

RESEARCH ARTICLE

# Identification of a robust subpathway-based signature for acute myeloid leukemia prognosis using an miRNA integrated strategy

Huijuan Chang<sup>1</sup>, Qiuying Gao<sup>2</sup>, Wei Ding<sup>3</sup>, Xueqin Qing<sup>4\*</sup>

**1** Department of Pneumology, Children's Hospital Affiliated to Zhengzhou University, Henan Children's Hospital, Zhengzhou Children's Hospital, Zhengzhou, Henan, China, **2** Department of Haematology, Shaanxi Provincial People's Hospital, Xi'an, Shanxi, China, **3** Department of Gastroenterology, 18 Bureau Gaobeidian Hospital, Gaobeidian, Hebei, China, **4** Department of Pediatric, Shanghai General Hospital, Shanghai, China

\* [qingdoctor99@163.com](mailto:qingdoctor99@163.com)



## Abstract

Acute myeloid leukemia (AML) is a heterogeneous disease, and survival signatures are urgently needed to better monitor treatment. MiRNAs displayed vital regulatory roles on target genes, which was necessary involved in the complex disease. We therefore examined the expression levels of miRNAs and genes to identify robust signatures for survival benefit analyses. First, we reconstructed subpathway graphs by embedding miRNA components that were derived from low-throughput miRNA-gene interactions. Then, we randomly divided the data sets from The Cancer Genome Atlas (TCGA) into training and testing sets, and further formed 100 subsets based on the training set. Using each subset, we identified survival-related miRNAs and genes, and identified survival subpathways based on the reconstructed subpathway graphs. After statistical analyses of these survival subpathways, the most robust subpathways with the top three ranks were identified, and risk scores were calculated based on these robust subpathways for AML patient prognoses. Among these robust subpathways, three representative subpathways, path: 05200\_10 from Pathways in cancer, path: 04110\_20 from Cell cycle, and path: 04510\_8 from Focal adhesion, were significantly associated with patient survival in the TCGA training and testing sets based on subpathway risk scores. In conclusion, we performed integrated analyses of miRNAs and genes to identify robust prognostic subpathways, and calculated subpathway risk scores to characterize AML patient survival.

## OPEN ACCESS

**Citation:** Chang H, Gao Q, Ding W, Qing X (2018) Identification of a robust subpathway-based signature for acute myeloid leukemia prognosis using an miRNA integrated strategy. PLoS ONE 13(3): e0194245. <https://doi.org/10.1371/journal.pone.0194245>

**Editor:** Xia Li, College of Bioinformatics Science and Technology, CHINA

**Received:** December 4, 2017

**Accepted:** February 27, 2018

**Published:** March 23, 2018

**Copyright:** © 2018 Chang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All expression files are available from the TCGA database at <https://portal.gdc.cancer.gov/>.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Acute myeloid leukemia (AML) is a biological heterogeneous disease, and its major characteristic is clonal proliferation of immature myeloid cells [1]. Currently, chemotherapy-based regimens fail to cure the majority of AML patients. And stem cell transplantation is a common treatment choice for these patients [2, 3]. However, the stem cell transplantation is not an option for all AML patients, which further illustrates the genetic heterogeneous nature of this disease. The present staging system is the common patient prognosis index in clinical analysis,

and the patients with late stage disease usually display worse prognoses [1, 4]. However, in most cases, patients with the same clinical stage may experience totally different clinical outcomes, especially for the patients with late stage. In addition, some patients with AML develop early metastasis, while some may not develop this complication. Therefore, there is a critical need to develop a novel strategy for exploring biological mechanisms involved in AML to identify markers for clinical characterization of this disease.

