



# A Novel Immune-Related Competing Endogenous RNA Network Predicts Prognosis of Acute Myeloid Leukemia

Shujuan Wang<sup>1</sup>, Lu Yang<sup>1</sup>, Yajun Liu<sup>2</sup>, Yan Xu<sup>1</sup>, Danfeng Zhang<sup>1</sup>, Zhongxing Jiang<sup>1\*</sup>, Chong Wang<sup>1\*</sup> and Yanfang Liu<sup>1\*</sup>

<sup>1</sup> Department of Hematology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, <sup>2</sup> Department of Orthopaedics, Rhode Island Hospital, Warren Alpert Medical School, Brown University, Providence, RI, United States

## OPEN ACCESS

### Edited by:

Naval Daver,  
University of Texas MD Anderson  
Cancer Center, United States

### Reviewed by:

Antonio Curti,  
University of Bologna, Italy  
Ciprian Tomuleasa,  
Iuliu Hatieganu University of Medicine  
and Pharmacy, Romania

### \*Correspondence:

Zhongxing Jiang  
jiangzx@zzu.edu.cn  
Chong Wang  
fccwangc@zzu.edu.cn  
Yanfang Liu  
fccliyuf1@zzu.edu.cn

### Specialty section:

This article was submitted to  
Hematologic Malignancies,  
a section of the journal  
Frontiers in Oncology

**Received:** 11 May 2020

**Accepted:** 21 July 2020

**Published:** 11 August 2020

### Citation:

Wang S, Yang L, Liu Y, Xu Y,  
Zhang D, Jiang Z, Wang C and Liu Y  
(2020) A Novel Immune-Related  
Competing Endogenous RNA  
Network Predicts Prognosis of Acute  
Myeloid Leukemia.  
*Front. Oncol.* 10:1579.  
doi: 10.3389/fonc.2020.01579

**Background:** Acute myeloid leukemia (AML) is a genetically, biologically and clinically heterogeneous hematopoietic malignancy that is highly dependent on the bone marrow (BM) microenvironment. Infiltrated immune cells and stromal cells are an important part of the BM microenvironment and significantly affect the progression of AML. Recently, the competing endogenous RNA hypothesis has gained great interests in the study of molecular and biological mechanisms of tumor occurrence and progression. However, research on how competing endogenous RNA relates to leukemia tumor microenvironment remains uninvestigated.

**Methods:** In this study, mRNA, miRNA and lncRNA data and clinical information of the AML cohort were obtained from The Cancer Genome Atlas (TCGA) database, and the immune and stromal scores were calculated using the ESTIMATE algorithm.

**Results:** We found that immune scores were significantly correlated with cytogenetic risk and overall survival, and also identified microenvironment-related mRNAs, miRNAs, and lncRNAs based on the immune and stromal scores. Differentially expressed mRNAs and lncRNAs were applied to weighted correlation network analysis (WGCNA) to identify the modules most relevant to the immune microenvironment of AML. Using miRNA database to predict miRNA-targeted genes, we established the immune-related competing endogenous RNA network consisting of 33 lncRNAs, 21 miRNAs and 135 mRNAs. Prognostic analysis was performed on the genes in the immune-related competing endogenous RNA network to screen out 15 lncRNAs, 2 miRNAs and 31 mRNAs with prognostic values.

**Conclusion:** In summary, we identified a novel immune-related mRNA-miRNA-lncRNA competing endogenous RNA network associated with the prognosis of AML, which may contribute to better understanding of the development and progression of AML and to serve as novel therapeutic targets.

**Keywords:** acute myeloid leukemia, immune microenvironment, weighted gene coexpression network analysis, competing endogenous RNA network, prognosis.

## INTRODUCTION

Acute myeloid leukemia (AML) is the most common type of acute leukemia in adults, caused by the clonal expansion of undifferentiated myeloid progenitor cells (1). Although most patients with AML can achieve complete remission by induction chemotherapy, the recurrence rate remains high and thus is the main factor affecting the outcomes of AML patients. Relapse often develops from minimal residual disease in the protective bone marrow (BM) microenvironment (2, 3). A comprehensive understanding of the BM microenvironment is conducive to the diagnosis and personalized treatment of AML (3).

Cytogenetics and molecular aberrations are the main factors for risk stratification in patients with AML (4). In addition, the BM microenvironment also plays a very important role in the homing and survival of leukemic cells (3). The BM microenvironment contains various components such as immune cells, stromal cells, endothelial progenitor cells, extracellular matrix, growth factors, and cytokines (5). The interaction between leukemic cells and BM microenvironment affects resistance to chemotherapy in AML (3). The modulation of BM microenvironment in AML is currently undergoing preclinical research and early clinical trials. Molecular inhibitors such as CXCR4 inhibitors, VLA-4 inhibitors and E-selectin inhibitors are currently undergoing clinical trials (6, 7). Immune cells and stromal cells are important components of the BM microenvironment, and are also the main influencing factors of leukemia development (8, 9). The ESTIMATE program is a common method to explore the microenvironment of many tumors (10). Recently, it has also been used to explore the prognostic genes in the microenvironment of AML patients (11–13). Most studies have focused on differentially expressed genes (DEGs). However, the interaction and relationship between genes are open to investigate. Moreover, the coding genes were extensively explored, but regions that encoded lncRNAs and miRNAs were less well-studied.

Weighted gene coexpression network analysis (WGCNA) is an algorithm commonly used in systems biology to explore the correlation between gene sets and the clinic (14). Functionalization is achieved by constructing a free-scale coexpression gene network (14). WGCNA can identify highly related genes and aggregate them into the same genetic module. Currently, WGCNA is used in multiple fields such as cancer and nervous system, or to identify potential biomarker candidates or new therapeutic targets from genetic data (15–17).

The competition endogenous RNA (ceRNA) hypothesis was a new regulatory mechanism between non-coding RNA (ncRNA), and messenger RNA (mRNA) proposed by Salmena et al. in 2011 (18). In this theory, crosstalk between ceRNAs is achieved by

competitively combining shared miRNAs (19). In recent years, the ceRNAs hypothesis has attracted widespread attention in the study of molecular and biological mechanisms of tumorigenesis and development (18). For example, previous studies have found that lncRNA-related ceRNAs were involved in the biological processes of glioblastoma and breast cancer (15, 20). The research on ceRNAs of leukemia was generally based on the differential genes screened by leukemia and normal controls (21), but no known module based on ceRNAs network related to microenvironment in leukemia has been set up.

In this study, mRNAs, miRNAs and lncRNAs data and clinical information of the AML cohort from The Cancer Genome Atlas (TCGA) database were used to calculate the immune and stromal scores of these AML cases using the ESTIMATE algorithm. Differentially expressed mRNAs and lncRNAs were applied to WGCNA to identify the modules most relevant to the AML immune microenvironment. Then the immune-related lncRNA-miRNA-mRNA ceRNA network was established to screen genes with clinical significance. These findings will help to better understand the role of tumor microenvironment in AML and shed light on the development and progression of AML.

## MATERIALS AND METHODS

### Data Acquisition

All data sets of AML patients were downloaded from TCGA database<sup>1</sup>. The data used in this study met the following criteria: (1) Excluding samples combined with other malignancies; (2) Samples with lncRNAs and miRNAs and mRNAs detection data. Finally, all lncRNAs, mRNAs, and miRNAs expression profiles of 138 AML specimens and the corresponding clinical follow-up data were downloaded.

### Identification of Differentially Expressed Genes

The ESTIMATE algorithm<sup>2</sup> was used to calculate the immune scores and stromal scores of 138 AML samples. Differentially expressed genes (DEGs), such as lncRNAs, miRNAs, and mRNAs, were identified between high and low score groups stratified by the median value of immune scores and stromal scores using limma package (22). All q values use FDR to correct the statistical significance of the multiple test. lncRNAs and mRNAs with  $|\log FC| > 1.3$  and  $FDR < 0.05$  were regarded as differentially expressed while miRNAs with  $|\log FC| > 2$  and  $FDR < 0.05$  were regarded as differentially expressed. Then all the DEGs were entered into R (version 3.5.1, Auckland, NZ, United States) for cluster analysis based on the expression value of each sample in its respective data set. The results were expressed in a clustergram. Each column represents a sample, and each row represents the expression level of a given gene.

