Received: 16 January 2018 | Accepted: 21 August 2018

DOI: 10.1002/jcb.27679

RESEARCH ARTICLE

WILEY Journal of Cellular Biochemistry

A 4-microRNA signature for survival prognosis in pediatric and adolescent acute myeloid leukemia

Ruiqi Zhu[1] | **Wenyi Lin**[1] | **Weiwei Zhao**[2] | **Fengjuan Fan**[1] | **Liang Tang**[1] | **Yu Hu**[1]

[1]Institute of Hematology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Department of Epidemiology and Biostatistics, School of Public Health, Harbin Medical University, Harbin, China

Correspondence
Yu Hu, Institute of Hematology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China.
Email: dr_huyu@126.com

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 31620103909, 81500109; Clinical Research Physician Program of Tongji Medical College, HUST; Integrated Innovative Team for Major Human Diseases Program of Tongji Medical College, HUST

Abstract

Acute myeloid leukemia (AML) is a hematologic malignancy with significant molecular heterogeneity. MicroRNAs (miRNAs) play a critical role in AML diagnosis, pathogenesis, and prognosis of AML. Little has been done to identify a miRNA signature in pediatric and adolescent patients for predicting overall survival. This study aims to identify a panel of miRNA signature that could predict the prognosis of all younger AML patients with all subtypes of AML by analyzing data from The Cancer Genome Atlas (TCGA). A total of 229 patients under 23 years with miRNA data and corresponding clinical data from TCGA database were enrolled in this study. Through conducting multivariate analysis in the training test, it was identified that the high expression of hsa-miR-509 and hsa-miR-542 were independent poor prognostic factors, whereas that of hsa-miR-146a and hsa-miR-3667 had a trend to be favorable factors. A 4-miRNA signature was constructed by these miRNAs considering the weight of each. In testing group and all 229 patients' cohort as well as 59 cytogenetically normal AML (CN-AML) patients' cohort, higher risk score was associated with shorter overall survival (OS). All results were confidential by using powerful statistical analysis. Gene ontology and Kyoto Encyclopedia of Genes and Genomes analysis were carried out to further develop leukemia-relevant mechanisms supporting the model. The results indicate that the 4-miRNA-based signature is a reliable prognostic biomarker for pediatric and adolescent AML patients.

KEYWORDS

a 4-miRNA signature, pediatric and adolescent AML, prognosis

1 | INTRODUCTION

Acute myeloid leukemia (AML) is rare in children and young adults, with an incidence of approximately seven cases per million children, annually.[1] MicroRNAs (miRNAs) are small noncoding RNA molecules which regulate gene expression at the posttranscriptional level[2] and thus show huge potential in predicting the prognosis of various kinds of cancers including AML.[3-7] Although there are broad similarities in pathogenesis and genetic background between younger and older AML patients, differences still exist. For example, standard-risk (patients with nonhigh-risk AML) group now accounts for 30% to 40% in pediatric AML, which is considerably larger than in adult AML.[1] Moreover, the frequency of AML-associated mutations differ between pediatric and adult AML.[8] Since the different frequency of molecular genetic alterations was also shown to affect miRNA expression,[9] an miRNA signature for young AML patients is urgently needed. So far, most studies have

Ruiqi Zhu, Wenyi Lin, and Yu Hu have contributed equally.

focused mainly on older AML patients while few studies have investigated the prognostic value of miRNAs in pediatric and adolescent AML patients. In this study, we developed and validated a 4-miRNA expression-based prognostic signature in pediatric and adolescent AML patients using data set extracted from The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/). Bioinformatic analysis was also applied to develop further mechanisms underlying the signature.

2 | METHODS AND MATERIALS

2.1 | AML data set from TCGA

The expression levels of miRNA and corresponding clinical information of pediatric and adolescent AML patients were downloaded from TCGA database. Original data included 1539 miRNA expression data and clinical information of 491 patients. MiRNAs whose expression level equal 0 reads per million in more than 25% of all observations were eliminated using R Bioconductor (version 3.3.2) (Fred Hutchinson Cancer Research Center, Seattle, WA). Some patients had two rows of miRNA information and these two rows were merged into one row by taking the average. Patients were selected according to the following criteria. First, patients without miRNA information were excluded. Second, patients who were alive and whose last contact days were unavailable were discarded. Third, patients with less than 1-month follow-up were excluded in the subsequent analysis. Finally, 229 patients were enrolled in this study. The selected patients were divided into a training set ($n = 153$) and testing set ($n = 76$) randomly. RNA-Seq data of TARGET-AML were also downloaded from TCGA database. As the data was extracted from TCGA, ethical approval and informed consent were not needed.

2.2 | Statistical analysis

Each miRNA was transformed into a binary variable and was put into univariate Cox model in the training set. The *P* value and false discovery rate (FDR) adjustment were both used to select the miRNAs which were significantly associated with overall survival (OS). Significant parameters were filtered out using 0.01 as the cutoff in both *P* value and FDR adjustment. Clinical variables including gender, ethnicity, white blood cell (WBC) counts, peripheral blasts karyotype, and genetic mutations reported to be associated with prognosis (FLT3-ITD, NPM1, CEBPA) were also put into univariate Cox model under the same standard. Each significant miRNA identified by univariate proportional hazards regression model in the training set was further evaluated in multivariate Cox model (Akaike's Information Criterion). Then, miRNAs that were significantly associated with OS in the multivariate tests ($P < 0.05$) were selected to generate the risk scoring system. To determine whether the miRNA signature can independently predict the prognosis, the risk score of each patient in the testing set was calculated using the risk scoring system, Kaplan-Meier curve was used out to check out the efficacy of the signature. In addition, the Kaplan-Meier curve was also used in cytogenetically normal AML (CN-AML) patients.

