



# Identification of Early Recurrence Factors in Childhood and Adolescent B-Cell Acute Lymphoblastic Leukemia Based on Integrated Bioinformatics Analysis

Yan Huang<sup>†</sup>, Jiazheng Li<sup>†</sup>, Yanxin Chen, Peifang Jiang, Lingyan Wang and Jianda Hu\*

Fujian Provincial Key Laboratory of Hematology, Fujian Institute of Hematology, Fujian Medical University Union Hospital, Fuzhou, China

## OPEN ACCESS

### Edited by:

Annalisa Lonetti,  
University of Bologna, Italy

### Reviewed by:

Leila Jahangiri,  
Birmingham City University,  
United Kingdom  
Claudio Cerchione,  
Romagnolo Scientific Institute  
for the Study and Treatment  
of Tumors (IRCCS), Italy

### \*Correspondence:

Jianda Hu  
drjiandahu@163.com

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

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

Received: 25 May 2020

Accepted: 04 September 2020

Published: 29 September 2020

### Citation:

Huang Y, Li J, Chen Y, Jiang P,  
Wang L and Hu J (2020) Identification  
of Early Recurrence Factors  
in Childhood and Adolescent B-Cell  
Acute Lymphoblastic Leukemia  
Based on Integrated Bioinformatics  
Analysis. *Front. Oncol.* 10:565455.  
doi: 10.3389/fonc.2020.565455

Over the past 50 years, great progress has been made in the diagnosis and treatment of acute lymphoblastic leukemia (ALL), especially in pediatric patients. However, early recurrence is still an important threat to the survival of patients. In this study, we used integrated bioinformatics analysis to look for biomarkers of early recurrence of B-cell ALL (B-ALL) in childhood and adolescent patients. Firstly, we obtained gene expression profiles from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database and the Gene Expression Omnibus (GEO) database. Then, we identified differentially expressed genes (DEGs) based on whether the disease relapsed early. LASSO and Cox regression analysis were applied to identify a subset of four genes: *HOXA7*, *S100A11*, *S100A10*, and *IFI44L*. A genetic risk score model was constructed based on these four optimal prognostic genes. Time-dependent receiver operating characteristic (ROC) curves were used to evaluate the predictive value of this prognostic model (3-, 5-, and 10-year AUC values >0.7). The risk model was significantly associated with overall survival (OS) and event-free survival in B-ALL (all  $p < 0.0001$ ). In addition, a high risk score was an independent poor prognostic risk factor for OS ( $p < 0.001$ ; HR = 3.396; 95% CI: 2.387–4.832). Finally, the genetic risk model was successfully tested in B-ALL using an external validation set. The results suggested that this model could be a novel predictive tool for early recurrence and prognosis of B-ALL.

**Keywords:** acute lymphoblastic leukemia, early recurrence, genetic risk score model, differential gene expression, microRNA

## INTRODUCTION

Acute lymphoblastic leukemia is the most common cancer in children, and includes both B-cell and T-cell lineages. B-ALL accounts for about 85% of pediatric ALL (1). With advances in chemotherapy and hematopoietic stem cell transplantation, the cure rate in childhood ALL is currently around 90% (2). The prognosis for adolescent patients aged 15 to 20 years receiving pediatric protocols is

similar to that of children, with a 5-year OS of 87.9% (3). In spite of significant progress in long-term survival, 15–20% of patients will suffer a recurrence, which is an important factor for increases in mortality (4). Relapses occurring less than 18 months after diagnosis are defined as very early relapses, relapses appearing between 18 months of initial diagnosis and 6 months after cessation of frontline therapy is regarded as an early relapse, and relapses developing after 6 months of cessation of frontline treatment are classified as late relapses (5). The time from diagnosis to relapse is an independent risk factor for overall survival (6), and the risk ratios for adverse outcome in patients with late, early, and very early relapse were 1, 2.4, and 2.9, respectively ( $p < 0.001$ ) (5). Patients with late recurrence had better survival rates than those who relapsed earlier (OS: 45–73% vs. 22–38%) (7–9). Thus, identifying patients with a high risk of earlier relapse is crucial to improving prognosis.

Genomic abnormalities have been confirmed to be related to treatment response and disease relapse in acute leukemia (10–12). An understanding of genetic abnormalities can guide the selection of subsequent regimens after the induction of remission. Technological progress has led to more detailed genetic profiling of leukemia, including DNA sequence abnormalities, gene expression abnormalities, chromosomal rearrangements, and abnormal epigenetic modifications. It has been suggested that the development of gene expression profiles for acute leukemia can improve the prediction of prognosis (13). For example, Ng et al., established a 17-gene stemness score for rapid determination of risk in acute myeloid leukemia (AML) (14). Overexpression of Wilms tumor 1 gene indicates an adverse prognosis and poor response to treatment in AML (15). In Ismail et al.'s study, the expression level of the gene *BIRC6* was an adverse risk factor in acute childhood leukemia (16). Therefore, we aimed to investigate the relationship between genetic profiling and early relapse of childhood B-ALL.

In this study, we identified differentially expressed genes (DEGs) between individuals with early recurrence and no early recurrence of B-cell ALL, using data from the GEO and TARGET public databases. We constructed a prognostic risk model based on DEGs, using regression analysis. We verified the prognostic value of this risk model using an external validation set from the TARGET database. Using the potential biomarkers of early relapse in B-cell ALL, we provided a new approach to early diagnosis and treatment of patients at high risk of recurrence.

## MATERIALS AND METHODS

### Database

We downloaded one dataset (GSE13576) from the GEO database<sup>1</sup>. The data from GSE13576 were based on the GPL570 platforms (Affymetrix Human Genome U133 Plus 2.0 Array) and included data on 197 pediatric ALL patients with mRNA expression information. The mRNA expression dataset of 486 B-cell ALL patients and their corresponding clinical information were obtained from the TARGET database<sup>2</sup>. This dataset

included 370 microarray data points and 116 mRNA-seq data points. The data of 370 cases were based on the GPL570 platform. miRNA-seq was performed in 53 of 486 patients from the TARGET dataset.

### Analysis of Differentially Expressed Genes

The definition of early recurrence in this study was less than 24 months from diagnosis. Data from 370 childhood and adolescent B-cell ALL patients from the TARGET database, and 197 pediatric patients from GSE13576 were classified into early relapse and non-early relapse groups. Of the 370 patients from the TARGET database, 127 had early recurrence and 243 did not have early recurrence. In the GSE13576 dataset, 24 patients were in the early relapse group, and 173 patients belonged to the no early recurrence group. We used the limma R package to normalize the data and performed gene differential analysis by comparing the two groups in the R computing environment (Version 3.6.2). The differences in the gene expression levels

**TABLE 1** | Characteristic of patients.

Characteristic	Number of cases	Percentages (%)
<b>Patients</b>	370	100
<b>Median age, years (range)</b>	10 (1–20)	/
<b>Sex</b>		
Female	145	39.2
Male	225	60.8
<b>WBC (<math>\times 10^9/L</math>)</b>		
$\geq 50$	152	41.1
$< 50$	218	58.9
<b>Cytogenetics</b>		
t(9;22)/BCR-ABL fusion gene	2	0.5
t(4;11)/AF4 fusion gene	14	3.8
Trisomy 4/10/17	38	10.3
Others	228	61.6
Unknown	88	23.8
<b>Down's Syndrome</b>		
Positive	32	8.7
Negative	272	73.5
Unknown	66	17.8
<b>CNS<sup>#</sup></b>		
CNS1	291	78.6
CNS2	57	15.4
CNS3	22	6.0
<b>MRD (Day 29)</b>		
$\geq 0.01\%$	164	44.3
$< 0.01\%$	199	53.8
Unknown	7	1.9
<b>Protocol</b>		
POG9906	200	54.1
AALL <sup>&amp;</sup>	170	45.9

<sup>#</sup>CNS, Central Nervous System, CNS1 = CNS negative, CNS2  $\leq 5$  WBC/ $\mu L$  in CSF with blasts, CNS3 = CNS positive; <sup>&</sup>AALL includes protocol AALL0232, 0331 and 03B1. '/' means the characteristic of median age is not suitable for percentage calculation.

<sup>1</sup><https://www.ncbi.nlm.nih.gov/geo/>

<sup>2</sup><https://ocg.cancer.gov/programs/target/>

were adjusted for multiple testing using the Benjamini-Hochberg method. Genes with a  $p$ -value  $< 0.05$  and  $|\text{Log}_2(\text{fold change})| > 0.5$  were defined as DEGs in this study. Finally, a volcano plot and heatmap of the DEGs was drawn using the ggrepel and heatmap R packages, respectively. The Venn diagrams of the DEGs were drawn using the website tool Venny 2.1.0<sup>3</sup>.

## Selection of Candidate Genes

The 370 samples from the TARGET database were randomly separated into training and test (6:4) sets for constructing and validating the prognostic models, using the base R package. Univariate Cox regression analyses were used to explore the correlations between the overall survival of B-cell ALL patients and the expression level of DEGs in training set ( $n = 222$ ). The genes with a  $p$ -value  $< 0.01$  were selected as candidate prognosis-related genes. To further investigate the candidate genes, least absolute shrinkage and selection operator (LASSO) regression analysis was applied using the glmnet R package. This algorithm minimizes the usual sum of squared errors with tenfold cross-validation, and identifies the optimal lambda value. The genes with coefficient are related to prognosis. The optimal prognostic genes derived from the LASSO regression model were validated using the Oncomine website<sup>4</sup>. We selected genes associated with early recurrence as hub genes with which to construct the prognostic model. Unpaired  $t$ -tests were used to compare

the levels of gene expression between groups, using GraphPad Prism (Version 7.03).