**Abbreviations:** AML, acute myeloid leukemia; AUC, area under curve; BM, bone marrow; CAM, cell adhesion molecules; ceRNA, competing endogenous RNAs; DEG, differentially expressed gene; DAVID, The Database for Annotation, Visualization and Integrated Discovery; GO, gene ontology; ICAM1, Intercellular Adhesion Molecule 1; IL10RA, Interleukin 10 Receptor Subunit Alpha; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction; TCGA, The Cancer Genome Atlas; TLR6, Toll Like Receptor 6; TLR8, Toll Like Receptor 8; TOM, topological overlap matrix; WGCNA, weighted correlation network analysis.

<sup>1</sup><https://portal.gdc.cancer.gov/>

<sup>2</sup><https://sourceforge.net/projects/estimateproject/>

## GO and Pathway Enrichment Analyses

The Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>3</sup> was applied to analyze enriched biological themes of DEGs functions, particularly GO (gene ontology) terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment (23).  $P < 0.05$  was set as the cut-off criterion.

## Weighted Gene Coexpression Network Analysis

WGCNA is an algorithm for identification of gene coexpression networks through high-throughput expression profiles of mRNAs or lncRNAs with different characteristics (14). Pairwise Pearson correlation analysis was used to evaluate the weighted coexpression relationship between all dataset topics in the adjacency matrix. In this study, WGCNA was used to analyze mRNAs and lncRNAs to obtain the mRNAs or lncRNAs most relevant to AML immune microenvironment.

## ceRNA Network Construction and Analysis

According to the results of WGCNA, we selected all mRNAs and lncRNAs in the most relevant module (turquoise) and differentially expressed miRNAs to construct a ceRNA network. Briefly, the associated ceRNA network in AML was constructed following three stages. (1) Prediction of lncRNA-miRNA: in order to make lncRNAs and miRNAs map into the interactions successfully, we used the miRanda<sup>4</sup> and PITA<sup>5</sup> to get targeted lncRNAs that miRNAs may regulate. (2) Prediction of miRNA-mRNA: three online databases, miRanda<sup>4</sup>, Targetscan<sup>6</sup>, and miRWalk<sup>7</sup> were used simultaneously for target mRNA prediction. (3) Construction of lncRNA-miRNA-mRNA ceRNA network: the ceRNA network was constructed based on the negatively regulating target relationships of miRNA-mRNA and miRNA-lncRNA correlation pairs.

## PPI Network Construction

The Retrieval of Interacting Genes (STRING) database<sup>8</sup> was utilized to construct a protein-protein interaction (PPI) network of the DEGs identified in the ceRNA network. The interacting pairs with a confidence score greater than 0.3 were considered as significant and were retained. The degree represents the number of interaction partners.

## Survival Analysis

Kaplan-Meier plots were constructed to illuminate the relationship between the overall survival of AML patients and the expression level of mRNAs, lncRNAs, and miRNAs. The statistical significance of the correlation was tested by the

log-rank test and  $P < 0.05$  was considered significant. The analysis was conducted using the R package of survival.

## Statistical Analysis

Graphpad Prism<sup>TM</sup> 5.01 (San Diego, CA, United States) or R (version 3.5.1, Auckland, NZ, United States) software was used for all data analyses. Differences across groups were compared using *Kruskal-Wallis* test for continuous variables. Differences were considered significant when  $P < 0.05$ .

## RESULTS

### Immune and Stromal Scores Are Associated With AML Clinical Parameters

We obtained gene expression profiles and clinical information of 138 AML patients from TCGA database (**Supplementary Material 1**). Among them, 74 (53.6%) were male and 64 (46.4%) were female, with a median age 55.5 (range 21–88) years old. According to the FAB classification, there were 13 cases of M0, 32 M1 cases, 34 M2 cases, 15 M3 cases, 27 M4 cases, 14 M5 cases, 2 M6 cases, and 1 M7 case. The ESTIMATE algorithm was used to calculate the immune scores and stromal scores of AML patients. The median immune score was 3027.01 (range from 1686.31 to 4291.79) and the median stromal score was -1024.97 (range from -1910.13 to 297.36).

We analyzed the relationship between immune scores and stromal scores and clinical parameters of AML patients. Cases with M4 subtype AML had the highest immune scores while cases with M3 subtype had the lowest immune scores ( $P < 0.0001$ ; **Figure 1A**). Similarly, M4 cases had the highest stromal scores, whereas M0 subtypes had the lowest ( $P < 0.0001$ ; **Figure 1B**). According to cytogenetics, AML patients were divided into three groups: favorable, intermediate/normal and poor. There was an obvious correlation between the cytogenetic risk and the immune scores ( $P = 0.0141$ ; **Figure 1C**; favorable vs. intermediate:  $P = 0.0041$ ; favorable vs. poor:  $P = 0.1142$ ; intermediate vs. poor:  $P = 0.2605$ ), but no significant correlation between the cytogenetic risk and the stromal scores was observed ( $P = 0.6192$ ; **Figure 1D**).

Using the median immune or stromal score as a threshold, AML patients were divided into two groups with low immune/stromal score and high immune/stromal score. Survival analysis showed that the survival rate of AML patients with low immune scores was significantly higher than that of patients with high immune scores ( $P = 0.0224$ ; **Figure 1E**). However, there was no significant difference in survival between patients with low stromal scores and those with high stromal scores ( $P = 0.4166$ ; **Figure 1F**).

### Identification of Differentially Expressed Genes Based on Immune Scores and Stromal Scores

Setting the cut-off criteria as  $|\log FC| > 1.3$  and  $FDR < 0.05$ , we identified 1399 mRNAs (**Figure 2A**) and 258 lncRNAs

