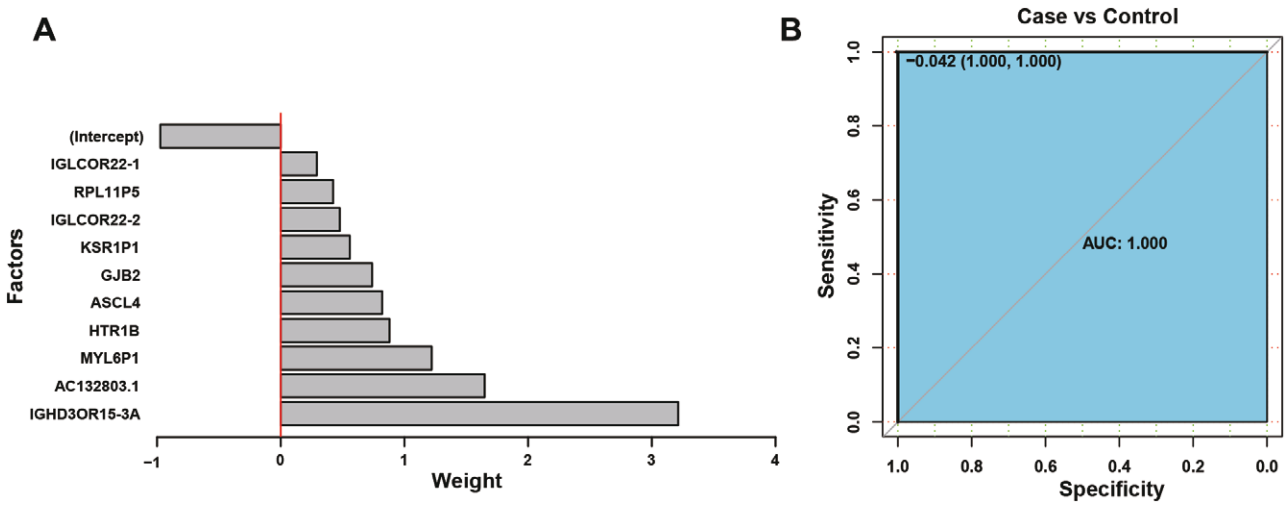

Figure 7. Establishment of stratification formula to distinguish patients between PG1 and PG2 based on a Lasso regression. (A) Bar graph showing 10 key feature genes identified with this approach and their weight coefficients, as well as the constant (Intercept), which is used to serve as the baseline level of the model. These numbers will be used in the prediction formula: $y = a1 \times x1 + a2 \times x2 + a3 \times x3 + \ldots\ldots + a10 \times x10 + b$ (see main text for details). (B) ROC curve used to evaluate the specificity and sensitivity of the above formula. AUC = area under the curve, PG1 = prognostic group 1, PG2 = prognostic group 2, ROC = receiver operating characteristic.

(**Fig. 6A**), likely influenced by multiple factors, including complex and regulatory processes in both transcriptional and translational levels in "Mut" group, as well as biased by other mutations in "WT" group. Similarly, significant changes in PFS were not generally observed in patients with these mutations (**Fig. 6B**). Among the top mutated genes, both *NPM1* and *TTN* mutations are associated with better PFS ($p = 0.019$ and 0.027, respectively), while there is a trend that MUC16 mutations are associated with a better prognosis ($p = 0.092$) (**Fig. 6B**).

