An Immunogenic Cell Death-Based Risk model Predicts Prognosis and Indicates Immune Infiltration Landscape in Acute Myeloid Leukemia

Guangxin Ma  
Qilu Hospital of Shandong University

Nana Wang  
Qilu Hospital of Shandong University

Xiaoran Bai  
Qilu Hospital of Shandong University

Dongmei Wang  
Qilu Hospital of Shandong University

Jingjing Ye  
Qilu Hospital of Shandong University

Fei Lu  
Qilu Hospital of Shandong University

Chunyan Ji (✉ jichunyan@sdu.edu.cn)  
Qilu Hospital of Shandong University

Research Article

Keywords: Immunogenic cell death, IncRNA, prognosis, acute myeloid leukemia.

Posted Date: November 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2247524/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

**Background:** Acute myeloid leukemia (AML) is the most common acute leukemia in adults with a high mortality rate. Immunogenic cell death (ICD) plays a crucial role in activation of adaptive immune response and may contribute to the efficacy of cancer immunotherapy. However, the relationship between ICD and AML prognosis is unveiled.

**Methods and materials:** A Pearson correlation analysis was utilized to identified ICD-related IncRNAs. Univariate cox regression analysis and subsequent LASSO analysis were performed to construct an ICD-associated IncRNAs signature. Survival analysis, ROC analysis, univariate and multivariate cox regression were applied to assess the predictive capacity and evaluate prognostic value for AML patients. ESTIMATE, CIBERSOT, and single sample gene set enrichment analysis (ssGSEA) algorithms were performed to estimate the immune infiltration landscape. Enrichment analysis was used to investigate the biological processes and pathways of the ICD-associated IncRNAs.

**Results:** A predictive risk signature was constructed based on seven ICD-associated IncRNAs (AFF2-IT1, AL5924292, LINC00987, MIR133A1HG, AC022182.2, NORAD and AC244502.1). High risk score was verified as an independent prognostic predictor for poor clinical outcomes in AML patients. Notably, we observed a remarkable difference in immune infiltration landscape, immunotherapy response and drug susceptibility related to risk stratification. In addition, functional enrichment analysis established that immune-related signaling pathways might mediate the role of ICD-related IncRNAs in AML.

**Conclusions:** The signature based on ICD-related IncRNAs can provide guidance to the accurate prediction of AML prognosis and also offer a novel perspective for individualized and precise treatment strategies for AML patients.

Introduction

AML is characterized by impaired hematopoietic function and bone marrow (BM) failure, generating fatal outcomes with acquired somatic cell genetic damage and the accumulation of hematopoietic progenitor cells [1]. Although great achievements have been achieved in novel targeted agents for AML personalized precise treatment, the long-term survival and complete remission (CR) rates remains unsatisfying[2]. Thus, identifying new reliable biomarkers for AML diagnosis, prognostic stratification and the possibility of targeted drug development is urgently required[3].

ICD is a type of regulated cell death, which can enhance tumor-specific immune responses by the release of tumor specific antigen (TSA), tumor associated antigen (TAA), as well as the danger-associated molecular patterns (DAMPs) from dying cancer cells [4, 5]. The DAMPs are consisted of adenosine triphosphate (ATP), calreticuline (CRT), high mobility group box protein B1 (HMGB1), heat shock protein (HSP), type interferon (IFN ) and Annexin 1 (ANXA)[4, 6]. Certain chemotherapeutic drugs (mitoxantrone and oxaliplatin), physicochemical therapies, photodynamic therapy or radiotherapy can lead to the occurrence of ICD, which in turn stimulates or enhances the anti-tumor immune response by immune cells.
recruitment and activation [7–9]. Accumulated preclinical and clinical evidences show ICD process is a promising effective therapy target for a variety of tumor types[10, 11]. However, the regulation of ICD remains unclear and is far from clinical appliance in AML.