## Genetic Risk Score Model Construction

The gene-related prognosis model in this study was established using the training set. We applied a multivariate COX regression model to assess the role of hub genes as independent prognostic risk factors. Then, a four-gene-based genetic risk score model was constructed. The risk score formula is as follows:

$$\text{Risk score} = \sum_{i=1}^n \beta_i \times \text{Exp}_i$$

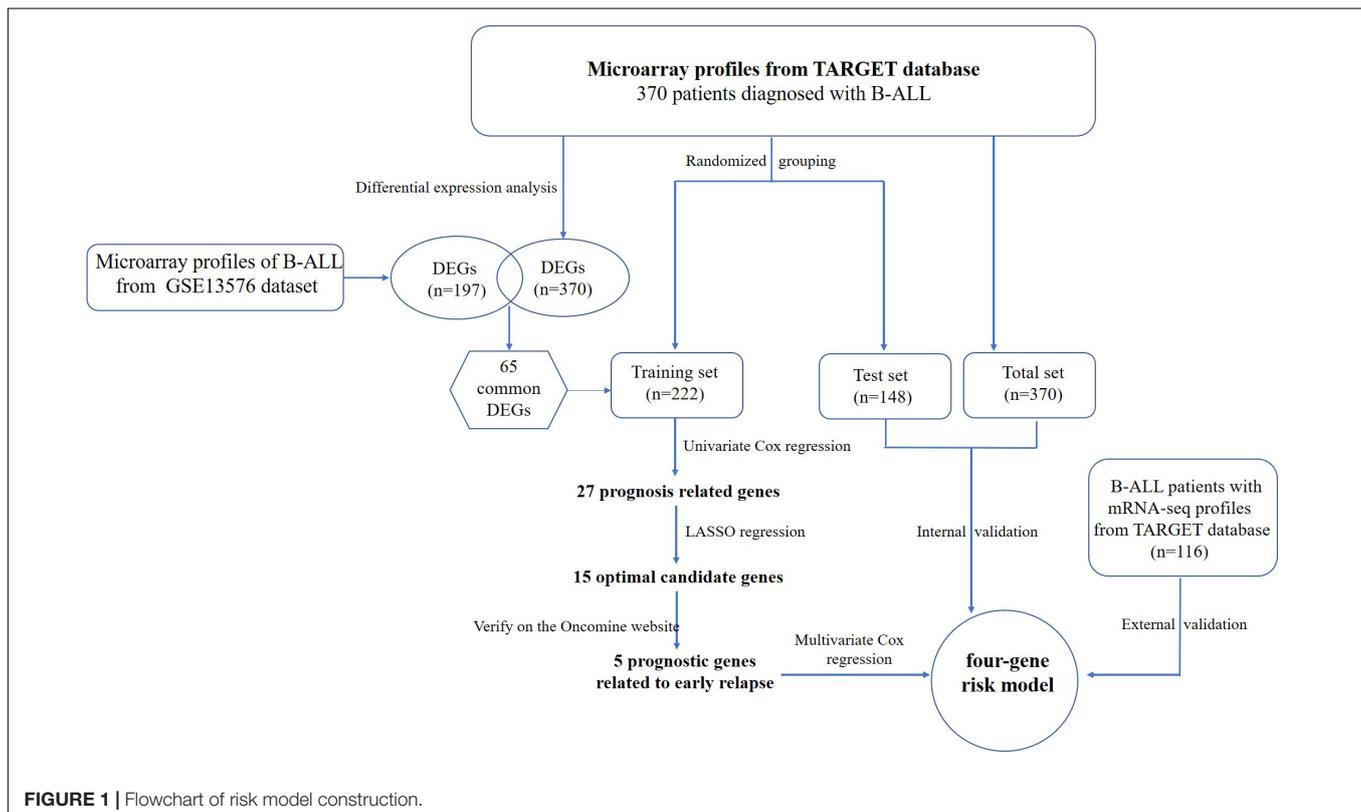
The  $\beta$  in the above formula is the regression coefficient of each gene, and the Exp is the expression level of the gene. The cut-off value of the risk score was defined as the median score. The 222 patients were separated into low- and high-risk groups based on the threshold. To evaluate the predictive power of this risk score model, we used the survival ROC R package to plot time-dependent receiver operating characteristic (ROC) curves, and calculated the area under the ROC curves (AUC).

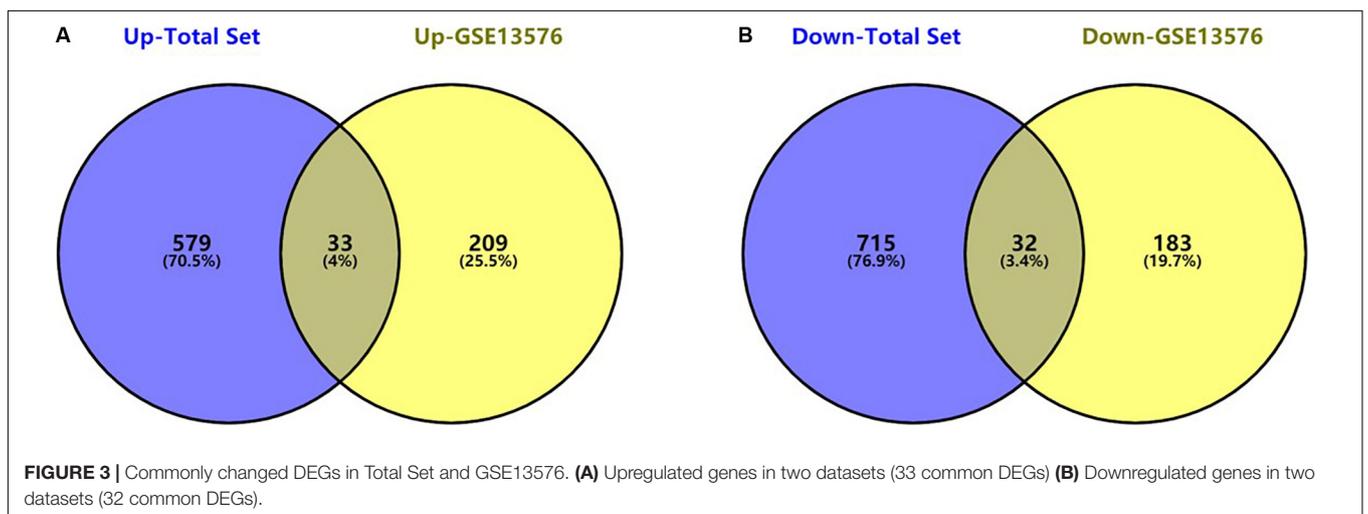
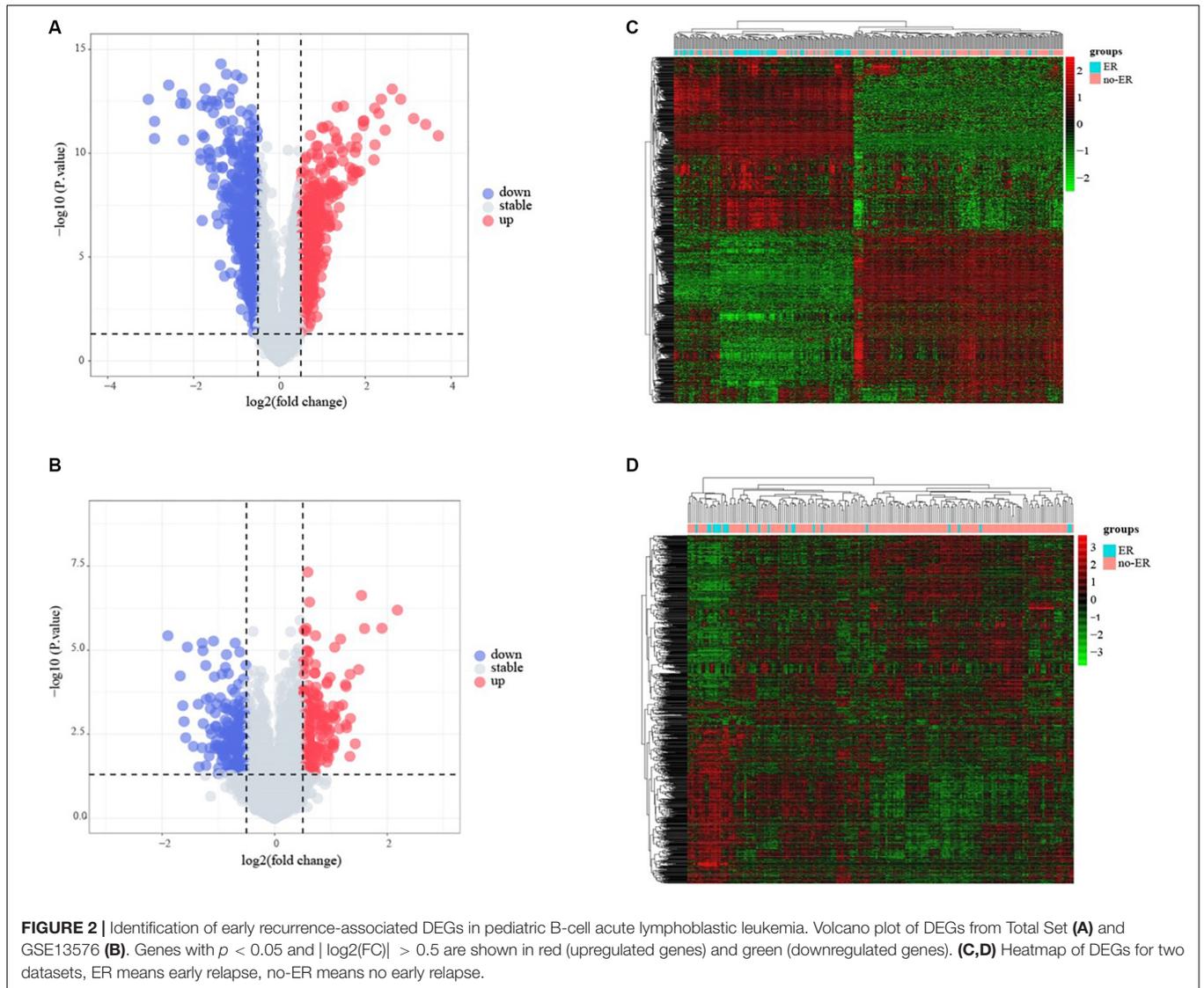
## Validation of the Genetic Risk Score Model

For internal and external validation, the testing set ( $n = 148$ ), total set ( $n = 370$ ) and external validation set ( $n = 116$ ) were used to assess the predictive power of the genetic