<sup>3</sup><http://david.ncifcrf.gov/>

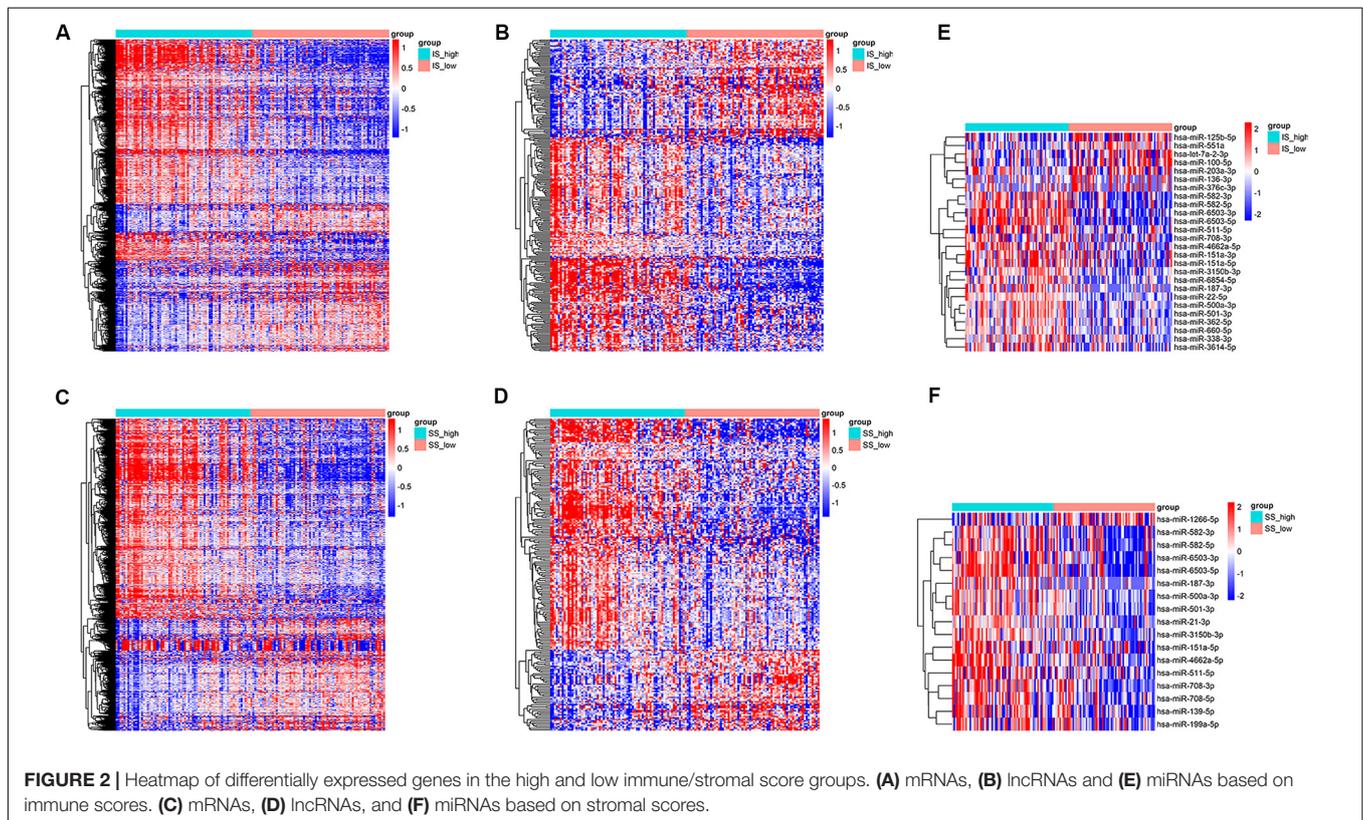
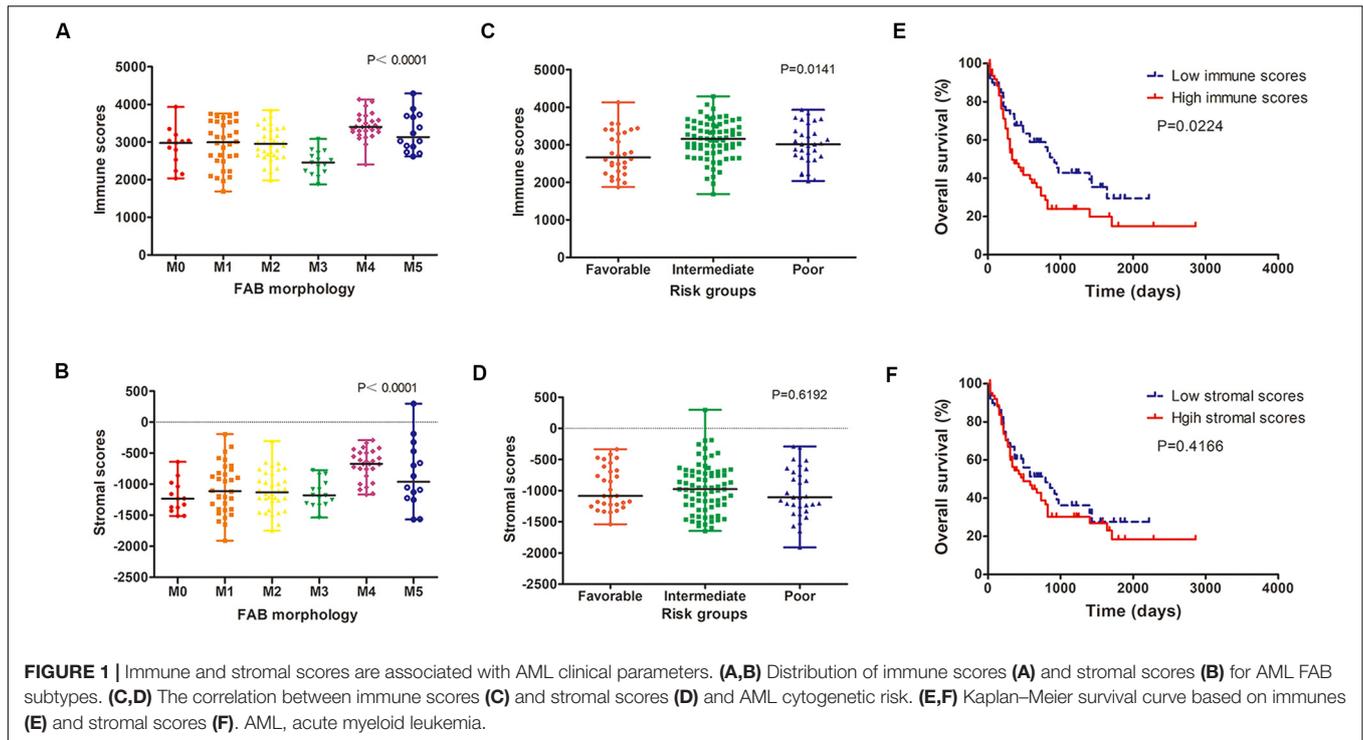
<sup>4</sup><http://www.microrna.org/>

<sup>5</sup>[https://genie.weizmann.ac.il/pubs/mir07/mir07\\_exe.html](https://genie.weizmann.ac.il/pubs/mir07/mir07_exe.html)

<sup>6</sup><http://www.targetscan.org/>

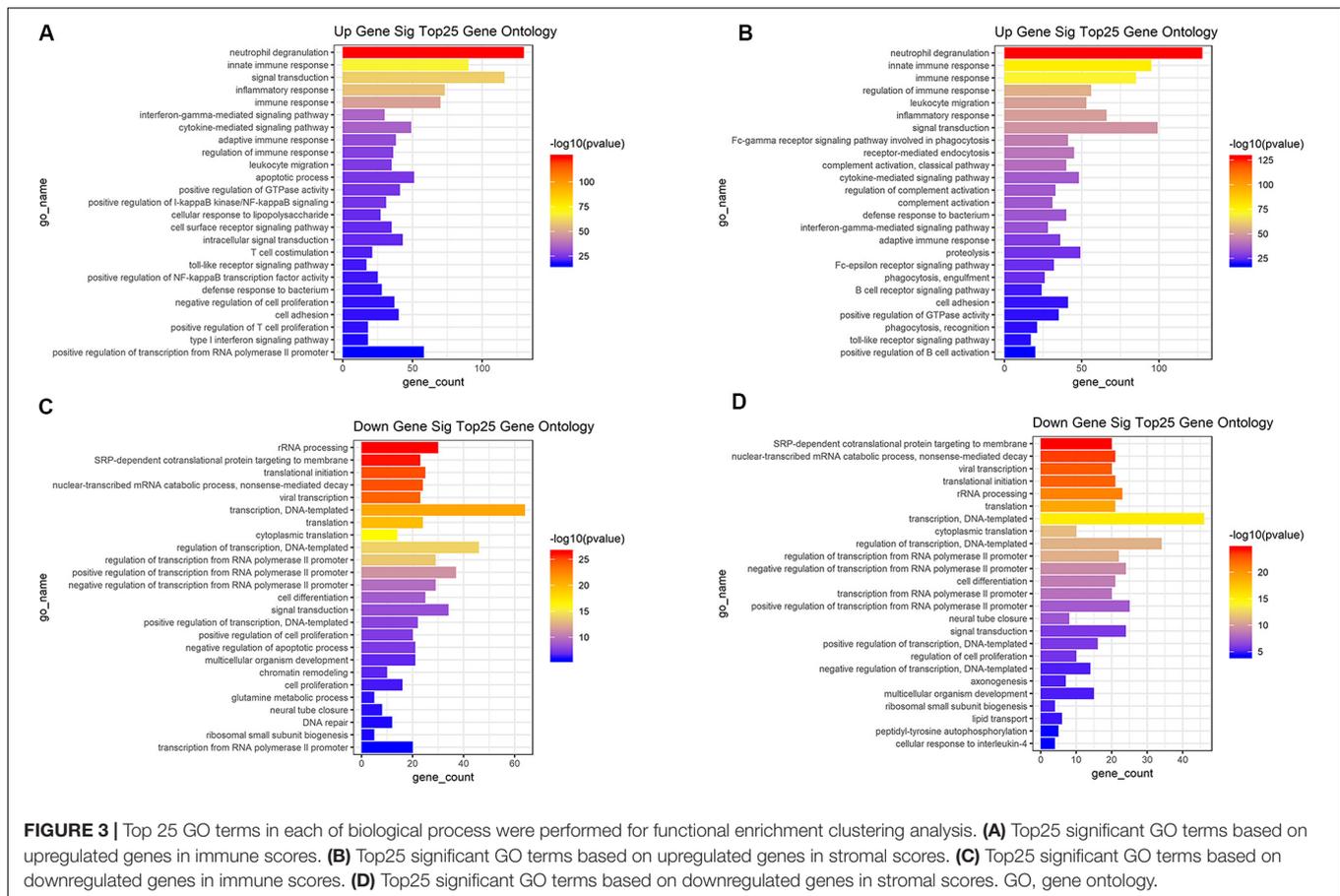
<sup>7</sup><http://129.206.7.150/>

<sup>8</sup><http://string-db.org>



(Figure 2B) based on immune scores, and 1166 mRNAs (Figure 2C) and 212 lncRNAs (Figure 2D) based on stromal scores. Setting the cut-off criteria as  $|\log FC| > 2$

and  $FDR < 0.05$ , we identified 26 and 17 miRNAs based on immune scores (Figure 2E) and stromal scores (Figure 2F), respectively. The DEGs of the low vs. high immune



score or stromal score groups were illustrated in the heat map (Figure 2).