Long non-coding RNAs (lncRNAs) are RNA molecules with a transcript length of more than 200 nucleotides and cannot encode proteins owing to the lack of an open reading frame[12]. LncRNA functions to regulate gene expression at various levels (epigenetic, transcriptional or post-transcriptional regulation, etc.)[13]. The dysregulation of some lncRNAs may be associated with the recurring mutations, clinical features and prognosis of AML, such as IncRNA UCA1, LINC00152 and LncRNA HOTTIP [14–16]. According to the recent literature, ICD-related IncRNAs including AC131391.1, PVT1, LINC00592, AL139147.1, and VCAN-AS1 have been illustrated to be involved in the process of prognostic prediction and individual treatment guidance of stomach adenocarcinoma patients [17]. Furthermore, specific lncRNAs combined with chemotherapeutic agents can improve the treatment effectiveness [18]. Thus, lncRNAs can be used as new therapeutic targets and biomarkers for AML treatment. Whereas the ICD-related lncRNA functions have not been fully discovered so far in AML.

In this study, we systematically investigated the prognostic significance of ICD-related lncRNAs in AML and constructed a corresponding prognostic risk signature with seven ICD-related IncRNAs. We further explored the underlying mechanism of prognostic signature in AML by function and pathway enrichment analysis. Immune infiltration and chemotherapy drug sensitivity were also investigated. The results indicated the risk signature not only provided a preliminary exploration of the ICD related mechanisms but also offered a new perspective and insight for individualized and precise treatment strategies for AML patients.

**Methods And Materials**

**Data collection**

The transcriptome information and relevant clinical data from a total of 151 patients with AML were downloaded from the TCGA (https://portal.gdc.cancer.gov/) database. The transcriptome information was collected in the fragment per kilobase million (FPKM) format that has been normalized through the Perl programming language (version Strawberry-Perl-5.30.0; https://www.perl.org). Meanwhile, participants with incomplete clinical information were excluded from the study.

**Identification of ICD-related IncRNAs (IRLs)**

A total of 33 ICD-related genes (IRGs) were collected from the previous research ([19], Supplementary Table 1). To identify IRLs, Pearson’s correlation analysis was subsequently conducted to investigate the relationship between IRLs and IRGs according to the criteria of |Pearson R| > 0.5 and \( p < 0.001 \).

**Construction an ICD-associated IncRNAs model**
Using the univariate Cox regression analysis to identify the prognostic IRLs and the least absolute shrinkage and selection operator (LASSO) algorithm was employed to calculate the minimum lambda of the prognostic IRLs via R package “glmnet”. Then, multivariate Cox regression analysis was conducted to identify the candidate prognostic IRLs and calculated the ICD score for AML patients. The ICD score of each sample was calculated according the following formula risk score = AFF2 – IT1 * (-0.283) + AL5924292 * (0.094) + LINC00987 * (0.272) + MIR133A1HG * (-0.231) + AC022182.2 * (0.08) + NORAD * (-0.030) + AC244502.1 * (-0.024). According to the median risk score, all AML samples were divided into high-risk and low-risk groups. Kaplan–Meier analysis was adopted to evaluate the difference of survival outcomes between the high- and low-risk subsets using the “survival” R package. $p < 0.05$ was considered as statistical difference. The principal component analysis (PCA) was used to investigate the distribution pattern of patients in low- and high-risk group based on the prognostic IRLs using R package “ggplot2”.

**Independent prognostic analysis and ROC curve plotting**

The risk score and clinical characteristics were merged to screen independent prognostic variable factors through univariate and multivariate Cox regression analysis. Additionally, the time-dependent receiver operating characteristic (ROC) curves were adopted to compare different factors in predictive accuracy via R package “timeROC”.

**Construction and assessment of nomogram**

A hybrid nomogram involved risk signature and several clinicopathological characters was created to estimate the 1-, 3-, and 5-year OS rates for AML patients by “rms” package. The calibration curve was performed to assess the predictive ability of the nomogram.