A 1000 times random permutation test to ensure the performance of our scoring system was performed in the testing set. In brief, we take the combination of overall survival time and vital status of patients in the research as a label, then each individual in our study has a label, and a risk score which is calculated using the proposed miRNA-scoring system. A random system was constructed by assigned labels randomly to individuals while the risk score keeps consistent with each individual. The random system was tested for survival significance. If the model performs well, a random system cannot predict the prognosis of patients and the area under the receiver operating characteristic curve (AUC) was supposed to equal 0.5. A thousand random systems were created by R. After all iterations, significance between AUC of random systems and the right label system was measured by *P* value with a cutoff of 0.05. The *P* value calculated greater than 0.05 means the 4-miRNA signature have no effect on the outcome.

2.3 | Bioinformatic analysis

We used TargetScan, miRanda, and miRTarBase to find the target genes of four miRNAs. RNA-Seq data were used to validate the target genes of the four miRNAs. A linear correlation was applied to test whether the target genes predicted by tools were correlated with particular miRNAs, with a cutoff *P* value of 0.05.

Gene ontology (GO) analysis was carried out by The Database for Annotation, Visualization, and Integrated Discovery (DAVID) online and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was carried out by KOBAS online.

3 | RESULTS

3.1 | TCGA data set and patient characteristics

In this study, a total of 229 pediatric and adolescent AML patients from TCGA database were enrolled. To increase the efficiency and accuracy of the training set, the number of patients enrolled in the training set ($n = 153$) was twice as that in the testing set ($n = 76$). No significant

TABLE 1 Characteristics of the study population of the training and the testing sets

Variable	Total, $n = 229$	Training set, $n = 153$	Testing set, $n = 76$	P
Gender				0.400
Male	119 (52.0%)	83 (54.3%)	36 (47.4%)	
Female	110 (48.0%)	70 (45.8%)	40 (52.6%)	
Average age, y	9.93 (0.03-22.55)	10.02 (0.03-22.55)	9.74 (0.38-18.96)	0.726[a]
Ethnicity				
Hispanic or Latino	44 (19.2%)	31 (20.3%)	13 (17.1%)	0.817
Not Hispanic or Latino	165 (72.0%)	115 (75.2%)	60 (78.9%)	
Unknown	10 (4.4%)	7 (4.6%)	3 (3.9%)	
Vital status				
Alive	132 (57.6%)	88 (57.5%)	44 (57.9%)	>0.999
Dead	97 (42.4%)	65 (42.5%)	32 (42.1%)	
Lab data				
WBC ($10^3/\mu L$)	76.74 (0.9-519)	80.34 (0.9-446)	69.48 (1.6-519)	0.387[a]
Peripheral blasts ($10^3/\mu L$)	57.24 (0-97)	58.33 (0-97)	55.06 (0-95)	0.408[a]
CNS disease				>0.999
Yes	16 (7.0%)	11 (7.2%)	5 (6.6%)	
No	213 (93.0%)	142 (92.8%)	71 (93.4%)	
FAB				0.949
M0	4 (1.7%)	2 (1.3%)	2 (2.6%)	
M1	29 (12.7%)	20 (13.1%)	9 (11.8%)	
M2	57 (24.9%)	37 (24.2%)	20 (26.3%)	
M4	58 (25.3%)	39 (25.5%)	19 (25.0%)	
M5	44 (19.2%)	30 (19.6%)	14 (18.4%)	
M6	2 (0.9%)	1 (0.7%)	1 (1.3%)	
M7	6 (2.6%)	3 (2.0%)	3 (3.9%)	
NOS	11 (4.8%)	7 (4.6%)	4 (5.3%)	
Unknown	18 (7.9%)	14 (9.2%)	4 (5.3%)	
Karyotype				0.793
Favorable	90 (39.3%)	60 (39.2%)	30 (39.5%)	
Intermediate	46 (20.1%)	29 (19.0%)	17 (22.4%)	
Unfavorable	93 (40.6%)	64 (41.8%)	29 (38.2%)	

Abbreviations: AML, acute myeloid leukemia; CNS, central nervous system; FAB, French-American-British; NOS, not otherwise specified; WBC, white blood cells.

The χ^2 test was used in Table 1 except items described above.

[a]The t test was used to test difference between average age, WBC, and peripheral blasts in the two sets.

difference in clinical variants was observed between the two sets (Table 1).

3.2 | Identification of a 4-miRNA risk scoring signature

A univariate Cox regression analysis was carried out to evaluate the prognostic value of each miRNA in the training set. Thirteen miRNAs were significantly associated with OS ($P < 0.01$), including hsa-miR-509, hsa-miR-542,

TABLE 2 Locations of miRNAs in the signature on chromosome

miRNAs	Location on chromosome
has-miR-509	Xq
has-miR-542	Xq
has-miR-3667	22q
hsa-miR-146a	5q

Abbreviation: miRNA, microRNA.

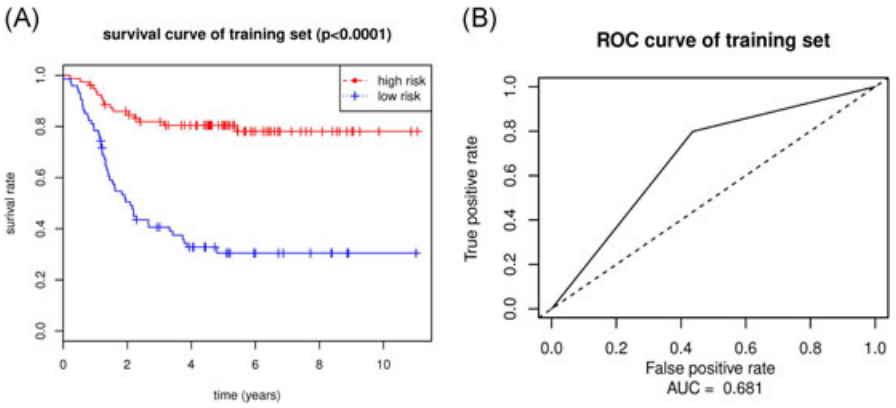

FIGURE 1 A, Kaplan-Meier curve for high score and low score in training set. B, ROC curve of the signature in training set, AUC = 0.681. AUC, area under receiver operating characteristic curve; ROC, receiver operating characteristic