3.5. Transcriptome-based prediction model

Based on the expression of 921 genes that were differentially expressed between any 2 supergroups (SC1–SC3), we first selected 781 genes with cutoff criteria (Wilcoxon test p value <.001; |log2FoldChange| >1) and removing outliers (values of $|x - \mu| > 3\sigma$). We next conducted multivariate Cox regression analysis and identified 96 genes with significant correlation between altered gene expression and clinical prognosis in 87 patients (Wald test p value <0.050). Based on Lasso regression, we identified 10 genes with their corresponding weight coefficients bigger than 0, while the weight of the remaining genes is 0 (**Fig. 7A**; Table S6, https://links.lww.com/BS/A163). We thus were able to establish a classification formula: $y = a1 \times x1 + a2 \times x2 + a3 \times x3 + \ldots\ldots + a10 \times x10 + b$ (y: score; ai: weight coefficient; xi: z score of the expression of gene 1 to 10; b: constant; see Table S6, https://links.lww.com/BS/A163, for details). Filling in the parameters from Table S6, https://links.lww.com/BS/A163, we were able to obtain the classification formula A (y = 1.6484**AC132803.1* + 0.8201**ASCL4* + 0.7387**GJB2* + 0.8795**HTR1B* + 3.2131**IGHD3OR15-3A* + 0.2923**IGLCOR22-1* + 0.478**IGLCOR22-2* + 0.5585**KSR1P1* + 1.2203**MYL6P1* + 0.4234**RPL11P5* – 0.9733) to distinguish patients between PG1 and PG2 subgroups.

As evaluated using ROC curves (**Fig. 7B**), both the specificity and sensitivity reached 100% when the cutoff was set to –0.042. With this formula, patients with transcriptome data are likely to be classified into either PG1 or PG2 group. While patients in PG2 usually show better prognosis, patients in PG1 usually show both good and bad prognosis (**Fig. 3D**).

To predict the prognosis of the patients classified into PG1, we performed a similar analysis as the above with modification for 55 patients within PG1. We manually divided these patients into 2 groups containing 20 relapsed patients (PFS = 1) and 35 patients in remission (PFS = 0) and selected 10 characteristic genes showing significant differences in altered expression between 2 groups. By Lasso regression analysis, we obtain 6 genes with their corresponding weight coefficients bigger than 0 and come up with a formula B (y = 0.0993**AL035106.1* + 0.0386**HOXA10* + 0.1727**ID1* * 0.6411**IGHD1-1* + 0.1337**IGHD3-16* + 0.1028**MBLAC1* – 0.5728) for prognostic stratification in PG1 patients (**Fig. 8A**; Table S7, https://links.lww.com/BS/A163). Using ROC curves, the specificity and sensitivity are 91.4% and 70.0%, respectively, with cutoff at –0.711 (**Fig. 8B**). We also identified highly mutated genes, including *DNMT3A* (21%), *GATA2* (11.5%), *NCOR2* (5.7%), and *RUNX1* (9.2%), that are associated with altered clinical outcomes ($p < 0.050$, long-rank test) (**Fig. 8C**), which may be added for outcome prediction in PG1 patients. This 2-step prediction model is summarized in **Figure 9**.

4. DISCUSSION

Current diagnostic and risk assessment mainly rely on cytogenetic assays for structural genomic lesions and targeted screening for genomic alterations. However, it has been noticed that genetic alterations are not always translated into aberrations in transcriptomic levels, which are functional products of gene expression and are believed to be more relevant in association with cancer pathogenically or clinically.[38,39] This discrepancy underscores the imperative to explore transcriptome signatures that can complement or even transcend genetic information for more accurate prognostic prediction.

Intensive attempts have been made to establish RNA-seq and multi-omics integration as improved strategies for molecular classification and biomarker discovery.[19,21,40–42] For instance, a recent multi-omics study on FLT3-mutated AML successfully deciphered immune hallmarks and biomarkers linked to FLT3 inhibitor sensitivity, demonstrating the power of integrated analysis to inform both biology and therapy.[43] Despite significant progress, accurate prognostic prediction and risk stratification remain challenging in clinical practice. It has been shown by a study that RNA-seq can be employed to detect somatic mutations robustly in AML.[40] Compared to DNA-based whole genome and exome sequencing, RNA-seq offers greater return enabling simultaneous detection of small variants, fusion genes,