**Infiltrated immune cells and tumor microenvironment (TME) analysis**

The CIBERSORTx algorithm was performed to quantify the proportions of immune cells in AML samples. Immune functions, immune cell expression and immune checkpoints populations were compared between the high and low-risk groups. ESTIMATE algorithm was adopted to evaluate the ratio of immune-stromal components in each sample in the tumor microenvironment and calculated the stromal score, immune score, ESTIMATE score (stromal score + immune Score) and tumor purity via R package “estimate”. Finally, tumor immune dysfunction and exclusion (TIDE) algorithm(http://tide.dfci.harvard.edu/) was conducted to predict the reactions to immunotherapy in AML patients. Statistical significance was set at $p$-value $< 0.05$.

**Drug sensitivity**

Drug sensitivity (IC50) is a significant marker for evaluating drug response to treatment, which was conducted to predict the drug response of each sample in low- and high-risk groups based on the Genomics of Drug Sensitivity in Cancer (GDSC) database, using R package “pRRophetic”.

**Function enrichment analysis.**
The R package “limma” was used to analyze the differently expressed genes (DEGs) between the high- and low-risk score groups. Genes with adjusted |Fold Change| ≥ 2 and p-value < 0.05 were considered statistically significantly different. Then, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were conducted for functional enrichment analysis using the “org.Hs.eg.db” package and “clusterProfiler package”. When p < 0.05 and FDR < 0.05, the biological processes and pathways were considered significant. Additionally, metascape (https://metascape.org) also provided a tool to further validate the function enrichment.

Statistical analysis

All statistical analysis were performed using R software (version 4.1.0) and Perl software. Spearman-ranked correlation analysis was applied to investigate the correlation between ICD score and IC50, with p-value < 0.05 was considered significantly different. Differential functions were analyzed using the Wilcoxon rank-sum test between the two groups, and statistical significance was set at p-value < 0.05.

Results

Identification of prognostic IRLs in AML

A total of 455 IncRNAs associated with IRGs were identified as IRLs in this study (Fig. 1A, Supplementary Table 2). Based on the univariate Cox regression analysis, 65 prognostic IncRNAs in AML cohort was screened (Supplementary Table 3) and subsequently subjected to the LASSO model to calculate the optimal coefficients and selected for the optimal prognostic IncRNAs for next analysis (Fig. 1B-C). Multivariate Cox regression analysis result suggested that 7 prognostic IRLs which could independently evaluate prognosis of AML were selected to construct the risk model.

Risk model construction based on IRLs prognostic signature in AML

According to the 7 prognostic IRLs, the risk score of each AML patient were calculated, and divided into the low- and high-risk group based on the median risk score. The risk curve and scatterplot results illustrated the mortality occurrence was positively associated with risk score (Fig. 2A). Kaplan-Meier survival analysis showed shorter overall survival (OS) in the high-risk group compared to low-risk group (p = 2.575e-07) (Fig. 2B), suggesting that the risk score has prognostic value for AML patients. The results of PCA illustrated that AML patients could be greatly distinguished based on 7 IncRNAs of the prognostic signature (Fig. 2C–D). The result of heatmap diagram displayed increased expression of AL5924292, AC022182.2 and LINC00987 in the high-risk group, while NORAD, AC244502.1, MIR133A1HG and AFF2–IT1 were at higher expression level in the low-risk group (Fig. 2E). The detailed statistical expression of the 7 prognostic IRLs in AML samples were further validated in Fig. 2.

Independent prognostic analysis of risk score and pathological features
To investigate whether the risk model based on the IRLs prognostic signature was an independent prognosis indicator for AML patients, univariate and multivariate Cox regression analyses were performed on the TCGA cohort. Result from univariate Cox regression analysis indicated age and risk score (all \( p < 0.001 \)) predicted dismal OS (Fig. 3A). The multivariate Cox regression analysis validated the independence of risk score and age for predicting AML prognosis (Fig. 3B). To optimize the predictive accuracy of the risk model, we then created a hybrid nomogram included risk score and clinicopathological features to predict OS rates for AML patients at 1-, 3-, and 5-year (Fig. 3C). The 1-, 3-, and 5-year calibration curves of nomogram exhibited the good consistency with the actual observation, suggesting that the proposed model performed similarly to an ideal model (Fig. 3D). The time-dependent ROC analysis was also used to evaluate the predictive reliability of our prognostic signature. The area under curve (AUC) at the 1-, 3-, and 5-year was 0.803, 0.804, and 0.879, respectively (Fig. 3E). At the 1-year ROC of the model, the AUC of risk score was 0.803, demonstrating the risk signature was more predictive than other pathological features, such as age and gender (Fig. 3F).