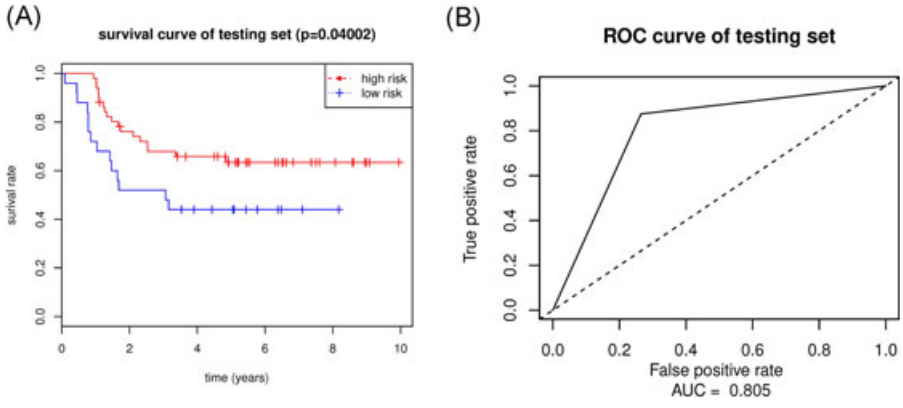

FIGURE 2 A, Kaplan-Meier curve for high score and low score in testing set. B, ROC curve of the signature in testing set, AUC = 0.805. AUC, area under receiver operating characteristic curve; ROC, receiver operating characteristic

hsa-miR-181a, hsa-miR-3667, hsa-miR-429, hsa-miR-185, hsa-miR-196b, hsa-miR-146a, hsa-miR-362, hsa-miR-196a, and hsa-miR-20b (ranked by increasing Cox *P* values). Next, a multivariate Cox model was introduced to pinpoint the microRNAs that could independently predict survival. Finally, we identified four miRNAs which were significantly related with patient survival: including hsa-miR-509, hsa-miR-542, hsa-miR-3667, and hsa-miR-146a. The chromosomal locations of the miRNAs are listed in Table 2. According to the result, it can be known that the high expression of hsa-miR-509 and hsa-miR-542 were independently associated with poor OS, while that of hsa-miR-3667 and hsa-miR-146a had association with favorable OS, with a multivariate Cox *P* value equal to 0.001, 0.003, 0.002, and 0.001, respectively. Then, a formula was constructed using the four miRNA expression status of the four miRNAs and their weights on OS which is represented by β coefficient in multivariate Cox model. The risk score = (0.914* status of hsa-miR-509) + (0.759* status of hsa-miR-542) – (0.837* hsa-miR-3667) – (0.856* status of hsa-miR-146a), where * is the product sign. In this formula, the high miRNA expression (expression level greater than the median) status equals 1 and the low expression status equals 0. Next, risk score was calculated in 153 patients of the training set and 76 patients of the testing set. The ones whose risk scores were greater than the median in the training set were assigned to a high-risk group and the others belong to the low-risk group.

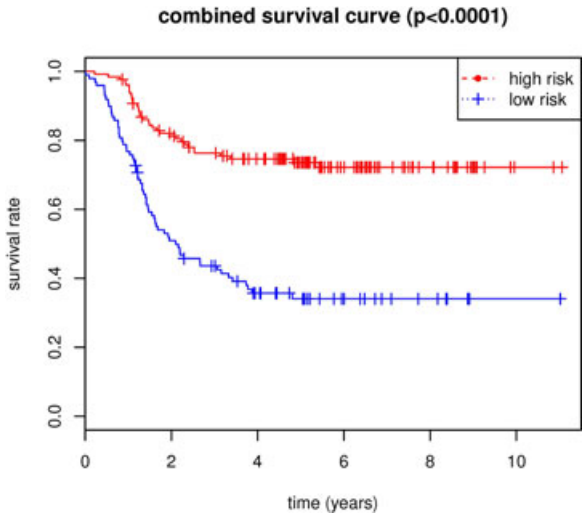

FIGURE 3 Kaplan-Meier curve for high score and low score in all 229 patients

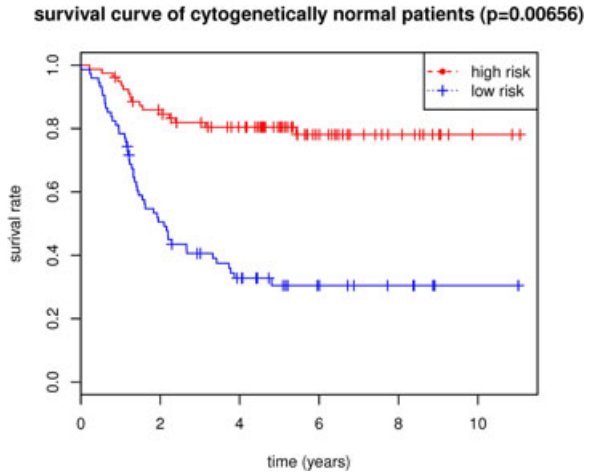

FIGURE 4 Kaplan-Meier curve for high score and low score in CN-AML patients. AML, acute myeloid leukemia; CN-AML, cytogenetically normal AML

3.3 | Survival analysis and performance of the 4-miRNA signature

Pediatric and adolescent AML patients in high-risk group had a significantly shorter OS compared with those in low-risk group ($P = 0.000$; Figure 1A). The survival analysis of the 4-miRNA signature was further conducted in the testing set and the entire 229 pediatric and adolescent AML cohorts. Both patients in high-risk group showed significant shorter OS, with a P value of 0.040 and 0.000 (Figures 2A and 3), respectively. To figure out how the signature performs in patients with normal karyotype, the Kaplan-Meier curve was used for 59 CN-AML patients. It was found in CN-AML group, patients with higher risk scores were still associated with shorter OS ($P = 0.007$; Figure 4). Relapse in AML patients always represents unfavorable prognosis. We wonder whether the signature could predict relapse. Kaplan-Meier curve was used in 229 pediatric and adolescent AML cohort to estimate the performance of 4-miRNA signature in predicting relapse. As shown in Figure 5A, patients with higher risk scores were significantly associated with a shorter relapse-free survival ($P = 0.0334$), which indicates patients with higher risk scores are easier to succumb to relapse.