<sup>3</sup><https://bioinfogp.cnb.csic.es/tools/venny/index.html>

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





**TABLE 2 |** The common DEGs identified among GSE13576 and total set.

**Gene Symbol (upregulated DEGs)**

TOMM22, COTL1, S100A11, TSPO, HOXA7, CTSC, BIK, TUBB2A, RAPGEF5, SUCLG2, SNX10, TOP1MT, CRIP1, S100A10, ANXA2P2, CXXC5, S100A4, CAMK2D, IGFBP7, S100A6, DAD1, CPVL, PKIG, TUBB6, HK2, WIP1, MEIS1, LGALS1, TST, CHN2, MAP7, PPP1R14A, CLEC11A

**Gene Symbol (downregulated DEGs)**

ZNF91, GFOD1, SFMBT2, ZNF704, SIPA1L2, GBP4, CD69, HERC5, IFI44, TCF4, SHANK3, DAPK1, IFI44L, MYO5C, MRC1, ST3GAL6, MAN1A1, NAV1, FHIT, CYP46A1, STK32B, COL5A1, MX1, PDE4B, S100Z, RASD1, POU4F1, ITGA6, MDK, P2RY14, DDIT4L, DNNT

**TABLE 3 |** The result of univariate Cox regression in training set.

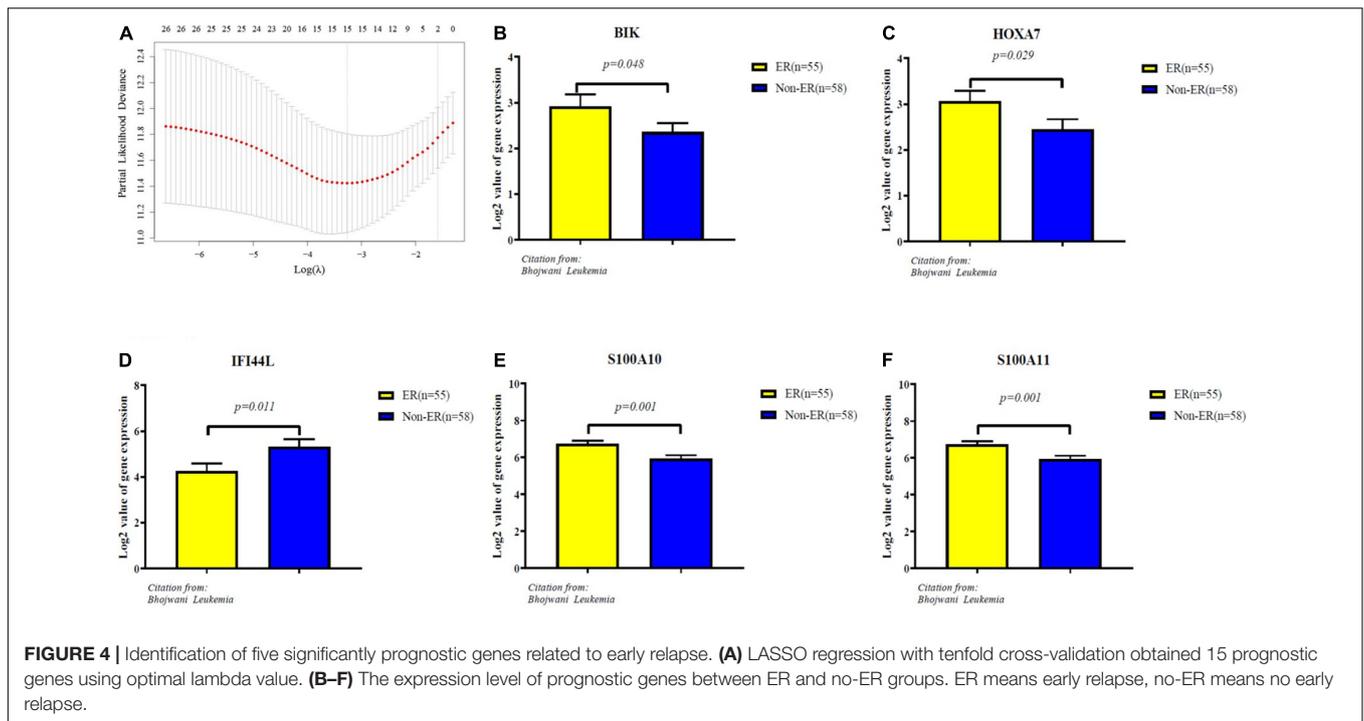
Gene	HR	z	p-value
HOXA7	1.538	5.805	6.43E-09
S100A11	1.476	5.324	1.02E-07
TOMM22	1.500	5.195	2.05E-07
SFMBT2	0.825	-5.175	2.27E-07
ZNF704	0.731	-5.113	3.18E-07
COTL1	1.384	5.050	4.42E-07
S100A6	1.498	4.947	7.55E-07
S100A4	1.524	4.885	1.04E-06
TSPO	1.514	4.609	4.05E-06
S100A10	1.350	4.455	8.41E-06
MDK	0.794	-4.058	4.95E-05
ANXA2P2	1.342	3.983	6.8E-05
HERC5	0.639	-3.938	8.21E-05
CYP46A1	0.808	-3.767	0.000165
CRIP1	1.239	3.686	0.000227
CTSC	1.288	3.668	0.000244
BIK	1.202	3.437	0.000588
HK2	1.322	3.369	0.000754
CXXC5	1.226	3.250	0.001155
GFOD1	0.713	-3.029	0.002457
TOP1MT	1.279	2.915	0.003553
CHN2	1.173	2.906	0.003658
IFI44L	0.896	-2.881	0.003967
LGALS1	1.136	2.750	0.005964
GBP4	0.780	-2.688	0.007177
MX1	0.837	-2.648	0.008097
IFI44	0.856	-2.639	0.008306
ZNF91	0.787	-2.410	0.015962
IGFBP7	1.144	2.329	0.019885
DAD1	1.208	2.260	0.023824
TUBB6	1.124	2.244	0.024804
RAPGEF5	1.303	2.182	0.029118
WIP1	1.168	2.104	0.035405
PPP1R14A	1.103	2.032	0.042198
SIPA1L2	0.918	-1.968	0.049121
MEIS1	1.100	1.952	0.050895
TUBB2A	1.128	1.951	0.051024
DAPK1	0.891	-1.945	0.051819
CPVL	1.100	1.890	0.058766
CAMK2D	1.118	1.776	0.075672
COL5A1	0.918	-1.752	0.079833
MAP7	1.111	1.743	0.081344

(Continued)

**TABLE 3 |** Continued

Gene	HR	z	p-value
TST	1.120	1.713	0.08663
DDIT4L	0.926	-1.668	0.09529
FHIT	0.892	-1.665	0.095844
ST3GAL6	0.908	-1.643	0.10029
SUCLG2	1.143	1.625	0.104209
PKIG	1.113	1.591	0.111563
SNX10	1.128	1.506	0.132115
CLEC11A	1.091	1.503	0.132803
MYO5C	0.873	-1.485	0.137488
MRC1	0.937	-1.359	0.174022
P2RY14	0.943	-1.083	0.278976
SHANK3	0.970	-0.942	0.346027
POU4F1	0.957	-0.855	0.39244
S100Z	0.954	-0.833	0.404886
NAV1	0.949	-0.787	0.431054
STK32B	1.039	0.662	0.508211
TCF4	0.964	-0.577	0.564058
PDE4B	0.973	-0.505	0.613274
RASD1	0.972	-0.469	0.638796
ITGA6	1.023	0.438	0.661388
DNNT	0.981	-0.402	0.687464
MAN1A1	0.984	-0.264	0.791807
CD69	1.012	0.155	0.876639

prognostic model for B-ALL. The risk score of each patient was calculated using the regression coefficients of four genes in the training set. The patients were divided into high- and low-risk groups according to the median risk score of the training set. Time-dependent ROC curves were plotted for the validation sets. Survival analysis of the patients was constructed using the R software, with the survival and survminer packages. We created Kaplan-Meier plots to illuminate the correlations between the genetic risk scores and the survival index of patients, including overall survival and event-free survival (17). Univariate and multivariate COX regression analyses were used to test whether the genetic risk score model is an independent prognostic risk factor relative to the clinical characteristics of the total set. Statistical significance was tested using log-rank tests. The prognostic risk factors were analyzed using Cox Regression with the survival R packages. A *p*-value < 0.05 was considered to be statistically significantly different.



**TABLE 4 |** The result of LASSO regression in training set.

Gene Symbol	Coefficients
ANXA2P2	0
BIK	0.065
CHN2	0.059
COTL1	0
CRIP1	0
CTSC	0
CXXC5	0.016
CYP46A1	0
GBP4	-0.020
GFOD1	0
HERC5	-0.027
HK2	0
HOXA7	0.239
IFI44	-0.055
IFI44L	-0.048
LGALS1	0
MDK	-0.069
MX1	-0.052
S100A10	0.006
S100A11	0.148
S100A4	0
S100A6	0.159
SFMBT2	-0.001
TOMM22	0
TOP1MT	0
TSPO	0
ZNF704	-0.068

### Analysis of Differentially Expressed MicroRNAs

For studying the regulation of gene expression, we obtained miRNA-seq data for 53 B-ALL patients from the TARGET dataset. Eight patients relapsed within 24 months, and were included in the early recurrence group, and the remaining 45 patients were included in the no early recurrence group. The edgeR R package was used to identify differentially expressed miRNAs (DEMs) between the two groups. Genes with a adjusted  $p$ -value  $< 0.05$  and  $|\text{Log}_2(\text{fold change})| > 0.5$  were defined as DEMs. Volcano plots were performed using R packages.

### mRNA-miRNA Network Construction

The miRNAs which were target DEGs were predicted by the miRWalk<sup>5</sup> and DIANA-Tarbase v.8 databases (18). We selected the miRNAs that existed in both prediction databases to construct an mRNA-miRNA network using Cytoscape software (Version 3.7.2).

