


A Novel 85-Gene Expression Signature Predicts Unfavorable Prognosis in Acute Myeloid Leukemia

Technology in Cancer Research & Treatment
 Volume 20: 1-10
 © The Author(s) 2021
 Article reuse guidelines:
sagepub.com/journals-permissions
 DOI: 10.1177/15330338211004933
journals.sagepub.com/home/tct


Yanli Lai¹, Lixia Sheng¹, Jiaping Wang¹, Miao Zhou¹ ,
 and Guifang OuYang¹

Abstract

Aim: Acute myeloid leukemia (AML) is a heterogeneous disorder with complex genetic basis and adverse prognosis. Cytogenetics risk, somatic mutations and gene expression profiles are important prognostic factors for AML patients. However, accurate stratification of patient prognosis remains an unsolved problem in AML. This study was to develop a novel gene profile to accurately classify AML patients into subgroups with different survival probabilities. **Methods:** Survival-related genes were determined by Kaplan–Meier survival analysis and multivariate analysis using the expression and clinical data of 405 AML patients from Oregon Health & Science University (OHSU) dataset and validated in The Cancer Genome Atlas (TCGA) database. Feature selection was performed by using the Least Absolute Shrinkage and Selection Operator (LASSO) method. With the LASSO model, a prognostic 85-gene score was established and compared with 2 known gene-expression risk scores. The stratification of AML patients was performed by unsupervised hierarchical clustering of 85 gene expression levels to identify clusters of AML patients with different survival probabilities. **Results:** The LASSO model comprising 85 genes was considered as the optimal model based on relatively high area under curve value (0.83) and the minimum mean squared error. The 85-gene score was associated with increased mortality in AML patients. Hierarchical clustering analysis of the 85 genes revealed 3 subgroups of AML patients in the OHSU dataset. The cluster1 AML patients were associated with more female cases, higher percent of bone marrow blast cells, 85-gene score, cytogenetics risk, more frequent FLT3-ITD, DNMT3A, NPI mutations, less frequent TP53, RUNX1 mutations, poorer overall survival than cluster2 tumors. The 85-gene score had higher AUC (0.75) than the 5-gene risk score and LSC17 score (0.74 and 0.65). **Conclusions:** The 85-gene score is superior to the 2 established prognostic gene signatures in the prediction of prognosis of AML patients.

Keywords

acute myeloid leukemia, 85-gene signature, overall survival

Received: January 24, 2021; Revised: January 24, 2021; Accepted: February 25, 2021.

Introduction

AML is a cancer of the myeloid line of blood cells characterized by acquired gene mutations, abnormalities in bone marrow, morphology, karyotype and alterations in gene expression.¹ Over the last 3 decades, the incidence rate of AML rose by 87.3%, with 119.57×10^3 new cases diagnosed in 2017.² Despite significant progresses in the therapy of AML, the median survival time of AML is as short as 8.5 months. The disease has a poor prognosis, with 2-year and 5-year overall survival (OS) rates less than 35%.³ Elderly AML patients are more likely to have a relatively poor survival, with above 70% of patients die from the disease within 1 year of diagnosis.^{4,5}

European Leukemia-Net (ELN) has been commonly used in clinical settings for the diagnosis and patient prognosis stratification, AML patients are classified into 3 distinct subgroups with different survival probabilities based on presence or absence of specific chromosomal aberrations.⁶ In recent years,

¹ Department of Hematology, Ningbo First Hospital, Ningbo, Zhejiang Province, China

Corresponding Author:

Guifang OuYang, Department of Hematology, Ningbo First Hospital, 59 Liuting Road, Ningbo, Zhejiang Province 315000, China.
 Email: nbhematology@163.com



Table 1. Association Between the Clinical Features and Patients' Mortality in 405 AML Patients of the OHSU Dataset.

Variables	Group	Alive	Dead	P value	Statistical method
Age		49.18	61.82	<0.001	Student t test
PBMBC		58.84	58.6	0.95	Student t test
Gender	Female	26	54	0.09	Fisher's exact test
	Male	77	94		
European Leukemia Net classification	Favorable	69	40	<0.001	Fisher's exact test
	Intermediate	48	94		
	Poor	42	103		
TP53 mutation	Mutant	3	29	<0.001	Fisher's exact test
	Wild-type	164	209		
ASXL1 mutation	Mutant	11	20	0.57	Fisher's exact test
	Wild-type	156	218		
RUNX1 mutation	Mutant	15	31	0.27	Fisher's exact test
	Wild-type	152	207		
FLT3-IDT	Negative	130	176	0.09	Fisher's exact test
	Positive	29	61		
CEBPA mutation	Negative	61	82	0.83	Fisher's exact test
	Positive	12	14		
IDH1 mutation	Negative	61	88	0.19	Fisher's exact test
	Positive	14	11		
DNMT3A mutation	Negative	35	53	0.61	Fisher's exact test
	Positive	25	31		
NPI mutation	Negative	116	178	0.56	Fisher's exact test
	Positive	43	57		
Chemotherapy	Yes	157	211	0.01	Fisher's exact test
	No	1	14		
Bone marrow transplant	Yes	70	42	<0.001	Fisher's exact test
	No	88	183		
Targeted therapy	Yes	15	55	<0.001	Fisher's exact test
	No	143	170		

the prognostic values of somatic mutations have been characterized, for instance, certain mutations are negative prognostic markers, such as internal tandem duplication in Fms-like tyrosine kinase 3-internal tandem duplication (FLT3-ITD), and Tumor protein p53 (*TP53*) mutations. In contrast, mutations in other genes are associated with favorable outcome, such as CCAAT Enhancer Binding Protein Alpha (*CEBPA*) and isocitrate dehydrogenase 2 (*IDH2*).⁷ Furthermore, a number of gene expression profiles have been established for prognostic stratification, such as the 5-gene risk score⁸ and the 17-gene leukemia stem cells (LSCs) score (LSC17).⁹ These risk scores are computed based on linear combination of a set of gene expression and promising for clinical application. However, accurate stratification of patient prognosis remains an unsolved problem in AML.