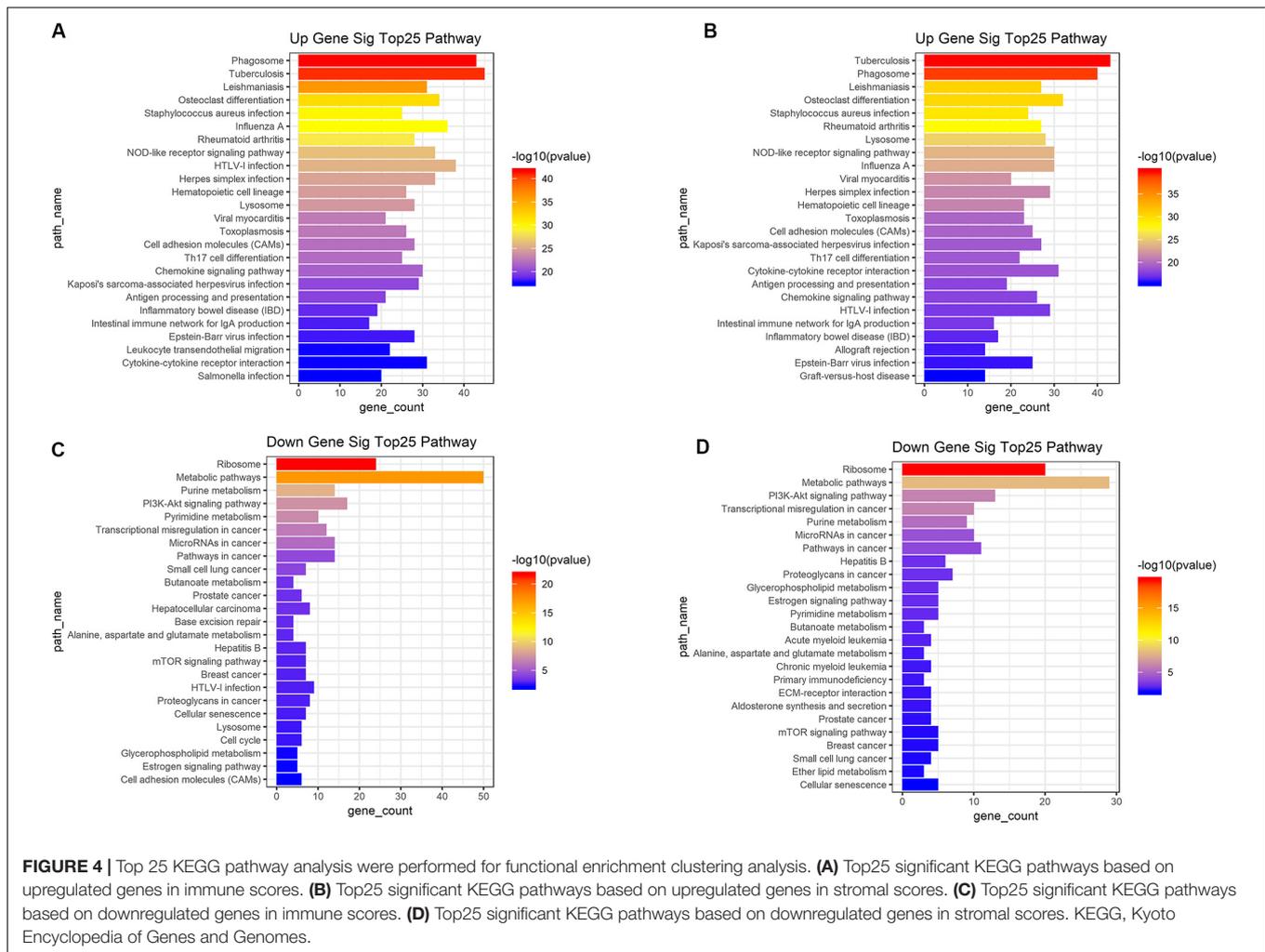
## Functional Enrichment Analysis of DEGs

Based on the DAVID (The Database for Annotation, Visualization and Integrated Discovery) gene annotation tool, we performed GO analyses of both upregulated and downregulated DEGs. The top 25 GO biological process indicated that the upregulated DEGs based on immune or stromal scores were primarily enriched in neutrophil degranulation, regulation of immune response, signal transduction and inflammatory response (Figures 3A,B), while the downregulated DEGs based on immune/stromal scores were primarily enriched in rRNA processing, regulation of translation, regulation of transcription and cell differentiation (Figures 3C,D). Subsequently, we performed KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment and interrelation analysis. KEGG analysis revealed that the upregulated DEGs were mainly enriched in infection, osteoclast differentiation, NOD-like receptor signaling pathway, hematopoietic cell lineage and cell adhesion molecules (CAMs) pathways (Figures 4A,B), while the downregulated DEGs were mainly enriched in ribosome, metabolism, PI3K-Akt signaling pathway, transcriptional dysregulation in cancer and miRNAs in cancer (Figures 4C,D). Above analyses revealed that these DEGs play a vital role in the development of

AML and thus require further research to determine their biological contribution.

## Construction of Weighted Correlation Network Analysis and Identification of Key Modules

Based on the results of survival analysis, DEGs based on immune scores were selected for subsequent analysis. The best  $\beta$  value in the lncRNA/mRNA coexpression network was 5, which was calculated using the 258 differential lncRNAs, 1399 differential mRNAs, and their expression data in leukemia samples. Next, the method of dynamic tree cutting was used to produce coexpression modules. Finally, 8 modules of lncRNA-mRNA coexpression networks were generated and the heat map plot of topological overlap matrix (TOM) was shown (Figure 5A). Each module was calculated and plotted with its corresponding clinical characteristics. Correlation analysis showed that Turquoise module displayed the highest relationship with AML immune scores ( $r = 0.77$ ), which included 760 mRNAs and 86 lncRNAs (Figure 5B). These 760 mRNAs were further used to perform the gene enrichment analysis. The genes were most related to neutrophil degranulation, immune response, inflammatory response, signal transduction and toll-like receptor signaling pathway (Figure 5C). In addition, genes were highly enriched in infection, osteoclast differentiation, NOD-like receptor signaling



pathway, metabolic pathways and hematopoietic cell lineage by KEGG analysis (Figure 5D).