## RESULTS

### Characteristics of B-ALL Patients in the Total Set From the TARGET Database

To determine the association of genetic profiles with early recurrence in B-ALL, we obtained microarray profiles from

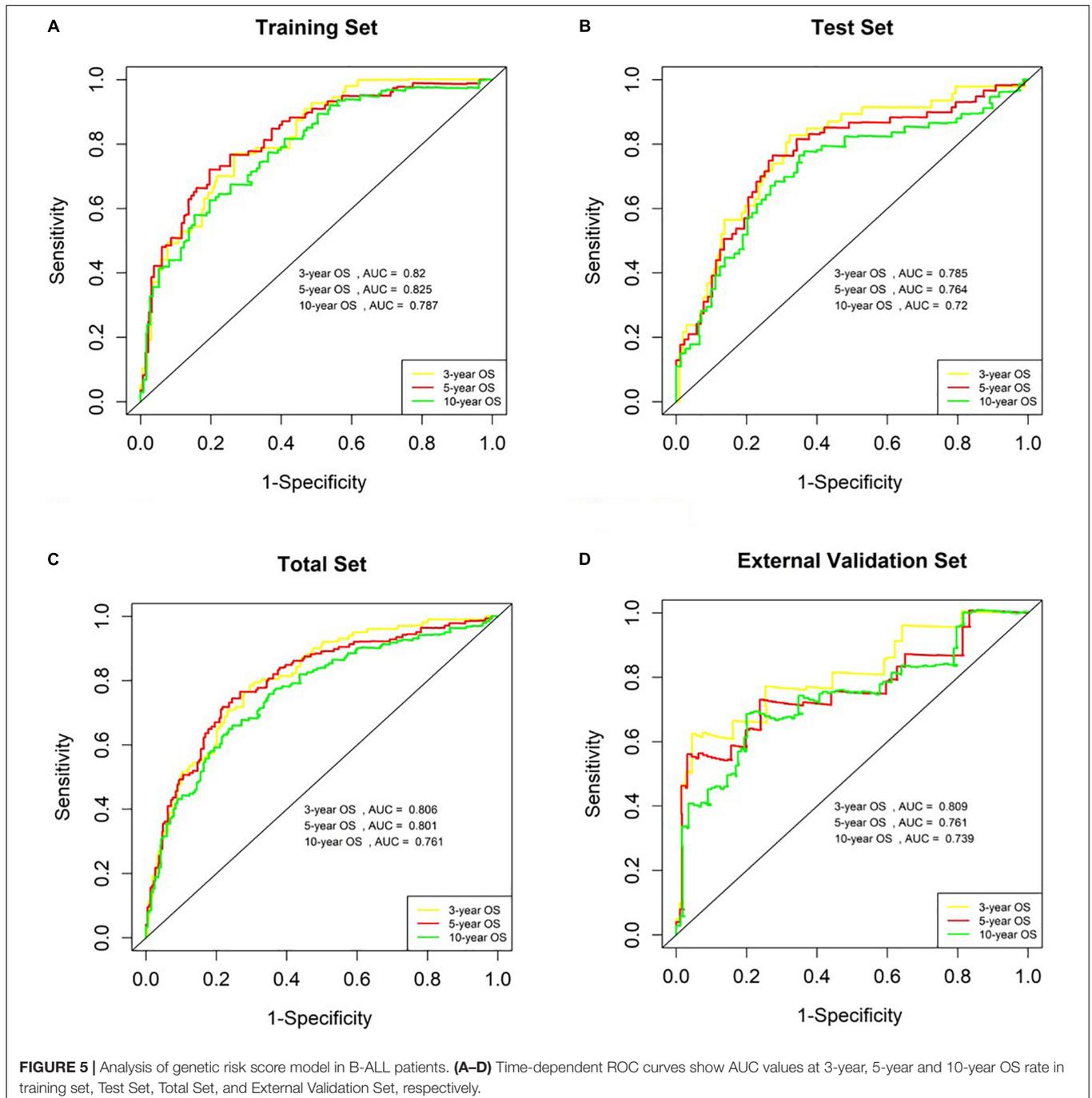
<sup>5</sup> <http://mirwalk.umm.uni-heidelberg.de/>

370 patients, and their clinical information, from the TARGET database. Among them, 225 (60.8%) patients were male and 145 (39.2%) were female. The age at initial diagnosis ranged from 1 to 20 years, with a median age of 10 years. The clinical characteristics of the patients are listed in **Table 1**, including cytogenetics, minimal residual leukemia (MRD) at day 29 after induction, and evaluation of the central nervous system (CNS) at diagnosis. The median follow-up time was 8.9 years for all 370 B-ALL patients. At the end of follow-up, 246 patients had relapsed (127 relapsed within 24 months since diagnosis) and 169 patients had died. The

median OS and EFS time were 9.6 years (range, 0.5–15.7 years) and 3.3 years (range, 0.1–15.7 years), respectively.

## Identification of DEGs Based on Disease Recurrence

The flowchart of the construction of our risk model is shown in **Figure 1**. To identify the DEGs, we analyzed the gene expression data of 370 precursor B-cell ALL patients obtained from TARGET, and 197 pediatric ALL patients from GSE13576.



Using a threshold of  $p < 0.05$  and  $|\text{fold change}| > 0.5$ , we identified a total of 1359 DEGs in the TARGET dataset, including 747 upregulated genes and 612 downregulate genes. There were 457 DEGs in GSE13576, including 215 upregulated genes and 242 downregulate genes. The DEGs of the early relapse vs. non-early relapse groups in these two datasets are shown in the volcano plot and heatmap, separately (Figure 2). Through integrated analysis, we identified 33 commonly upregulated genes and 32 downregulated genes from the TARGET and GEO datasets (Figure 3 and Table 2).

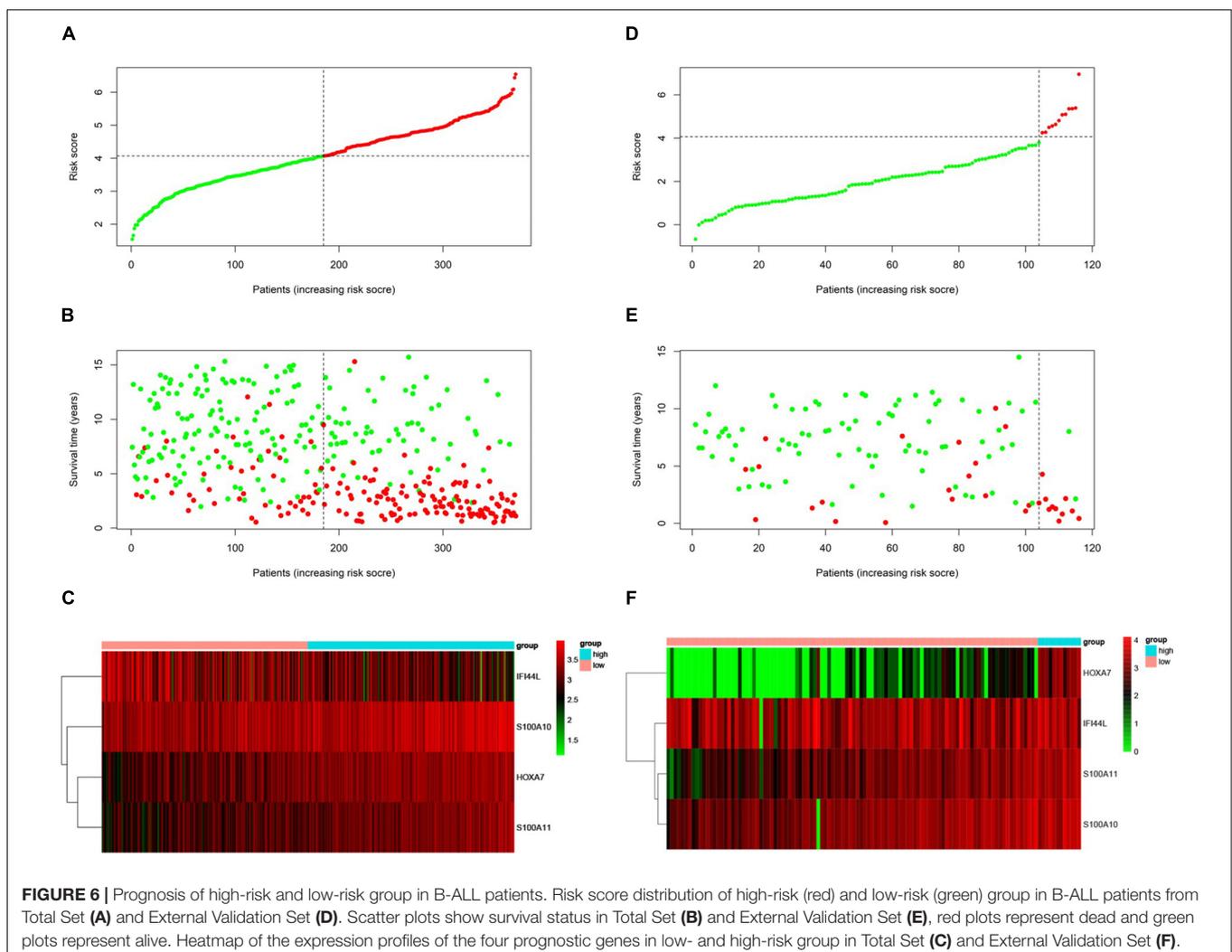
## Construction of the Prognostic Risk Score Model Using the Training Set

The 65 common DEGs (including 33 upregulated and 32 downregulated genes) were analyzed for prognostic value using univariate Cox regression on the training set. As shown in Table 3, 27 common DEGs had a  $p$ -value  $< 0.01$ , including 17 with a Hazard Ratio (HR)  $> 1$  and 10 with HR  $< 1$ . Using the LASSO algorithm, the optimal lambda value was 0.038, and 15 of the 27 DEGs were selected as potential

prognosis-related genes (Figure 4A and Table 4). Subsequently, we searched for the relationship between 15 common DEGs and early relapse using the Oncomine database. As shown in Figures 4B–F, the gene expression of *BIK*, *HOXA7*, *S100A10* and *S100A11* were significantly higher in B-ALL patients who relapsed within 24 months (both  $p < 0.05$ ). The gene *IFI44L* had lower expression levels in early relapse patients ( $p < 0.05$ ). These five genes associated with early recurrence were selected to build a genetic risk score model for B-ALL patients, using multivariate Cox regression. The risk model was constructed from the four optimal genes: *HOXA7*, *S100A10*, *S100A11*, and *IFI44L*. The genetic risk score model is as follows:

$$\begin{aligned} \text{Risk score} = & (0.329 * \text{expression value of } HOXA7) + \\ & (0.136 * \text{expression value of } S100A10) + \\ & (0.275 * \text{expression value of } S100A11) + \\ & (-0.151 * \text{expression value of } IFI44L) \end{aligned}$$

The AUC values for OS at 3 years, 5 years and 10 years were 0.820, 0.825 and 0.787 in the training set, respectively