In the present study, our goal was to identify a novel gene profile to accurately classify AML patients into subgroups with different survival probabilities. We performed various survival analyses to detect survival-related genes in the OHSU dataset¹⁰ and validated the results in the AML patients of the TCGA database.¹¹ We created a novel 85-gene score which is a linear regression model with 85 gene expression levels as explanatory variables to accurately predict the OS of AML patients. Finally, we performed hierarchical clustering of 85 genes and identified

3 distinct subsets of AML patients with significant differences in overall survival.

Methods and Materials

Data Acquisition

We obtained clinical data and gene expression of AML patients from 2 different sources. The first is the Tyner's study comprising 405 AML patients (hereafter referred to as the OHSU dataset). The second source comes from the TCGA database which provides researchers with RNA-seq expression data, and detailed clinical data of 173 AML patients (hereafter referred to as the TCGA dataset).¹¹ For the TCGA dataset, we removed those genes which have no expression values in more than 90% AML samples, leaving the final set of 18366 genes for the downstream analysis.¹⁰

Survival Analyses

We utilized the pROC package to determine the optimal cut-off value for each gene¹² and divided AML patients into 2 subgroups: the "high-expression" and "low-expression" groups according to the cut-off value of the gene. Then we performed Kaplan–Meier survival analysis and logistic

regression model to investigate the prognostic value of gene expression using the survival package.^{13,14} Survival-related genes were further stratified into risk genes with odd ratio (OR) greater than 1 and protective genes with OR ranging from 0 and 1.

Development and Validation of the 85-Gene Score

We performed 10-fold cross-validation of the LASSO model to select the optimal combination of genes for the prediction of OS in the OHSU dataset using the R package glmnet.¹⁵ Using the 85 gene expression levels as explanatory variables, a prognostic 85-gene score formula was created. 85-gene score = $7.76 + \text{expression of gene 1} \times \beta_1 + \text{expression of gene 2} \times \beta_2 + \dots + \text{expression of gene n} \times \beta_n$. β values were the coefficients generated by the LASSO model of the OHSU dataset. We performed Kaplan–Meier survival analysis and logistic regression analysis to analyze the association of 85-gene score with OS following the same method as described in the survival analysis. The associations between clinical characteristics and 85-gene score were analyzed by linear regression model. In order to compare the performance of our 85-gene score with those of 5-gene risk score and LSC17 score, we first conducted multivariate survival analysis using overall survival as response variable, prognostic scores and survival-related clinical features as prediction variables. Then, we computed area under curve (AUC) values accordingly by the R pROC package for the 3 prognostic scores. $P < 0.05$ was predefined to be statistically significant.

Unsupervised Hierarchical Clustering Analysis

We performed hierarchical clustering of 85 genes using the R package pheatmap¹⁶ and identified distinct subsets of AML patients with significant differences in overall survival. We utilized different statistical methods to compare the differences in clinical characteristics between subgroups of AML patients. For quantitative variables, the student t test was used. Fisher exact test was applied to the comparison of categorical variables. With respects to the comparison of OS, we used the Kaplan-Meier survival analysis method as described in the survival analyses section. $P < 0.05$ was predefined as statistically significant.

Gene Set Enrichment Analysis

In order to understand why 85-gene score is predictive of AML patients' survival, we partitioned the AML samples into 2 distinct groups: the high and low risk groups according to the cutoff value of 85-gene score determined by the pROC package. Gene set enrichment analysis (GSEA)¹⁷ was performed to analyze the altered signaling pathways between the 2 different risk groups. The default parameters were used in the GSEA analysis.

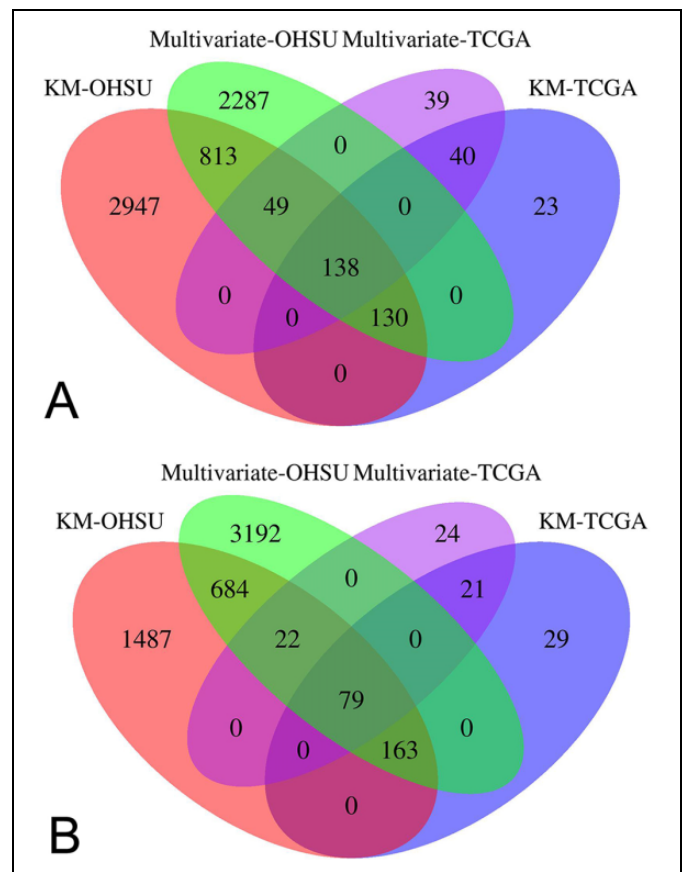


Figure 1. The overlap of prognosis-associated genes between OHSU and TCGA datasets. A. The overlap of protective type genes between OHSU and TCGA datasets. B. The overlap of risk type genes between OHSU and TCGA datasets.

