



Novel Prognostic Signature for Acute Myeloid Leukemia: Bioinformatics Analysis of Combined CNV-Driven and Ferroptosis-Related Genes

Chunjiao Han^{1†}, Jiafeng Zheng^{2†}, Fangfang Li^{3†}, Wei Guo^{1,2*} and Chunquan Cai^{4*}

OPEN ACCESS

Edited by:

Ying Li,
The University of Texas Health Science
Center at San Antonio, United States

Reviewed by:

Xianzhe Li,
The Sixth Affiliated Hospital of Sun
Yat-sen University, China
Zi-Jun Xu,
Jiangsu University Affiliated People's
Hospital, China

*Correspondence:

Wei Guo
guowei79656@126.com
Chunquan Cai
15122656313@126.com

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

Specialty section:

This article was submitted to
Human and Medical Genomics,
a section of the journal
Frontiers in Genetics

Received: 06 January 2022

Accepted: 22 March 2022

Published: 26 April 2022

Citation:

Han C, Zheng J, Li F, Guo W and Cai C
(2022) Novel Prognostic Signature for
Acute Myeloid Leukemia:
Bioinformatics Analysis of Combined
CNV-Driven and Ferroptosis-
Related Genes.
Front. Genet. 13:849437.
doi: 10.3389/fgene.2022.849437

¹Clinical School of Paediatrics, Tianjin Medical University, Tianjin, China, ²Department of Pulmonology, Tianjin Children's Hospital/Tianjin University Children's Hospital, Tianjin, China, ³Department of Rheumatology and Immunology, Tianjin Children's Hospital/Tianjin University Children's Hospital, Tianjin, China, ⁴Department of Institute of Pediatrics, Tianjin Children's Hospital/Tianjin University Children's Hospital, Tianjin, China

Background: Acute myeloid leukemia (AML), which has a difficult prognosis, is the most common hematologic malignancy. The role of copy number variations (CNVs) and ferroptosis in the tumor process is becoming increasingly prominent. We aimed to identify specific CNV-driven ferroptosis-related genes (FRGs) and establish a prognostic model for AML.

Methods: The combined analysis of CNV differential data and differentially expressed genes (DEGs) data from The Cancer Genome Atlas (TCGA) database was performed to identify key CNV-driven FRGs for AML. A risk model was constructed based on univariate and multivariate Cox regression analysis. The Gene Expression Omnibus (GEO) dataset was used to validate the model. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted to clarify the functional roles of DEGs and CNV-driven FRGs.

Results: We identified a total of 6828 AML-related DEGs, which were shown to be significantly associated with cell cycle and immune response processes. After a comprehensive analysis of CNVs and corresponding DEGs and FRGs, six CNV-driven FRGs were identified, and functional enrichment analysis indicated that they were involved in oxidative stress, cell death, and inflammatory response processes. Finally, we screened 2 CNV-driven FRGs (DNAJB6 and HSPB1) to develop a prognostic risk model. The overall survival (OS) of patients in the high-risk group was significantly shorter in both the TCGA and GEO (all $p < 0.05$) datasets compared to the low-risk group.

Conclusion: A novel signature based on CNV-driven FRGs was established to predict the survival of AML patients and displayed good performance. Our results may provide potential targets and new research ideas for the treatment and early detection of AML.

Keywords: acute myeloid leukemia, copy number variations, ferroptosis, prognosis, gene signature

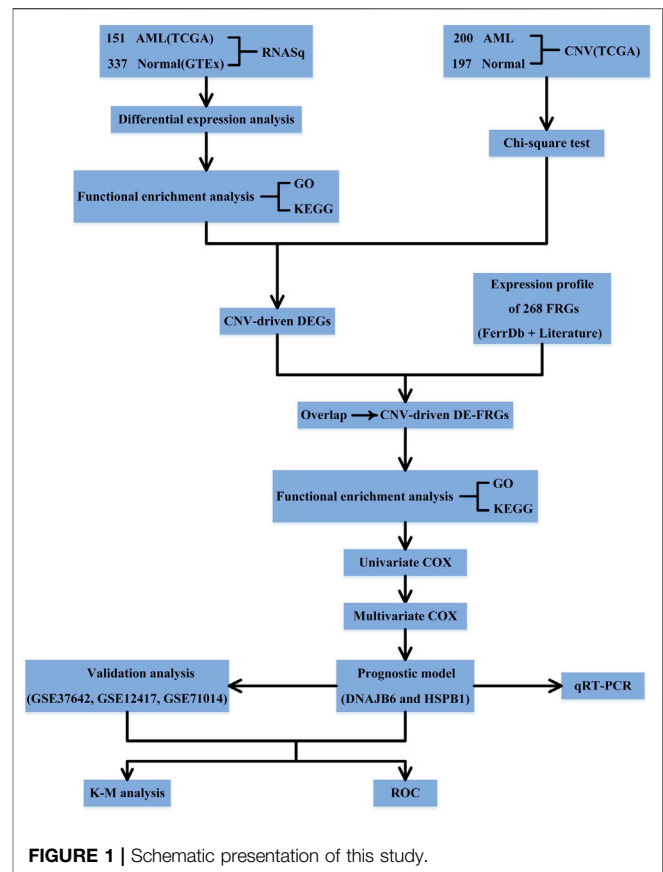
INTRODUCTION

AML is a heterogeneous malignancy that is characterized by imbalanced hematopoietic stem cells and uncontrolled differentiation. It accounts for 15–20% of leukemia in children and has a higher incidence during adolescence (Arber et al., 2016). Despite intensive treatment, the long-term survival rate of children with AML is still only 45–55% (Carver et al., 2003). Cytogenetic outcomes and molecular is considered the most important prognostic factors at present (Döhner et al., 2017), however, the current genetic methods still can't achieve accurate prognosis prediction. Our study is aimed to construct a novel prognostic model as well as an improvement of risk-adapted therapy for patients.

Ferroptosis is a newly discovered iron-dependent programmed cell death, which is unlike apoptosis, autophagy, necrosis, pyroptosis, and other cell death forms (Grimwade et al., 2016). It results in cell death by inducing excessive membrane lipid peroxidation (Herold et al., 2018). It is reported that ferroptosis induction can inhibit the growth of tumor cells especially with resistance to traditional therapies (Herold et al., 2014; Döhner et al., 2017). For refractory and relapsed AML, resistance to apoptosis is an important therapeutic measure (Hong et al., 2021; Huang et al., 2021). Currently, as a potential therapeutic target for cancer treatment, ferroptosis has attracted worldwide attention. Some studies have suggested that some ferroptosis-related genes can be prognosis biomarkers (Huot et al., 1996; Herold et al., 2014; Döhner et al., 2017). The upregulation of phospholipid hydroperoxidase GPx4 is associated with poor prognosis of AML (Jakob et al., 1993). In addition, a study found that low AKR1C2 and SOCS1 expression was highly correlated with more favorable overall survival and disease-free survival in AML patients (Jiang et al., 2020).

CNV refers to the duplication, inversion, or deletion of a DNA sequence of more than one kilobase (Jiang et al., 2020). Recently, CNV has been recognized as an important source of genetic variation which was found to play an important role in many cancers. The previous study suggested that the presence of CNV affects not only protein expression but also long non-coding RNAs and miRNAs (Kerr et al., 1994). Cytogenetic CNV abnormalities have been included in WHO classification (2016) (Kim et al., 2012) and other risk stratification strategies (Liang et al., 2009; Li et al., 2013; Kuett et al., 2015), and constitute the single strongest prognostic factor for complete remission (CR) and overall survival (OS) of AML. A study has found CNVs in 23 patients (76.7%) with NK-AML in Korea (Jiang et al., 2020). It showed that CNV increase is an independent predictive factor for shorter event-free survival and may affect the success of Ara-C and anthracycline-based chemotherapy. However, most previous studies focused on CNV or transcriptome changes, there is still a lack of comprehensive research on how CNV drives AML.

Although FRGs and CNV can be used as prognostic markers of AML respectively, the relationship between FRGs and CNV has not been reported at present, which aroused our interest. In



the present study, we used transcriptomics and CNVs profiles to identify CNV driven FRGs and aimed to construct a prognostic model of AML. Our study may contribute to a better understanding of the underlying mechanisms and provide a new therapeutic target for the treatment of AML.