Receiver operating characteristic (ROC) curves were also applied to evaluate the performance of miRNA signature. The area under ROC curve (AUC) in the training set and the testing set was 0.681 and 0.805, respectively (Figures 1B and 2B). When the 4-miRNA signature was used to predict relapse at 1 year, the AUC in 229 pediatric and adolescent AML cohort was 0.693 (Figure 5B).

The figure and heatmap in 229 AML cohort were also created to evaluate the signature. It can be clearly seen that the high-risk group had more death cases than low-risk group (Figure 6). To validate whether the 4-miRNA signature is able to apply to other pediatric and adolescent AML patients, we did permutation test in the testing group and found that the AUC of random systems showed great significance with that of our studied cohort ($P = 0.000$; Figure 7). These results indicate that our model could successfully predict the prognosis of young AML patients.

3.4 | Correlation between miRNAs signature and clinical characteristics

The association between 4-miRNA signature and clinical variants as well as that between 4-miRNA signature and genetic mutations were examined in 229 AML cohort (Tables 3 and 4). Of all variants, FAB category, Karyotype, and CEBPA mutation were related with risk

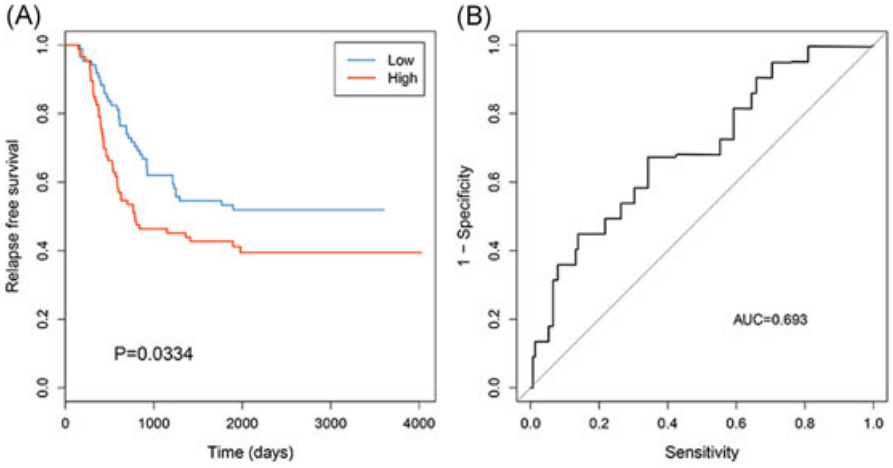

FIGURE 5 A, Kaplan-Meier curve for high score and low score in all 220 patients (*Y*, relapse-free survival). B, ROC curve of the signature in all 229 patients (at 1-year relapse; AUC = 0.693). AUC, area under receiver operating characteristic curve; ROC, receiver operating characteristic