Results

Characteristics of AML Patients

In the OHSU dataset, we found 3 risk factors for overall survival, including older patient's age, higher ELN classification and *TP53* mutation ($P < 0.05$ for all cases, student t test or fisher exact test, Table 1). As expected, treatments such as chemotherapy, bone marrow transplant and targeted therapy were protective factors for OS ($P < 0.05$ for all cases, fisher exact test, Table 1). Results in the TCGA dataset validated that patient's age, ELN classification and *TP53* mutation were risk factors for OS ($P < 0.05$ for all cases, student t test or fisher exact test, supplementary Table 1). No significant correlation was found between other characteristics and OS in the 2 datasets (P values > 0.05 for all cases, student t test or fisher exact test, Table 1 and supplementary Table 1).

Survival Analyses Between Patient Mortality and Gene Expression in AML

Kaplan-Meier survival analysis exhibited that high expression levels of 4077 genes and 2435 genes were indicative of improved and inferior prognosis respectively ($P < 0.05$ for all

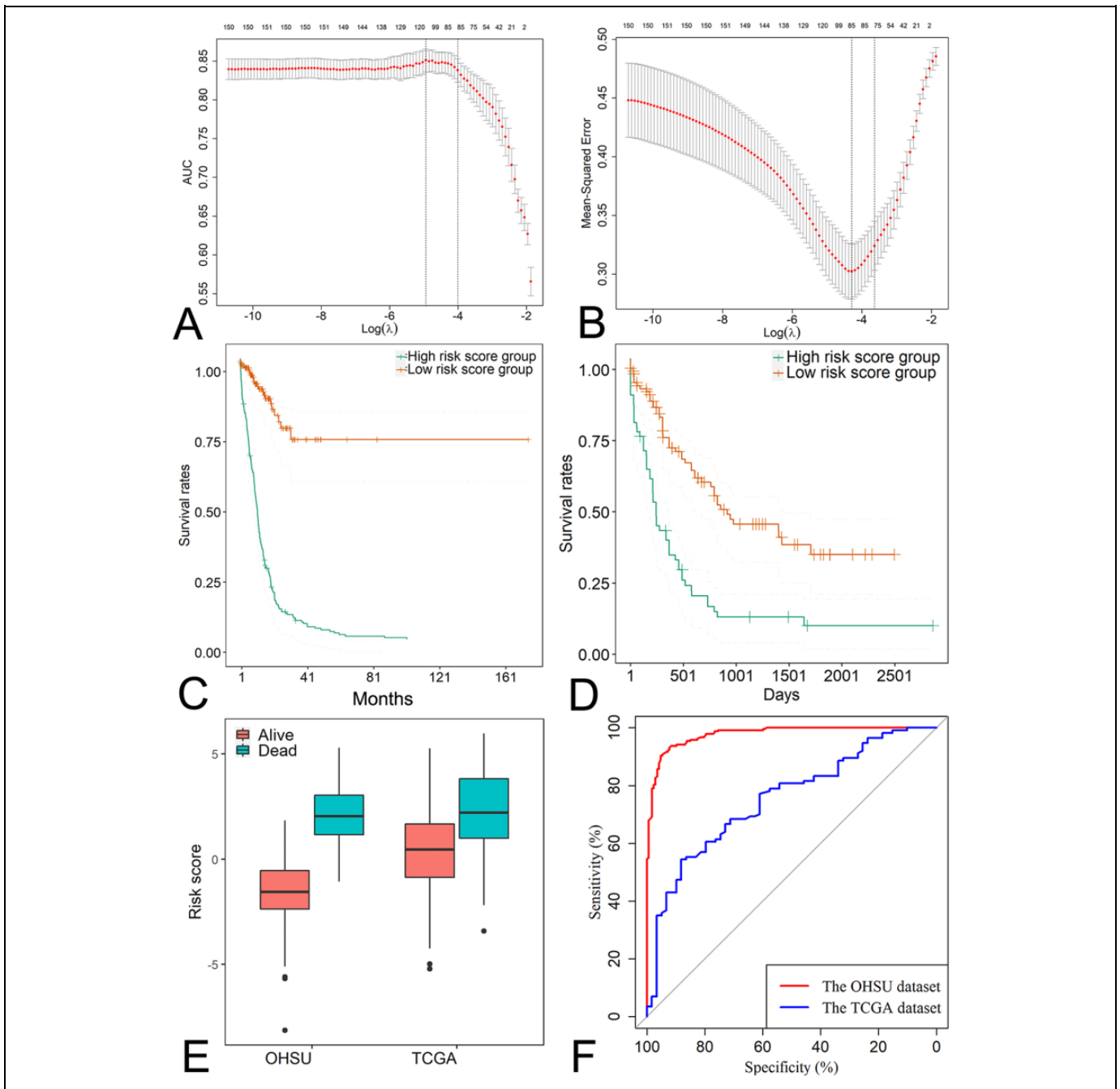


Figure 2. The 85-gene score is a risk prognostic factor in AML. A. The relationship between AUC and log scaled lambda values and number of genes with non-zero coefficients in the LASSO model. The x and y labels denoted log scaled lambda values and AUC, respectively. The numbers on the top were the number of genes with non-zero coefficients reserved in the LASSO model. The left and right vertical dotted lines indicated the lambda.min and lambda.1se for λ , respectively. The former is the one which minimizes out-of-sample loss in cross validation. The latter is the one which is the largest lambda value within 1 standard error of the minimum. B. The relationship between mean squared error and log scaled lambda values and number of genes in the LASSO model. C. Kaplan-Meier survival analysis between patients' OS and the 85-gene score in the OHSU cohort. D. Kaplan-Meier survival analysis between patients' OS and the 85-gene score in the TCGA dataset. E. The difference of the 85-gene scores between deceased and living AML patients in the OHSU and TCGA cohorts. F. The ROC curves of the 85-gene scores in the OHSU and TCGA datasets.

cases, log rank test, Figure 1). Then logistic regression model was performed between patients' OS and the survival-associated features, including ELN classification, patients' age, bone marrow transplant, chemotherapy, targeted therapy and

6512 gene expression levels. Multivariate analyses confirmed 1130 protective genes and 948 risk genes after the adjustment of prognosis-associated features. Furthermore, Kaplan-Meier survival analysis and multivariate analysis confirmed 138

Table 2. Multivariate Analyses Between OS and the Risk Score in the TCGA and OHSU Datasets.