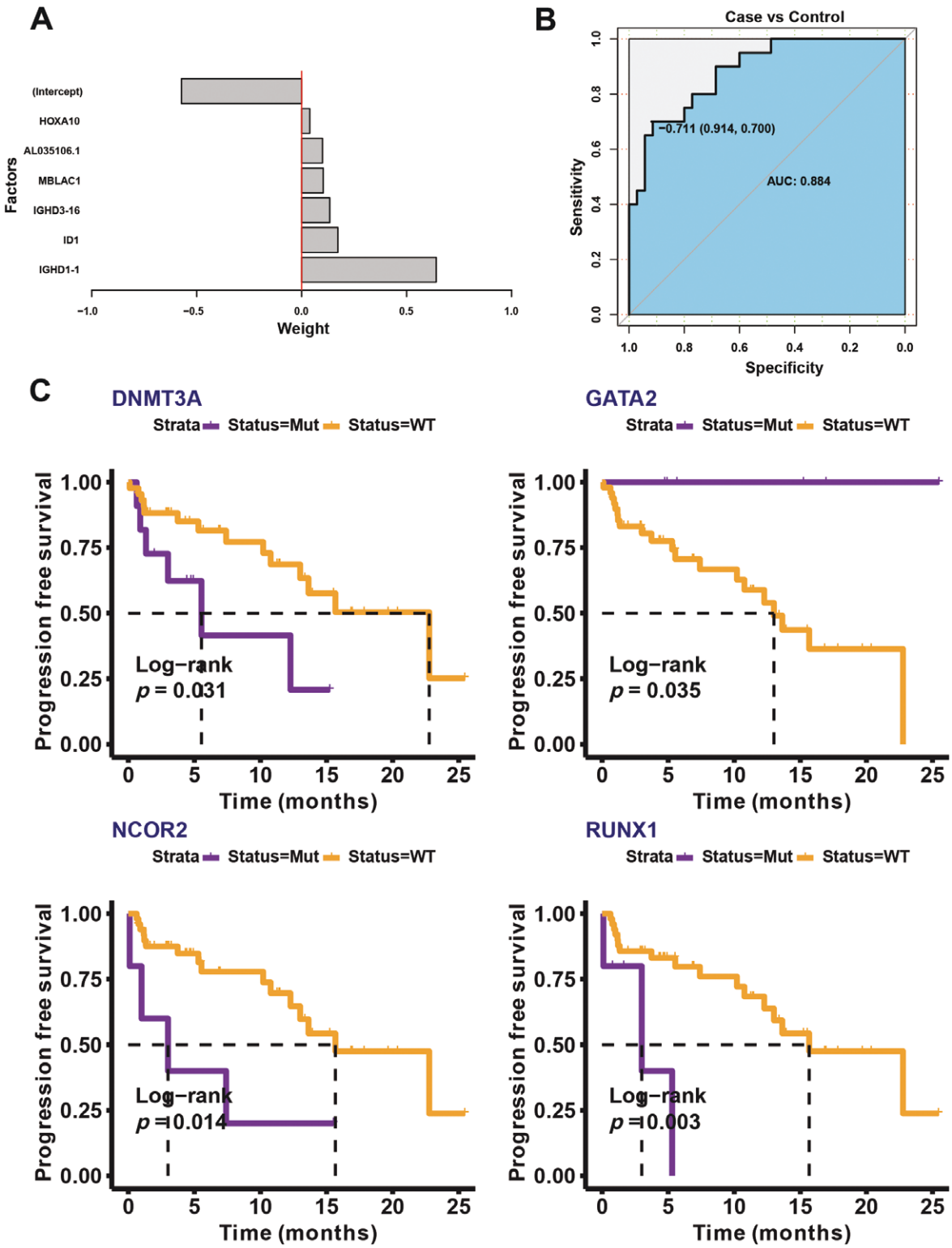

Figure 8. Establishment of a stratification formula to predict prognosis within PG1 patients. (A) Bar graph showing 10 key feature genes identified based on Lasso regression and their weights and constant, which are employed in the prediction formula as in Figure 7A. (B) ROC curve used to evaluate the specificity and sensitivity of the above formula. These numbers will be used in the prediction formula: $y = a1 \times x1 + a2 \times x2 + a3 \times x3 + \ldots\ldots + a10 \times x10 + b$ (see main text for details). (C) Surviving curves of 4 top mutated genes that were associated with altered clinical outcomes. The results of log-rank tests were also presented in each graph. AUC = area under the curve, ROC = receiver operating characteristic.