## ceRNA Network Construction

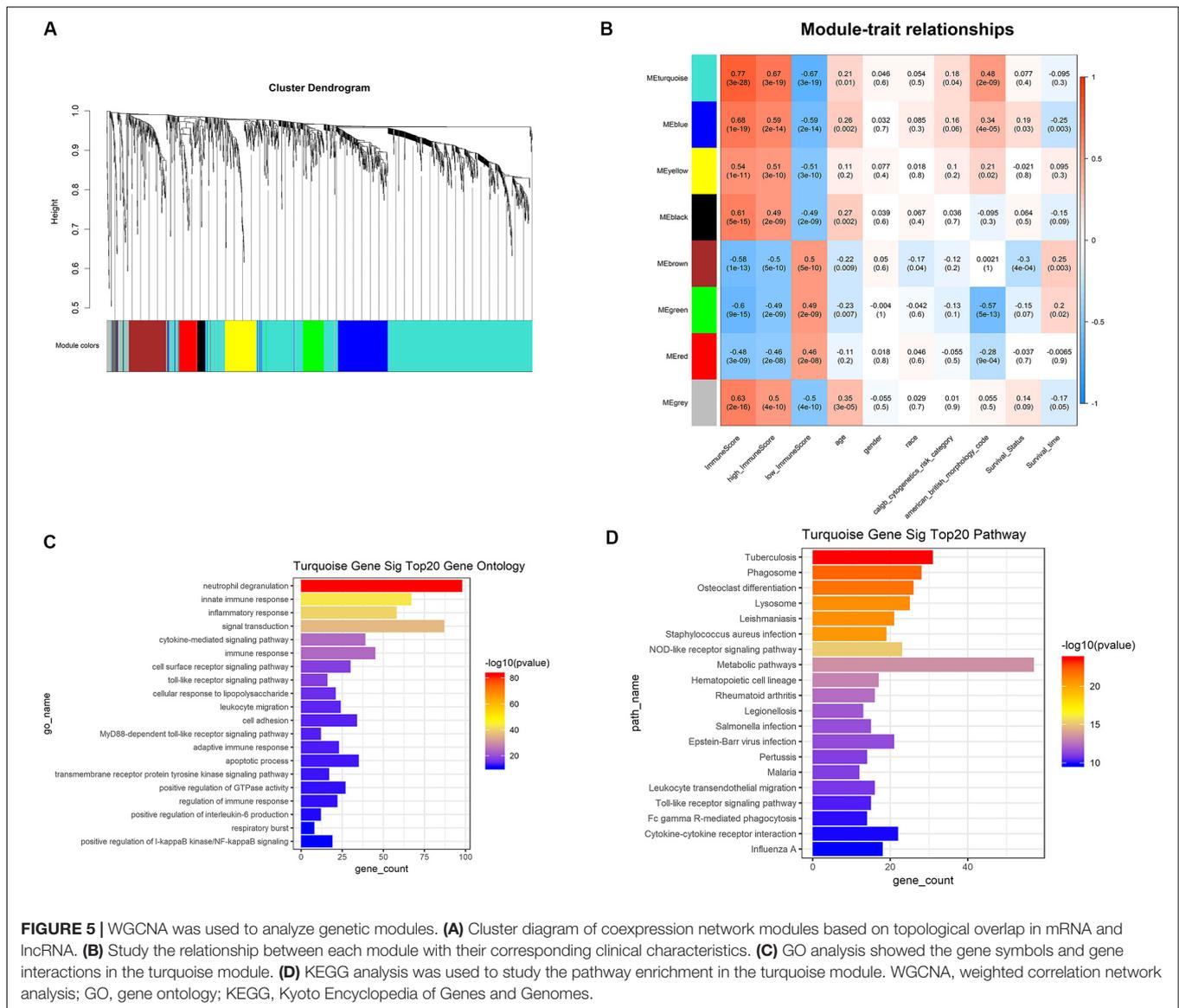
Since the Turquoise module showed the highest relationship with AML immune scores, we selected lncRNAs and mRNAs in the Turquoise module and 26 differentially expressed miRNAs based on immune scores to construct a ceRNA network. Firstly, based on the PITA and miRcode online database that matches potential miRNAs with lncRNAs, a total of 315 lncRNA-miRNA pairs contained 63 lncRNAs and 25 miRNAs. Then, we searched for the mRNAs targeted by the differentially expressed miRNAs using three target gene prediction websites, miRanda, miRWalk, and TargetScan. Using these websites, we detected 664, 671, and 655 target mRNAs, respectively. Based on the Venn intersection analysis, 289 target mRNAs were selected. Subsequently, we matched the predicted target genes with the mRNAs in the Turquoise module. Then, we constructed the ceRNA network by integrating the miRNA-lncRNA-mRNA interactions. At last, a final lncRNA-miRNA-mRNA ceRNA network was constructed with 33 lncRNAs, 21 miRNAs and 135 mRNAs (Figure 6A).

## Protein-Protein Interaction (PPI) Network Analysis

To further explore the interplay among the mRNAs in ceRNA, we constructed a PPI network based on the STRING (The Retrieval of Interacting Genes) online database (Figure 6B). In the network, *TLR8* (Toll Like Receptor 8), *ICAM1* (Intercellular Adhesion Molecule 1), *TLR6* (Toll Like Receptor 8), and *IL10RA* (Interleukin 10 Receptor Subunit Alpha) had higher degrees (16, 13, 10, and 10, respectively) (Supplementary Table 1). The genes encoding these proteins have been confirmed to be associated with immune microenvironment and leukemia progression (24–27).

## Association Between mRNAs, miRNAs, and lncRNAs in ceRNA and Overall AML Survival

We further analyzed the prognostic values of mRNAs, miRNAs, and lncRNAs in the ceRNA network. Subjects were divided into high-expression and low-expression cohorts according to the median value of these genes. For overall survival,



**FIGURE 5 |** WGCNA was used to analyze genetic modules. **(A)** Cluster diagram of coexpression network modules based on topological overlap in mRNA and lncRNA. **(B)** Study the relationship between each module with their corresponding clinical characteristics. **(C)** GO analysis showed the gene symbols and gene interactions in the turquoise module. **(D)** KEGG analysis was used to study the pathway enrichment in the turquoise module. WGCNA, weighted correlation network analysis; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

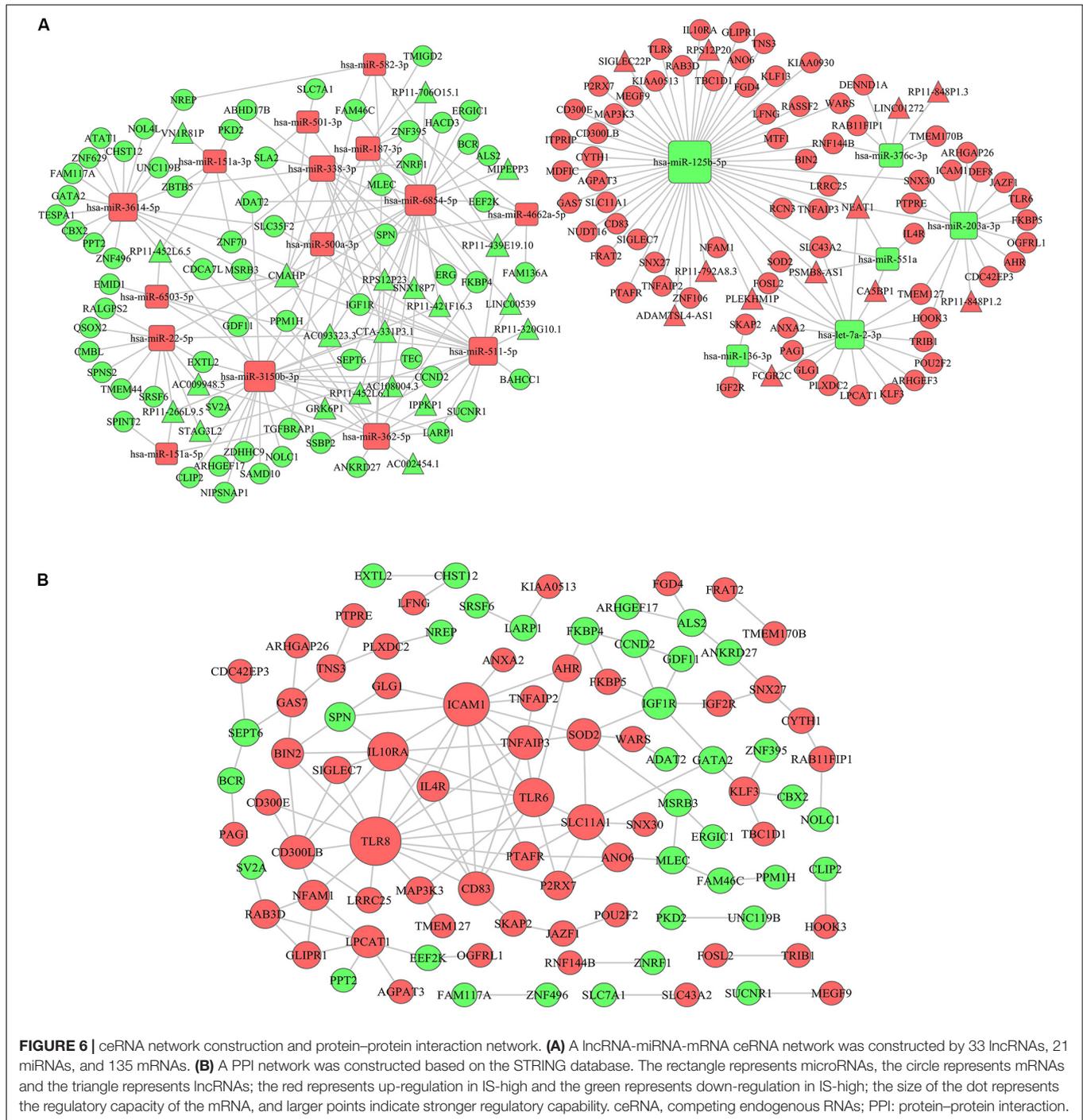
the high-expression and low-expression cohorts were split for the log-rank test through the package survival in R software. 15 out of 33 lncRNAs (*AC009948.5*, *CMAHP*, *CTA-331P3.1*, *FCGR2C*, *GRK6P1*, *LINC00539*, *LINC01272*, *MIPEPP3*, *PSMB8-AS1*, *RP11-266L9.5*, *RP11-320G10.1*, *RP11-421F16.3*, *RP11-439E19.10*, *RP11-792A8.3* and *STAG3L2*), 2 out of 21 miRNAs (*hsa-miR-125b-5p* and *hsa-miR-338-3p*) and 31 out of 135 mRNAs (*AGPAT3*, *ANKRD27*, *CBX2*, *CCND2*, *CD300LB*, *CYTH1*, *ERG*, *GDF11*, *IGF1R*, *KIAA0513*, *KIAA0930*, *LAR1*, *LFNG*, *LPCAT1*, *NREP*, *NUDT16*, *POU2F2*, *PPM1H*, *PTAFR*, *QSOX2*, *RAB3D*, *RALGPS2*, *SIGLEC7*, *SLC43A2*, *SRSF6*, *TNFAIP2*, *TNS3*, *TRIB1*, *ZBTB5*, *ZNF70*, and *ZNRF1*) were associated with overall survival according to the log-rank test ( $P < 0.05$ ). Representative genes are shown in **Figure 7**.