Variable	OHSU dataset			Variable	TCGA dataset		
	OR	2.5%-97.5%CI	P value		OR	2.5-97.5%CI	P value
Age	1.03	1.00-1.06	0.06	Age	1.04	1.02-1.07	<0.001
Cytogenetics risk	0.35	0.16-0.72	0.01	Cytogenetics risk	1.27	0.68-2.41	0.46
Chemotherapy	1.31	0.03-25.2	0.88	TP53	10447140	3.40E-21-NA	0.99
Transplant	0.16	0.05-0.46	<0.001	Risk score	1.61	1.33-1.99	<0.001
Targeted therapy	2.84	0.72-12.87	0.15				
TP53.mutation	5.05	0.66-50.91	0.14				
Risk score	16.79	8.75-38.8	<0.001				

Notably, OR and CI refer to odds ratio and confidence interval, respectively.

genes and 79 genes were positively and negatively correlated with overall survival in the TCGA cohort respectively ($P < 0.05$ for all cases, log rank test, Figure 1).

Eight-Five Score Is a Risk Factor for Prognosis in AML

We performed 10-fold cross validation of LASSO model to determine the optimal model for the prediction of OS in the OHSU dataset. When the log (lambda) was equal to -4.2 and the number of genes with non-zero coefficients was 85, the AUC value of LASSO model was 0.83 and the mean squared error was minimum (Figure 2A-B). Therefore, the LASSO model comprising 85 genes was considered as the optimal model. The association of 85 genes with OS, intercept and coefficients of 85 genes were presented in the supplementary Tables 2-4. A prognostic 85-gene score formula was created using the coefficients of 85 genes generated by the optimal LASSO model. Kaplan-Meier survival analysis showed the AML patients with high 85-gene scores exhibited higher mortality rates than those with low 85-gene scores in the OHSU cohort ($P < 0.001$, log rank test, Figure 2C). Logistic regression model analysis verified that the 85-gene score was a risk factor for prognosis in AML patients ($P < 0.001$, OR: 16.79, 95% confidence interval [CI]: 8.75-38.8, Table 2). The negative correlation between OS and 85-gene score was validated in the TCGA dataset (Table 2 and Figure 2D). Furthermore, the dead AML patients showed significantly higher 85-gene scores than those living patients in the 2 cohorts ($P < 0.05$ for all cases, student t test, Figure 2E). The AUC values were 0.92 and 0.75 in OHSU and TCGA datasets respectively (Figure 2F), indicating the the 85-gene score performs well in predicting OS in AML patients.

Eight-Five Score Is Associated With Clinical Factors in AML

Linear regression model was used to investigate the association between 85-gene score and each clinical factor in the OHSU and TCGA cohorts. In the OHSU cohort, 85-gene score was significantly positively correlated with ELN Classification, age, TP53 mutation, targeted molecular therapy, FLT3-ITD,

gender, *RUNX1* mutation and negatively correlated with chemotherapy and transplant ($p < 0.05$ for all cases, Figure 3A). Furthermore, the 85-gene score exhibited significantly positive correlation with *NP1*, *FLT3*, *DNMT3A*, ELN Classification, BMBPC, neoadjuvant treatment ($p < 0.05$ for all cases, Figure 3B). The 85-gene score also showed negative correlation with *RUNX1* and *ASXL1* mutation, however, the association was not statistically significant ($p > 0.05$ for all cases, Figure 3B).

Unsupervised Hierarchical Clustering Analysis

Three subsets of AML patients were identified by hierarchical clustering of the 85 genes in the OHSU dataset (Figure 4A). The cluster1 AML tumors exhibited more female cases, higher BMBPC, 85-gene score, cytogenetics risk, more frequent FLT3-ITD, *DNMT3A*, *NP1* mutations, less frequent *TP53*, *RUNX1* mutations, poorer OS than cluster2 tumors (P values < 0.05 for all cases, student t test, fisher exact test or log-rank test, Figure 4B and supplementary Table 5). The other factors somatic mutations in *IDH1*, *IDH2*, *CEBPA*, *ASXL1* genes and treatment didn't exhibit significant difference between subgroups of AML patients in the OHSU cohort (P values > 0.05 for all cases, fisher exact test, supplementary Table 5). We also found 3 clusters of AML patients in the TCGA dataset (Figure 4C). Cluster1 tumors were significantly associated with higher 85-gene score, higher cytogenetics risk, lower frequencies of *NP1* and *FLT3* mutations, inferior OS than cluster2 and cluster3 tumors (P values < 0.05 for all cases, student t test, fisher exact test or log-rank test, Figure 4D and supplementary Table 6).

Eight-Five Score Related Pathway Analysis

Eleven signaling pathways were significantly enriched in the high 85-gene score group of the OHSU cohort, with long term depression, glycerolipid metabolism, vascular endothelial growth factor (VEGF) signaling pathway, phosphatidylinositol signaling system and gap junction the top 5 most enriched pathways (Figure 5, $p < 0.05$ for all cases). In contrast, genes in the pathways of glycosaminoglycan degradation, RNA polymerase were significantly enriched in the low 85-gene score