MATERIALS AND METHODS

Data Source

Gene expression profiles of 151 bone marrow samples from AML patients were downloaded from the TCGA database. Since sequencing data of control samples were not available from the TCGA database, we obtained RNA sequencing data of 337 normal peripheral blood samples from the GTEX database. The GSE37642 (Mitra et al., 2008; Liang et al., 2016; Meng et al., 2016; Nibourel et al., 2017) was obtained from the GEO database. The GSE37642 dataset (platform: GPL570; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE37642>) contains transcriptional data from 136 AML patients with complete survival information, which was used for external validation analysis of the prognostic signature. Furthermore, the GSE12417 (platform: GPL570; n = 73; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12417>) and GSE71014 (platform: GPL10558; n = 104; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE71014>) datasets were downloaded as complementary

external validation sets to perform validation analyses on a larger sample base.

Figure 1 shows the schematic presentation of this study.

Collection of FRGs

We downloaded 259 FRGs from the FerrDb online database (<http://www.zhounan.org/ferrdb/>). Meanwhile, a total of 60 FRGs were also obtained from the report of Yingkai Hong et al. (Herold et al., 2014). After de-duplication (retention of unique values), a total of 268 FRGs were obtained for our study (**Supplementary Table S1**).

Differential Expression Analysis

Gene expression profiles of AML ($n = 151$) and normal ($n = 337$) samples from TCGA and GTEx were subjected to normalize Between Arrays function for normalization, and subsequently, differential expression analysis (AML vs. normal) was performed using the R package limma. The significance threshold was set to $|\log_2 \text{fold change (FC)}| \geq 1$ and $p < 0.05$. A volcano map showing the distribution of the identified DEGs was plotted based on the R package ggplot2. In addition, we also verified that the identified DEGs were able to distinguish between normal and AML samples by Principal Component Analysis (PCA) to imply the applicability of the samples.

Functional Enrichment Analysis

To reveal the functions of target genes, R package clusterProfiler (Papaemmanuil et al., 2016) was used to conduct Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. The GO terms were comprised of the following three divisions: biological process (BP), cellular component (CC), and molecular function (MF). $p < 0.05$ was regarded as statistically significant.

Integrative Analysis of Gene Expression and CNVs

The CNV data of AML patients ($n = 200$) and normal subjects ($n = 197$) were downloaded from the TCGA database (sequencing platform Affymetrix SNP 6.0). Genes in CNV regions were annotated using Genome Research Consortium Human build 38 (GRCh38) as the reference genome. CNVs alteration rates between normal and tumor samples were then compared using the Chi-square test, and CNVs data with $p < 0.05$ were chosen for the next analysis. Then the CNVs data and DEGs data of the same sample were merged to construct a matrix. By using the Kolmogorov-Smirnov test, those genes showing the same tendency both in CNVs and differential gene expression (CNV increase-upregulated; CNV decrease-downregulated) were selected as CNV-driven DEGs. Moreover, CNV-driven DEGs and identified FRGs were analyzed for overlap, and the common genes in both gene lists were defined as CNV-driven DE-FRGs.

Construction, Evaluation, and Validation of the Prognostic Model

The TCGA-AML dataset containing 132 AML samples with complete survival information was used for prognostic gene

screening and prognostic model construction and evaluation. The GSE37642 ($n = 136$), GSE12417 ($n = 73$), and GSE71014 ($n = 104$) datasets were used as an independent external validation set for prognostic model validation analysis. Prognostic genes were screened by Cox regression analysis. Briefly, target genes were included in univariate Cox regression analysis, and variables satisfying $p < 0.1$ were included in stepwise regression multivariate Cox analysis. The variables obtained from multivariate Cox regression analysis were identified as the optimal variables for the construction of the prognostic model. The risk score for each AML patient was calculated using the regression coefficient (coef) calculated from the multivariate Cox analysis and the expression of prognostic genes. The formula for calculating the risk score as shown below:

$$\begin{aligned} \text{risk score} = & \text{coef}_{\text{gene}_1} \times \text{expression}_{\text{gene}_1} \\ & + \text{coef}_{\text{gene}_2} \times \text{expression}_{\text{gene}_2} + \dots \\ & + \text{coef}_{\text{gene}_n} \times \text{expression}_{\text{gene}_n} \end{aligned}$$

The samples were classified into high- and low-risk groups based on the median value of the risk score in each dataset (TCGA dataset and independent external validation set). The difference in OS between the two risk subgroups was assessed based on the R package survival using Kaplan-Meier (K-M) analysis. Receiver operating characteristic (ROC) curves were plotted by survROC package to assess the accuracy of the risk score for prognostic prediction in patients with TCGA-AML and GSE37642/GSE12417/GSE71014 datasets-AML.

Independent Prognostic Analysis of the Risk Score

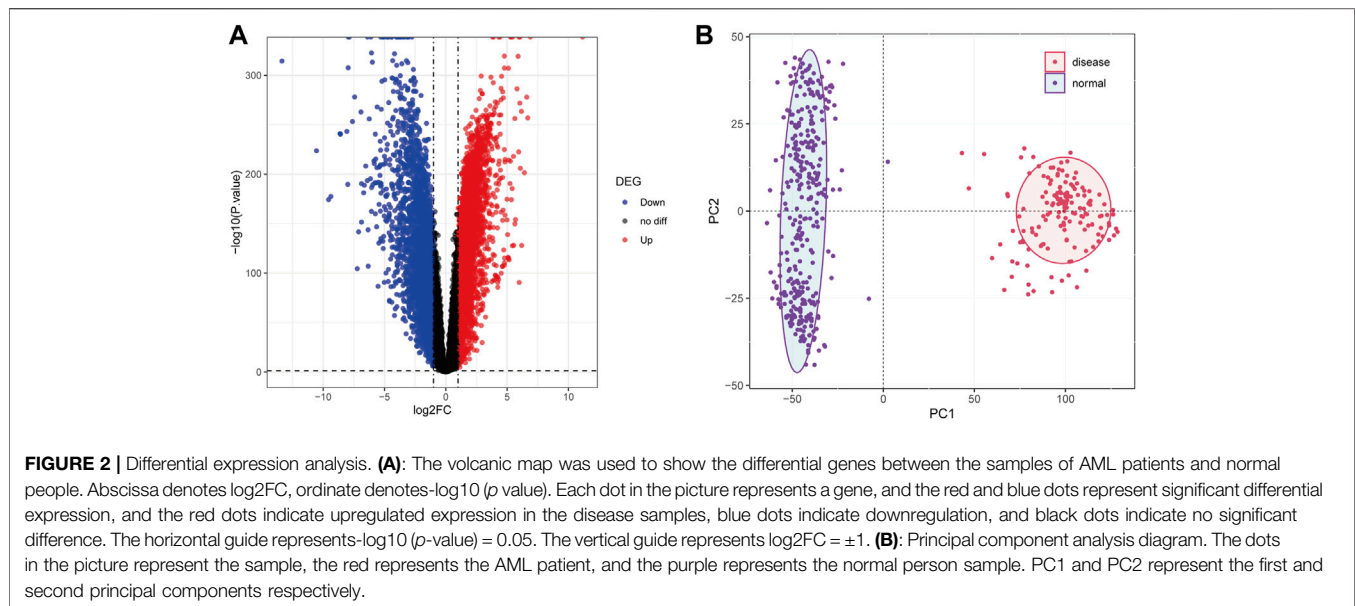
Clinical characteristics of AML (age and sex) available in the TCGA database were included in the Cox regression analysis along with the risk score. Univariate Cox regression analyses with $p < 0.05$ output would be performed further in multivariate Cox analyses. Ultimately, variables with $p < 0.05$ generated by multivariate Cox regression analysis were considered as independent prognostic factors for AML.

Patient and Tissue Preparation

We selected the blood of 10 AML patients and 10 healthy people to carry on the quantitative PCR test to the genes screened in this study. The experimental verification data are obtained from the sample database of the Institute of Pediatrics of Tianjin Children's Hospital, and no samples are obtained directly from the children's body, so they do not need to be approved by the institutional ethics committee. All experimental operations are carried out in accordance with the relevant guidelines and regulations, and have been repeatedly verified by professional laboratory researchers.

RNA Isolation and Quantitative Real-Time Polymerase Chain Reaction

Whole-cell RNA was extracted via the RNAiso Plus (TaKaRa, Japan). Quantitative real-time polymerase chain reaction (qRT-



PCR) was performed to detect DNAJB6 and HSPB1 expression using American Bio rad Bole T100 gradient PCR instrument. The following primers were used for qRT-PCR: DNAJB6 primers: upstream primer F: 5'-TATGAAGTGCTGTCGGATGCTAAG A-3'; downstream primer R: 5'-GAAGACATCATCTGGGTT ACGGA-3'. The conditions for PCR to DNAJB6 were: the size of the amplified product was 144bp, annealing temperature was 55–60°C. The sequences of HSPB1 primers used were: upstream primer F: 5'-GCAGTCCAACGAGATCACCA-3' and downstream primer R: 5'-TACTTGGCGGCAGTCTCATC-3'. The conditions for PCR to HSPB1 were: the size of the amplified product was 97bp, annealing temperature was 55–60°C. qRT-PCR was conducted in triplicate for each sample. All gene expression levels were normalized to that of GAPDH using the $2^{-\Delta\Delta Ct}$ method. Unpaired *t*-test (two-tailed) was used for the comparison analyses.