Age ( $\geq 60$  years vs.  $< 60$  years) and risk group (favorable vs. intermediate/normal vs. poor) were also associated with overall survival according to the log-rank test ( $P = 0.000$  and

0.002, respectively). The above mentioned 15 lncRNAs, 31 mRNAs and 2 miRNAs were brought into further multivariate cox proportional hazard regression analysis with age and risk group. Finally, 7 mRNAs (*CCND2*, *ERG*, *LPCAT1*, *NUDT16*, *RALGPS2*, *TNFAIP2* and *ZNF70*), 6 lncRNAs (*CMAHP*, *FCGR2C*, *PSMB8-AS1*, *RP11-266L9.5*, *RP11-320G10.1*, *RP11-792A8.3*, and *STAG3L2*) and 2 miRNAs (*hsa-miR-125b-5p* and *hsa-miR-338-3p*) were independently associated with overall survival (**Table 1**).

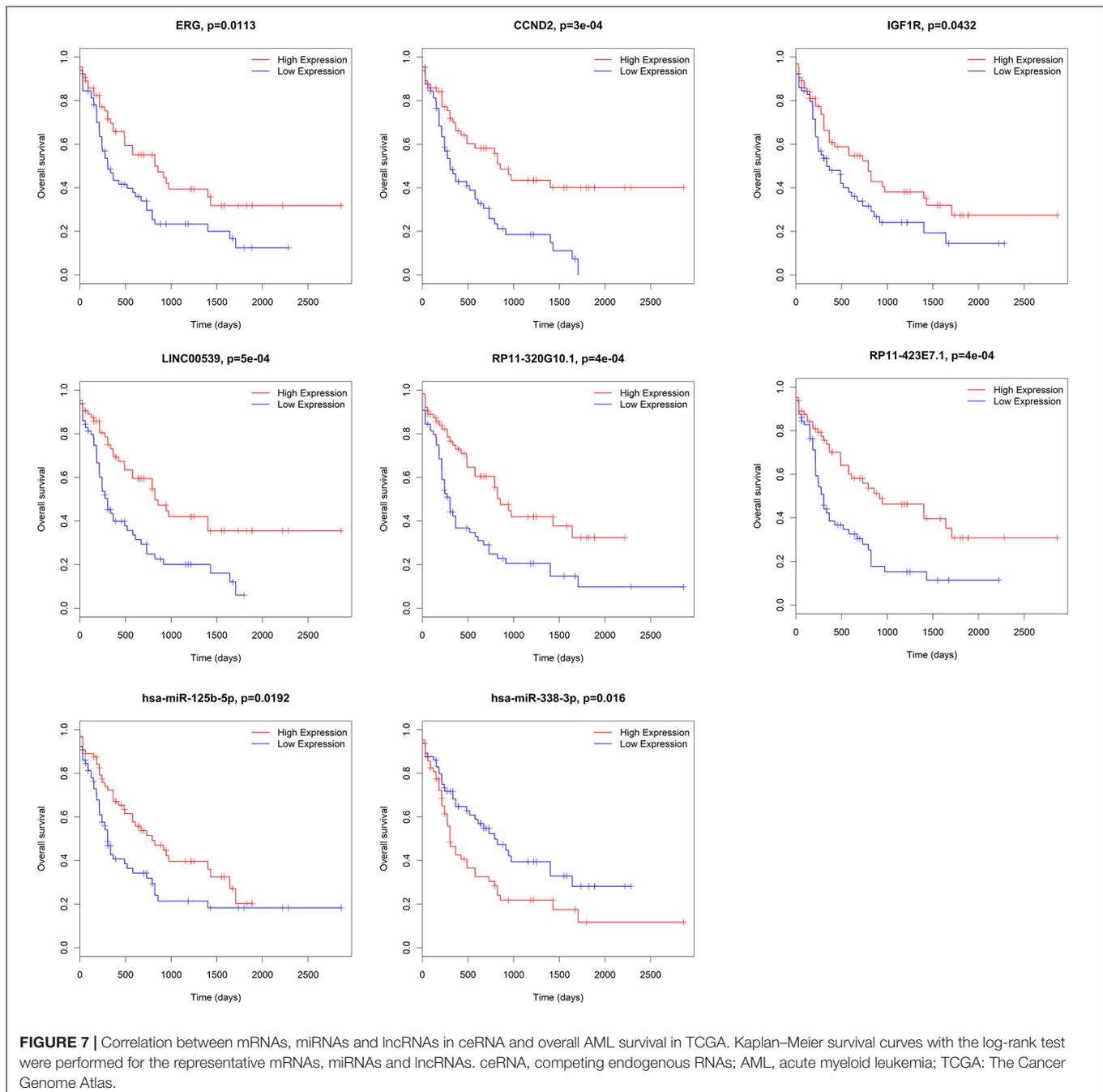
## DISCUSSION

In recent years, studies about the roles of gene mutations and chromosomal translocations in the occurrence and development of AML and their prognostic values have made significant progress (4). However, the BM microenvironment which also plays an important role in the pathophysiology



process in AML are poorly understood (5). Therefore, most treatments previously targeted only tumor cells, but few targeted the tumor microenvironment. In-depth research on the microenvironment of leukemia will help to further understand the mechanism of leukemia development, and may find new targets for microenvironment treatment (7). This study screened microenvironment-related genes based on the TCGA database, and further established microenvironment-related lncRNA-miRNA-mRNA ceRNA networks through WGCNA.

First, we calculated the immune scores and stromal scores of AML patients based on the ESTIMATE algorithm, and found that these scores were related to the FAB typing of AML. In addition, immune scores were significantly correlated with cytogenetic risk and overall survival. ESTIMATE is a new algorithm to infer the level of stromal and immune cells in tumor tissues and tumor purity using gene expression data (10). High immune scores in the BM samples from patients with poor prognosis indicated that more immune cells were recruited in their BM



microenvironment. This may be due to that AML cells actively shape the BM environment and immune cells to promote disease progression through cellular, structural, and functional changes (11). However, there was no significant correlation between stromal scores and cytogenetic risk or survival of AML patients, suggesting that the proportion of stromal cells were comparable in different group. Possible explanation may be that stromal cells play an important role in solid tumors (28, 29), while its role in leukemia is not as strong as in solid tumors (11–13).

Then we identified differentially expressed mRNAs, miRNAs, and lncRNAs based on the immune scores or stromal

scores. Functional enrichment analysis indicated that these DEGs were mainly involved in immune and inflammatory responses. Consistent with these results, previous studies have shown that the biology of the immune system is essential for the formation of a complex BM microenvironment (5). In recent years, the understanding of the immunological characteristics of AML has increased, and the development of effective AML immunotherapy strategies has attracted widespread attention (30).

In recent years, important advances in ceRNA coexpression network research have developed rapidly. The disruption of