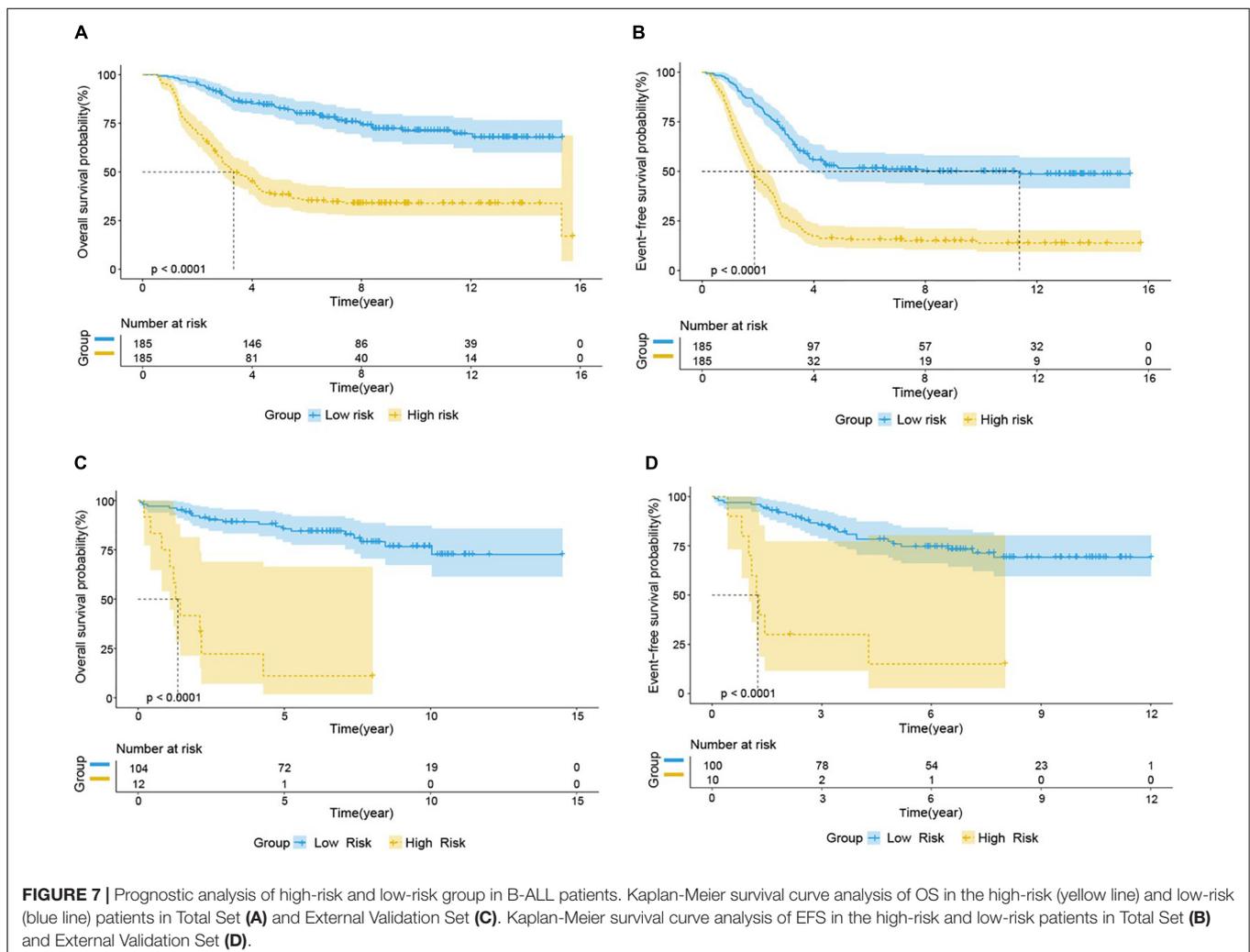


(Figure 5A). The risk score ranged from 1.535 to 6.539, and we selected the median risk score as the cut-off value (4.045). If the risk score was greater than the cut-off value, the patient was defined as belonging to the high-risk group, otherwise the patient was classified in the low-risk group.

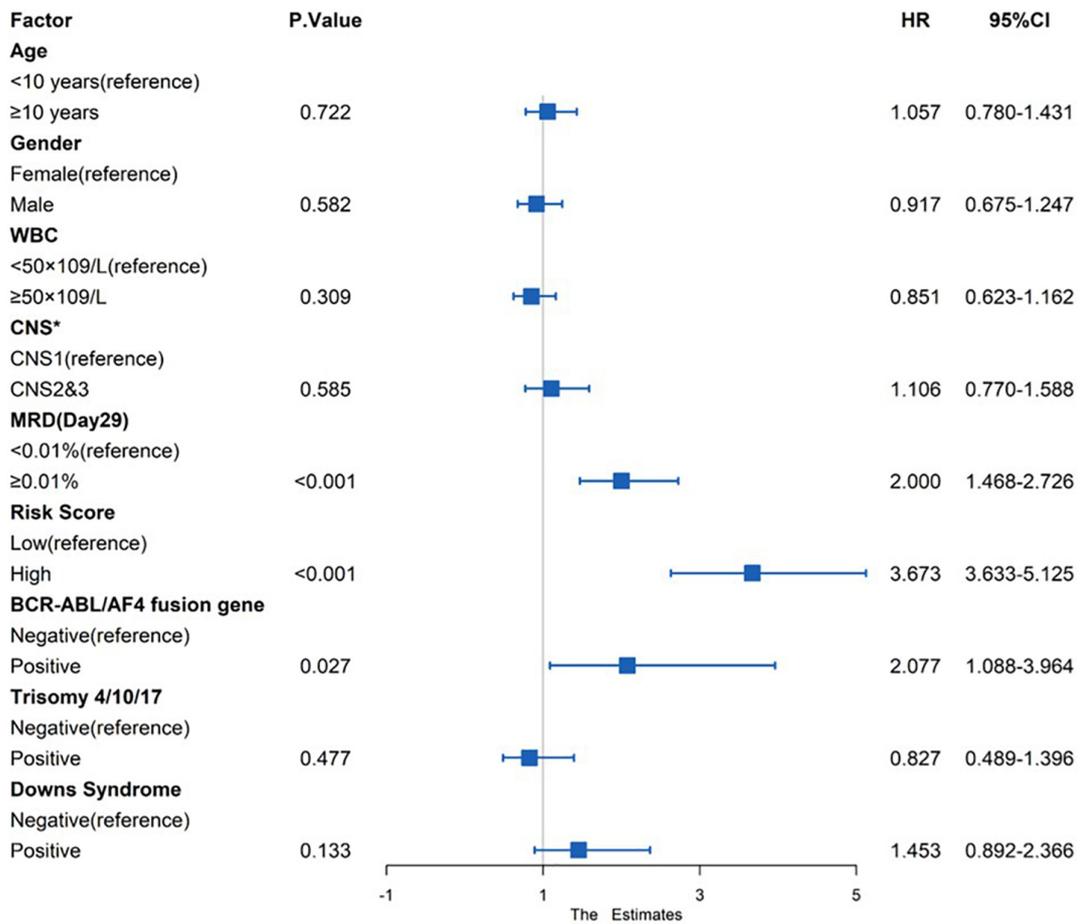
### Verification of the Prognostic Risk Score Model in the Validation Set

We used the test set ( $n = 148$ ), the total set ( $n = 370$ ) and the external validation set ( $n = 116$ ) to assess the predictive value of the risk score model constructed using the training set. We used the median value of the risk score formula defined using the training set as the threshold for distinguishing high- and low risk groups. The AUC curve showed that the 3-year, 5-year and 10-year OS of the test set, using the total set and the external validation set were 0.785, 0.764, 0.720, 0.806, 0.801, 0.761, 0.809, 0.761, and 0.739, respectively (Figures 5B–D). The distribution of risk score, survival status, and the heatmap of risk genes in internal total set and external validation set are shown in

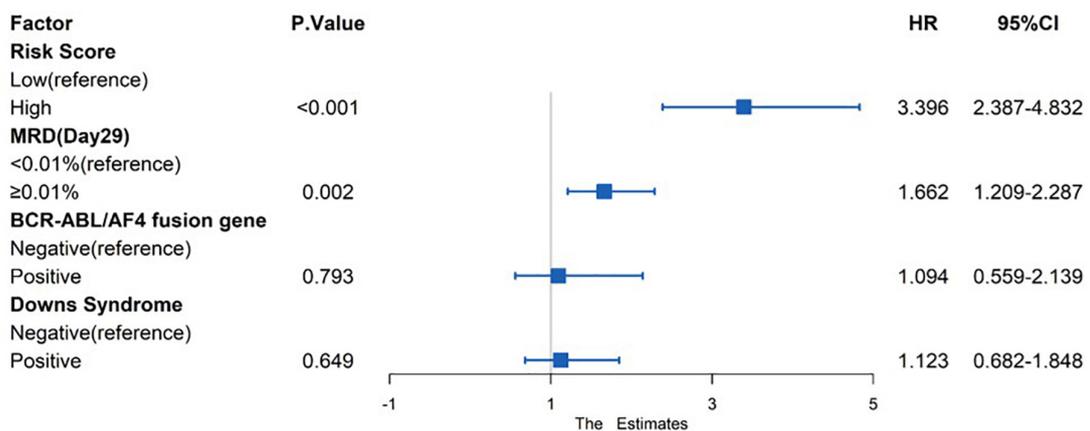
Figure 6. Among these four prognosis-related genes, *HOXA7*, *S100A10*, and *S100A11* were overexpressed in high risk group and *IFI44L* was downregulated. The prognosis of patients was worse with the increase in risk scores. In the total set, the mortality of patients in the high-risk group (64.3%) was higher than that in low-risk group (27.0%,  $p < 0.0001$ ). Similar conclusions were drawn for the external validation set (mortality in high- and low-risk groups: 33.3% vs. 2.3%,  $p < 0.0001$ ). Kaplan-Meier curves showed that patients in the high-risk group had significantly poorer OS (3-year OS for total set: 55.0% vs. 89.6%; 95% CI: 48.3–62.7%, 85.3–94.2%; 3-year OS for external validation set: 22.2% vs. 89.2%; 95% CI: 7.17–68.9%, 83.3–95.4%; 5-year OS for total set: 38.3% vs. 82.7%; 95% CI: 31.8–46.0%, 77.3–88.4%; 5-year OS for external validation set: 11.1% vs. 85.7%; 95% CI: 1.86–66.5%, 79.0–93.0%; all  $p < 0.0001$ ) and EFS (3-year EFS for total set: 26.5% vs. 68.6%; 95% CI: 20.8–33.7%, 62.2–75.6%; 3-year EFS for external validation set: 30.0% vs. 85.5%; 95% CI: 11.6–77.3%, 78.8–92.9%; 5-year EFS for total set: 15.7% vs. 51.6%; 95% CI: 11.2–21.9%, 44.8–59.5%; 5-year EFS for external validation set: 15.0% vs. 75.9%; 95% CI: 2.80–80.4%, 67.6–85.3%; all  $p < 0.0001$ ) (Figure 7).