Statistical Analysis

All analyses in this study were performed in R software. A log-rank test was used to check the significant difference in OS between groups. An area under the ROC curve (AUC) served as an indicator of prognostic accuracy. Unless otherwise specified, a *p*-value less than 0.05 was considered statistically significant.

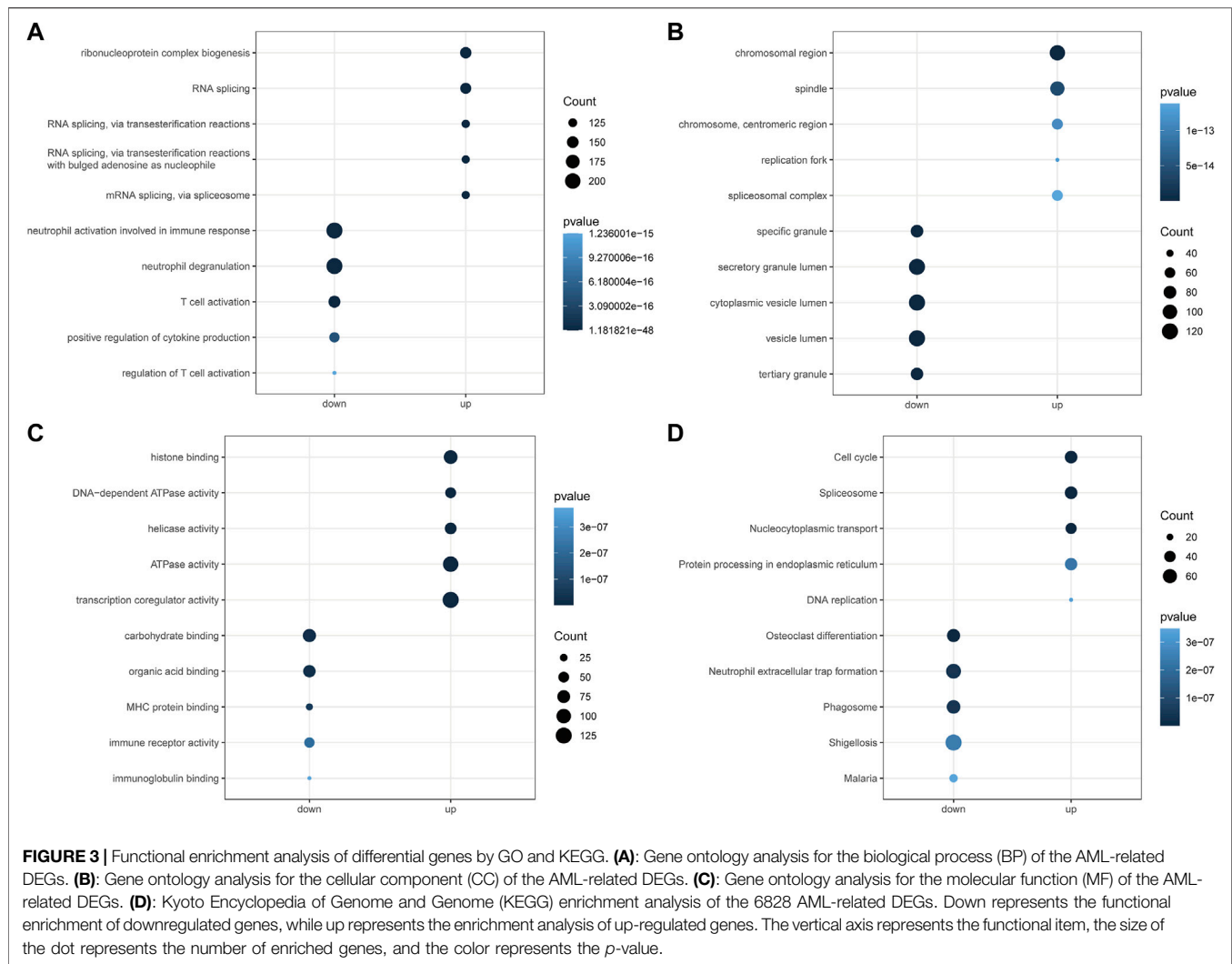
RESULTS

Exploration of AML-Related DEGs

The normalized expression profiles of TCGA-AML ($n = 151$) and GTEx-normal ($n = 337$) were selected as the basis for differential expression analysis. By R package limma, we identified 6,828 DEGs between AML and normal samples. Among them, a total of 3,330 met $\log_2 FC \geq 1$ and $p < 0.05$ and 3,498 matched $\log_2 FC \leq -1$ and $p < 0.05$ (**Figure 2A**; **Supplementary Table S2**). Furthermore, PCA performed

based on the obtained DEGs demonstrated that samples from different groups were clustered in the same category (**Figure 2B**).

To explore the potential mechanisms of the above AML-related DEGs, we performed GO and KEGG function enrichment analysis for upregulated DEGs and downregulated DEGs, respectively, using the clusterProfiler package. **Figures 3A–C** illustrated the top5 terms that were significantly enriched in the three categories of GO, BP, CC, and MF, based on upregulated and downregulated DEGs. In the BP category, specifically, upregulated DEGs were significantly associated with “ribonucleoprotein complex biogenesis”, “RNA splicing”, “RNA splicing, via transesterification reactions” (**Figure 3A**); besides, these genes were found to be closely correlated with the cell cycle (“DNA replication”, “chromosome segregation”, “mitotic nuclear division”, etc.). Moreover, upregulated DEGs might be played mainly in CCs such as “chromosomal region”, “spindle”, “chromosome, centromeric region” (**Figure 3B**) for MFs such as “histone binding”, “DNA-dependent ATPase activity”, “helicase activity” (**Figure 3C**). The results of the detailed GO analysis for the upregulated DEGs could be reviewed in **Supplementary Table S3**. For the down-regulated DEGs, they were significantly correlated with immune responses (“regulation of immune effector process”, “negative regulation of immune system process”, “immune response-activating signal transduction”, etc.) and biological processes of immune cells (“neutrophil degranulation”, “T cell activation”, “T-cell differentiation”, “lymphocyte proliferation”, etc.) (**Figure 3A**). Also, these genes were able to perform the molecular functions of “carbohydrate binding”, “organic acid binding”, and “MHC protein binding” (**Figure 3C**) in “specific granule”, “secretory granule lumen”, and “cytoplasmic vesicle lumen” (**Figure 3B**). Detailed GO analysis results for downregulated DEGs were displayed in **Supplementary Table S4**. The top5 pathways