and whole-transcriptome expression information. Therefore, it is likely that RNA-seq can be established as a flexible and comprehensive platform for patient stratification, risk assessment, and therapy selection in AML.[19,41] As noticed, a study by Cheng et al[21] demonstrated that 8 molecular subtypes (G1–G8) with both subtypes showing significant overlapping with the existing AML subgroups as well as those being redefined based on their enhanced consensus method and clustering algorithms

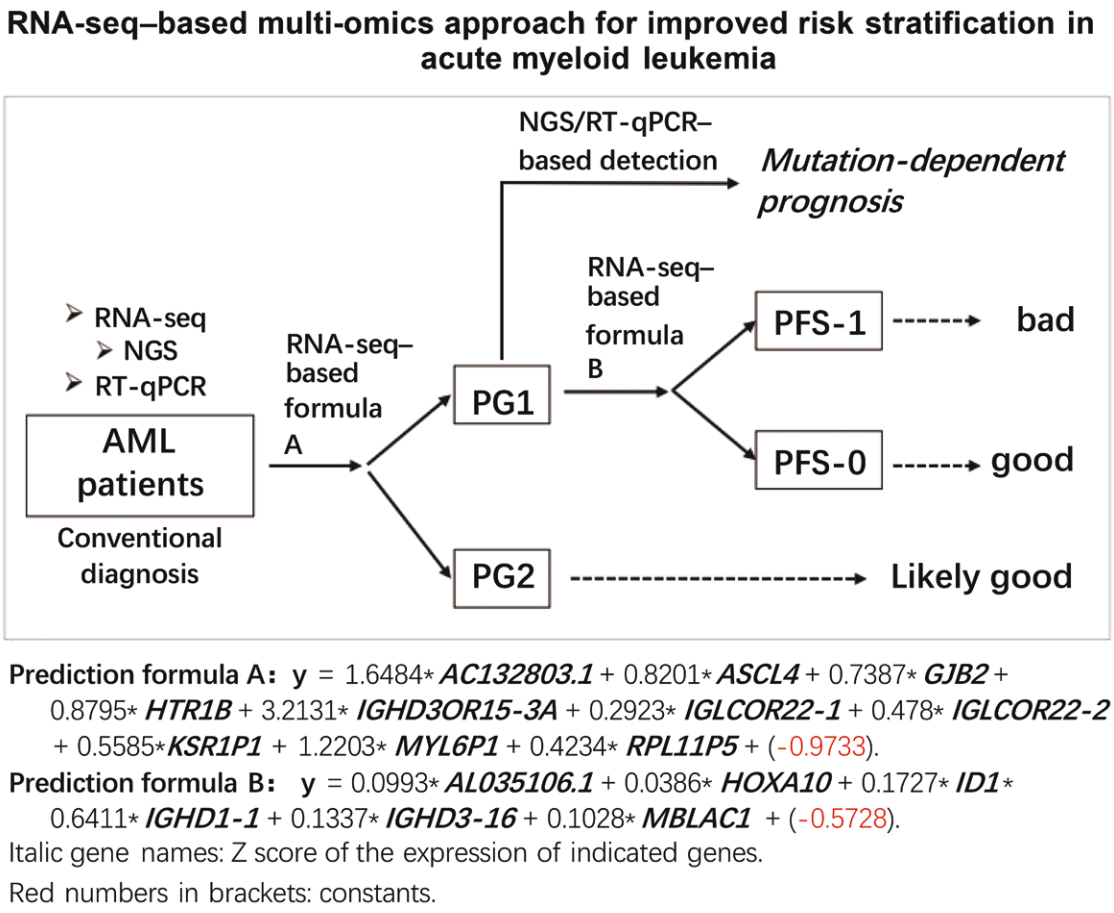

Figure 9. Prognostic prediction model based on transcriptomics-based multi-omics approach for improved risk stratification in acute myeloid leukemia. The prediction formulae A and B are based on this formula: "$y = a1 \times x1 + a2 \times x2 + a3 \times x3 + \ldots\ldots + a10 \times x10 + b$ (y: score; ai: weight coefficient; xi: z score of the expression of gene 1 to 10; b: constant)," which was established from a Lasso regression model. AML = acute myeloid leukemia, NGS = next-generation sequencing, PG1 = prognostic group 1, PG2 = prognostic group 2, RT-qPCR = reverse transcription quantitative polymerase chain reaction.