**Association of IRLs prognostic signature and immune microenvironment landscape**

To analyze the immune cell infiltration of AML patients in low- and high-risk group, we adopted IBERSORTx algorithm to quantify the fraction of immune cells and enrichment scores of immune-related functional pathways between the high- and low-risk groups. Result from immune cell bubble graphs showed that samples from the high-risk group were significantly positively correlated with 18 types of immune cells, such as regulatory T cells (Treg), T cell follicular helper cells, myeloid derived suppressor cells (MDSC), neutrophils and others (Fig. 4A). Immune function result suggested that the patients with high-risk score had higher immune function score compared to those with low-risk score, such as APC co-inhibition, chemokine receptors (CCR) and type I and II interferon responses (Fig. 4B). Moreover, the differences in immune checkpoints also were analyzed between different groups due to the significance of checkpoint-based immunotherapy (Fig. 4C). *BTLA, LAG3, CTLA4, PD-1, PDCD1LG2* and *PD-L1* were expressed higher in the high-risk group, which may provide a rational explanation for the poorer outcome in the high-risk group. In terms of TME scores, high-risk patients showed higher immune, stromal, and ESTIMATE scores than low-risk patients, while tumor purity were higher in low-risk patients (Fig. 4D–G), suggesting the levels in immune cells infiltration and stroma component in the TME were positively correlated with risk score. Additionally, we adopted TIDE algorithm to demonstrate the predictive power of risk scores for immunotherapy. Compared to the low-risk group, the result showed that TIDE scores were dramatically higher in the high-risk group (Fig. 4H), indicating low-risk group was more responsive to immunotherapy.

**Correlation analysis of ICD score and drug sensitivity**

Despite of drug resistance, targeted drug therapy and chemotherapy are the main treatments for AML patients. Therefore, we investigated the association between the antineoplastic drug sensitivity and risk score by comparing the IC50 values. The IC50 of WZ-1-84 and TGX221 in high-risk group were significantly lower than in low-risk group, whereas the IC50 of BI-2536 and VX-680 were higher in high-risk
The correlation of risk score and drug sensitivity indicated that the risk score was significantly negatively correlated with WZ-1-84 (R = -0.34, p = 7.3e-05) and TGX221 (R = -0.29, p = 0.00079), but positively associated with BI-2536 (R = 0.29, p = 0.00071) and VX-680 (R = 0.31, p = 0.00031) (Fig. 5E-H). These results indicated that the response to antineoplastic drugs differed according to the two risk groups. **Fig. 5.** Drug sensitivity analysis in ICD-high and ICD-low group. IC50 values shows a significant difference in ICD-high and ICD-low group among **A** WZ-1-84, **B** TGX221, **C** BI-2536, **D** VX-680. 

**Function and pathway enrichment analysis**

To explore potential biological process and signaling pathways of the ICD-associated risk signature, DEGs between the high- and low-risk subgroups were screened and exhibited in Fig. 6A. Then, these DEGs were used for functional enrichment analysis by GO analysis and KEGG pathway analysis. Results of GO showed that DEGs in the high-risk and low-risk groups were associated with many immune-related functions such as leukocyte mediated immunity, antigen processing and presentation of peptide or antigen processing and polysaccharide antigen via MHC class II, immune receptor activity, cytokine-mediated signaling pathway and some MHC class II and I-related function (Fig. 6B). The KEGG enrichment analysis suggested that DEGs tend to be enriched in many pathways such as cell adhesion molecules, hematopoietic cell lineage, antigen processing and presentation, as well as the allograft rejection (Fig. 6D). Moreover, the enrichment analysis by Metascape dataset showed that these ICD-related DEGs are mainly enriched in immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell, negative regulation of immune system process, immune response-regulating signaling pathway, response to immune effector process, regulation of interleukin-12, interferon alpha/beta signaling, and chemokine production and regulation of immune effector process (Fig. 6C,6E). In conclusion, these results illustrate that the ICD-associated genes may have an impact on AML patients through immune-related processes and pathways.