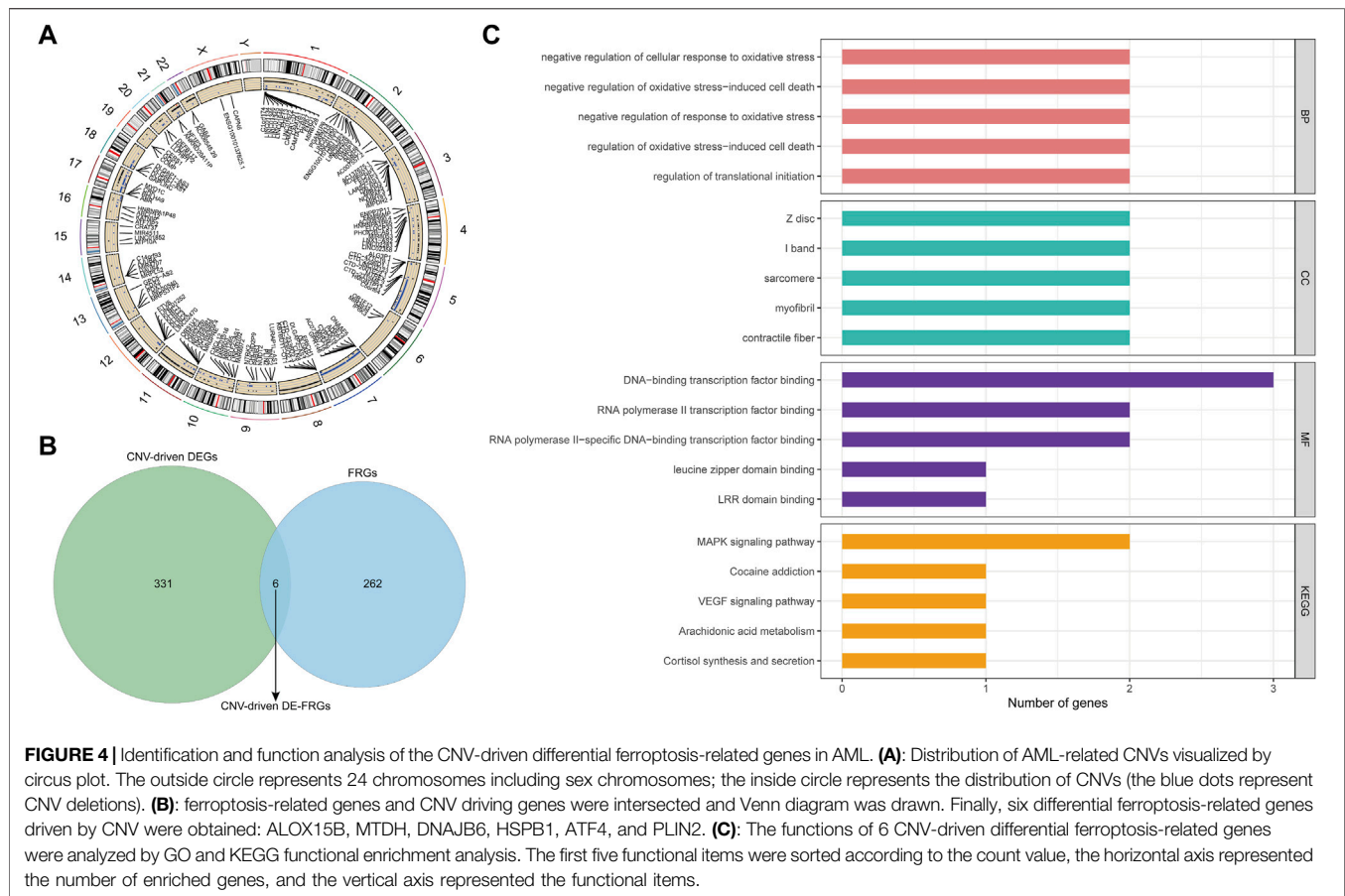
significantly enriched in KEGG analysis were shown in **Figure 3D**, with upregulated DEGs involved in “Cell cycle”, “Spliceosome”, “Nucleocytoplasmic transport”, “Protein processing in endoplasmic reticulum”, and “DNA replication” (**Supplementary Table S5**); downregulated DEGs were closely associated with “Osteoclast differentiation”, “Neutrophil extracellular trap formation”, “Phagosome”, “Shigellosis”, and “Malaria” (**Supplementary Table S6**).

Identification of CNV-Driven DE-FRGs in AML Patients

By applying the Chi-square test, 4637 CNV genes associated with AML were identified ($p < 0.05$; **Supplementary Table S7**). The distribution of AML-related CNVs in chromosomes was presented in **Figure 4A**. Then CNV-driven DEGs were screened using the Kolmogorov-Smirnov test. We selected 337 CNV-driven DEGs, of which 127 were upregulated in AML with increased CNV (**Supplementary Table S8**) and the remaining 210 were downregulated with decreased CNV

(**Supplementary Table S9**). Subsequently, by overlap analysis (**Figure 4B**), a total of 6 CNV-driven DE-FRGs, namely ALOX15B, MTDH, DNAJB6, HSPB1, ATF4, and PLIN2, were identified in the above list of CNV-driven DEGs and the list of 268 FRGs (**Supplementary Table S1**).

To illustrate the potential functional and biological effects of these CNV-driven DE-FRGs, GO (**Supplementary Table S10**) and KEGG (**Supplementary Table S11**) analyses were performed (**Figure 4C**). Results showed that CNV-driven DE-FRGs were significantly enriched in the BP category in terms related to oxidative stress response and its mediated cell death, such as “negative regulation of cellular response to oxidative stress”, “negative regulation of oxidative stress-induced cell death”, “negative regulation of response to oxidative stress”, and “regulation of oxidative stress-induced cell death”. KEGG analysis showed that these genes were involved in “MAPK signaling pathway”, “Cocaine addiction”, “VEGF signaling pathway”, “Arachidonic acid metabolism”, and “Cortisol synthesis and secretion”. These results suggested that CNV-driven DE-FRGs were probably implicated in the cell cycle



dysregulation and the inflammatory immune response during the disease process.

Construction of a Prognostic Signature Based on CNV-Driven DE-FRGs

We matched the expression profiles of the identified CNV-driven DE-FRGs in 132 TCGA-AML samples containing complete survival data. Univariate Cox regression analysis was applied to identify the CNV-driven DE-FRGs associated with survival time in AML patients. $p < 0.1$ was set as the cut-off value, and a total of 2 variables associated with survival in TCGA-AML patients were screened, namely DNAJB6 and HSPB1 (**Figure 5A**). HR > 1 for HSPB1 (HR = 1.2222, 95% CI: 0.992–1.506, $p = 0.059$) may be an oncogenic gene in AML, whereas DNAJB6 (HR = 0.514, 95% CI: 0.345–0.767, $p = 0.001$) with HR < 1 was expected to be a protective factor for AML. Further, the Cox model consisting of DNAJB6 and HSPB1 (**Figure 5B**) was identified as the optimal prognostic signature for AML by sophisticated calculations of multivariate Cox analysis with stepwise regression.

Evaluation of a Two-Gene Prognostic Signature-based Risk System

We assessed the efficacy of prognostic signature consisting of DNAJB6 and HSPB1 for predicting AML prognosis by risk

scoring system. The risk score for each TCGA-AML patient was calculated according to the following formula: risk score = $(-0.65 * \text{expression of DNAJB6}) + (0.19 * \text{expression of HSPB1})$. All TCGA-AML samples were divided into high- ($n = 66$) and low- ($n = 66$) risk groups according to the median risk score (median = 1.029) (**Supplementary Table S12**). **Figure 5C** demonstrated the risk profile and AML survival distribution in the TCGA database, suggesting that low-risk AML patients had relatively longer OS than high-risk patients. Moreover, the heatmap showed that DNAJB6 was relatively less expressed in the high-risk group compared to the low-risk group, while HSPB1 tended to be more highly expressed in the high-risk group (**Figure 5C**). K-M survival analysis confirmed that the low-risk group had better OS and the high-risk score was linked to poor outcomes (**Figure 5D**). The ROC curve assessed the accuracy of the risk score to predict OS in TCGA-AML patients at 1, 3, and 5 years 1 year OS had an AUC of 0.726, 3 years of 0.634, and 5 years of 0.741 (**Figure 5E**). These results indicated that our risk score had a more reliable performance in predicting the prognosis of AML patients.

Validation of the Two-Gene Prognostic Signature in the GEO Database

The risk scores for AML patients in the GSE37642 ($n = 136$), GSE12417 ($n = 73$), and GSE71014 ($n = 104$) datasets were

calculated using the above equations, and the samples were divided into high-risk and low-risk groups based on the optimal threshold for risk scores (Figures 6A–C). The predictive validity of the risk scores in the independent external validation set was consistent with that in TCGA database. Patients in the low-risk group had significantly

longer survival times compared with the high-risk group (Figures 6D–F). Meanwhile, ROC curve analysis showed that the AUC of risk score in predicting patients’ OS at 1, 3, and 5 years was 0.631, 0.682, and 0.675 in the GSE37642 dataset (Figure 6G). 0.640, 0.621, and 0.589, respectively, in the GSE12417 dataset (Figure 6H). 0.652, 0.782, and 0.766,