High-throughput gene expression levels can be utilized to study the biological heterogeneity issues of complex disease. For certain disorders, studies have identified different genes as risk markers, and have confirmed the usefulness of using these genes in their studies. Studies have reported the statistically significant association of gene expressions with AML. For example, Costa et al. performed the meta-analysis and identified the immunophenotypic markers at the gene expression levels. As a result, twelve antigens exhibited a negative impact with AML prognosis, and provided further guidance for new therapeutic targets [5]. By analyzing the available data sets, Huang et al. identified 11 genes as a potential prognostic marker of AML, and demonstrated the good performances for predicting overall survival in the validation cohorts [6]. More importantly, functionally-based analyses could be used for better interpretation of disease formation and progression [7, 8]. At a functional level, different gene signatures may play similar biological roles, with robust common functions. Thus, a number of functionally-based methods have been developed as signatures for better biological interpretations [9–11], and have displayed more robust results than gene-based analyses. Multiple genes involved in the same biological processes often act together during disease formation and progression. In addition, functionally-based analyses can reduce the number of dimensions of variables (from genes to functions), and the representative functions can increase the biological robustness necessary for exploring the heterogeneous diseases.

Through post-transcriptional regulation, microRNAs (miRNAs) can further affect the functional activities by inhibiting the expression levels of messenger RNAs (mRNAs) [12–14]. High-throughput technologies can simultaneously measure the expression levels of multiple genes or miRNAs, and can be used to characterize functional conditions in complex diseases. During disease formation and progression, the dysregulated genes usually interact with each other within biological processes or pathways, and have close regulatory relationships with miRNAs [15, 16]. For AML survival, the study of Shah et al. identified and characterized the NRF2-regulated miRNAs. Their findings have revealed the NRF2 regulation of miR-125B and miR-29B acted to promote AML cell survival, which further indicated the involvement of miRNA-gene regulation in AML prognoses [17]. In addition, by miRNA-gene network analysis, Zhang et al. found novel biological functions and pathways involved in AML survival, and mir-425 and CD44 were both identified as risk factors [18]. However, most function-level methods mentioned above were performed based only on gene expression, and the regulatory roles of miRNAs were not determined.

Studies of biological pathways display an advantage because of their topological structures. Because most pathways contain numerous genes, the concept of a subpathway (pathway region) was defined in previous studies [19, 20]. Containing a smaller scale of components, subpathways provide more information in terms of interpreting functional conditions. Abnormalities of subpathways are associated with the etiology of multiple types of diseases [21], further illustrating the advantages of analyses of subpathways compared to more complex entire pathways. We therefore considered both miRNA expression and the subpathway scale to comprehensively identify functional signatures for AML.

In this study, we performed a robust bioinformatics analysis to determine the expression levels of miRNAs and genes to identify prognostic subpathway-based signatures, and the subpathway contained the miRNA and gene components. First, we obtained the gene-based

subpathway graphs and embedded the miRNA molecules into these subpathway graphs by considering experimentally validated miRNA-gene relationships. Then, we randomly divided the total data set from The Cancer Genome Atlas (TCGA) into one training set and one testing set. Based on the training set, we obtained 100 subsets, and then identified survival-related miRNAs and genes using the univariable Cox method. Survival related subpathways were further identified by considering miRNAs and genes in each training subset. Using statistical counting, we finally identified six robust subpathways as a prognostic signature and verified their predictive power using an independent testing set.

## Materials and methods

### Training and testing sets from TCGA

We downloaded and obtained the AML data set, including expression data sets and patient clinical information from TCGA data portal (<https://portal.gdc.cancer.gov/>), which is an interactive data system for researchers to search, download, upload, and analyze harmonized cancer genomic data sets. For the AML expression data sets, the gene expression was generated using HTseq-FPKM, and the miRNA expression was generated using BCGSC miRNA profiling. We directly obtained the expression values (FPKM values) for each miRNA and gene per AML sample, and the average expression values for an miRNA or gene was calculated for duplicated samples. In addition, we eliminated the samples with survival times < 1 month, because these AML patients might have died for other reasons than the disease [22]. Finally, the expression data sets (including the miRNAs and genes) and clinical information of a total of 112 patients were utilized in this study, and we further randomly divided them into a training set (n = 76) and a testing set (n = 36). These two data sets resulted in the same ratio of good and poor survival samples with median survival times as a cutoff, which was also the same as the original data set. The training set and the testing set were totally independent of each other. We utilized the training set to identify robust survival-related subpathways and the testing set to verify the signature predictive power.