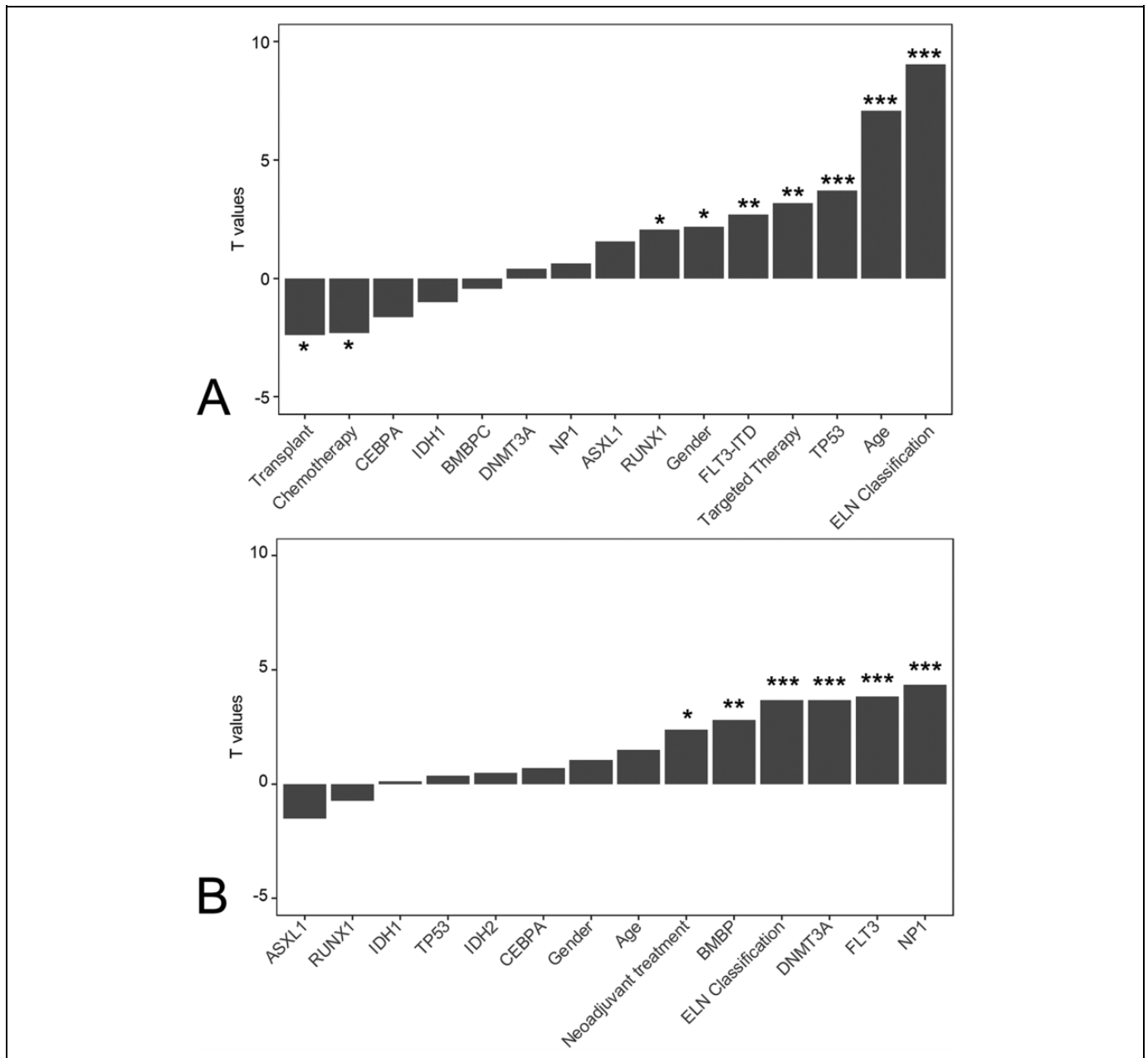


Figure 3. The associations of clinical characteristics with the 85-gene score. A. The associations between clinical characteristics with the 85-gene score in the OHSU cohort. B. The associations between clinical characteristics with the 85-gene score in the TCGA cohort. Of note, *, ** and *** stand for P value <0.05, <0.01 and 0.001, respectively.

group of the OHSU cohort (supplementary Figure 1, $p < 0.05$ for all cases). These results suggest that the overall survival of AML patients could be accurately predicted by the 85-gene score, the above-mentioned pathways might play a critical role in the association of 85-gene score with survival.

Comparisons of Prognostic Significance of the 85-Gene Score With Established Prognostic Gene Signatures

We compared the survival impact of the 85-gene score with other established gene expression-based prognostic

signatures. We performed multivariate analysis of the 85-gene score, 5-gene risk score and LSC17 score as well as prognosis-associated features in the TCGA cohort. The 85-gene score and 5-gene risk score remained significant prognostic factors independently of prognosis-associated features. Notably, 85-gene score achieved a higher OR than the 5-gene risk score and LSC17 score in the multivariate survival analysis (supplementary Table 7). Furthermore, ROC analysis showed the 85-gene score had higher AUC (0.75) than the 5-gene risk score and LSC17 score (0.74 and 0.65, Figure 6). Our data suggested the 85-gene score is superior to the 2