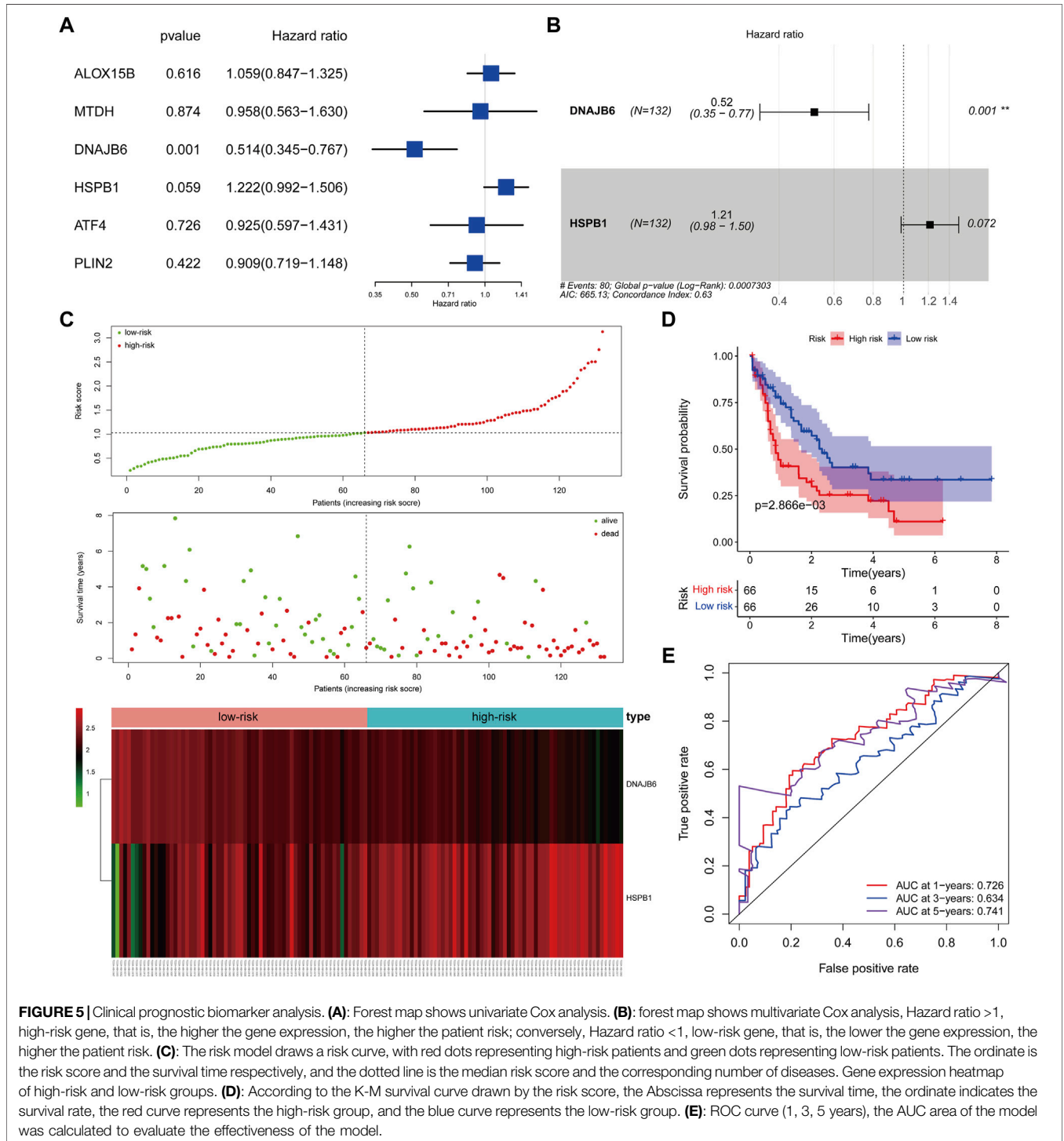


FIGURE 5 | Clinical prognostic biomarker analysis. **(A):** Forest map shows univariate Cox analysis. **(B):** forest map shows multivariate Cox analysis, Hazard ratio > 1, high-risk gene, that is, the higher the gene expression, the higher the patient risk; conversely, Hazard ratio < 1, low-risk gene, that is, the lower the gene expression, the higher the patient risk. **(C):** The risk model draws a risk curve, with red dots representing high-risk patients and green dots representing low-risk patients. The ordinate is the risk score and the survival time respectively, and the dotted line is the median risk score and the corresponding number of diseases. Gene expression heatmap of high-risk and low-risk groups. **(D):** According to the K-M survival curve drawn by the risk score, the Abscissa represents the survival time, the ordinate indicates the survival rate, the red curve represents the high-risk group, and the blue curve represents the low-risk group. **(E):** ROC curve (1, 3, 5 years), the AUC area of the model was calculated to evaluate the effectiveness of the model.

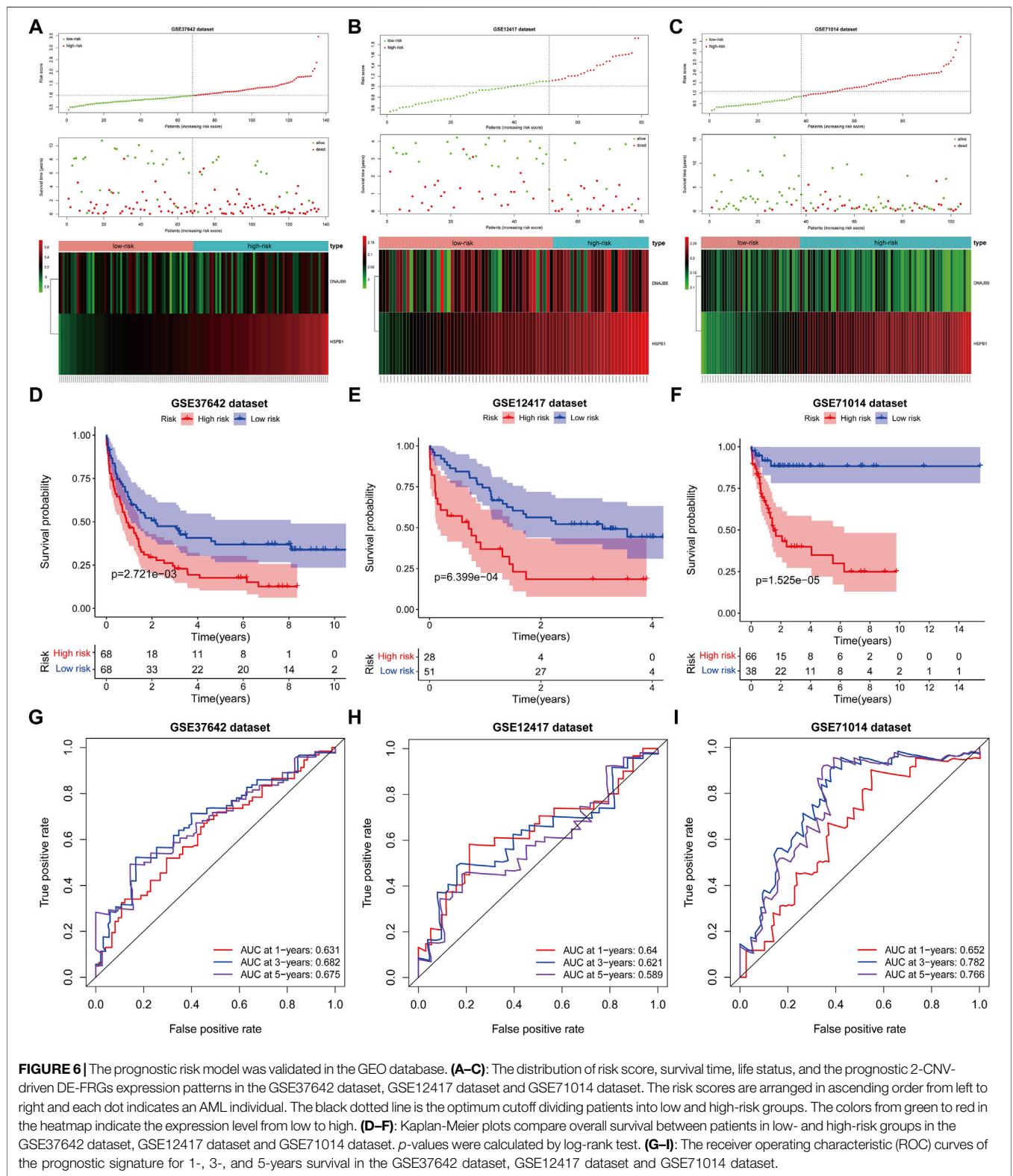


FIGURE 6 | The prognostic risk model was validated in the GEO database. **(A–C)**: The distribution of risk score, survival time, life status, and the prognostic 2-CNV-driven DE-FRGs expression patterns in the GSE37642 dataset, GSE12417 dataset and GSE71014 dataset. The risk scores are arranged in ascending order from left to right and each dot indicates an AML individual. The black dotted line is the optimum cutoff dividing patients into low and high-risk groups. The colors from green to red in the heatmap indicate the expression level from low to high. **(D–F)**: Kaplan-Meier plots compare overall survival between patients in low- and high-risk groups in the GSE37642 dataset, GSE12417 dataset and GSE71014 dataset. *p*-values were calculated by log-rank test. **(G–I)**: The receiver operating characteristic (ROC) curves of the prognostic signature for 1-, 3-, and 5-years survival in the GSE37642 dataset, GSE12417 dataset and GSE71014 dataset.

TABLE 1 | Univariate Cox analysis.