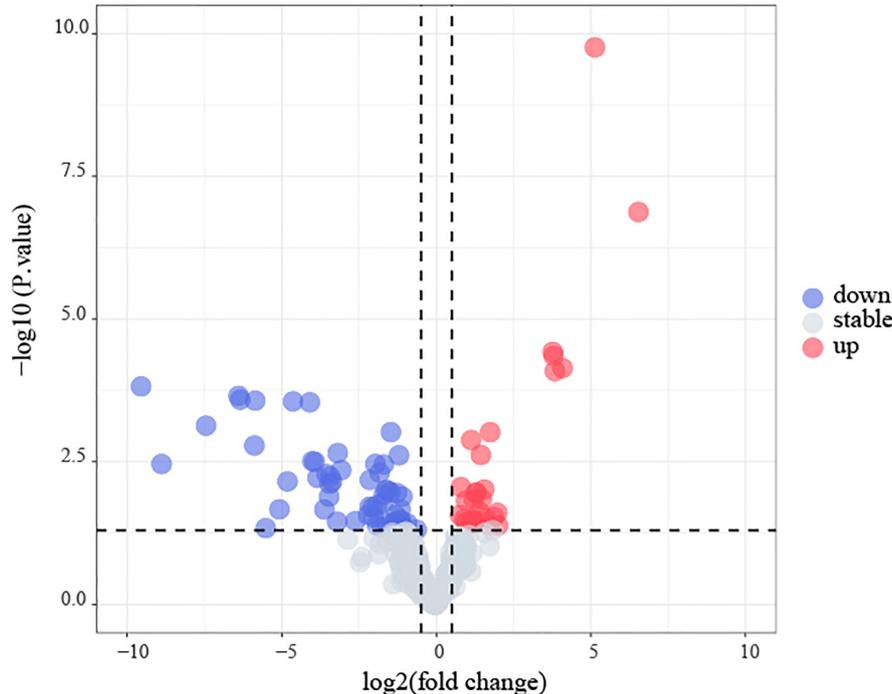
**A**



**B**



**FIGURE 8 |** Prognostic prediction value of four-gene prognostic risk models in Total Set. **(A)** Univariate Cox regression analysis of OS. **(B)** Multivariate Cox regression analysis of OS. “\*” means CNS, Central Nervous System, CNS1 = CNS negative, CNS2 ≤ 5 WBC/μL in CSF with blasts, CNS3 = CNS positive.



**FIGURE 9 |** Identification of early recurrence-associated DEMs in B-cell acute lymphoblastic leukemia from TARGET database. Volcano plot of DEGs from TARGET. Genes with  $p < 0.05$  and  $|\log_2(\text{FC})| > 0.5$  are shown in red (upregulated microRNAs) and blue (downregulated microRNAs).

**TABLE 5 |** The DEMs identified in 53 B-ALL patients from TARGET database.

**Gene Symbol (upregulated DEMs)**

hsa-mir-885, hsa-mir-4465, hsa-mir-3919, hsa-mir-124-3, hsa-mir-124-1, hsa-mir-124-2, hsa-mir-3662, hsa-mir-27b

**Gene Symbol (downregulated DEMs)**

hsa-mir-26a-1, hsa-mir-16-2, hsa-mir-30c-1, hsa-mir-486-2, hsa-let-7f-1, hsa-mir-103a-2, hsa-mir-1247, hsa-mir-324, hsa-mir-10b

## The Genetic Risk Score Model Is an Independent Prognostic Factor

Risk factors such as age, gender, white blood count (WBC), central nervous system (CNS) status at diagnosis, minimal residual disease (MRD) at day 29, risk score and genetic abnormalities were evaluated using univariate Cox model analysis (**Figure 8A**) in the total set. Only factors with a  $p$ -value of  $< 0.2$  (risk score, MRD at day 29, BCR-ABL/AF4 fusion gene and Down's Syndrome) in the univariate analysis were included in the multivariate analysis. Multivariate Cox regression analysis showed that high risk score and  $\text{MRD} \geq 0.01\%$  at day 29 were independent prognostic risk factors ( $p < 0.001$  and  $= 0.002$ ; HR = 3.396 and 1.662; 95% CI were 2.387–4.832 and 1.209–2.287, respectively) (**Figure 8B**).

## Identification of DEMs Based on Disease Recurrence

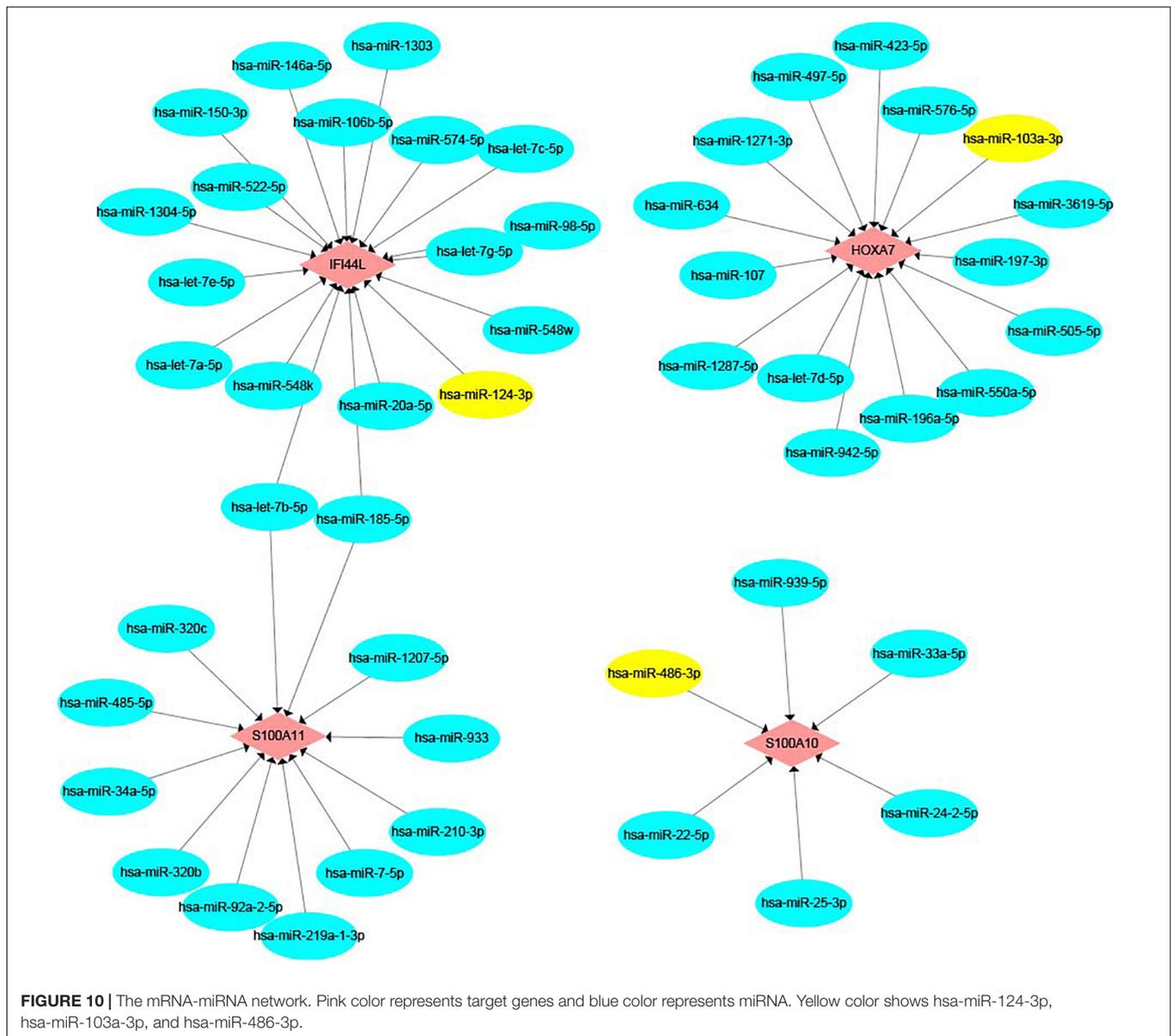
Amongst the 53 B-ALL patients from the TARGET database, 17 differential microRNA precursors (pre-miRNAs) were identified based on early disease recurrence (**Figure 9**), including eight upregulated miRNAs and nine downregulated miRNAs (**Table 5**).

## MiRNA Targeting DEGs

We built an mRNA-miRNA network using data from two prediction websites and Cytoscape. As shown in **Figure 10**, hsa-miR-124-3p, which is a pre-miRNA (hsa-mir-124-1, hsa-mir-124-2, and hsa-mir-124-3) were upregulated in early recurrence patients, targets the gene *IFI44L*. Hsa-miR-103a-3p and hsa-miR-486-3p have the target genes *HOXA7* and *S100A10*, respectively. These two miRNAs could be produced by shear processing of mir-103-2 and mir-486-2, both of which are downregulated in patients who relapsed within 24 months.