**TABLE 1** | Multivariate cox proportional hazard regression analysis of 6 lncRNAs, 7 mRNAs, and 2 miRNAs.

Genes	HR (95%CI)	P
<i>CCND2</i>	0.640 (0.373–0.977)	0.040
<i>ERG</i>	0.603 (0.377–0.967)	0.036
<i>LPCAT1</i>	1.888 (1.179–3.023)	0.008
<i>NUDT16</i>	1.826 (1.143–2.917)	0.012
<i>RALGPS2</i>	0.414 (0.258–0.663)	0.000
<i>TNFAIP2</i>	2.010 (1.263–3.199)	0.003
<i>ZNF70</i>	0.526 (0.328–0.842)	0.007
<i>CMAHP</i>	0.580 (0.367–0.918)	0.020
<i>FCGR2C</i>	1.741 (1.103–2.748)	0.017
<i>PSMB8-AS1</i>	1.776 (1.116–2.828)	0.015
<i>RP11-266L9.5</i>	0.597 (0.373–0.957)	0.031
<i>RP11-320G10.1</i>	0.540 (0.338–0.862)	0.010
<i>RP11-792A8.3</i>	0.608 (0.390–0.948)	0.028
<i>STAG3L2</i>	0.597 (0.375–0.950)	0.030
<i>hsa-miR-125b-5p</i>	0.562 (0.353–0.897)	0.016
<i>hsa-miR-338-3p</i>	1.609 (1.010–2.564)	0.045

the ceRNA network balance is a major cause for tumorigenesis (31). Therefore, understanding the complex interactions between different ceRNA networks will lead to an in-depth understanding of gene regulatory networks and has implications for cancer treatment (31). In addition, the lncRNA-miRNA-mRNA ceRNA network can predict the prognosis of the disease. For example, a lncRNA-miRNA-mRNA ceRNA network was established based on RNA-Seq data of breast cancer from TCGA, which consists of 8 miRNAs, 48 lncRNAs, and 10 mRNAs (32). Multiplex cox regression analyses showed that four of these lncRNAs (*ADAMTS9-AS1*, *LINC00536*, *AL391421.1*, and *LINC00491*) have significant prognostic value (32). Wang et al. identified 108 lncRNAs, 10 miRNAs, and 8 mRNAs from a database to construct a lncRNA-miRNA-mRNA ceRNA network (21). In the network, a univariate and multivariate cox proportional hazard regression analysis was used to establish a survival model with 8 target mRNAs (*HOXA9*, *INSR*, *KRIT1*, *MYB*, *SPRY2*, *UBE2V1*, *WEE1*, and *ZNF711*), where AUC (area under curve) is 0.831, indicating the sensitivity and specificity of prognostic prediction (21). However, the screening of differential genes in this study was based on leukemia patients and normal people, and did not focus on tumor microenvironment. So far, there is no ceRNA research based on the leukemia microenvironment. In our research, the screening of differential genes was based on the immune score. Then WGCNA was used to identify the modules most relevant to the AML immune microenvironment. Then, using WGCNA and miRNA prediction websites, a lncRNA-miRNA-mRNA ceRNA network consisting of 33 lncRNAs, 21 miRNAs and 135 mRNAs was constructed.

Subsequently, we built a PPI network predicting the interaction among the proteins encoded by the 135 DEGs in the ceRNA network. *TLR8*, *ICAM1*, *TLR6*, and *IL10RA* had higher degrees. *TLR8* and *TLR6* are members of the Toll-like receptor family which is upstream to the transcription

factor NF $\kappa$ B and part of the innate immune system and plays an important role in progression of AML (24). *ICAM1* is one of the CAMs, a large class of transmembrane proteins, involved in the binding of cells to another cell or extracellular matrix, and involved in cell proliferation, differentiation, movement, transportation, apoptosis and tissue structure (25). The protein encoded by *IL10RA* is a receptor for interleukin 10, which has been shown to mediate the immunosuppressive signal of interleukin 10, and thus inhibits the synthesis of proinflammatory cytokines and is reported to promote survival of progenitor myeloid cells through the insulin receptor substrate-2/PI3K/AKT pathway (26, 27). These results indicated that this novel ceRNA network were closely associated with immune microenvironment and progression of AML.

Furthermore, 15 lncRNAs, 2 miRNAs, and 31 mRNAs with prognostic significance were screened out, which could be used as biomarkers for prognosis. Among the genes with prognostic significance in our module of immune-related ceRNA network, there were AML related reports about *CBX2*, *CCND2*, *ERG*, *IGF1R*, *LARP1*, *LFNG*, *NUDT16*, *POU2F2*, *PTAFR*, *RAB3D*, *SIGLEC7*, *SRSF6*, *TNFAIP2*, *TRIB1*, *ZBTB5*, and *ZNRF1*, the most reported of which were *ERG*, *CCND2*, and *IGF1R*. *ERG* translocation was involved in the occurrence and development of AML, and high expression of *ERG* was a poor prognostic factor for patients with normal karyotype AML (33, 34). *CCND2* mutations were more common in CBF-AML, and it was also a frequent mutation event in *t*(8, 21) AML (35, 36). *CCND2* led to increased phosphorylation of retinoblastoma proteins, leading to significant cell cycle changes and increased proliferation of AML cell lines (36). Nicolas Chapuis et al. found that IGF-1 spontaneous lesions played a key role in PI3K/AKT activation of AML cells, providing strong evidence for targeting IGF1R as a potential new therapy for AML (37). The functions of *AGPAT3*, *ANKRD27*, *CD300LB*, *CYTH1*, *GDF11*, *KIAA0513*, *KIAA0930*, *LPCAT1*, *NREP*, *PPM1H*, *QSOX2*, *RALGPS2*, *SLC43A2*, *TNS3*, and *ZNF70* in AML have not been reported. We identified 15 lncRNAs with clinical significance. Among them, only *CMAHP* was reported to be related to MLL-positive AML (38), and other lncRNAs have not been reported in leukemia. Two miRNAs, including *hsa-miR-125b-5p* and *hsa-miR-338-3p*, have been reported to be associated with a variety of cancers (39–42), but no studies have been reported related to AML. All these unreported mRNAs, miRNAs and lncRNAs may be potential novel biomarkers or therapeutic targets for AML.

It is important to note that our current research has limitations. We selected the target data from the TCGA public database only through the biological algorithm method. We should further verify the results of this article in clinical patients in further study.

## CONCLUSION

In summary, a comprehensive bioinformatics analysis was performed on the AML dataset in TCGA, with an emphasis on the immune microenvironment. Using WGCNA and

miRNA prediction programs, an immune-related lncRNA-miRNA-mRNA ceRNA network was established, and DEGs with prognostic value were further identified. Further studies of these genes are needed in the clinic and may provide new insights into the pathogenesis of AML. This study increases our understanding of the complex interactions between AML tumor cells and the BM microenvironment and may provide novel prognostic factors and therapeutic targets.

## DATA AVAILABILITY STATEMENT

All data sets of AML patients were downloaded from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>).

## AUTHOR CONTRIBUTIONS

YFL, CW, and ZJ: conceptualization and design. SW: data acquisition and writing – original draft. SW, LY, YX, and DZ: methodology. SW and CW: data analysis and interpretation. YJL and YFL: writing – review and editing. YFL, CW, and ZJ:

project administration. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Numbers 81800137 and U1804191) and the Henan Medical Science and Technology Research Project (Grant Number 2018020068).

## ACKNOWLEDGMENTS

We thank the TCGA database for the availability of the data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.01579/full#supplementary-material>