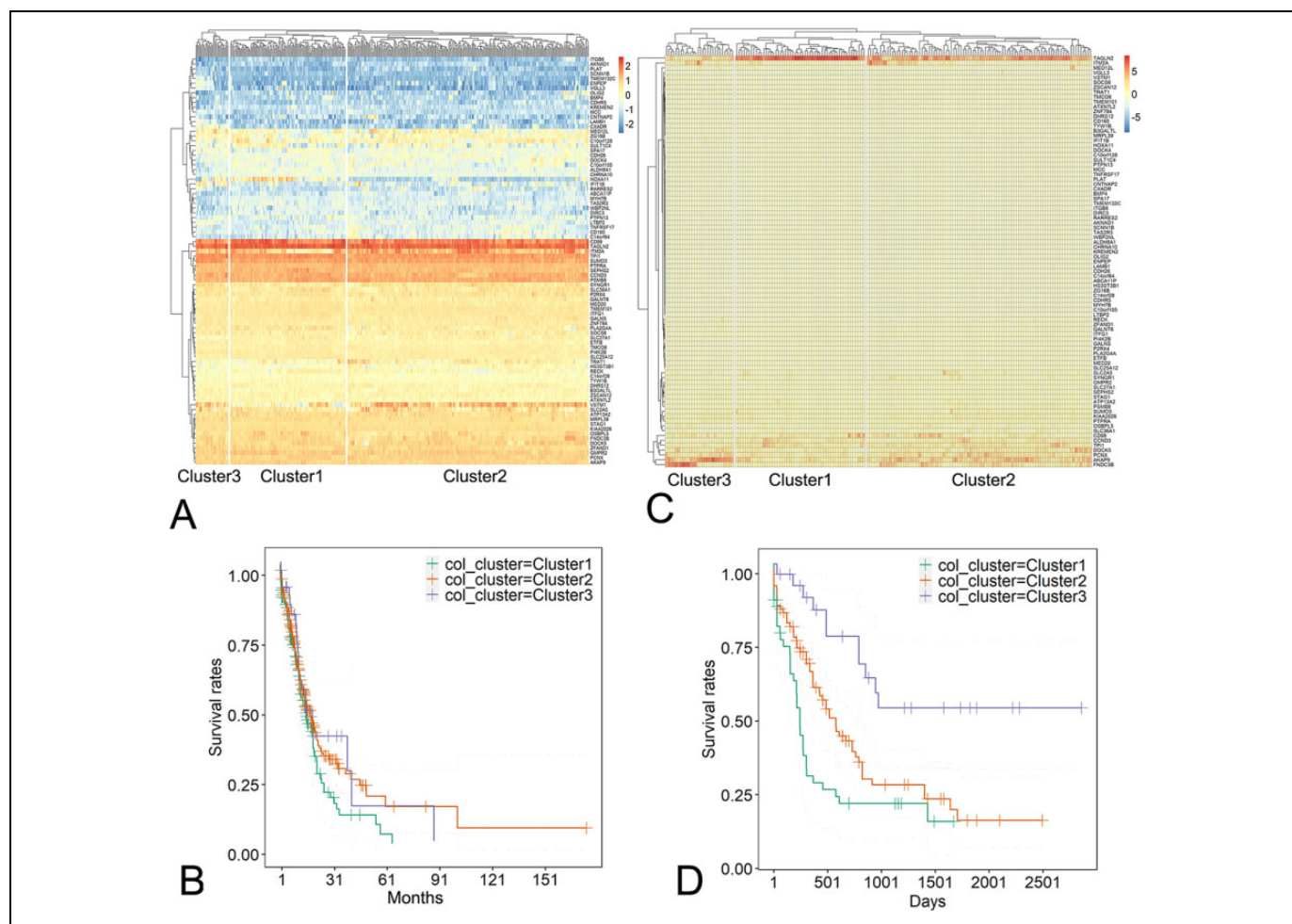


Figure 4. Hierarchical clustering of the 85 genes uncovered 3 classes of AML patients. A. Hierarchical clustering of the 85 genes uncovered 3 classes of AML patients in the OHSU cohort. B. The difference in overall survival between the 3 subsets of AML patients in the OHSU cohort. C. Hierarchical clustering of the 85 genes uncovered 3 classes of AML patients in the TCGA cohort. D. The difference in overall survival between the 3 clusters of AML patients in the TCGA cohort.

established prognostic gene signatures in the prediction of prognosis of AML patients.

Discussion

The 2017 ELN guidelines which incorporate cytogenetic abnormalities and driver gene mutations are widely applied to the evaluation of prognostic risk.¹⁸ In recent years, several gene expression profiles have been demonstrated to be potential prognostic biomarkers in AML. Sha *et al* developed a 5-gene risk score based on the linear combination of expression levels of 5 genes, including *PLA2G4A*, *CALCRL*, *DOCK1*, *FCHO2* and *LRCH4* and found the 5-gene score was effectively predictive of inferior prognosis in AML patients.⁸ Stanley developed a LSC17 score on the basis of 17 differentially expressed genes between 138 LSC+ and 89 LSC-cell fractions. The LSC17 score was highly prognostic and accurately predict initial therapy resistance.⁹ Though the risk classification of

AML has remarkably advanced, the accuracies of these methods are still needed to be improved.

In this study, the 85-gene score remained significantly associated with inferior OS after adjustment of survival-related clinical characteristics. Moreover, we have demonstrated that the 85-gene score performed better than the 5-gene risk score and the LSC17 score in the estimation of patient prognosis. The mechanisms by which higher 85-gene score is implicated in the poor prognosis of AML patients remain to be characterized. The GESA analysis revealed the VEGF signaling pathway and gap junction were significantly enriched in the high 85-gene score group. Gap junctions consist of clusters of intercellular channels that are critical to the direct communication between adjacent cells. The pathway plays pivotal roles in the regulation of cell growth, invasion, metastasis and differentiation and in the maintenance of tissue homeostasis.¹⁹ The VEGF family of soluble protein growth factors are implicated in the angiogenesis and lymphangiogenesis.²⁰ We believe the gap junctions and VEGF pathways in part contribute to the prognostic