with extra feature filtration steps. Strikingly, this classification shows strong correlation with the WHO classification, including significant overlapping between 2 approaches for the subtypes defined by the fusion genes *PML::RARA* (G1), *CBFB::MYH11* (G2), and *RUNX1::RUNX1T1* (G3). However, such clear-cut genetic clustering was not observed in our study, which may reflect differences in methodology, patient characteristics, and other unrecognized differences between 2 studies. While Cheng et al[21] used Consensus Clustering, we employed hierarchical clustering with Canberra distance, which may explain the differences in genetic clustering results between the 2 studies. It should be pointed out that, by our analysis, we clearly see a strong association between molecular clustering and prognostic outcomes. We also demonstrated the importance of disease relevance reflected by aberrations on transcriptome levels, as genomic aberrations may not always be reflected on transcriptome level and even on protein level in other cases.[16,44,45] This is also consistent with the observations that targeted therapies may not always be effective on patients with targeted mutations due to tumor heterogenicity[46] and other challenges, but showing rare efficacy in patients without targeted mutations.[4,47–49]

By GO enrichment analysis, we identified several gene clusters that were highly regulated in AML. The GPCR signaling pathway was identified as top hit in both MF and BP categories by GO enrichment analysis (**Fig. 4**). GPCR is one of the largest families of human membrane proteins, playing a significant role in various BP, including sensory perception, neurotransmission, and endocrine processes.[50,51] Up to now, there are approximately 34% of the Food and Drug Administration (FDA)-approved drugs in the US targeting GPCRs.[52,53] GPCRs have also been suggested to be important in controlling tumor growth, invasion, migration, metastasis, and other aspects of cancer development and progression.[31] While 800 GPCRs have been identified in human, half of them are associated with olfactory function (about 400) and other sensory functions.[54] Intriguingly, olfactory receptor activity was also identified as one of the top regulated MF categories by GO and olfactory transduction was identified by KEGG, suggesting that genes in both GPCR or OR clusters are strongly associated with AML.[55–59] Interestingly, accumulating evidence has demonstrated aberrant expression of OR genes in various cancers, suggesting an important role of this group of genes in tumorigenesis and potentials in tumor diagnostics and drug discovery.[60–62] In addition, we also identified other sensory-related biological groups that were disrupted in AML. Moreover, gene clusters associated with DNA binding, RNA polymerase, as well as cell differentiation that are fundamental to cancer were also identified as top hits by 3 different types of assays, supporting the authenticity of our approach.

Our classification correlates strongly with prognosis: PG2 patients had favorable outcomes overall, whereas PG1 patients showed mixed outcomes (**Fig. 3D**). Adding additional steps in risk stratification, we were able to predict these patients' treatment outcome with high accuracy and specificity (**Figs. 7–9**). The main limitations of this study are that the single-center design of this study may introduce selection bias, the sample size of patients is not large enough to cover and conduct statistical analysis on rare molecular alterations, and there is a lack of additional independent cohorts to validate this prediction

model. Moreover, although genomic and transcriptomic data were integrated in this study, the final model was constructed primarily based on transcriptomic data, with a lack of in-depth multi-omics integration analysis. In addition, the underlying molecular mechanisms and mechanistic insights associated with the identified transcriptomic signatures and pathways remain largely unelucidated. All these limitations should be addressed thoroughly in the future.

In conclusion, this study improves risk stratification or outcome prediction by employing RNA-seq–based multi-omics approach. We not only confirm the independent value of RNA sequencing in molecular subtyping and prognostic prediction but also reveal the potential significance of the olfactory receptor pathway in AML, providing new clues for future mechanistic studies and target discovery.

ACKNOWLEDGMENTS

Noncommunicable Chronic Diseases-National Science and Technology Major Project (2025ZD0545804). This study was supported by grants from the National Natural Science Foundation (No. 82170212, 82370181). The Second Affiliated Hospital of the Army Medical University is a special project for clinical research (No. 2024F010).