	HR	HR.95L	HR.95H	p-value
Age	1.03482198014829	1.01882519224958	1.05106993696686	1.66E-05
Gender	0.962366615022004	0.619557528552139	1.49485634348315	0.864443099593092
Risk score	2.55281810175952	1.70387284566633	3.82474565355407	5.54E-06

TABLE 2 | Multivariate Cox analysis.

	HR	HR.95L	HR.95H	p-value
Age	1.03132276756969	1.01480193820568	1.04811255365579	0.000181619864716146
Risk score	2.26714423194939	1.47977186168049	3.47346986489139	0.000169704374503001

respectively, in the GSE71014 dataset (Figure 6I). This evidence suggested that our risk score had more satisfactory predictive validity and some general applicability.

Risk Score was an Independent Prognostic Factor for AML

We explored the relationship between risk score and clinical characteristics (age and gender) in the TCGA-AML dataset by the Wilcoxon test. The results revealed that the risk score was significantly different between age subgroups ($p = 0.02$), and the risk score was positively correlated with age, with higher risk levels at older ages (Supplementary Figure 1A). Besides, the risk scores were higher in the male subgroup than in the female group, but which was not statistically significant ($p = 0.084$; Supplementary Figure 1B).

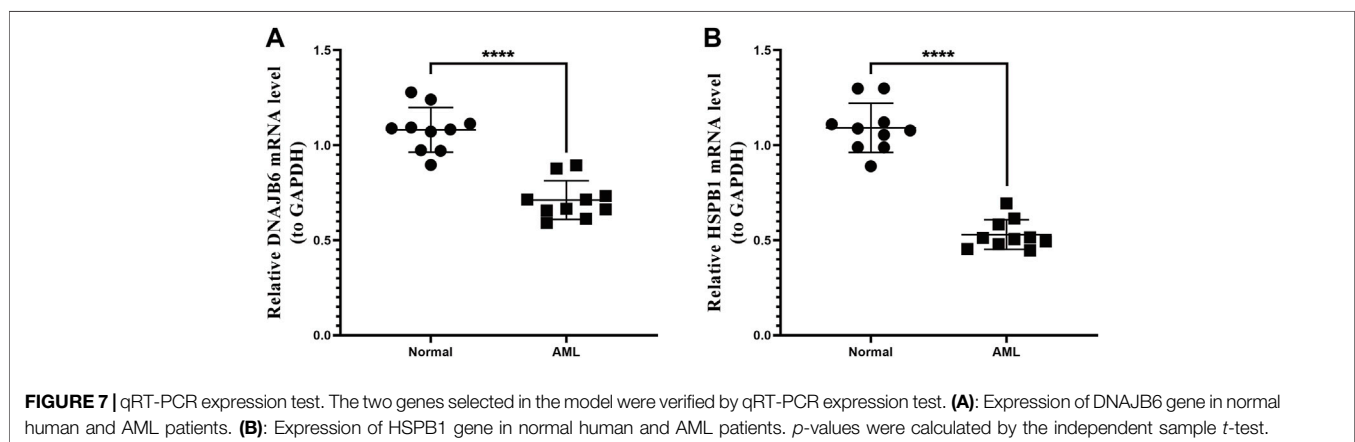
Further, Cox regression analysis was utilized to evaluate whether the risk score could predict the outcome of AML patients independently of clinical characteristics (age and gender). Univariate Cox analysis noted that age and risk score were significantly associated with prognosis in AML patients ($p < 0.05$; Table 1). Ultimately, multivariate Cox analysis indicated that age and risk score were the independent factors of AML prognosis (Table 2).

Validation of Prognostic Gene Expression in PCR Expression Test

We collected the blood of 10 AML patients and 10 healthy individuals and analyzed the mRNA expression levels of two prognostic genes by PCR. The results showed that the expression of the DNAJB6 gene and the HSPB1 gene in AML patients were both lower than that in healthy people ($p < 0.0001$; Figures 7A, P < 0.0001; Figure 7B). There was a significant statistical difference in the expression level after statistical analysis.

DISCUSSION

With the progress of AML treatment such as the combination of chemotherapy and stem cell transplantation, the outcomes of AML patients have great improvement. However, the prognosis of AML patients still cannot be accurately judged. The defect of apoptosis is a common cause of chemoresistance (Poeta et al., 2008). Ferroptosis is different from apoptosis which can provide us with new ideas for inducing cancer cell death (Radtke et al., 2009). Some reports have determined that leukemia cells are more sensitive to the ferroptosis inducer erastin than other cancer cell types (Seth and Singh, 2015; Reyna et al., 2017). In addition, CNV has been reported to be associated with chemotherapy response



and many studies on CNV in AML have been carried out (Song et al., 2016; Shen et al., 2018; Jiang et al., 2020; Shao et al., 2021). However, none of them has comprehensively evaluated the role of gene expression derived from CNV. Both ferroptosis and CNV can be used as prognostic markers, but the relationship between them in AML has not been reported. Therefore, we combined ferroptosis and CNV aiming to improve the prognostic prediction and management efficiency in AML patients.

We combined analysis of CNV differential data and differentially expressed genes (DEGs) data to identify key CNV-driven FRGs for AML by using publicly available AML datasets. A total of 6 CNV-driven DE-FRGs (ALOX15B, MTDH, DNAJB6, HSPB1, ATF4, and PLIN2) were identified by overlap analysis, and functional enrichment analysis indicated that they were involved in oxidative stress, cell death, and inflammatory response processes. Finally, 2 CNV-driven FRGs (DNAJB6 and HSPB1) were identified by Univariate analysis and COX model. The OS of patients in the high-risk group was significantly shorter in the datasets compared to the low-risk group.

DNAJB6 encodes a highly conserved DNAJ/Hsp40 family chaperone protein and interacts with Hsp70 chaperone protein (Stevens et al., 2019). Glutathione peroxidase 4 (GPX4, an antioxidant enzyme) is a defensive protein which is a kind of the GSH peroxidase (Sun et al., 2015). GPx4 inhibits lipid reactive oxygen species (ROS) production by decreasing phospholipid hydroperoxide and plays an important role in inhibiting iron cell apoptosis (Reyna et al., 2017). Overexpressing DNAJB6a can has the ability to downregulate GPX4 and promote ferroptosis (Tang et al., 2011). It is reported that DNAJB6 expression was downregulated in ESCC tissues and it acts as an anti-oncogene in ESCC (Tang et al., 2011). In addition, Mitra A et al. also reported that DNAJB6a can weaken malignant activity of breast carcinoma (Vosberg et al., 2016). However, Zhang et al. found that DNAJB6 is an oncogene that can aggravate the invasion of colorectal cancer (Wei et al., 2020). In our study, we found the expression of DNAJB6 oncogene was lower in the low-risk group compared with the high-risk group associated with poor prognosis. At the same time, according to the hazard ratio (HR) of univariate analysis, the HR of DNAJB6 is 0.514, indicating that DNAJB6 is a protective factor in AML. Our result is consistent with previous study about ESCC and breast cancer.

HSPB1 (also named Hsp27), a member of the small heat shock protein family, is involved in regulating cytoskeletal tissue or stabilizing abnormally folded proteins to prevent aggregation (Yang et al., 2013; Yan et al., 2021). Its abnormal expression in cancer is associated with aggressive tumor behavior, increased chemoresistance, and poor prognosis (Carver et al., 2003). It is overexpressed in many cancers such as prostate, breast, gastric, ovarian, bladder and pancreas (Carver et al., 2003). However, our study showed that HSPB1 is downregulated in AML. In addition, the HR of univariate analysis indicated that HSPB1 is a negative factor in AML. The result revealed that the higher the gene expression, the higher the risk of patients. A study demonstrated that HSPB1 is a negative regulator of ferroptotic cancer cell death (Yang et al., 2014). Phosphorylated HSPB1 can not only inhibit apoptosis and induce autophagy (Yu et al., 2012; Yang and Stockwell, 2016), but also reduce cellular iron uptake and lipid

ROS production (Yang et al., 2014). It is of great significance for us to study its treatment for ferroptosis-mediated cancer.

Currently, some studies have found ferroptosis-related gene signatures which can predict prognosis genes that can predict AML (Yu et al., 2015). Huang et al. (Yu et al., 2015) developed a 12 FRG-based prognostic risk model comprised of 10 risk-related genes (GPX4, CD44, FH, CISD1, SESN2, LPCAT3, AIFM2, ACSL5, HSPB1, and SOCS1) and 2 protective genes (ACSL6 and G3BP1) to predict clinical outcomes. The 12 FRGs are divided into 4 categories according to their functions: lipid metabolism (GPX4, LPCAT3, ACSL5, ACSL6), antioxidant (CD44, SESN2, AIFM2), iron metabolism (CISD1, HSPB1), and cancer metabolism (SOCS1, FH, G3BP1). Although their prediction value is better than our result by comparing ROC, it is easy to exclude some prognostic genes of AML by using a single marker to construct a prognostic model. In our study, two genes are CNV-driven ferroptosis-related genes which were obtained by CNV and ferroptosis binding analysis. The prognostic model composed of these two genes is more conducive to clinical analysis and judgement.