### The miRNA-gene interaction data

The miRNA-gene interactions were obtained from the miRTarBase [23], mir2Disease [24], miRecords [25], and TarBase [26] databases. By integrating these four databases, a total of 55146 miRNA-gene interactions between 1110 miRNAs and 20186 genes were obtained. Among these interactions, 6459 miRNA-gene interactions involving 358 miRNAs and 3452 genes that had been verified by low-throughput experiments were utilized in this study.

### The reconstructed subpathways by embedding miRNAs

First, we utilized the available R package [27] to obtain all biological pathways from the KEGG pathway database, which included 152 metabolic pathways and 191 non-metabolic pathways. Then, we utilized the K-clique method [19] to define and locate subpathway regions based on the distance similarity strategy in each metabolic and non-metabolic pathway, and the default K value (3) was used. After these processes, we obtained the gene-based subpathway regions. Then, by considering the miRNA-gene interactions from the verified data, we further embedded the miRNA components into biological subpathways, and the miRNA that regulated at least one gene within one subpathway was included in the corresponding subpathway. Finally, we considered the scale of the miRNA-embedded subpathway graphs, and the small scale graphs with less than one miRNA or three genes were eliminated in further analyses.

## Identification of survival-related subpathways

Based on the reconstructed subpathways containing both gene and miRNA components, we further identified survival-related subpathways by integrating miRNA and gene levels. First, using the expression data sets derived from TCGA training set, we identified survival miRNAs and genes using the univariable Cox method with a significance value of  $P < 0.05$ . Then, we mapped these survival miRNAs and genes into the reconstructed subpathway graphs. To evaluate whether one subpathway was related with patient survival at both the gene and miRNA levels, we utilized the hypergeometric test to evaluate the integrated statistical significance as follows:

$$P = 1 - \sum_{k=0}^{r_g+r_{mir}-1} \frac{\binom{t_g+t_{mir}}{k} \binom{m_g+m_{mir}-t_g-t_{mir}}{n_g+n_{mir}-k}}{\binom{m_g+m_{mir}}{n_g+n_{mir}}}$$

Where  $m_g$  and  $m_{mir}$  were the numbers of all genes and miRNAs in the TCGA data sets, and  $n_g$  and  $n_{mir}$  were the numbers of survival genes and miRNAs obtained from univariate Cox analysis, of which  $r_g$  genes and  $r_{mir}$  miRNAs were also included in one subpathway graph. And this subpathway graph contained  $t_g$  genes and  $t_{mir}$  miRNAs. The subpathways with hypergeometric p-value  $< 0.05$  were regarded as prognostic signatures at the miRNA-gene integrated levels.