ETHICAL APPROVAL

The study was conducted in accordance with the International Conference on Harmonisation good clinical practice guidelines, the Declaration of Helsinki, and applicable local regulations. All patients enrolled provided written informed consent before study participation. The samples of patients involved in this study were approved by the Ethics Review Committee of the Chinese Registration Clinical Trial under registration number ChiECRCT-20180145 on September 17, 2018.

AUTHOR CONTRIBUTIONS

Y.Y., L.Y., and J.F. drafted the manuscript. J.L., H.L., X.T., Y.M., X.H., G.C., Y.C., and Q.W. collected clinical samples. D.G. and S.T. are responsible for bioinformatics analysis. J.L., X.Z., and C.Z designed and supervised the studies and reviewed the manuscript. All authors read and approved the final manuscript.

REFERENCES

[1] Venugopal S, Sekeres MA. Contemporary management of acute myeloid leukemia: a review. *JAMA Oncol* 2024;10(10):1417–1425.
[2] Kantarjian HM, DiNardo CD, Kadia TM, et al. Acute myeloid leukemia management and research in 2025. *CA Cancer J Clin* 2025;75(1):46–67.
[3] Li M, Li Y, Qu Q, et al. Clinical features and prognostic nomogram for therapy-related acute myeloid leukemia after allogeneic hematopoietic stem cell transplantation. *Cancer Lett* 2025;612:217460.
[4] Bhansali RS, Pratz KW, Lai C. Recent advances in targeted therapies in acute myeloid leukemia. *J Hematol Oncol* 2023;16(1):29.
[5] Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. *Leukemia* 2022;36(7):1703–1719.
[6] Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood* 2022;140(12):1345–1377.
[7] Wang Y, Chang YJ, Chen J, et al. Consensus on the monitoring, treatment, and prevention of leukaemia relapse after allogeneic haematopoietic stem cell transplantation in China: 2024 update. *Cancer Lett* 2024;605:217264.
[8] Passaro A, Al Bakir M, Hamilton EG, et al. Cancer biomarkers: emerging trends and clinical implications for personalized treatment. *Cell* 2024;187(7):1617–1635.
[9] Zhou Y, Tao L, Qiu J, et al. Tumor biomarkers for diagnosis, prognosis and targeted therapy. *Signal Transduct Target Ther* 2024;9(1):132.
[10] Döhner H, Wei AH, Löwenberg B. Towards precision medicine for AML. *Nat Rev Clin Oncol* 2021;18(9):577–590.
[11] Levin M, Stark M, Ofran Y, Assaraf YG. Deciphering molecular mechanisms underlying chemoresistance in relapsed AML patients: towards precision medicine overcoming drug resistance. *Cancer Cell Int* 2021;21(1):53.
[12] Denny JC, Collins FS. Precision medicine in 2030-seven ways to transform healthcare. *Cell* 2021;184(6):1415–1419.
[13] Yan N, Wang ZL, Wang XJ, et al. Measurable residual disease testing by next generation sequencing is more accurate compared with multiparameter flow cytometry in adults with B-cell acute lymphoblastic leukemia. *Cancer Lett* 2024;598:217104.
[14] Sager R. Expression genetics in cancer: shifting the focus from DNA to RNA. *Proc Natl Acad Sci U S A* 1997;94(3):952–955.
[15] Rassy E, Pavlidis N. Progress in refining the clinical management of cancer of unknown primary in the molecular era. *Nat Rev Clin Oncol* 2020;17(9):541–554.
[16] Barbeira AN, Bonazzola R, Gamazon ER, et al. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol* 2021;22(1):49.