## DISCUSSION

Recurrence is an important factor for the increase of mortality in ALL, and the early recurrence of leukemia predicts a worse prognosis. Therefore, the identification of factors associated with early recurrence is vital for the identification of new risk subgroups in ALL patients, and also offers new directions for targeted therapies in high-risk patients. MLL rearranged, hypodiploid, BCR/ABL fusion gene and high end-induction MRD are the common risk factors for relapse in childhood ALL



patients (19). And several published articles showed that the biomarkers for early recurrence in B-ALL include the persistence of MRD, chromosome 19p13 translocations, upregulation of nucleotide excision repair genes, deletion of CDKN2A/B and overexpression of LINC00152 (**Supplementary Table S1**) (20–24). In this study, we looked for genes involved in early relapse of B-ALL on a larger sample. DEGs were identified from early recurrence vs. non-early recurrence groups both in the total set from the TARGET database and in the GSE13576 dataset. The common DEGs were subjected to univariate COX regression and LASSO regression in the training set. The optimal candidate DEGs were validated using the Oncomine website. We constructed a genetic risk score model of B-ALL using stepwise multivariate COX regression analysis, and verified its prognosis predictive value.

In this study, a four-gene risk score model was identified from 65 common DEGs. The AUCs of the risk model in the training and verification sets were all greater than 0.7, indicating that the risk score had a good prognosis predictive value in >70% of patients. The survival indexes OS and EFS in the high-risk group were confirmed to be worse than those in low-risk group with statistically significant differences. To further verify the relationship between the model and prognosis, we included clinical factors for univariate and multivariate COX analyses. The results showed that high genetic risk score and MRD  $\geq$  0.01% at day 29 were independent adverse prognostic factors.

The *HOXA7* gene is a member of the homeobox (HOX) gene family, which plays vital roles in hematopoiesis and cell differentiation (25). Previous research has shown that the gene *HOXA7* is overexpressed in MLL-fusion leukemias, which

commonly have very poor outcomes (26–29). The proteins encoded by *S100A10* and *S100A11* are members of the S100 family. S100 proteins are involved in many cellular processes, such as cell growth and motility, cell cycle progression and differentiation (30). Dysregulated expression of S100 proteins is a common feature of human cancers, including pediatric ALL (31, 32). High *S100A10* expression has a link to poor outcomes and chemoresistance in several types of cancer, including leukemia (33). The overexpression of *S100A11* is also associated with cancer progression and poor survival. The *IFI44L* gene is a type I interferon-stimulated gene. Although the function of *IFI44L* is unclear, some studies have revealed that it is involved in the inflammatory response, and may be a tumor suppressor (34–36). According to data from a study by Bhojwani et al., the expression of the genes *HOXA7*, *S100A10*, and *S100A11* was significantly higher in B-ALL patients who relapsed within 24 months than in those who did not (4). In contrast, *IFI44L* was highly expressed in patients with no recurrence or late recurrence. This finding is consistent with those of our research. These results suggest the possibility of the use of the four genes as early prognostic biomarkers for childhood and adolescent B-ALL.

MiRNAs are dysregulated in cancers, including hematological malignancies (37, 38). MiR-124 is regarded as a tumor suppressor in some solid tumors (39–41). Previous studies indicated that upregulation of miR-124 can reduce the invasion ability of tumor cells (42, 43). However, the role miR-124 plays in hematological malignancies remains controversial (44). Chen et al., suggested that miR-124-1 deregulation might have a favorable impact on prognosis in AML (45). Liang found that miR-124 is overexpressed in pediatric prednisone-poor response ALL, and contributes to glucocorticoid resistance (46). MiR-103a-3p is considered to be an oncogene in gastric cancer, colorectal cancer, and breast cancer, while it suppresses cell proliferation and invasion in prostate cancer (47–49). In leukemia, the function of miR-103 is also controversial. Upregulation of miR-103 was found by Yefenof et al., to sensitize leukemia cells to glucocorticoids (50). However, a study by Zhang et al., showed that miR-103 is upregulated in adriamycin-resistant cells (51). miR-486-3p could act as a tumor-suppressive miRNA in several cancers (52–54). In our study, pre-miR-124s (miR-124-1, miR-124-2 and miR-124-3) were upregulated in early recurrence B-ALL, and miR-124-3p was predicted to target the gene *IFI44L*. pre-miR-103a (miR-103a-2) and pre-miR-486 (miR-486-2), which target the genes *HOXA7* and *S100A10*, respectively, were downregulated in patients who relapsed within 24 months. In summary, the dysregulation of microRNA may be a potential mechanism for the abnormal expression of these prognostic genes.

## CONCLUSION

Our study identified four genes related to early recurrence of B-ALL in childhood and adolescent patients, using integrated bioinformatics analysis. The four-gene risk score model is

an independent prognostic risk indicator. Detection of the expression levels of these four genes provides a new signature for the early identification of high-risk patients.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by all data are from public database on the internet. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

YH was mainly responsible for the idea design and writing of the article. JL was in charge of the writing of the manuscript. YC was in charge of the data collection. PJ was responsible for the chart making. LW provided the statistical assistance. JH contributed to the idea design and article modification for this article. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (81870135), the Special Fund from the Fujian Provincial Department of Finance (2016B041), the Cooperation Project of University and Industry in Fujian Province (2017Y4005), the Joint Research Project of Health and Education (WKJ2016-2-06), the Joint Funds for the Innovation of Science and Technology, Fujian Province (2016Y9032), the Startup Fund of Scientific Research, Fujian Medical University (2019QH1022, 2019QH2012, and 2019QH2016), and the Construction Project of the Fujian Medical Center of Hematology (Min201704).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.565455/full#supplementary-material>

**TABLE S1** | Risk factors of early relapse in B-ALL in published literatures.

## REFERENCES

- Tasian SK, Loh ML, Hunger SP. Childhood acute lymphoblastic leukemia: integrating genomics into therapy. *Cancer*. (2015) 121:3577–90. doi: 10.1002/cncr.29573
- Schmiegelow K, Forestier E, Hellebostad M, Heyman M, Kristinsson J, Söderhäll S, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia*. (2010) 24:345–54. doi: 10.1038/leu.2009.251
- Pui CH, Pei D, Campana D, Bowman WP, Sandlund JT, Kaste SC, et al. Improved prognosis for older adolescents with acute lymphoblastic leukemia. *J Clin Oncol*. (2011) 29:386–91. doi: 10.1200/jco.2010.32.0325
- Bhojwani D, Kang H, Moskowitz NP, Min DJ, Lee H, Potter JW, et al. Biologic pathways associated with relapse in childhood acute lymphoblastic leukemia: a Children's oncology group study. *Blood*. (2006) 108:711–7. doi: 10.1182/blood-2006-02-002824
- Tallen G, Rätei R, Mann G, Kaspers G, Niggli F, Karachunsky A, et al. Long-term outcome in children with relapsed acute lymphoblastic leukemia after time-point and site-of-relapse stratification and intensified short-course multidrug chemotherapy: results of trial ALL-REZ BFM 90. *J Clin Oncol*. (2010) 28:2339–47. doi: 10.1200/jco.2009.25.1983
- Nguyen K, Devidas M, Cheng SC, La M, Raetz EA, Carroll WL, et al. Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's oncology group study. *Leukemia*. (2008) 22:2142–50. doi: 10.1038/leu.2008.251
- Parker C, Krishnan S, Hamadeh L, Irving JAE, Kuiper RP, Révész T, et al. Outcomes of patients with childhood B-cell precursor acute lymphoblastic leukaemia with late bone marrow relapses: long-term follow-up of the ALLR3 open-label randomised trial. *Lancet Haematol*. (2019) 6:e204–16. doi: 10.1016/s2352-3026(19)30003-1
- Roy A, Cargill A, Love S, Moorman AV, Stoneham S, Lim A, et al. Outcome after first relapse in childhood acute lymphoblastic leukaemia—lessons from the United Kingdom R2 trial. *Br J Haematol*. (2005) 130:67–75. doi: 10.1111/j.1365-2141.2005.05572.x
- Malempati S, Gaynon PS, Sather H, La MK, Stork LC. Outcome after relapse among children with standard-risk acute lymphoblastic leukemia: children's oncology group study CCG-1952. *J Clin Oncol*. (2007) 25:5800–7. doi: 10.1200/jco.2007.10.7508
- Gregers J, Christensen IJ, Dalhoff K, Lausen B, Schroeder H, Rosthøj S, et al. The association of reduced folate carrier 80G>A polymorphism to outcome in childhood acute lymphoblastic leukemia interacts with chromosome 21 copy number. *Blood*. (2010) 115:4671–7. doi: 10.1182/blood-2010-01-256958
- Davies SM, Borowitz MJ, Rosner GL, Ritz K, Devidas M, Winick N, et al. Pharmacogenetics of minimal residual disease response in children with B-precursor acute lymphoblastic leukemia: a report from the Children's oncology group. *Blood*. (2008) 111:2984–90. doi: 10.1182/blood-2007-09-114082
- Wesolowska-Andersen A, Borst L, Dalgaard MD, Yadav R, Rasmussen KK, Wehner PS, et al. Genomic profiling of thousands of candidate polymorphisms predicts risk of relapse in 778 Danish and German childhood acute lymphoblastic leukemia patients. *Leukemia*. (2015) 29:297–303. doi: 10.1038/leu.2014.205
- Schuurhuis GJ. Cancer: a gene-expression profile for leukaemia. *Nature*. (2016) 540:346–8. doi: 10.1038/nature20488
- Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. (2016) 540:433–7. doi: 10.1038/nature20598
- Marjanovic I, Karan-Djurasevic T, Ugrin M, Virijevic M, Vidovic A, Tomin D, et al. Use of wilms tumor 1 gene expression as a reliable marker for prognosis and minimal residual disease monitoring in acute myeloid leukemia with normal karyotype patients. *Clin Lymphoma Myeloma Leukemia*. (2017) 17:312–9. doi: 10.1016/j.clml.2016.12.006
- Ismail EA, Mahmoud HM, Tawfik LM, Habashy DM, Adly AA, El-Sherif NH, et al. BIRC6/Apollon gene expression in childhood acute leukemia: impact on therapeutic response and prognosis. *Eur J Haematol*. (2012) 88:118–27. doi: 10.1111/j.1600-0609.2011.01734.x
- Holme JA, Morrison E, Samuelsen JT, Wiger R, Lag M, Schwarze PE, et al. Mechanisms involved in the induction of apoptosis by T-2 and HT-2 toxins in HL-60 human promyelocytic leukemia cells. *Cell Biol Toxicol*. (2003) 19:53–68. doi: 10.1023/a:1022069715399
- Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res*. (2018) 46:D239–45. doi: 10.1093/nar/gkx1141
- Ceppi F, Cazzaniga G, Colombini A, Biondi A, Conter V. Risk factors for relapse in childhood acute lymphoblastic leukemia: prediction and prevention. *Expert Rev Hematol*. (2015) 8:57–70. doi: 10.1586/17474086.2015.978281
- Katz F, Ball L, Gibbons B, Chessells J. The use of DNA probes to monitor minimal residual disease in childhood acute lymphoblastic leukaemia. *Br J Haematol*. (1989) 73:173–80. doi: 10.1111/j.1365-2141.1989.tb00249.x
- Khalidi HS, O'Donnell MR, Slovak ML, Arber DA. Adult precursor-B acute lymphoblastic leukemia with translocations involving chromosome band 19p13 is associated with poor prognosis. *Cancer Genet Cytogenet*. (1999) 109:58–65. doi: 10.1016/s0165-4608(98)00153-8
- Kathiravan M, Singh M, Bhatia P, Trehan A, Varma N, Sachdeva MS, et al. Deletion of CDKN2A/B is associated with inferior relapse free survival in pediatric B cell acute lymphoblastic leukemia. *Leukemia Lymphoma*. (2019) 60:433–41. doi: 10.1080/10428194.2018.1482542
- Ibrahim OM, As Sobeai HM, Grant SG, Latimer JJ. Nucleotide excision repair is a predictor of early relapse in pediatric acute lymphoblastic leukemia. *BMC Med Genom*. (2018) 11:95. doi: 10.1186/s12920-018-0422-2
- Bárceñas-López DA, Núñez-Enríquez JC, Hidalgo-Miranda A, Beltrán-Anaya FO, May-Hau DI, Jiménez-Hernández E, et al. Transcriptome analysis identifies LINC00152 as a biomarker of early relapse and mortality in acute lymphoblastic leukemia. *Genes*. (2020) 11:302. doi: 10.3390/genes11030302
- van Oostveen J, Bijl J, Raaphorst F, Walboomers J, Meijer C. The role of homeobox genes in normal hematopoiesis and hematological malignancies. *Leukemia*. (1999) 13:1675–90. doi: 10.1038/sj.leu.2401562
- Kettle LM, Lebert-Ghali C, Grishagin IV, Dickson GJ, O'Reilly PG, Simpson DA, et al. Pathways, processes, and candidate drugs associated with a Hoxa cluster-dependency model of leukemia. *Cancers*. (2019) 11:2036. doi: 10.3390/cancers11122036
- Ferrando AA, Armstrong SA, Neuberger DS, Sallan SE, Silverman LB, Korsmeyer SJ, et al. Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. *Blood*. (2003) 102:262–8. doi: 10.1182/blood-2002-10-3221
- Mann G, Attarbaschi A, Schrappe M, De Lorenzo P, Peters C, Hann I, et al. Improved outcome with hematopoietic stem cell transplantation in a poor prognostic subgroup of infants with mixed-lineage-leukemia (MLL)-rearranged acute lymphoblastic leukemia: results from the Interfant-99 Study. *Blood*. (2010) 116:2644–50. doi: 10.1182/blood-2010-03-273532
- Balagobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A, et al. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood*. (2009) 114:2489–96. doi: 10.1182/blood-2009-04-215152
- Halawi A, Abbas O, Mahalingam M. S100 proteins and the skin: a review. *J Eur Acad Dermatol Venereol JEADV*. (2014) 28:405–14. doi: 10.1111/jdv.12237
- Bresnick AR, Weber DJ, Zimmer DB. S100 proteins in cancer. *Nat Rev Cancer*. (2015) 15:96–109. doi: 10.1038/nrc3893
- Szczepanek J, Pogorzala M, Jarzab M, Oczko-Wojciechowska M, Kowalska M, Tretny A, et al. Expression profiles of signal transduction genes in ex vivo drug-resistant pediatric acute lymphoblastic leukemia. *Anticancer Res*. (2012) 32:503–6.
- Saiki Y, Horii A. Multiple functions of S100A10, an important cancer promoter. *Pathol Int*. (2019) 69:629–36. doi: 10.1111/pin.12861
- Huang WC, Tung SL, Chen YL, Chen PM, Chu PY. IFIG4L is a novel tumor suppressor in human hepatocellular carcinoma affecting cancer stemness, metastasis, and drug resistance via regulating met/Src signaling pathway. *BMC Cancer*. (2018) 18:609. doi: 10.1186/s12885-018-4529-9
- Karimifard M, Pakzad B, Karimzadeh H, Mousavi M, Kazemi M, Salehi A, et al. Interferon-induced protein 44-like gene promoter is differentially methylated in peripheral blood mononuclear cells of systemic lupus erythematosus patients. *J Res Med Sci*. (2019) 24:99. doi: 10.4103/jrms.JRMS\_83\_19
- Zhao M, Zhou Y, Zhu B, Wan M, Jiang T, Tan Q, et al. IFI44L promoter methylation as a blood biomarker for systemic lupus erythematosus. *Ann Rheum Dis*. (2016) 75:1998–2006. doi: 10.1136/annrheumdis-2015-208410