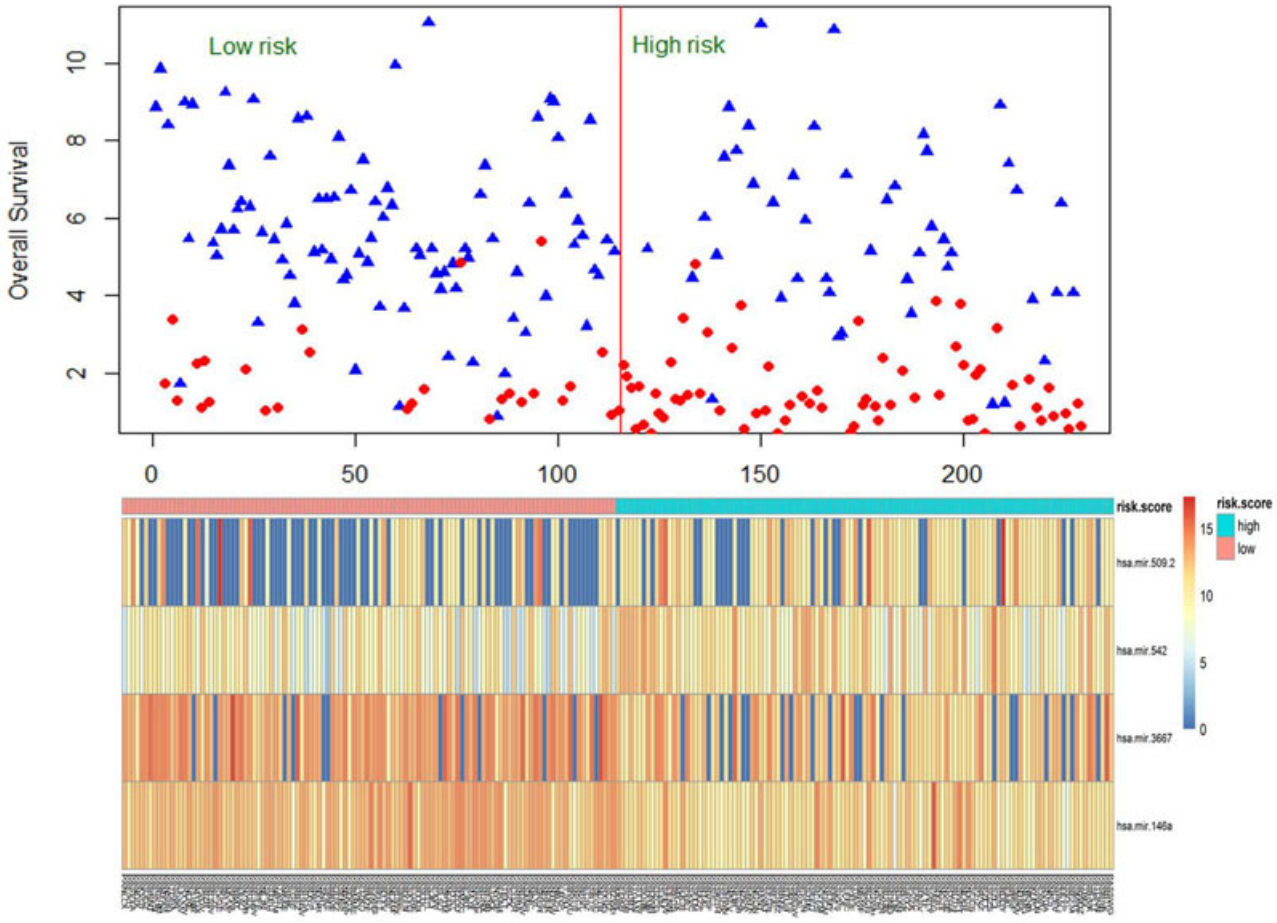

FIGURE 6 4-miRNA expression along with risk score in heatmap and outcome of 229 pediatric and adolescent patients. The figure above: survival status and survival time together with risk score were shown in this graph. Red points represent dead patients and blue triangles represent alive patients. Red line in the middle of the picture separate patients into high risk and low risk through risk score. The figure below: heatmap of miRNA expression in all patients. Labels in the bottom of heatmap represent patient id. AML, acute myeloid leukemia; CN-AML, cytogenetically normal AML; miRNA, microRNA

score. The high-risk score group was more likely to have M5 subtype ($P = 0.000$) and unfavorable karyotype ($P = 0.001$) while the low-risk score group was more likely to often possess M1 ($P = 0.040$), M2 ($P = 0.002$) subtype, favorable karyotype ($P = 0.000$), and CEBPA mutation ($P = 0.048$). Although the difference in NPM1 mutation between the patients with lower scores and those with higher scores was not statistically significant, patients with lower risk scores tend to have NPM1 mutations (12 vs 5; $P = 0.140$).

3.5 | Multivariate analysis in 4-miRNA signature and clinical variants

Whether the proposed scoring signature could function as an independent prognostic factor needs to be detected. Univariate analysis was conducted for clinical variants and genetic mutations in the 229 AML

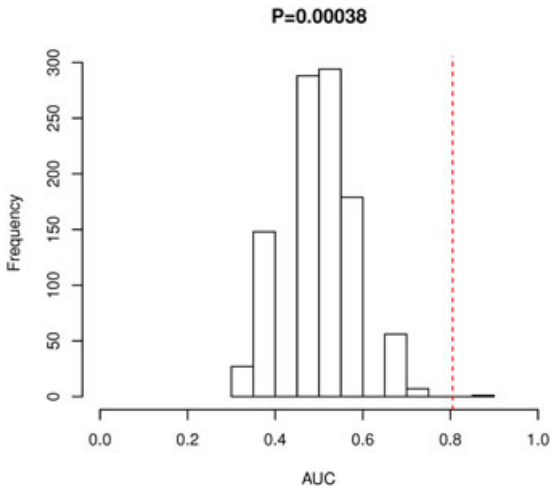

FIGURE 7 Outcome of permutation test. Red line represents AUC of the correct label of the testing set and columns represent AUC of random systems. Thousand random systems were created for the permutation test. AUC, area under receiver operating characteristic curve; ROC, receiver operating characteristic

TABLE 3 Correlation between microRNA score and clinical data, FAB subtypes and chromosomal abnormalities in all 229 pediatric and adolescent AML patients

Variant	Total, $n = 229$	MicroRNA score		P
		Low, $n = 115$	High, $n = 114$	
Gender				0.945
Male	119 (52.0%)	59	60	
Female	110 (48.0%)	56	54	
Average age, y	9.92 (0.03-22.55)	9.72 (0.38-18.85)	10.13 (0.03-22.55)	0.592[a]
Ethnicity				0.421
Hispanic or Latino	44 (19.2%)	26 (22.6%)	18 (15.8%)	
Not Hispanic or Latino	175 (76.4)	84 (73.0%)	91 (79.8%)	
Unknown	10 (4.4%)	5 (4.3%)	5 (4.4%)	
Laboratory data				
WBC ($10^3/\mu L$)	76.73 (0.9-519)	76.42 (1.6-519)	77.05 (0.9-446)	0.985[a]
Peripheral blasts ($10^3/\mu L$)	57.27 (0-97)	60.27 (0-97)	54.24 (0-97)	0.097[a]
CNS disease				0.810
Yes	16 (7.0%)	9 (7.8%)	7 (6.1%)	
No	213 (93.0%)	106 (92.2%)	107 (93.9%)	
FAB category				
M0	4 (1.7%)	2 (1.7%)	2 (1.8%)	>0.999
M1	29 (12.7%)	21 (18.3%)	8 (7.0%)	**0.040**
M2	57 (24.9%)	39 (33.9%)	18 (15.8%)	**0.002**
M4	58 (25.3%)	31 (27.0%)	27 (23.7%)	0.676
M5	44 (19.2%)FAB, M	10 (8.7%)	34 (29.8%)	**0.000**
M6	2 (0.9%)	1 (0.9%)	1 (0.9%)	>0.999
M7	6 (2.6%)	2 (1.7%)	4 (3.5%)	0.671
NOS	11 (48.0%)	4 (3.5%)	7 (6.1%)	0.563
Unknown	18 (7.9%)	5 (4.3%)	13 (11.4%)	0.082
Karyotype				
Favorable	90 (39.3%)	64 (55.6%)	26 (22.8%)	**0.000**
Intermediate	46 (20.1%)	17 (14.8%)	29 (25.4%)	0.065
Unfavorable	93 (40.6%)	34 (29.6%)	59 (51.8%)	**0.001**

Abbreviation: AML, acute myeloid leukemia; CNS, central nervous system; FAB, French-American-British; NOS, not otherwise specified; WBC, white blood cells.