## REFERENCES

- Ferrara F, Schiffer CA. Acute myeloid leukaemia in adults. *Lancet*. (2013) 381:484–95. doi: 10.1016/S0140-6736(12)61727-9
- Zeng Z, Shi YX, Samudio IJ, Wang R-Y, Ling X, Frolova O, et al. Targeting the leukemia microenvironment by CXCR4 inhibition overcomes resistance to kinase inhibitors and chemotherapy in AML. *Blood*. (2009) 113:6215–24. doi: 10.1182/blood-2008-05-158311
- Shafat MS, Gnanaswaran B, Bowles KM, Rushworth SA. The bone marrow microenvironment—Home of the leukemic blasts. *Blood Rev*. (2017) 31:277–86. doi: 10.1016/j.blre.2017.03.004
- Bullinger L, Döhner K, Döhner H. Genomics of acute myeloid leukemia diagnosis and pathways. *J Clin Oncol*. (2017) 35:934–46. doi: 10.1200/JCO.2016.71.2208
- Ayala F, Dewar R, Kieran M, Kalluri R. Contribution of bone microenvironment to leukemogenesis and leukemia progression. *Leukemia*. (2009) 23:2233–41. doi: 10.1038/leu.2009.175
- Uy GL, Rettig MP, Motabi IH, McFarland K, Trinkaus KM, Hladnik LM, et al. A phase 1/2 study of chemosensitization with the CXCR4 antagonist plerixafor in relapsed or refractory acute myeloid leukemia. *Blood*. (2012) 119:3917–24. doi: 10.1182/blood-2011-10-383406
- Rashidi A, Uy GL. Targeting the microenvironment in acute myeloid leukemia. *Curr Hematol Malig Rep*. (2015) 10:126–31. doi: 10.1007/s11899-015-0255-4
- Austin R, Smyth MJ, Lane SW. Harnessing the immune system in acute myeloid leukaemia. *Crit Rev Oncol Hematol*. (2016) 103:62–77. doi: 10.1016/j.critrevonc.2016.04.020
- Yehudai-Resheff S, Attias-Turgeman S, Sabbah R, Gabay T, Musallam R, Fridman-Dror A, et al. Abnormal morphological and functional nature of bone marrow stromal cells provides preferential support for survival of acute myeloid leukemia cells. *Int J Cancer*. (2019) 144:2279–89. doi: 10.1002/ijc.32063
- Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. (2013) 4:2612. doi: 10.1038/ncomms3612
- Yan H, Qu J, Cao W, Liu Y, Zheng G, Zhang E, et al. Identification of prognostic genes in the acute myeloid leukemia immune microenvironment based on TCGA data analysis. *Cancer Immunol Immunother*. (2019) 68:1971–8. doi: 10.1007/s00262-019-02408-7
- Huang S, Zhang B, Fan W, Zhao Q, Yang L, Xin W, et al. Identification of prognostic genes in the acute myeloid leukemia microenvironment. *Aging*. (2019) 11:10557–80. doi: 10.18632/aging.102477
- Ni J, Wu Y, Qi F, Li X, Yu S, Liu S, et al. Screening the cancer genome atlas database for genes of prognostic value in acute myeloid leukemia. *Front Oncol*. (2019) 9:1509. doi: 10.3389/fonc.2019.01509
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. (2008) 9:559. doi: 10.1186/1471-2105-9-559
- Yao Y, Zhang T, Qi L, Zhou C, Wei J, Feng F, et al. Integrated analysis of co-expression and ceRNA network identifies five lncRNAs as prognostic markers for breast cancer. *J Cell Mol Med*. (2019) 23:8410–9. doi: 10.1111/jcmm.14721
- Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CCY, O'Donovan MC, et al. Methyloomic trajectories across human fetal brain development. *Genome Res*. (2015) 25:338–52. doi: 10.1101/gr.180273.114
- Liu Q, Jiang C, Xu J, Zhao M-T, Van Bortle K, Cheng X, et al. Genome-wide temporal profiling of transcriptome and open chromatin of early cardiomyocyte differentiation derived from hiPSCs and hESCs. *Circulat Res*. (2017) 121:376–91. doi: 10.1161/CIRCRESAHA.116.310456
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. (2011) 146:353–8. doi: 10.1016/j.cell.2011.07.014
- Karreth FA, Pandolfi PP. ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Discov*. (2013) 3:1113–21. doi: 10.1158/2159-8290.CD-13-0202
- Zhang K, Li Q, Kang X, Wang Y, Wang S. Identification and functional characterization of lncRNAs acting as ceRNA involved in the malignant progression of glioblastoma multiforme. *Oncol Rep*. (2016) 36:2911–25. doi: 10.3892/or.2016.5070
- Wang J-D, Zhou H-S, Tu X-X, He Y, Liu Q-F, Liu Q, et al. Prediction of competing endogenous RNA coexpression network as prognostic markers in AML. *Aging*. (2019) 11:3333–47. doi: 10.18632/aging.101985
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. (2015) 43:e47. doi: 10.1093/nar/gkv007
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. (2000) 28:27–30.
- Brenner AK, Bruserud Ø. Functional toll-like receptors (TLRs) Are expressed by a majority of primary human acute myeloid leukemia cells and inducibility

- of the TLR signaling pathway is associated with a more favorable phenotype. *Cancers (Basel)*. (2019) 11:973. doi: 10.3390/cancers11070973
25. Dore AI, Santana-Lemos BAA, Coser VM, Santos FLS, Dalmazzo LF, Lima ASG, et al. The association of ICAM-1 Exon 6 (E469K) but not of ICAM-1 Exon 4 (G241R) and PECAM-1 Exon 3 (L125V) polymorphisms with the development of differentiation syndrome in acute promyelocytic leukemia. *J Leukoc Biol*. (2007) 82:1340–3.
  26. Yanagi T, Mizuochi T, Takaki Y, Eda K, Mitsuyama K, Ishimura M, et al. Novel exonic mutation inducing aberrant splicing in the IL10RA gene and resulting in infantile-onset inflammatory bowel disease: a case report. *BMC Gastroenterol*. (2016) 16:10. doi: 10.1186/s12876-016-0424-5
  27. Glocker E-O, Kotlarz D, Boztug K, Gertz EM, Schäffer AA, Noyan F, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med*. (2009) 361:2033–45. doi: 10.1056/NEJMoa0907206
  28. Wan B, Liu B, Huang Y, Lv C. Identification of genes of prognostic value in the ccRCC microenvironment from TCGA database. *Mol Genet Genomic Med*. (2020) 8:e1159. doi: 10.1002/mgg3.1159
  29. Li S, Xu W. Mining TCGA database for screening and identification of hub genes in kidney renal clear cell carcinoma microenvironment. *J Cell Biochem*. (2019). 121:3952–3960. doi: 10.1002/jcb.29511
  30. Lichtenegger FS, Krupka C, Haubner S, Köhnke T, Subklewe M. Recent developments in immunotherapy of acute myeloid leukemia. *J Hematol Oncol*. (2017) 10:142. doi: 10.1186/s13045-017-0505-0
  31. Sanchez-Mejias A, Tay Y. Competing endogenous RNA networks: tying the essential knots for cancer biology and therapeutics. *J Hematol Oncol*. (2015) 8:30. doi: 10.1186/s13045-015-0129-1
  32. Fan C-N, Ma L, Liu N. Systematic analysis of lncRNA-miRNA-mRNA competing endogenous RNA network identifies four-lncRNA signature as a prognostic biomarker for breast cancer. *J Trans Med*. (2018) 16:264. doi: 10.1186/s12967-018-1640-2
  33. Martens JHA. Acute myeloid leukemia: a central role for the ETS factor ERG. *Int J Biochem Cell Biol*. (2011) 43:1413–6. doi: 10.1016/j.biocel.2011.05.014
  34. Metzeler KH, Dufour A, Benthaus T, Hummel M, Sauerland M-C, Heinecke A, et al. ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: a comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. *J Clin Oncol*. (2009) 27:5031–8. doi: 10.1200/JCO.2008.20.5328
  35. Faber ZJ, Chen X, Gedman AL, Boggs K, Cheng J, Ma J, et al. The genomic landscape of core-binding factor acute myeloid leukemias. *Nat Genet*. (2016) 48:1551–6. doi: 10.1038/ng.3709
  36. Eisfeld AK, Kohlschmidt J, Schwind S, Nicolet D, Blachly JS, Orwick S, et al. Mutations in the CCND1 and CCND2 genes are frequent events in adult patients with t(8;21)(q22;q22) acute myeloid leukemia. *Leukemia*. (2017) 31:1278–85. doi: 10.1038/leu.2016.332
  37. Chapuis N, Tamburini J, Cornillet-Lefebvre P, Gillot L, Bardet V, Willems L, et al. Autocrine IGF-1/IGF-1R signaling is responsible for constitutive PI3K/Akt activation in acute myeloid leukemia: therapeutic value of neutralizing anti-IGF-1R antibody. *Haematologica*. (2010) 95:415–23. doi: 10.3324/haematol.2009.010785
  38. Tuborgh A, Meyer C, Marschalek R, Preiss B, Hasle H, Kjeldsen E. Complex three-way translocation involving MLL, ELL, RREB1, and CMAHP genes in an infant with acute myeloid leukemia and t(6;19;11)(p22.2;p13.1;q23.3). *Cytogenet Genome Res*. (2013) 141:7–15. doi: 10.1159/000351224
  39. Zeybek A, Öz N, Kalemci S, Edgünlü T, Kızıltuğ MT, Tosun K, et al. Diagnostic value of MiR-125b as a potential biomarker for stage I lung adenocarcinoma. *Curr Mol Med*. (2019) 19:216–27. doi: 10.2174/1566524019666190314113800
  40. Zhou X-G, Huang X-L, Liang S-Y, Tang S-M, Wu S-K, Huang T-T, et al. Identifying miRNA and gene modules of colon cancer associated with pathological stage by weighted gene co-expression network analysis. *Oncotargets Therapy*. (2018) 11:2815–30. doi: 10.2147/OTT.S163891
  41. Xue Q, Sun K, Deng H-J, Lei S-T, Dong J-Q, Li G-X. MicroRNA-338-3p inhibits colorectal carcinoma cell invasion and migration by targeting smoothened. *Jpn J Clin Oncol*. (2014) 44:13–21. doi: 10.1093/jjco/hyt181
  42. Yang M, Liu R, Sheng J, Liao J, Wang Y, Pan E, et al. Differential expression profiles of microRNAs as potential biomarkers for the early diagnosis of esophageal squamous cell carcinoma. *Oncol Rep*. (2013) 29:169–76. doi: 10.3892/or.2012.2105

**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.

Copyright © 2020 Wang, Yang, Liu, Xu, Zhang, Jiang, Wang and Liu. 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.