[17] Koprulu M, Carrasco-Zanini J, Wheeler E, et al. Proteogenomic links to human metabolic diseases. *Nat Metab* 2023;5(3):516–528.
[18] Zhang Q, Ai Y, Abdel-Wahab O. Molecular impact of mutations in RNA splicing factors in cancer. *Mol Cell* 2024;84(19):3667–3680.
[19] Docking TR, Parker JDK, Jädersten M, et al. A clinical transcriptome approach to patient stratification and therapy selection in acute myeloid leukemia. *Nat Commun* 2021;12(1):2474.
[20] Jabbour E, Short NJ, Jain N, et al. The evolution of acute lymphoblastic leukemia research and therapy at MD Anderson over four decades. *J Hematol Oncol* 2023;16(1):22.
[21] Cheng WY, Li JF, Zhu YM, et al. Transcriptome-based molecular subtypes and differentiation hierarchies improve the classification framework of acute myeloid leukemia. *Proc Natl Acad Sci U S A* 2022;119(49):e2211429119.
[22] Archer KJ, Fu H, Mrózek K, et al. Improving risk stratification for 2022 European LeukemiaNet favorable-risk patients with acute myeloid leukemia. *Innovation (Camb)* 2024;5(6):100719.
[23] Li J, Zong S, Wan Y, et al. Integration of transcriptomic features to improve prognosis prediction of pediatric acute myeloid leukemia with KMT2A rearrangement. *Hemasphere* 2023;7(12):e979.
[24] Wang X, Wang J, Wei S, et al. The latest edition of WHO and ELN guidance and a new risk model for Chinese acute myeloid leukemia patients. *Front Med (Lausanne)* 2023;10:1165445.
[25] Cao XX, Cai H, Mao YY, et al. Next-generation sequencing-based genetic landscape and its clinical implications for Chinese acute myeloid leukemia patients. *Cancer Cell Int* 2018;18:215.
[26] Xu R, Wunsch D. Survey of clustering algorithms. *IEEE Trans Neural Netw* 2005;16(3):645–678.
[27] Bullinger L, Döhner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004;350(16):1605–1616.
[28] Sherman BT, Hao M, Qiu J, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 2022;50(W1):W216–W221.
[29] Chang Y, Li S, Li Z, et al. Non-invasive detection of lymphoma with circulating tumor DNA features and protein tumor markers. *Front Oncol* 2024;14:1341997.
[30] O'Hayre M, Degese MS, Gutkind JS. Novel insights into G protein and G protein-coupled receptor signaling in cancer. *Curr Opin Cell Biol* 2014;27:126–135.
[31] Chaudhary PK, Kim S. An insight into GPCR and G-Proteins as cancer drivers. *Cells* 2021;10(12):3288.
[32] Kalra S, Mittal A, Gupta K, et al. Analysis of single-cell transcriptomes links enrichment of olfactory receptors with cancer cell differentiation status and prognosis. *Commun Biol* 2020;3(1):506.
[33] Cramer P. Organization and regulation of gene transcription. *Nature* 2019;573(7772):45–54.
[34] Bonczek O, Wang L, Gnanasundram SV, et al. DNA and RNA binding proteins: from motifs to roles in cancer. *Int J Mol Sci* 2022;23(16):9329.
[35] Zhu K, Xia Y, Tian X, et al. Characterization and therapeutic perspectives of differentiation-inducing therapy in malignant tumors. *Front Genet* 2023;14:1271381.
[36] Zhang H, Kong Q, Wang J, Jiang Y, Hua H. Complex roles of cAMP-PKA-CREB signaling in cancer. *Exp Hematol Oncol* 2020;9(1):32.
[37] Chang RB, Beatty GL. The interplay between innate and adaptive immunity in cancer shapes the productivity of cancer immunosurveillance. *J Leukoc Biol* 2020;108(1):363–376.