## Calculation of single and combined subpathway risk scores

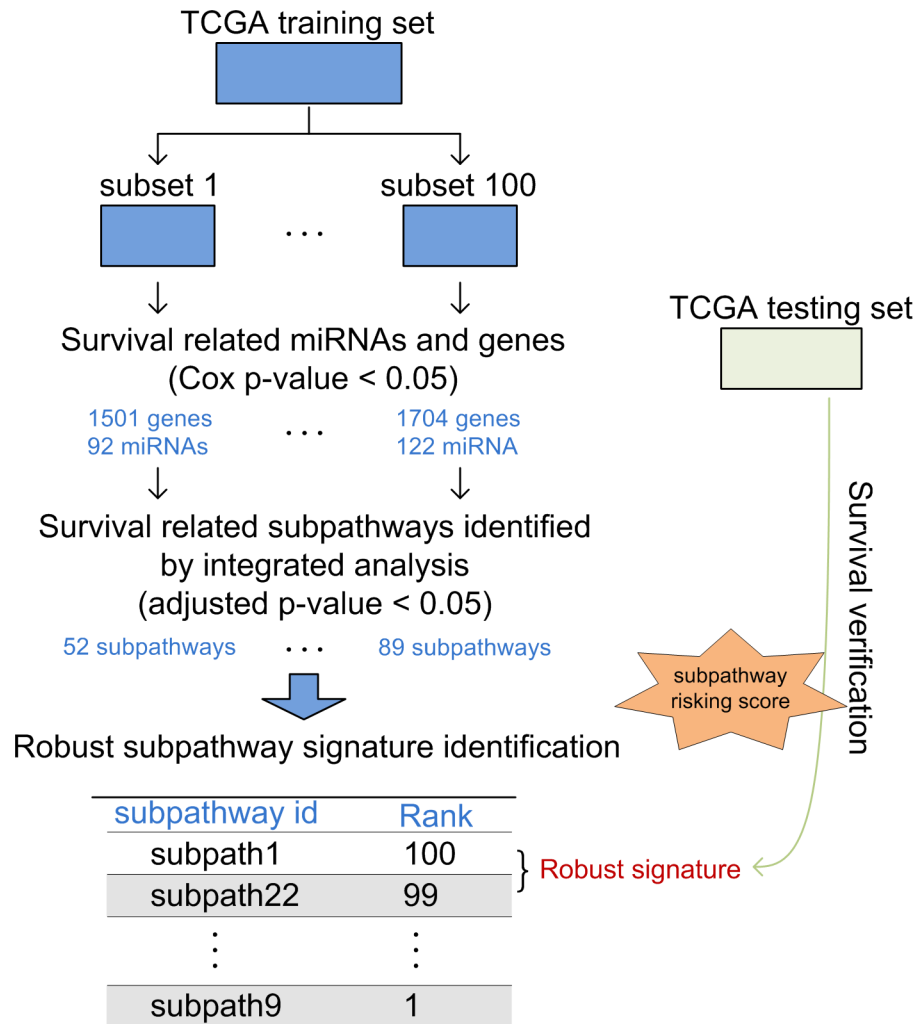
Based on each robust survival subpathway, we calculated the risk score by considering both miRNAs and genes that were involved in the subpathway. The risk score was a linear combination of the expression value for each miRNA and gene multiplied by a weighting value obtained from the univariate Cox analysis using the training set. For multiple subpathways, the combined risk score was also a linear combination of the risk score for these subpathways multiplied by a weighting value calculated by univariate Cox analyses based on the risk score in the training set.

## Results

### Construction of subpathway graphs and utilization of TCGA data sets

We utilized the available R package software to obtain subpathway graphs from the KEGG database by maintaining original topology information, and we embedded the regulatory roles of miRNAs by considering the low-throughput miRNA-gene interactions (see [Materials and Methods](#)). As a result, the constructed subpathway graphs contained an average of 21.8 miRNAs and 19.5 genes by the node level, and 29.2 miRNA-gene interactions by the edge level, and a total of 1773 subpathways with at least one miRNA and three genes were finally obtained for the identification of survival subpathways. Compared with entire pathways, the subpathway graphs contained smaller sets of components and smaller regulatory interactions between miRNAs and genes. The detailed information of these subpathways was provided in [S1 Table](#).

Based on the constructed subpathway graphs involving genes and miRNAs, we performed miRNA-gene integrated analyses to identify robust survival subpathways ([Fig 1](#)). First, we randomly divided the total TCGA data set into a training set and a testing set, which contained the same ratio of “good” and “poor” survival samples (median survival time as cutoff). There was no significant difference of patient characteristics between the training and testing data



**Fig 1. The study design.** The training set was used to identify robust prognostic subpathways, and the testing set was used to verify the predictive power of subpathway signatures.

<https://doi.org/10.1371/journal.pone.0194245.g001>

sets (S2 Table). Then, we utilized the training set to perform bioinformatics analyses to identify robust survival subpathways. Finally, we utilized the testing set to verify the predictive performance of the robust subpathway-based signatures.

### Identification of robust survival subpathways

Based on the training set, we performed a bootstrap process to identify robust survival subpathways. First, we randomly selected half of the total training set and formed 100 training subsets (from subsets 1–100). These subsets maintained the same ratio of “good” and “bad” survival samples (median survival time as cut-off) as the total training set. For each training subset, we then performed univariate Cox analyses to identify survival miRNAs and genes with a P-value < 0.05. The survival subpathways were also identified by integrated analyses of the levels of miRNAs and genes using the hypergeometric test (see Materials and Methods). Finally, based on 100 training subsets, we identified 100 survival subpathway lists and obtained a total of survival subpathways by merging all the results. Thus, each survival subpathway included was assigned to a count value from 1–100, and the subpathways with higher values

were more robust. As shown in [S1 Fig](#), the subpathway (path: 04662\_8) from the B cell receptor signaling pathway displayed a first value (48), which meant that this subpathway was significantly related with survival in a total of 48 training subsets. Moreover, the subpathways from the Pathways in cancer and the T cell receptor signaling pathway displayed high values, and some subpathways from the Cell cycle were also ranked in the top ten. These subpathways were involved in the biological mechanisms of tumors, so further exploration and dissection could provide guidance for AML survival implications. Detailed information of the top 50 robust subpathways is shown in [S3 Table](#).