**Discussion**

AML is the most common type of acute leukemia in adults, and its prevalence and mortality are increasing worldwide due to its unclear pathogenesis and complex genetic mutations that affect clinical prognosis and provide potential targets for drug development [20]. Early detection and risk stratification are essential to improve AML outcome[21]. In this study, we first constructed a risk model based on seven IRLs, and preliminarily investigated the possible mechanisms involved in this process. We also sought to explore the relationship between prognostic models that predict prognosis in AML and the immune landscape. Based on the different expressions of immune checkpoint markers and antineoplastic drugs sensitivities, we provide a new insight into AML individualizing therapy.

LncRNA plays an important role in occurrence, development, and prognosis of AML [22]. Among seven IncRNAs in the prognostic risk model, the IncRNA LINC00987 and MIR133A1HG deregulation were demonstrated significantly correlated with prognosis and tumor progression in AML [23, 24], indicated that LINC00987 and MIR133A1HG may serve as new promising individualized therapeutic targets for the
treatment of AML [25]. Our data reaffirmed its potential value as therapeutic targets in AML. NORAD is a newly characterized IncRNA activated by DNA damage and implicated in maintaining genomic stability and regular mitosis [26]. In non-M3 AML patients, upregulation of NORAD was illustrated significantly associated with poor overall survival and unfavorable cytogenetic risk[27]. However, the underlying mechanism of NORAD has not been fully clarified, and more studies are urgently needed to comprehensive explore precise molecular mechanisms. The expression pattern, clinical relevance and underlying mechanisms of AC244502.1, AFF2 – IT1, AL5924292 and AC022182.2 in AML patients is still poorly understood which need further validation in the future.

Immune regulation is markedly involved in progression and clinical outcomes of AML. Therefore, the use of immunotherapy to produce an effective immune response to the tumor to delay the development of cancer is considered to be effective in AML [28]. The number and proportion of infiltrating immune cells are regarded as important factors affecting cancer development and immunotherapy response and closely related to patient outcomes [29]. In our results, immunosuppressive cell types such as Treg and MDSCs were significantly increased in high-risk AML group, which could generate the immunosuppressive tumor microenvironment[30, 31]. Treg-mediated immune-suppression plays critical part in tumor immune evasion, may be responsible for the incidence and development of AML [32, 33]. MDSCs, as a significant role in tumor progression, is correlated with poor outcome and decreased efficiency of immunotherapy [34]. MDSCs mediated immunosuppression can contribute to immune evasion by depleting necessary nutrients for lymphocytes to function, inhibiting T cell proliferation and survival via generation of oxidative stress[35]. The above results suggested that the higher immunosuppression in the tumor microenvironment of high-risk patients contribute to tumor progression and poorer prognosis and response of immunotherapy. In addition, TIDE analysis between different groups was utilized to evaluate the response of immunotherapy [36]. The higher the TIDE score is more frequently associated with immune escape, which may suggest a limited response and shorter survival time for patients treated with immune checkpoint inhibitors [37]. The low-risk group show a better potential immunotherapy response, which is in line with previous report. The influence of ICD classification on prognosis was explained by combining with the significance of immunotherapy in clinical application.

In conclusion, we successfully constructed a predictive signature in AML based on seven IRLs, which may offer novel perspectives in seeking immunotherapy for AML patients. Meanwhile, we believe that these findings will increase the understanding about ICD and have the potential to promote more individual and precise therapeutic treatment in AML.

Declarations

Acknowledgements

Not applicable.
Funding

This work was supported by grants from the Distinguished Taishan Scholars in Climbing Plan (tspd20210321), the National Natural Science Foundation of China (82070160, 82000165, 82170182), the Major Research Plan of the National Natural Science Foundation of China (91942306), the 68th China postdoctoral Science Foundation (2020M682171), the Key Program of Natural Science Foundation of Shandong Province (ZR2020KH016, ZR2021MH302), the Fundamental Research Funds for the Central Universities (2022JC012), the Independently Cultivate Innovative Teams of Jinan, Shandong Province (2021GXRC050) and the Clinical Practical New Technology and Development Fund of Qilu Hospital, Shandong University (2019-5).