The χ^2 test was used in Table 1 except items described above. Bold values indicate $P < 0.05$.

[a]The t test was used to test difference between average age, WBC, and peripheral blasts in the two sets.

cohort. Clinical variants were reported to be associated with prognosis include gender, age, ethnicity, WBC, peripheral blasts, and karyotype (Table 5). It was found that patients with unfavorable karyotype were more likely to suffer shorter OS ($P = 0.000$) while patients with NPM1 mutation ($P = 0.025$) tended to have longer OS. After putting karyotype, NPM1 mutations, and miRNA risk score into multivariate analysis, the 4-miRNA signature was still proved to be a significant independent prognostic factor in younger AML patients ($P = 0.000$; Table 6).

TABLE 4 Correlation between microRNA score and gene alterations in 229 pediatric and adolescent AML patients

Mutations	Total, $n = 227$	MicroRNA score		P
		Low, $n = 114$	High, $n = 112$	
NPM1	17 (7.5%)	12 (10.5%)	5 (4.5%)	0.140
CEBPA	17 (7.5%)	13 (11.4%)	4 (3.6%)	**0.048**
FLT3-ITD	35 (15.4%)	16 (14.0%)	19 (17.0%)	0.671

Abbreviations: AML, acute myeloid leukemia. Bold values indicate $P < 0.05$.

TABLE 5 Univariate Cox regression analysis for overall survival in all 229 pediatric and adolescent AML patients

Variables	Hazard ratio	95% CI	*P*
Gender	1.320	0.885-1.967	0.173
Age	1.002	0.673-1.492	0.993
Ethnicity	1.232	0.869-1.748	0.241
WBC ($10^3/\mu L$)	0.846	0.565-1.267	0.417
Peripheral blasts ($10^3/\mu L$)	1.083	0.724-1.619	0.698
Karyotype	1.582	1.253-1.999	**0.000**
NPM1 mutation	0.201	0.050-0.818	**0.025**
CEBPA mutation	0.345	0.109-1.091	0.070
FLT3-ITD mutation	1.546	0.924-2.586	0.097

Abbreviations: AML, acute myeloid leukemia; CI, confidence interval; WBC, white blood cell. Bold values indicate $P < 0.05$.

3.6 | Bioinformatic analysis of target genes and pathways

Target genes of hsa-miR-509, hsa-miR-542, hsa-miR-3667, and hsa-miR-146a were predicted by three prediction tools including TargetScan, miRanda, and miRTarBase. A total of 148 target genes were extracted in this experiment (Supplementary Table S1). These targets predicted by the tools may not reflect the actual situation. This was mainly because the efficiency of suppression by the miRNAs was cell type-specific due to lncRNAs and other factors. RNA-Seq data was used to validate whether the expressions of predicted genes were associated with particular miRNAs in pediatric and adolescent AML patients. As shown in Table 7, the expressions of some targets predicted by the tools were indeed associated with miRNAs in signature. GO and KEGG analysis were also carried out and part of the results are displayed in Tables 8 and 9. All results can be seen in Supplementary Table S2.

4 | DISCUSSION

Accumulating evidence reveals that miRNAs play crucial roles in tumorigenesis and prognosis of pediatric and adolescent AML. Many scholars have devoted their effort into creating miRNA risk scoring system for cancer prognosis prediction.[4,7,10-12] In this study, a 4-miRNA signature was created for the first time to predict the prognosis of pediatric and adolescent AML patients regardless of karyotype. A 4-miRNA signature (hsa-miR-509, hsa-miR-542, hsa-miR-3667, and hsa-miR-146a) was constructed based on TCGA TARGET-AML program. The signature was constructed by the training set and was validated by Kaplan-Meier curve and permutation test of the testing set. Kaplan-Meier curve was further introduced into cytogenetically normal group. The high-risk score group was associated with unfavorable prognostic

TABLE 6 Multivariate Cox regression analysis for overall survival in all 229 pediatric and adolescent AML patients

Variables	Hazard ratio	95% CI	*P*
Karyotype	1.369	1.071-1.749	**0.012**
NPM1 mutation	0.244	0.060-0.995	**0.049**
miRNA risk score	2.700	1.705-4.274	**0.000**

Abbreviations: AML, acute myeloid leukemia; CI, confidence interval; miRNA, microRNA. Bold values indicate $P < 0.05$.

TABLE 7 Part of target genes of miRNAs predicted by the tools and validated by RNA-Seq

miRNA	Targets	*P*	*r*
hsa-miR-509	NR1D2	0.001	0.308
hsa-miR-509	BRWD3	0.001	0.295
hsa-miR-509	ZFX	0.001	0.268
hsa-miR-509	ANKS1A	0.009	−0.197
hsa-miR-509	ARHG EF7	0.010	0.196
hsa-miR-509	UHRF1BP1	0.034	0.161
hsa-miR-509	ZNF460	0.049	0.150
hsa-miR-146a	CCL5	0.001	−0.384
hsa-miR-146a	IRAK1	0.001	−0.383
hsa-miR-146a	SMAD4	0.001	0.379
hsa-miR-146a	BRCA1	0.001	−0.301
hsa-miR-146a	TLR4	0.001	−0.293
hsa-miR-146a	PPP1R11	0.001	−0.252
hsa-miR-3667	GRINA	0.001	−0.274
hsa-miR-3667	IRGQ	0.001	0.234
hsa-miR-3667	CLDN12	0.017	0.180
hsa-miR-3667	FOXK2	0.045	−0.152

Abbreviations: miRNA, microRNA; *r*, correlation coefficient.
Correlation coefficient $r > 0$, target genes are positively related with miRNAs; $r < 0$, target genes are negatively related with miRNAs.