37. Iorio MV, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med.* (2012) 4:143–59. doi: 10.1002/emmm.201100209
38. Rupaimoole R, Calin GA, Lopez-Berestein G, Sood AK. miRNA deregulation in cancer cells and the tumor microenvironment. *Cancer Discov.* (2016) 6:235–46. doi: 10.1158/2159-8290.cd-15-0893
39. Hatziapostolou M, Polytaichou C, Aggelidou E, Drakaki A, Poultsides GA, Jaeger SA, et al. An HNF4 $\alpha$ -miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. *Cell.* (2011) 147:1233–47. doi: 10.1016/j.cell.2011.10.043
40. Lang Q, Ling C. MiR-124 suppresses cell proliferation in hepatocellular carcinoma by targeting PIK3CA. *Biochem Biophys Res Commun.* (2012) 426:247–52. doi: 10.1016/j.bbrc.2012.08.075
41. Silber J, Lim DA, Petritsch C, Persson AI, Maunakea AK, Yu M, et al. miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. *BMC Med.* (2008) 6:14. doi: 10.1186/1741-7015-6-14
42. Xia H, Cheung WK, Ng SS, Jiang X, Jiang S, Sze J, et al. Loss of brain-enriched miR-124 microRNA enhances stem-like traits and invasiveness of glioma cells. *J Biol Chem.* (2012) 287:9962–71. doi: 10.1074/jbc.M111.332627
43. Zheng F, Liao YJ, Cai MY, Liu YH, Liu TH, Chen SP, et al. The putative tumour suppressor microRNA-124 modulates hepatocellular carcinoma cell aggressiveness by repressing ROCK2 and EZH2. *Gut.* (2012) 61:278–89. doi: 10.1136/gut.2011.239145
44. Liu YX, Wang L, Liu WJ, Zhang HT, Xue JH, Zhang ZW, et al. MiR-124-3p/B4GALT1 axis plays an important role in SOCS3-regulated growth and chemo-sensitivity of CML. *J Hematol Oncol.* (2016) 9:69. doi: 10.1186/s13045-016-0300-3
45. Chen XX, Lin J, Qian J, Qian W, Yang J, Ma JC, et al. Dysregulation of miR-124-1 predicts favorable prognosis in acute myeloid leukemia. *Clin Biochem.* (2014) 47:63–6. doi: 10.1016/j.clinbiochem.2013.09.020
46. Liang YN, Tang YL, Ke ZY, Chen YQ, Luo XQ, Zhang H, et al. MiR-124 contributes to glucocorticoid resistance in acute lymphoblastic leukemia by promoting proliferation, inhibiting apoptosis and targeting the glucocorticoid receptor. *J Steroid Biochem Mol Biol.* (2017) 172:62–8. doi: 10.1016/j.jsbmb.2017.05.014
47. Ge J, Mao L, Xu W, Fang W, Wang N, Ye D, et al. miR-103a-3p suppresses cell proliferation and invasion by targeting tumor protein D52 in prostate cancer. *J Invest Surg.* (2020) 13:1–9. doi: 10.1080/08941939.2020.1738602
48. Hu X, Miao J, Zhang M, Wang X, Wang Z, Han J, et al. miRNA-103a-3p promotes human gastric cancer cell proliferation by targeting and suppressing ATF7 in vitro. *Mol Cells.* (2018) 41:390–400. doi: 10.14348/molcells.2018.2078
49. Chen HY, Lin YM, Chung HC, Lang YD, Lin CJ, Huang J, et al. miR-103/107 promote metastasis of colorectal cancer by targeting the metastasis suppressors DAPK and KLF4. *Cancer Res.* (2012) 72:3631–41. doi: 10.1158/0008-5472.can-12-0667
50. Kfir-Erenfeld S, Haggiag N, Biton M, Stepensky P, Assayag-Asher N, Yefenof E. miR-103 inhibits proliferation and sensitizes hemopoietic tumor cells for glucocorticoid-induced apoptosis. *Oncotarget.* (2017) 8:472–89. doi: 10.18632/oncotarget.13447
51. Wan L, Tian Y, Zhang R, Peng Z, Sun J, Zhang W. MicroRNA-103 confers the resistance to long-treatment of adriamycin to human leukemia cells by regulation of COP1. *J Cell Biochem.* (2018) 119:3843–52. doi: 10.1002/jcb.26431
52. Chou ST, Peng HY, Mo KC, Hsu YM, Wu GH, Hsiao JR, et al. MicroRNA-486-3p functions as a tumor suppressor in oral cancer by targeting DDR1. *J Exp Clin Cancer Res.* (2019) 38:281. doi: 10.1186/s13046-019-1283-z
53. Wang M, Chao CC, Chen PC, Liu PI, Yang YC, Su CM, et al. Thrombospondin enhances RANKL-dependent osteoclastogenesis and facilitates lung cancer bone metastasis. *Biochem Pharmacol.* (2019) 166:23–32. doi: 10.1016/j.bcp.2019.05.005
54. Wang JX, Liu Y, Jia XJ, Liu SX, Dong JH, Ren XM, et al. Upregulation of circFLNA contributes to laryngeal squamous cell carcinoma migration by circFLNA-miR-486-3p-FLNA axis. *Cancer Cell Int.* (2019) 19:196. doi: 10.1186/s12935-019-0924-9

**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 Huang, Li, Chen, Jiang, Wang and Hu. 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.