[38] Adashek JJ, Kato S, Parulkar R, et al. Transcriptomic silencing as a potential mechanism of treatment resistance. *JCI Insight* 2020;5(11):e134824.
[39] Tsimberidou AM, Fountzilas E, Bleris L, Kurzrock R. Transcriptomics and solid tumors: the next frontier in precision cancer medicine. *Semin Cancer Biol* 2022;84:50–59.
[40] Gu M, Zwiebel M, Ong SH, et al. RNAmut: robust identification of somatic mutations in acute myeloid leukemia using RNA-sequencing. *Haematologica* 2020;105(6):e290–e293.
[41] Arindrarto W, Borràs DM, de Groen RAL, et al. Comprehensive diagnostics of acute myeloid leukemia by whole transcriptome RNA sequencing. *Leukemia* 2021;35(1):47–61.
[42] Brown LM, Lonsdale A, Zhu A, et al. The application of RNA sequencing for the diagnosis and genomic classification of pediatric acute lymphoblastic leukemia. *Blood Adv* 2020;4(5):930–942.
[43] Niu M, Wang N, Yang D, et al. Multi-omics integration reveals immune hallmarks and biomarkers associated with FLT3 inhibitor sensitivity in FLT3-mutated AML. *Blood Sci* 2025;7(2):e00227.
[44] Deng YT, You J, He Y, et al. Atlas of the plasma proteome in health and disease in 53,026 adults. *Cell* 2025;188(1):253–271.e7.
[45] Abell NS, DeGorter MK, Gloudemans MJ, et al. Multiple causal variants underlie genetic associations in humans. *Science* 2022;375(6586):1247–1254.
[46] Chen A, Hu S, Wang QF. Tumor heterogeneity of acute myeloid leukemia: insights from single-cell sequencing. *Blood Sci* 2019;1(1):73–76.
[47] Liu B, Zhou H, Tan L, Siu KTH, Guan XY. Exploring treatment options in cancer: tumor treatment strategies. *Signal Transduct Target Ther* 2024;9(1):175.
[48] Waarts MR, Stonestrom AJ, Park YC, Levine RL. Targeting mutations in cancer. *J Clin Invest* 2022;132(8):e154943.
[49] Jin J, Wu X, Yin J, et al. Identification of genetic mutations in cancer: challenge and opportunity in the new era of targeted therapy. *Front Oncol* 2019;9:263.
[50] Zhang M, Chen T, Lu X, Lan X, Chen Z, Lu S. G protein-coupled receptors (GPCRs): advances in structures, mechanisms, and drug discovery. *Signal Transduct Target Ther* 2024;9(1):88.
[51] Liu Y, An L, Wang X, et al. Engineering a controllable and reversible switch for CAR-based cellular immunotherapies via a genetic code expansion system. *J Hematol Oncol* 2024;17(1):122.
[52] Yang D, Zhou Q, Labroska V, et al. G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduct Target Ther* 2021;6(1):7.
[53] Casadó V, Casadó-Anguera V. What are the current trends in G protein-coupled receptor targeted drug discovery? *Expert Opin Drug Discov* 2023;18(8):815–820.
[54] Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. *Nat Rev Drug Discov* 2017;16(12):829–842.
[55] Huang D, Yu Z, Lu H, et al. Adhesion GPCR ADGRE2 maintains proteostasis to promote progression in acute myeloid leukemia. *Cancer Res* 2024;84(13):2090–2108.
[56] Saito Y, Kaneda K, Suekane A, et al. Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. *Leukemia* 2013;27(8):1637–1649.
[57] Alberich Jordà M, Rayman N, Tas M, et al. The peripheral cannabinoid receptor Cb2, frequently expressed on AML blasts, either induces a neutrophilic differentiation block or confers abnormal migration properties in a ligand-dependent manner. *Blood* 2004;104(2):526–534.
[58] Manteniotis S, Wojcik S, Brauhoff P, et al. Functional characterization of the ectopically expressed olfactory receptor 2AT4 in human myelogenous leukemia. *Cell Death Discov* 2016;2:15070.
[59] Manteniotis S, Wojcik S, Göthert JR, et al. Deorphanization and characterization of the ectopically expressed olfactory receptor OR51B5 in myelogenous leukemia cells. *Cell Death Discov* 2016;2:16010.
[60] Geng Y, Peveler WJ, Rotello VM. Array-based "chemical nose" sensing in diagnostics and drug discovery. *Angew Chem Int Ed Engl* 2019;58(16):5190–5200.
[61] Maßberg D, Hatt H. Human olfactory receptors: novel cellular functions outside of the nose. *Physiol Rev* 2018;98(3):1739–1763.
[62] Chen P, Wang W, Liu R, et al. Olfactory sensory experience regulates gliomagenesis via neuronal IGF1. *Nature* 2022;606(7914):550–556.