### Robust subpathways significantly predicted patient clinical outcomes

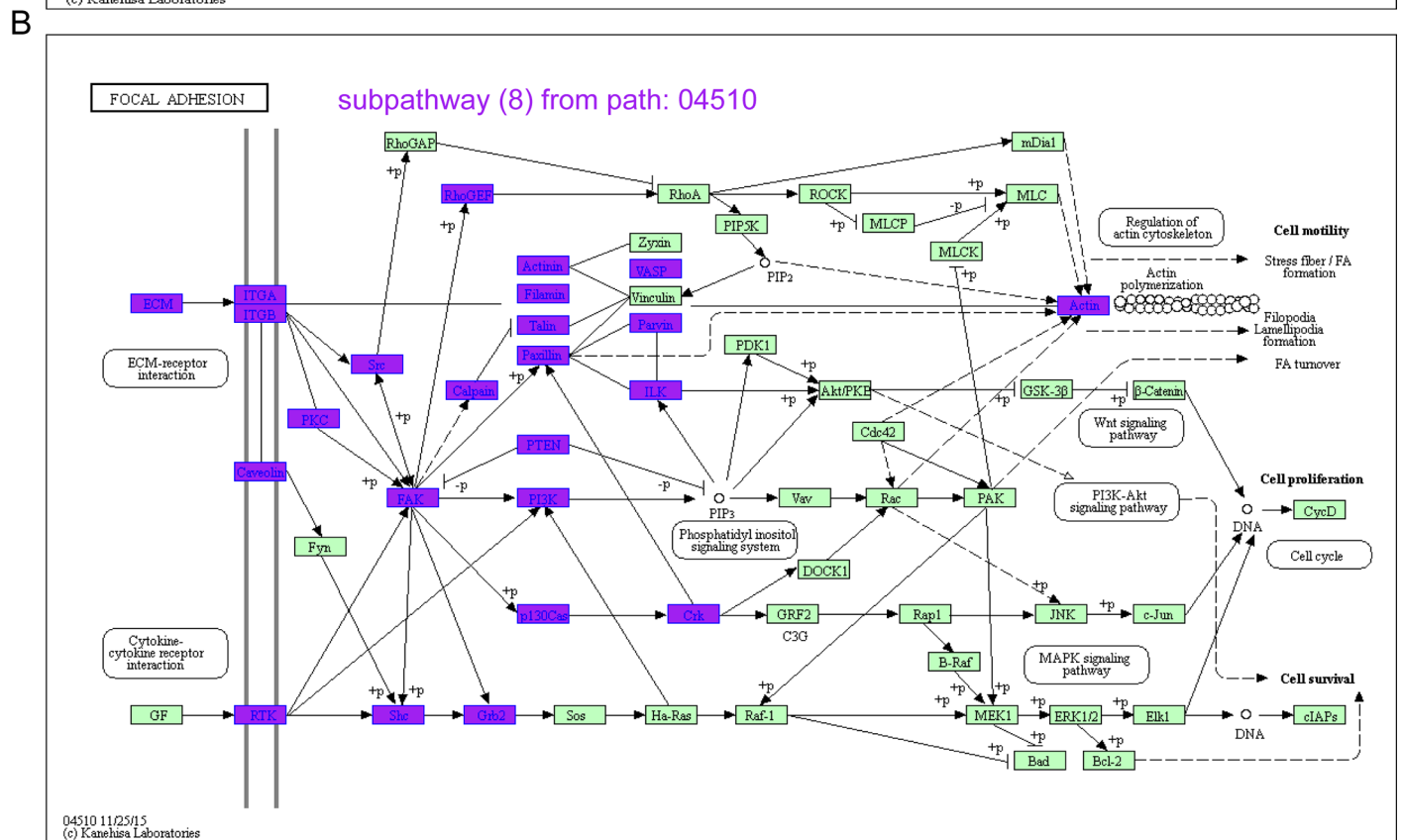
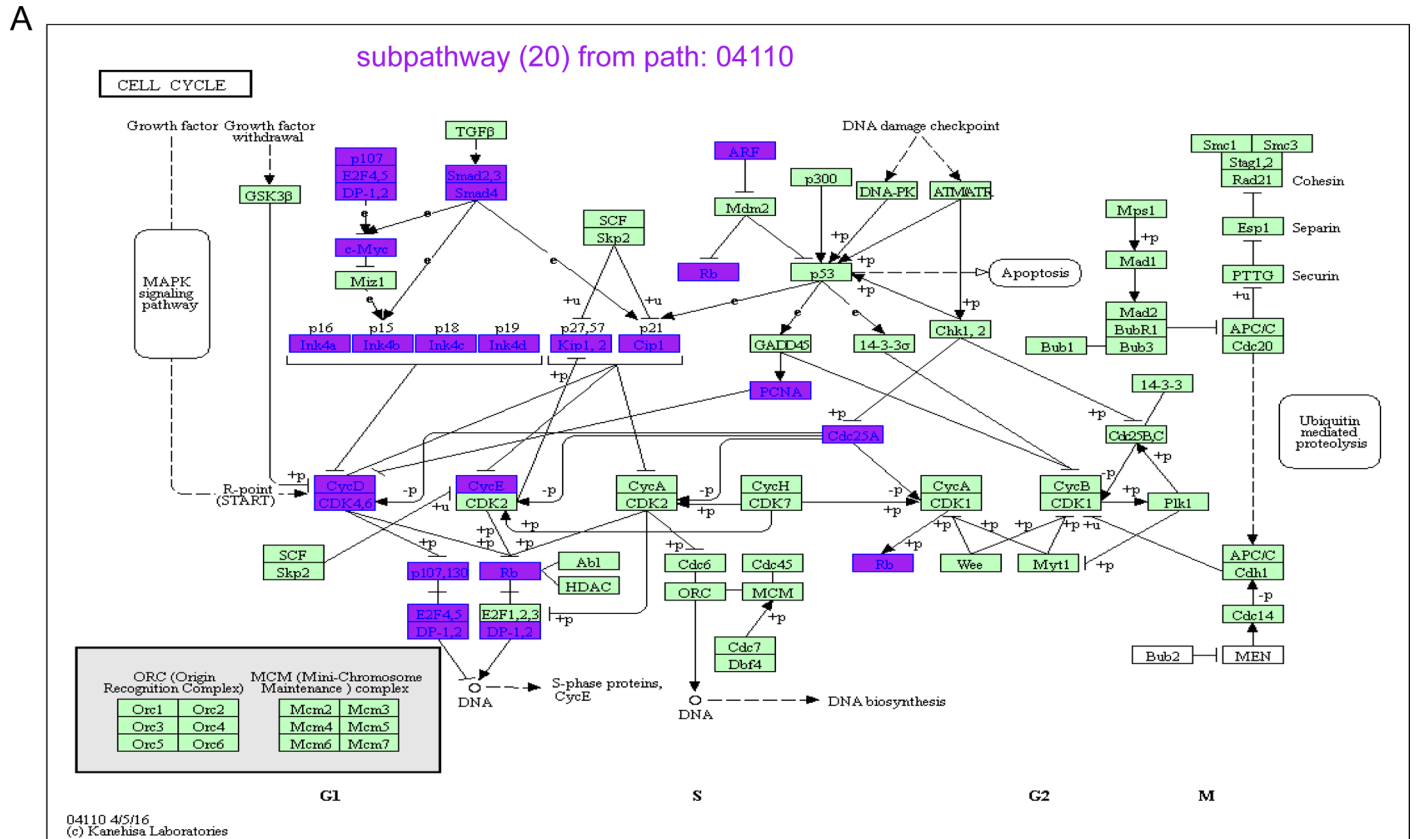
Based on each robust subpathway, we calculated the risk score for survival analyses (see [Materials and Methods](#)). The risk score simultaneously considered the contributions of miRNAs and genes. Furthermore, the median score was taken as the cutoff for the low risk and high risk group classification. The Kaplan-Meier method was used to generate survival curves and the difference between two survival curves was evaluated by using the log-rank test. All tests were two-tailed, and a value of  $P < 0.05$  was considered to indicate a significant result. Among the robust subpathways with rank value  $> 43$ , we observed that three representative subpathways, path: 05200\_10 from Pathways in cancer, path: 04110\_20 from the Cell cycle, and path: 04510\_8 from Focal adhesion were associated with the survival status of AML patients ([Fig 2](#) and [S2 Fig](#)). For TCGA training set, the survival P-value of the risk subpathway score for path: 05200\_10, path: 04110\_20, and path: 04510\_8 were 1.0E-08, 2.0E-08, and 0, respectively ([Fig 3](#)). Similarly, the risk scores were calculated for the samples in the testing set and two risk groups (high risk and low risk). In accordance with the results from the training set, patients with a high subpathway risk score belonged to the high risk group, and the patients with low subpathway risk scores belonged to the low risk group. The log-rank test showed that there was a significant difference of survival times between the low risk and high risk groups ( $P = 0.008$  for path: 04110\_20, see [Fig 3B](#);  $P = 0.002$  for path: 04510\_8, see [Fig 3D](#); and  $P = 0.023$  for path: 05200\_10, see [S2 Fig](#)). Based on the results for these three subpathways, we further defined the combined subpathway risk score for AML survival analyses (see [Materials and Methods](#)). For the test set, the combined risk score was also related to the survival status of AML patients with  $P = 0.021$ , further showing the predictive performance of the robust survival subpathways.