TABLE 8 Part of GO analysis results of 148 target genes

Category	ID	Term	Counts	*P*
Cellular components	0005737	Cytoplasm	66	0.000
	0005634	Nucleus	63	0.000
	0000790	Nuclear chromatin	8	0.001
	0005829	Cytosol	40	0.004
	0005654	Nucleoplasm	35	0.004
	0009986	Cell surface	11	0.010
	0000151	Ubiquitin ligase complex	5	0.010
Molecular function	0005515	Protein binding	93	0.000
	0042826	Histone deacetylase binding	7	0.000
	0004842	Ubiquitin-protein transferase activity	9	0.005
	0001076	Transcription factor activity, RNA polymerase II transcription factor binding	3	0.006
Biological process	0031663	Lipopolysaccharide-mediated signaling pathway	5	0.000
	0002755	MyD88-dependent toll-like receptor signaling pathway	5	0.000
	0007250	Activation of NF-κB-inducing kinase activity	4	0.000
	0042346	Positive regulation of NF-κB import into nucleus	4	0.001
	0048661	Positive regulation of smooth muscle cell proliferation	5	0.001
	0045944	Positive regulation of transcription from RNA polymerase II promoter	18	0.002
	0071260	Cellular response to mechanical stimulus	5	0.002
	0051092	Positive regulation of NF-κB transcription factor activity	6	0.004
	0000122	Negative regulation of transcription from RNA polymerase II promoter	14	0.005
	0032968	Positive regulation of transcription elongation from RNA polymerase II promoter	3	0.005
	0042981	Regulation of apoptotic process	7	0.007
	0051865	Protein autoubiquitination	4	0.007
	0007049	Cell cycle	7	0.008
	0051571	Positive regulation of histone H3-K4 methylation	3	0.008

Abbreviations: GO, gene ontology; NF-κB, nuclear factor κB.

factors such as unfavorable karyotype and less frequent favorable genetic mutations. Multivariate analysis in 229 AML cohort confirmed the independence of this 4-miRNA signature from other important prognosis factors.

There was also a significant difference in FAB subtypes also showed significant difference between high-risk and low-risk group. The low-risk group was more likely to have M1 and M2 subtypes while the high-risk group was more likely to have M5 subtypes. The FAB classification system relies primarily on the morphology of leukemia blasts in the bone marrow that indicates at which step the malignant progenitor cells stop differentiating.[13] It was reported that high expression of miR-146a was often accompanied with more M1 subtype rather than M5 subtype, and it was confirmed participating in the block of M1 development and maturation.[13] MiR-146a, with complicated biological function, was also found to act as a tumor suppressor in hematological malignancies.[14-17] Its biological function is very complicated. The mechanism why losing miR-146a could increase myeloproliferation was reported to be dependent on the activation of NF-κB.[18] Our result was consistent with previous studies. So far, miR-509, miR-542, and miR-3667 have not been frequently reported to be associated with AML. MiR-542 was reported to be a negative prognostic factor in osteosarcoma,[19] while other studies regarded it as a tumor suppressor in other cancer types.[20,21] MiR-509 was reported to be a solid tumor suppressor,[22,23] of which the function needs to be further studied in pediatric AML. There have been no report on miR-3667.

To gain a further insight into the functional role of the four miRNAs, we extracted their target genes for

TABLE 9 Part of KEGG pathway results of 148 target genes ($P = 0.05$)

Term	ID	*P*	Input
Toll-like receptor signaling pathway	hsa04620	0.000	TLR2, CCL5, STAT1, TLR4, TRAF6, IRAK1, RAC1
Pathways in cancer	hsa05200	0.000	DAPK2, CCDC6, SMAD4, STAT1, LAMC2, TRAF6, FAS, PTGS2, CRKL, BRCA2, RAC1
NF-κB signaling pathway	hsa04064	0.000	CD40LG, TRAF6, PTGS2, IRAK1, TLR4
Chemokine signaling pathway	hsa04062	0.000	CCL5, STAT1, BCAR1, CRKL, RAC1, CCR9
MAPK signaling pathway	hsa04010	0.000	MAP3K1, TRAF6, FAS, CRKL, RAC1, HSPA1A
HTLV-I infection	hsa05166	0.003	NFATC4, SMAD4, MAP3K1, CRTC1, HLA-A
TNF signaling pathway	hsa04668	0.009	FAS, PTGS2, CCL5
PI3K-Akt signaling pathway	hsa04151	0.010	TLR2, TLR4, RAC1, BRCA1, LAMC2
Apoptosis	hsa04210	0.016	LMNB1, FAS, PMAIP1
Wnt signaling pathway	hsa04310	0.017	NFATC4, SMAD4, RAC1
cGMP-PKG signaling pathway	hsa04022	0.026	NFATC4, MYLK, PRKCE
MicroRNAs in cancer	hsa05206	0.027	PTGS2, BRCA1, CRKL, PRKCE
p53 signaling pathway	hsa04115	0.029	FAS, PMAIP1
Chronic myeloid leukemia	hsa05220	0.032	SMAD4, CRKL
cAMP signaling pathway	hsa04024	0.040	RAC1, PDE4C, CFTR
TGF-β signaling pathway	hsa04350	0.041	SMAD4, SMURF2

Abbreviations: KEGG, Kyoto Encyclopedia of Genes and Genomes; MAPK, mitogen-activated protein kinase; NF-κB, nuclear factor κB; TGF-β, transforming growth factor β.

bioinformatics analysis. In KEGG pathway results, classical AML pathways including MAPK signaling pathway (MAP3K1, TRAF6, FAS, CRKL, RAC1, HSPA1A), PI3K-AKt signaling pathway (TLR2, TLR4, RAC1, BRCA1, LAMC2), and Wnt signaling pathway (NFATC4, SMAD4, RAC1) all showed significant relationship with the four miRNAs in our study.