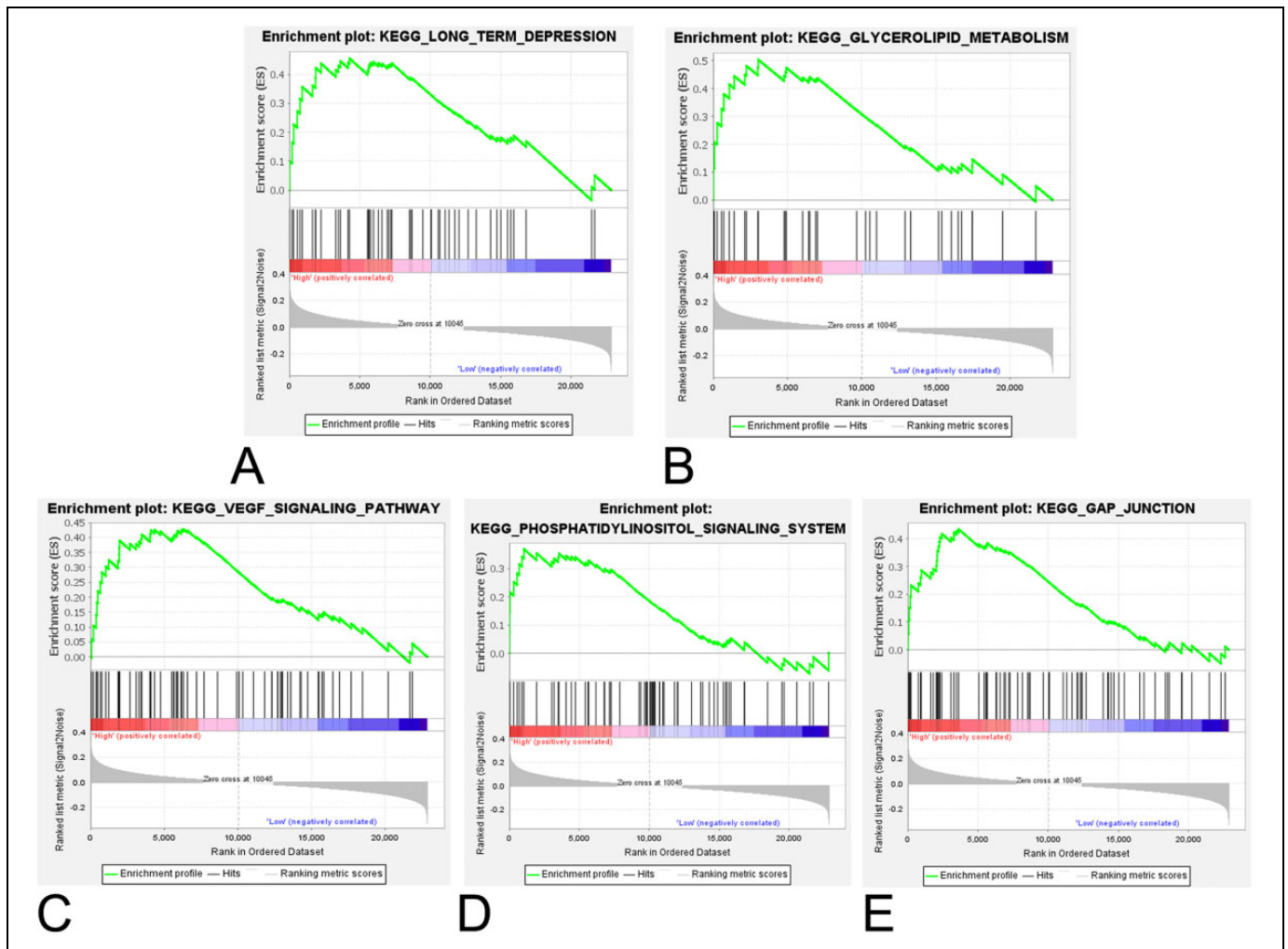


Figure 5. GSEA based on the expression of the OHSU dataset identified significantly up-regulated signaling pathways in the high 85-gene score group, including long term depression (A), glycerolipid metabolism (B), VEGF signaling pathway (C), phosphatidylinositol signaling system (D) and gap junction (E).

importance of 85-gene score in AML. Further mechanistic studies are needed to investigate its role in modulating poor prognosis in AML.

Of the 85 genes, many genes have oncogenic functions in cancers. Take *PLA2G4A* and *SLC2A5* for example, the *PLA2G4A* is over-expressed in various cancer types.²¹⁻²⁴ *PLA2G4A* depletion dramatically inhibits the proliferation and viability of glioblastoma cells,²¹ lung cancer cells, colon cancer cells.²⁴ *SLC2A5* encodes GLUT5 which plays an important role in the transportation of fructose in mammalian cells.²⁵ Up-regulated expression of *SLC2A5* has been reported in a wide range of cancer types.²⁶⁻²⁹ In consistent with this study, overexpression of *SLC2A5* is associated with poor prognosis in lung cancer²⁶ and AML.²⁸ Depletion of *SLC2A5* expression caused reduction in cellular proliferation, invasion and promotion of cellular apoptosis. In contrast, enhanced expression of *SLC2A5* facilitated cellular proliferation, invasion, and enhanced tumorigenicity in lung cancer.²⁶

Furthermore, considering the 85-gene expression signature effectively predicts AML patient prognosis independently of known prognosticators, such as somatic mutations in *DNMT3A*, *IDH1*, *IDH2* and *CEBPA*, the 5-gene expression signature may have applicability to the fraction of AML patients without somatic mutations in these driver genes. Though we have demonstrated the 85-gene score is a risk factor for prognosis in AML patients, the potential of prognosis prediction is needed to be validated in a large cohort of clinical samples. The verification of the efficacy of the 85-gene score will be the focus of our future studies.

Lastly, of the 85 prognosis-associated genes, some genes may become druggable targets for AML patients. Take the *PLA2G4A* and *SLC2A5* genes for example, knockdown of the 2 genes enabled significant inhibition of cellular proliferation, invasion and tumorigenic capability, indicating targeting these genes might make it possible for potential cure of AML patients.

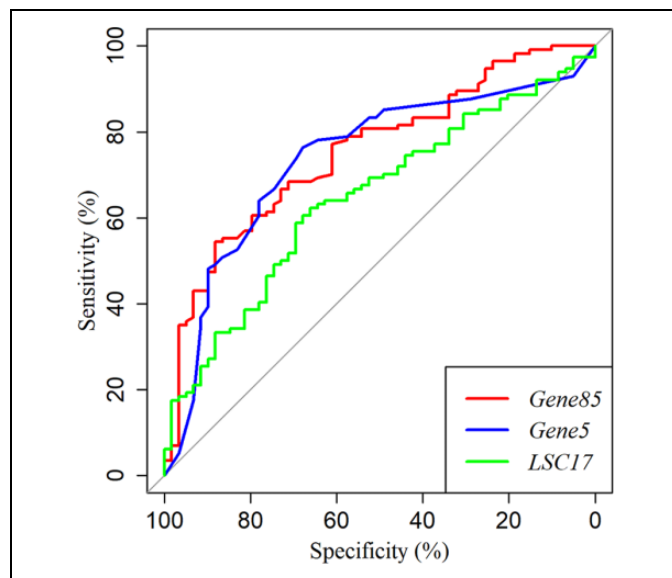


Figure 6. The ROC curves of the 85-gene score, 5-gene score and LSC17 score in the TCGA dataset.

Conclusion

In summary, this study presented a new 85 gene expression signature that has prognostic values and effectively stratifies AML patients into subgroups of AML patients. The 85-gene score is superior to established gene-expression risk scores and indicative of an unfavorable prognosis in AML patients.