Conflicts of interest

The authors declare that they have no conflicts of interest regarding the publication of this study.

Data availability

The datasets generated for this study can be found in the TCGA database (https://tcga-data.nci.nih.gov/tcga/).

Authors’ contributions

Nana Wang and Guangxin Ma designed the experiments. Guanxin Ma extracted the data and performed the statistical analysis. Xiaoran Bai and Jingjing Ye contributed to the study design. Nana Wang wrote the manuscript. Dongmei Wang, Fei Lu, and Chunyan Ji revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participates

This study was approved by the Human Ethics Committee of Qilu Hospital. Authors confirm that the experiments involving human data were conducted in accordance with the Declaration of Helsinki. All datasets were obtained from the online databases, and it was confirmed that informed written consent had already been obtained.

Consent for publication

All authors agree to publish.

References

   https://doi.org/10.1182/blood-2016-08-733196


Figures
**Figure 1**

Identification of prognostic ICD-related IncRNA in AML samples. **A** Sankey relationship diagram of ICD genes and ICD-related IncRNAs. **B** Distribution of the LASSO coefficients of ICD-related IncRNAs. **C** The 10-fold cross-validation of variable selection in the least absolute shrinkage and selection operator (LASSO) algorithm.
Figure 2

The prognostic performance of the ICD-related lncRNA risk model. A The risk score and survival status of each sample. The blue and red dots represent survival and death, respectively. B The Kaplan–Meier survival analysis of the high-risk and low-risk groups based on the risk model and median risk score. C-D PCA analysis between the low-risk and high-risk groups based on the risk model of the 7 ICD-related
IncRNAs expression profiles. **E-F** The heatmap and box plot of the expression levels of 7 ICD-related IncRNAs in the high-risk and low-risk groups. * $p < 0.01$, ** $p < 0.001$ and *** $p < 0.0001$.

Figure 3

Evaluation of the risk signature including 7 ICD-related IncRNAs. **A** Univariate Cox regression analysis of the correlation between OS and various clinicopathological features including age, gender and risk
Multivariate Cox regression analysis revealed that the age and the risk model were independent prognostic factors for predicting the OS of AML patients. The hybrid nomogram for predicting the OS of patients with AML at 1-, 3-, and 5-year. Calibration curves of the nomogram for OS prediction at 1-, 3-, and 5-year. Time-dependent ROC curve shows the AUC at 1-, 3-, and 5-year. ROC curve analysis of the clinical features and IRLs prognostic signature.
Examination of immune characteristics in high- and low-risk groups. A-B Single-sample gene set enrichment analysis (ssGSEA) scores of 23 immune cells and 13 immune-related functions in high- and low-risk groups. C Differences in expression of common immune checkpoints in the risk groups. D-G Box plots comparing immune score, stromal score, ESTIMATE score and tumor purity between the low- and high-risk groups, respectively. H TIDE score between high-risk group and low-risk group. *p < 0.05, **p < 0.01, and ***p < 0.001.

Figure 5

Drug sensitivity analysis in ICD-high and ICD-low group. IC50 values shows a significant difference in ICD-high and ICD-low group among A WZ-1-84, B TGX221, C BI-2536, D VX-680. E-H Correlation analysis of ICD score and drug sensitivity.
Figure 6

The functional enrichment analysis of the differential expressed genes (DEGs) in the ICD-low and ICD-high group. 

A The volcano diagram exhibits the DEGs between high- and low-risk groups with the threshold setting at $|FC| \geq 2$ and $p$-value < 0.05. 

B GO enrichment analysis of DEGs between low- and high-risk groups. 

D KEGG enrichment analysis of DEGs in low- and high-risk groups. 

C, E Graph showing
the enrichment analysis based on the Metascape Online, bar plot and network showing the distribution and relationship of the different functions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.7z