However, this study is also subject to limitations. First, the censored rate of TCGA TARGET-AML data set was relatively high, which affected the reliability of the survival analysis. Second, given the fact there was no other public miRNA database or studies with survival information in young AML patients for reference, we considered that further validation by large cohorts are needed, although the results were validated in the testing set.

In conclusion, this study identified a 4-miRNA signature as potential prognostic predictor for pediatric and adolescent AML patients using TCGA TARGET-AML data set. Further studies and functional investigations are required to explore the underlying mechanisms of these miRNAs in signature.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

Study design and manuscript preparation were done by Ruiqi Zhu. Statistical analysis was done by Ruiqi Zhu and Weiwei Zhao. Literature research was completed by Wenyi Lin. Manuscript editing was made by Liang Tang and Fengjuan Fan. Yu Hu revised the manuscript and funded this study.

ORCID

Yu Hu http://orcid.org/0000-0002-2815-4568

REFERENCES

1. Zwaan CM, Kolb EA, Reinhardt D, et al. Collaborative efforts driving progress in pediatric acute myeloid leukemia. *J Clin Oncol*. 2015;33:2949-2962.
2. Garzon R, Calin GA, Croce CM. MicroRNAs in cancer. *Annu Rev Med*. 2009;60:167-179.
3. Shivarov V, Dolnik A, Lang KM, et al. MicroRNA expression-based outcome prediction in acute myeloid leukemia: novel insights through cross-platform integrative analyses. *Haematologica*. 2016;101:e454-e456.
4. Xu J, Zhao J, Zhang R. Four microRNAs signature for survival prognosis in colon cancer using TCGA data. *Sci Rep*. 2016;6:38306.
5. Chuang MK, Chiu YC, Chou WC, Hou HA, Chuang EY, Tien HF. A 3-microRNA scoring system for prognostication in de

novo acute myeloid leukemia patients. *Leukemia*. 2015;29:1051-1059.
6. Du F, Yuan P, Zhao ZT, et al. A miRNA-based signature predicts development of disease recurrence in HER2 positive breast cancer after adjuvant trastuzumab-based treatment. *Sci Rep*. 2016;6:33825.
7. Gao X, Wu Y, Yu W, Li H. Identification of a seven-miRNA signature as prognostic biomarker for lung squamous cell carcinoma. *Oncotarget*. 2016;7:81670-81679.
8. Marjanovic I, Kostic J, Stanic B, et al. Parallel targeted next generation sequencing of childhood and adult acute myeloid leukemia patients reveals uniform genomic profile of the disease. *Tumour Biol*. 2016;37:13391-13401.
9. Marcucci G, Mrozek K, Radmacher MD, Garzon R, Bloomfield CD. The prognostic and functional role of microRNAs in acute myeloid leukemia. *Blood*. 2011;117:1121-1129.
10. Marziali G, Buccarelli M, Giuliani A, et al. A three-microRNA signature identifies two subtypes of glioblastoma patients with different clinical outcomes. *Mol Oncol*. 2017;11:1115-1129.
11. Hayes J, Thygesen H, Tumilson C, et al. Prediction of clinical outcome in glioblastoma using a biologically relevant nine-microRNA signature. *Mol Oncol*. 2015;9:704-714.
12. Lu M, Kong X, Wang H, Huang G, Ye C, He Z. A novel microRNAs expression signature for hepatocellular carcinoma diagnosis and prognosis. *Oncotarget*. 2017;8:8775-8784.
13. Lutherborrow M, Bryant A, Jayaswal V, et al. Expression profiling of cytogenetically normal acute myeloid leukemia identifies microRNAs that target genes involved in monocytic differentiation. *Am J Hematol*. 2011;86:2-11.
14. Boldin MP, Taganov KD, Rao DS, et al. miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *J Exp Med*. 2011;208:1189-1201.
15. Starczynowski DT, Morin R, McPherson A, et al. Genome-wide identification of human microRNAs located in leukemia-associated genomic alterations. *Blood*. 2011;117:595-607.
16. Yan W, Guo H, Suo F, Han C, Zheng H, Chen T. The effect of miR-146a on STAT1 expression and apoptosis in acute lymphoblastic leukemia Jurkat cells. *Oncol Lett*. 2017;13:151-154.
17. Zhao JL, Rao DS, Boldin MP, Taganov KD, O'Connell RM, Baltimore D. NF-kappaB dysregulation in microRNA-146a-deficient mice drives the development of myeloid malignancies. *Proc Natl Acad Sci USA*. 2011;108:9184-9189.
18. So AY, Zhao JL, Baltimore D. The Yin and Yang of microRNAs: leukemia and immunity. *Immunol Rev*. 2013;253:129-145.
19. Cheng D, Yu T, Hu T, Yao M, Fan C, Yang Q. MiR-542-5p is a negative prognostic factor and promotes osteosarcoma tumorigenesis by targeting HUWE1. *Oncotarget*. 2015;6:42761-42772.
20. Venkatadri R, Muni T, Iyer AK, Yakisich JS, Azad N. Role of apoptosis-related miRNAs in resveratrol-induced breast cancer cell death. *Cell Death Dis*. 2016;7:e2104.
21. Zhou M, Wang H, Zhou K, et al. A novel EGFR isoform confers increased invasiveness to cancer cells. *Cancer Res*. 2013;73:7056-7067.
22. Wang Y, Cui M, Cai X, et al. The oncoprotein HBXIP up-regulates SCG3 through modulating E2F1 and miR-509-3p in hepatoma cells. *Cancer Lett*. 2014;352:169-178.
23. Xing F, Sharma S, Liu Y, et al. miR-509 suppresses brain metastasis of breast cancer cells by modulating RhoC and TNF-alpha. *Oncogene*. 2015;34:4890-4900.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zhu R, Lin W, Zhao W, Fan F, Tang L, Hu Y. A 4-microRNA signature for survival prognosis in pediatric and adolescent acute myeloid leukemia. *J Cell Biochem*. 2019;120:3958-3968. https://doi.org/10.1002/jcb.27679