Authors' Note

As all the data used in the study were collected from public databases, the study didn't need to be approved by the ethical board of Ningbo First Hospital.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: The study was financially supported by Medical and Health Science and Technology Project of Zhejiang Province (2021KY273) and the National Science Foundation of Zhejiang Province (Grant No. LY17H160005).

ORCID iD

Miao Zhou  <https://orcid.org/0000-0002-5463-4784>

Supplemental Material

Supplemental material for this article is available online.

References

1. Estey E, Döhner H. Acute myeloid leukaemia. *Lancet*. 2006; 368(9550):1894-1907.
2. Yi M, Li A, Zhou L, Chu Q, Song Y, Wu K. The global burden and attributable risk factor analysis of acute myeloid leukemia in

- 195 countries and territories from 1990 to 2017: estimates based on the global burden of disease study 2017. *J Hematol Oncol*. 2020;13(1):1-16. doi:10.1186/s13045-020-00908-z
3. Shallis RM, Wang R, Davidoff A, Ma X, Zeidan AM. Epidemiology of acute myeloid leukemia: recent progress and enduring challenges. *Blood Rev*. 2019;36:70-87. doi:10.1016/j.blre.2019.04.005
4. Shah A, Andersson TML, Racht B, Björkholm M, Lambert PC. Survival and cure of acute myeloid leukaemia in England, 1971-2006: a population-based study. *Br J Haematol*. 2013;162(4): 509-516. doi:10.1111/bjh.12425
5. Meyers J, Yu Y, Kaye JA, Davis KL. Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. *Appl Health Econ Health Policy*. 2013;11(3): 275-286. doi:10.1007/s40258-013-0032-2
6. Mrózek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. *Blood Rev*. 2004;18(2):115-136. doi:10.1016/S0268-960X(03)00040-7
7. Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1079-1089. doi:10.1056/NEJMoa1112304
8. Sha K, Lu Y, Zhang P, et al. Identifying a novel 5-gene signature predicting clinical outcomes in acute myeloid leukemia. *Clin Transl Oncol*. 2021;23(3):648-656. Published online August 2020. doi:10.1007/s12094-020-02460-1
9. Ng SWK, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 2016;540(7633):433-437. doi:10.1038/nature20598
10. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature*. 2018;562(7728): 526-531. doi:10.1038/s41586-018-0623-z
11. Network TCGAR, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074. doi:10.1056/NEJMoa1301689
12. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77. doi:10.1186/1471-2105-12-77
13. Therneau T. Survival analysis. *Cran*. Published online 2016. doi: 10.1007/978-1-4419-6646-9
14. Fox J. Cox Proportional-Hazards Regression for Survival Data The Cox Proportional-Hazards Model. *Most*. 2002;2008:1-18. doi:10.1016/j.carbon.2010.02.029
15. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for cox's proportional hazards model via coordinate descent. *J Stat Softw*. 2011;39(5):1-13. doi:10.18637/jss.v039.i05
16. Warnes G, Bolker B, Bonebakker L, et al. *Gplots: Various R Programming Tools for Plotting Data*. R package version 2.17. 0.2015;2.
17. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102

18. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447. doi:10.1182/blood-2016-08-733196
19. Aasen T, Leithe E, Graham SV, et al. Connexins in cancer: bridging the gap to the clinic. *Oncogene*. 2019;38(23):4429-4451. doi:10.1038/s41388-019-0741-6
20. Stacker SA, Achen MG. The VEGF signaling pathway in cancer: the road ahead. *Chin J Cancer*. 2013;32(6):297-302. doi:10.5732/cjc.012.10319
21. Yang L, Zhang H. Expression of cytosolic phospholipase a2 alpha in glioblastoma is associated with resistance to chemotherapy. *Am J Med Sci*. 2018;356(4):391-398. doi:10.1016/j.amjms.2018.06.019
22. Runarsson G, Feltenmark S, Forsell PKA, Sjöberg J, Björkholm M, Claesson H-E. The expression of cytosolic phospholipase A2 and biosynthesis of leukotriene B4 in acute myeloid leukemia cells. *Eur J Haematol*. 2007;79(6):468-476. doi:10.1111/j.1600-0609.2007.00967.x
23. Sundarraj S, Kannan S, Thangam R, Gunasekaran P. Effects of the inhibition of cytosolic phospholipase A2 α in non-small cell lung cancer cells. *J Cancer Res Clin Oncol*. 2012;138(5):827-835. doi:10.1007/s00432-012-1157-7
24. Parhamifar L, Jeppsson B, Sjölander A. Activation of cPLA2 is required for leukotriene D4-induced proliferation in colon cancer cells. *Carcinogenesis*. 2005;26(11):1988-1998. doi:10.1093/carcin/bgi159
25. Douard V, Ferraris RP. The role of fructose transporters in diseases linked to excessive fructose intake. *J Physiol*. 2013;591(2):401-414. doi:10.1113/jphysiol.2011.215731
26. Weng Y, Fan X, Bai Y, et al. SLC2A5 promotes lung adenocarcinoma cell growth and metastasis by enhancing fructose utilization. *Cell Death Discov*. 2018;4(1):38. doi:10.1038/s41420-018-0038-5
27. Zhao P, Huang J, Zhang D, et al. SLC2A5 overexpression in childhood Philadelphia chromosome-positive acute lymphoblastic leukaemia. *Br J Haematol*. 2018;183(2):242-250. doi:10.1111/bjh.15580
28. Chen WL, Wang YY, Zhao A, et al. Enhanced fructose utilization mediated by SLC2A5 is a unique metabolic feature of acute myeloid leukemia with therapeutic potential. *Cancer Cell*. 2016;30(5):779-791. doi:10.1016/j.ccell.2016.09.006
29. Medina Villaamil V, Aparicio Gallego G, Valbuena Rubira L, et al. Fructose transporter GLUT5 expression in clear renal cell carcinoma. *Oncol Rep*. 2011;25(2):315-323. doi:10.3892/or.2010.1096