Recently, some scholars have studied AML and found that atorvastatin has activity on AML by up regulating comprehensive stress pathway and inhibiting oxidative phosphorylation (Yusuf et al., 2020). This study showed that atorvastatin inhibited the oxygen consumption rate of AML cells, which has specific significance for chemotherapy-resistant AML primordial cells dependent on oxidative phosphorylation. This study found that HSPB1 gene is enriched in the pathway of oxidative stress. The expression of HSPB1 gene can be considered to have an impact on the progression of AML. It can be speculated that there may be a compensation mechanism between oxidative phosphorylation of HSPB1 and atorvastatin quinone. In addition, studies have shown that leukemia cells rely on aldehyde dehydrogenase 3a2 (aldh3a2) enzyme to oxidize long-chain fatty aldehydes to prevent cell oxidative damage (Zhang et al., 2015), but do not rely on normal myeloid cell counterparts. At the same time, aldehyde is a by-product of oxidative phosphorylation and increased nucleotide synthesis in cancer. It is produced by lipid peroxide and is the basis of non-caspase dependent cell death and iron death. At present, the dependence of leukemia cells on aldh3a2 has been observed in a variety of mouse and human myeloid leukemia. In addition, the inhibition of aldh3a2 and GPx4 has comprehensive lethality. GPx4 inhibits lipid ROS and then block the ferroptosis process. GPx4 inhibition is a known trigger of iron death, but GPx4 inhibition itself has little effect on AML cells. Inhibition of aldh3a2 provides a therapeutic opportunity for the unique metabolic state of AML cells, and may become a new idea for the treatment of AML in the future.

HSPB1 gene is enriched in multiple oxidative stress pathways, and it is an iron death driving gene. It can be speculated that GPx4 inhibition and HSPB1 gene expression are related to the development of AML. Studies have shown that DNAJB6 can promote the iron death process of ESCC (Tang et al., 2011). In ESCC, the overexpression of DNAJB6 is accompanied by a significant decrease in GPx4 protein level. The study also shows that DNAJB6 plays an anticancer role in the process of ESCC through iron death mechanism. This

study also shows that DNAJB6 is a protective factor of AML. In addition, a study on bone marrow damage (Zhou and Chen, 2021) showed that Fanconi anemia complement group D2 (FANCD2), a nuclear protein involved in DNA damage repair, can prevent iron apoptosis mediated damage in bone marrow stromal cells (BMSC). Knockout of FANCD2 increases biochemical events related to iron death (e.g., ferrous accumulation, glutathione consumption and malondialdehyde production). Mechanistically, FANCD2 is involved in regulating the expression of genes and/or proteins of iron metabolism (such as fth1, TF, TFRC, HAMP, HSPB1, slc40a1 and steam3) and lipid peroxidation (such as GPx4). However, progressive BMF in FA patients is closely related to AML. Therefore, the mechanism of the linkage of HSPB1, GPx4 and FANCD2 in AML related diseases needs to be further explored.

This study also has some limitations. Our prognostic model was constructed by existing public datasets, although it is validated by PCR expression test, more prospective investigations are needed to validate its predictive power. In the future, we should continue to pay attention to the role of these two genes in AML pathology by experimental and clinical studies. We believe that our prognostic model based on CNV-driven FRGs is of great significance in predicting the survival of AML patients and will offer novel insight for AML research.

DATA AVAILABILITY STATEMENT

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

REFERENCES

- Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., et al. (2016). The 2016 Revision to the World Health Organization Classification of Myeloid Neoplasms and Acute Leukemia. *Blood* 127 (20), 2391–2405. doi:10.1182/blood-2016-03-643544
- Carver, J., Rekas, A., Thorn, D., and Wilson, M. (2003). Small Heat-Shock Proteins and Clusterin: Intra- and Extracellular Molecular Chaperones with a Common Mechanism of Action and Function? *IUBMB Life (International Union Biochem. Mol. Biol. Life)* 55 (12), 661–668. doi:10.1080/15216540310001640498
- Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., et al. (2017). Diagnosis and Management of AML in Adults: 2017 ELN Recommendations from an International Expert Panel. *Blood* 129 (4), 424–447. doi:10.1182/blood-2016-08-733196
- Grimwade, D., Ivey, A., and Huntly, B. J. P. (2016). Molecular Landscape of Acute Myeloid Leukemia in Younger Adults and its Clinical Relevance. *Blood* 127 (1), 29–41. doi:10.1182/blood-2015-07-604496
- Herold, T., Jurinovic, V., Batcha, A. M. N., Bamopoulos, S. A., Rothenberg-Thurley, M., Ksienzyk, B., et al. (2018). A 29-gene and Cytogenetic Score for the Prediction of Resistance to Induction Treatment in Acute Myeloid Leukemia. *Haematologica* 103 (3), 456–465. doi:10.3324/haematol.2017.178442
- Herold, T., Metzler, K. H., Vosberg, S., Hartmann, L., Röhl, C., Stölzel, F., et al. (2014). Isolated Trisomy 13 Defines a Homogeneous AML Subgroup with High Frequency of Mutations in Spliceosome Genes and Poor Prognosis. *Blood* 124 (8), 1304–1311. doi:10.1182/blood-2013-12-540716

AUTHOR CONTRIBUTIONS

CH and JZ contributed to the conception of the study. JZ and FL contributed significantly to analysis and manuscript preparation. CH performed the data analyses and wrote the manuscript. WG and CC helped perform the analysis with constructive discussions. All authors contributed to the article and approved the final manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China [Grant number 81771589], the Program of Tianjin Science and Technology Plan [Grant number 18ZXDBSY00170; 20JCZJC00170], Tianjin Natural Science Foundation [Grant number 21JCYBJC00460], the Public Health and Technology Project of Tianjin [grant number TJWJ2021ZD007] and General project of Tianjin Children's Hospital [grant number Y2020013]. We are grateful for the financial support from the "Tianjin Medical Key Discipline (Specialist) Construction Project".

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.849437/full#supplementary-material>

Supplementary Figure S1 | Clinical significance of the prognostic signature in TCGA set. Risk score in the different ages (A) and gender (B) *P*-values were calculated by the Wilcoxon test.

- Hong, Y., Lin, M., Ou, D., Huang, Z., and Shen, P. (2021). A Novel Ferroptosis-Related 12-gene Signature Predicts Clinical Prognosis and Reveals Immune Relevancy in clear Cell Renal Cell Carcinoma. *BMC Cancer* 21 (1), 831. doi:10.1186/s12885-021-08559-0
- Huang, X., Zhou, D., Ye, X., and Jin, J. (2021). Construction of a Novel Ferroptosis-Related Gene Signature for Predicting Prognosis and Immune Microenvironment in Acute Myeloid Leukemia. *Bosn J. Basic Med. Sci.* doi:10.17305/bjms.2021.6274
- Huot, J., Houle, F., Spitz, D. R., and Landry, J. (1996). HSP27 Phosphorylation-Mediated Resistance against Actin Fragmentation and Cell Death Induced by Oxidative Stress. *Cancer Res.* 56 (2), 273–279.
- Jakob, U., Gaestel, M., Engel, K., and Buchner, J. (1993). Small Heat Shock Proteins Are Molecular Chaperones. *J. Biol. Chem.* 268 (3), 1517–1520. doi:10.1016/s0021-9258(18)53882-5
- Jiang, B., Zhao, Y., Shi, M., Song, L., Wang, Q., Qin, Q., et al. (2020). DNAJB6 Promotes Ferroptosis in Esophageal Squamous Cell Carcinoma. *Dig. Dis. Sci.* 65 (7), 1999–2008. doi:10.1007/s10620-019-05929-4
- Kerr, J. F. R., Winterford, C. M., and Harmon, B. V. (1994). Apoptosis. Its Significance in Cancer and Cancer Therapy. *Cancer* 73 (8), 2013–2026. doi:10.1002/1097-0142(19940415)73:8<2013::aid-cnrcr2820730802>3.0.co;2-j
- Kim, K. I., Kim, T.-k., Kim, I.-W., Ahn, K.-S., Yoon, S.-S., Shin, W. G., et al. (2012). Copy Number Variations in normal Karyotype Acute Myeloid Leukaemia and Their Association with Treatment Response. *Basic Clin. Pharmacol. Toxicol.* 111 (5), 317–324. doi:10.1111/j.1742-7843.2012.00904.x
- Kuett, A., Rieger, C., Perathoner, D., Herold, T., Wagner, M., Sironi, S., et al. (2015). IL-8 as Mediator in the Microenvironment-Leukaemia Network in Acute Myeloid Leukemia. *Sci. Rep.* 5, 18411. doi:10.1038/srep18411

- Li, Z., Herold, T., He, C., Valk, P. J. M., Chen, P., Jurinovic, V., et al. (2013). Identification of a 24-gene Prognostic Signature that Improves the European LeukemiaNet Risk Classification of Acute Myeloid Leukemia: an International Collaborative Study. *Jco* 31 (9), 1172–1181. doi:10.1200/jco.2012.44.3184
- Liang, H., Yoo, S.-E., Na, R., Walter, C. A., Richardson, A., and Ran, Q. (2009). Short Form Glutathione Peroxidase 4 Is the Essential Isoform Required for Survival and Somatic Mitochondrial Functions. *J. Biol. Chem.* 284 (45), 30836–30844. doi:10.1074/jbc.M109.032839
- Liang, L., Fang, J.-Y., and Xu, J. (2016). Gastric Cancer and Gene Copy Number Variation: Emerging Cancer Drivers for Targeted Therapy. *Oncogene* 35 (12), 1475–1482. doi:10.1038/ncr.2015.209
- Meng, E., Shevde, L. A., and Samant, R. S. (2016). Emerging Roles and Underlying Molecular Mechanisms of DNAB6 in Cancer. *Oncotarget* 7 (33), 53984–53996. doi:10.18632/oncotarget.9803
- Mitra, A., Fillmore, R. A., Metge, B. J., Rajesh, M., Xi, Y., King, J., et al. (2008). Large Isoform of MRJ (DNAB6) Reduces Malignant Activity of Breast Cancer. *Breast Cancer Res.* 10 (2), R22. doi:10.1186/bcr1874
- Nibourel, O., Guihard, S., Roumier, C., Pottier, N., Terre, C., Paquet, A., et al. (2017). Copy-number Analysis Identified New Prognostic Marker in Acute Myeloid Leukemia. *Leukemia* 31 (3), 555–564. doi:10.1038/leu.2016.265
- Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* 374 (23), 2209–2221. doi:10.1056/NEJMoa1516192
- Poeta, G., Bruno, A., Del Principe, M., Venditti, A., Maurillo, L., Buccisano, F., et al. (2008). Deregulation of the Mitochondrial Apoptotic Machinery and Development of Molecular Targeted Drugs in Acute Myeloid Leukemia. *Cdt* 8 (3), 207–222. doi:10.2174/156800908784293640
- Radtke, I., Mullighan, C. G., Ishii, M., Su, X., Cheng, J., Ma, J., et al. (2009). Genomic Analysis Reveals Few Genetic Alterations in Pediatric Acute Myeloid Leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 106 (31), 12944–12949. doi:10.1073/pnas.0903142106
- Reyna, D. E., Garner, T. P., Lopez, A., Kopp, F., Choudhary, G. S., Sridharan, A., et al. (2017). Direct Activation of BAX by BSA1 Overcomes Apoptosis Resistance in Acute Myeloid Leukemia. *Cancer Cell* 32 (4), 490–505. doi:10.1016/j.ccell.2017.09.001
- Seth, R., and Singh, A. (2015). Leukemias in Children. *Indian J. Pediatr.* 82 (9), 817–824. doi:10.1007/s12098-015-1695-5
- Shao, R., Wang, H., Liu, W., Wang, J., Lu, S., Tang, H., et al. (2021). Establishment of a Prognostic Ferroptosis-related Gene Profile in Acute Myeloid Leukemia. *J. Cel Mol Med* 25 (23), 10950–10960. doi:10.1111/jcmm.17013
- Shen, Z., Song, J., Yung, B. C., Zhou, Z., Wu, A., and Chen, X. (2018). Emerging Strategies of Cancer Therapy Based on Ferroptosis. *Adv. Mater.* 30 (12), 1704007. doi:10.1002/adma.201704007
- Song, X., Xie, Y., Kang, R., Hou, W., Sun, X., Epperly, M. W., et al. (2016). FANCD2 Protects against Bone Marrow Injury from Ferroptosis. *Biochem. Biophysical Res. Commun.* 480 (3), 443–449. doi:10.1016/j.bbrc.2016.10.068
- Stevens, A. M., Xiang, M., Heppler, L. N., Tošić, I., Jiang, K., Munoz, J. O., et al. (2019). Atovaquone Is Active against AML by Upregulating the Integrated Stress Pathway and Suppressing Oxidative Phosphorylation. *Blood Adv.* 3 (24), 4215–4227. doi:10.1182/bloodadvances.2019000499
- Sun, X., Ou, Z., Xie, M., Kang, R., Fan, Y., Niu, X., et al. (2015). HSPB1 as a Novel Regulator of Ferroptotic Cancer Cell Death. *Oncogene* 34 (45), 5617–5625. doi:10.1038/ncr.2015.32
- Tang, D., Kang, R., Livesey, K. M., Kroemer, G., Billiar, T. R., Van Houten, B., et al. (2011). High-mobility Group Box 1 Is Essential for Mitochondrial Quality Control. *Cel Metab.* 13 (6), 701–711. doi:10.1016/j.cmet.2011.04.008
- Vosberg, S., Herold, T., Hartmann, L., Neumann, M., Opatz, S., Metzler, K. H., et al. (2016). Close Correlation of Copy Number Aberrations Detected by Next-Generation Sequencing with Results from Routine Cytogenetics in Acute Myeloid Leukemia. *Genes Chromosomes Cancer* 55 (7), 553–567. doi:10.1002/gcc.22359
- Wei, J., Xie, Q., Liu, X., Wan, C., Wu, W., Fang, K., et al. (2020). Identification the Prognostic Value of Glutathione Peroxidases Expression Levels in Acute Myeloid Leukemia. *Ann. Transl. Med.* 8 (11), 678. doi:10.21037/atm-20-3296
- Yan, H.-f., Zou, T., Tuo, Q.-z., Xu, S., Li, H., Belaidi, A. A., et al. (2021). Ferroptosis: Mechanisms and Links with Diseases. *Sig Transduct. Target. Ther.* 6 (1), 49. doi:10.1038/s41392-020-00428-9
- Yang, L., Cao, L., Yang, M., Tang, D., Kang, R., Min, X., et al. (2013). Hsp27: a Novel Therapeutic Target for Pediatric M4/M5 Acute Myeloid Leukemia. *Oncol. Rep.* 29 (4), 1459–1466. doi:10.3892/or.2013.2274
- Yang, W. S., SriRamaratnam, R., Welsch, M. E., Shimada, K., Skouta, R., Viswanathan, V. S., et al. (2014). Regulation of Ferroptotic Cancer Cell Death by GPX4. *Cell* 156 (1–2), 317–331. doi:10.1016/j.cell.2013.12.010
- Yang, W. S., and Stockwell, B. R. (2016). Ferroptosis: Death by Lipid Peroxidation. *Trends Cel Biol.* 26 (3), 165–176. doi:10.1016/j.tcb.2015.10.014
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Yu, Y., Xie, Y., Cao, L., Yang, L., Yang, M., Lotze, M. T., et al. (2015). The Ferroptosis Inducer Erastin Enhances Sensitivity of Acute Myeloid Leukemia Cells to Chemotherapeutic Agents. *Mol. Cell Oncol.* 2 (4), e1054549. doi:10.1080/23723556.2015.1054549
- Yusuf, R. Z., Saez, B., Sharda, A., van Gestel, N., Yu, V. W. C., Baryawno, N., et al. (2020). Aldehyde Dehydrogenase 3a2 Protects AML Cells from Oxidative Death and the Synthetic Lethality of Ferroptosis Inducers. *Blood* 136 (11), 1303–1316. doi:10.1182/blood.2019001808
- Zhang, T.-T., Jiang, Y.-Y., Shang, L., Shi, Z.-Z., Liang, J.-W., Wang, Z., et al. (2015). Overexpression of DNAB6 Promotes Colorectal Cancer Cell Invasion through an IQGAP1/ERK-dependent Signaling Pathway. *Mol. Carcinog.* 54 (10), 1205–1213. doi:10.1002/mc.22194
- Zhou, F., and Chen, B. (2021). Prognostic Significance of Ferroptosis-Related Genes and Their Methylation in AML. *Hematology* 26 (1), 919–930. doi:10.1080/16078454.2021.1996055

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

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

Copyright © 2022 Han, Zheng, Li, Guo and Cai. 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.