### Functional analyses of the subpathway signatures

To dissect the functional roles of the subpathway signatures, we first extracted the gene components within these subpathways and performed DAVID functional enrichment analyses. As shown in [S4 Table](#), some signaling terms, including epidermal growth factor receptor signaling pathway (GO: 0007173) and ERBB2 signaling pathway (GO: 0038128), were significantly enriched in the Path: 05200\_10 subpathway. Moreover, immune related terms were also included, including leukocyte migration. For the Path: 04110\_20 subpathway, some GO terms regarding cell cycle regulation were identified as significant. Similarly, the Path: 04510\_8 subpathway was also related with immune, angiogenesis, and cell adhesion, including cell adhesion (GO: 0007155) and cell-substrate adhesion (GO: 0031589). The cell adhesion and cell cycle functions were involved in different types of human tumors.

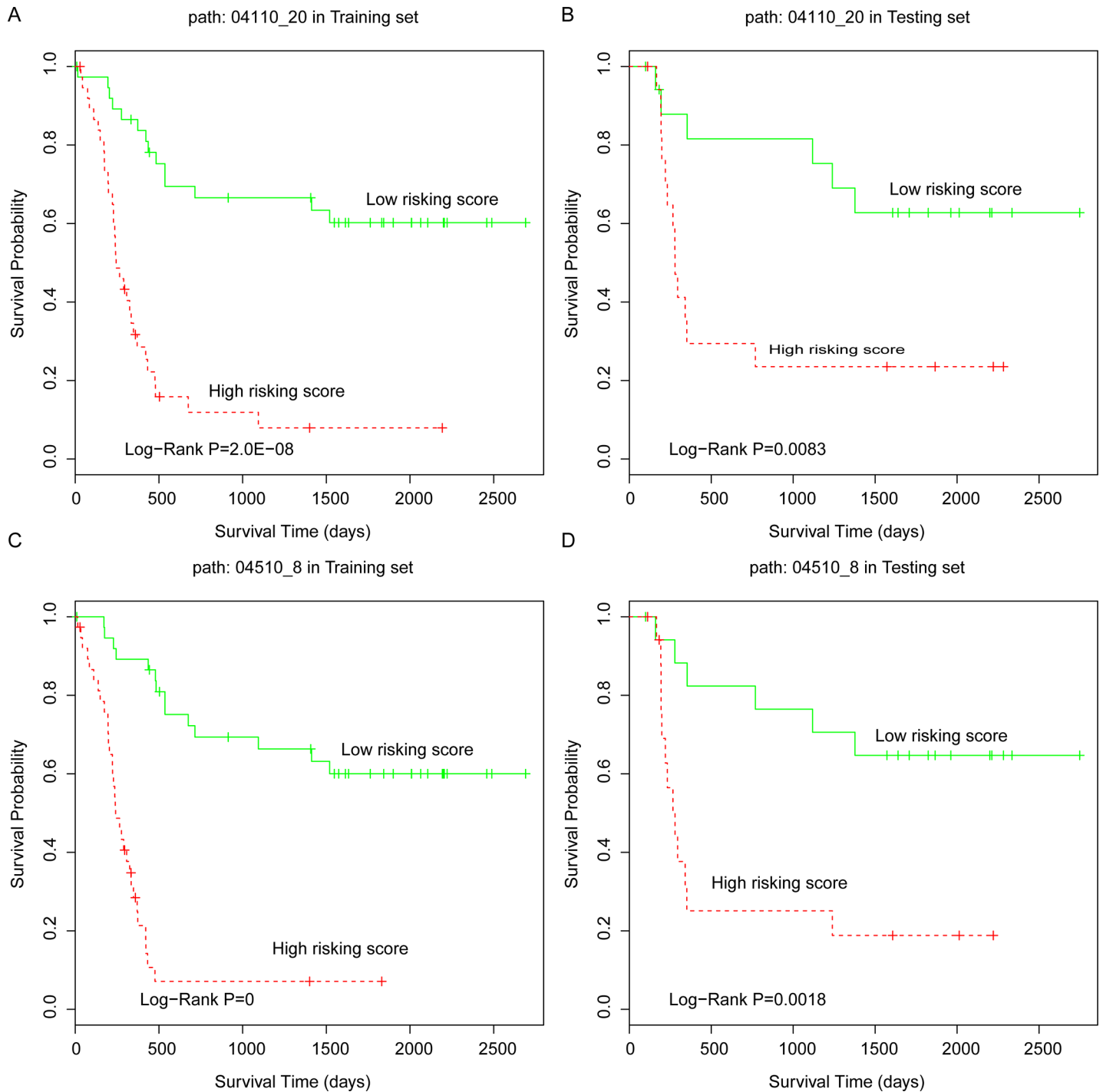
### Discussion

Previous studies have reported that the gene-based expression signatures derived from different studies shared poor overlaps, which suggested the necessity of a functionally-based



**Fig 2. The two subpathways.** (A) Subpathway\_20 from the Cell cycle (path: 04110). (B) Subpathway\_8 from Focal adhesion (path: 04510).

<https://doi.org/10.1371/journal.pone.0194245.g002>



**Fig 3. The K-M curves of robust subpathways.** (A) The 04110\_20 in the training set. (B) The 04110\_20 in the testing set. (C) The 04510\_8 in the training set. (D) The 04510\_8 in the testing set. The P-value was calculated using the log-rank test.

<https://doi.org/10.1371/journal.pone.0194245.g003>



analytical strategy. In this study, we identified robust survival subpathways for AML survival analysis using the sample-matched miRNA and gene expression data sets from TCGA database. In the previous studies, the regulatory relations between miRNAs and genes were considered in the AML prognosis analysis [17, 18]. However, in this study, we considered both the expression levels of miRNAs and genes into the survival subpathway identification based on the reconstructed graphs. Overall, this was the first study to identify AML survival signatures at the subpathway level by integrated analyses of both miRNAs and genes.

Many methods have been developed to analyze tumor biological mechanisms at the pathway level. Ooi *et al.* developed a computational approach to identify tumor molecular pathways, and the new driving pathway patterns they found could be used to define clinically relevant subgroups for gastric cancers [10]. Moreover, Huang *et al.* proposed a pathway-based model by including the Cox proportional hazard regression method and further applied this model to predict outcomes for breast cancer patients [11]. However, most current studies performed the function- or pathway-level analyses by considering only gene expression, and did not include the regulatory roles of small non-coding RNAs, for example miRNAs.

Although biological pathways exhibited advantages over Gene Ontology functional terms because of their topology information, the large amount of genes within an entire pathway scale presented challenges for previous analyses. To resolve this issue, the subpathway concept (pathway regions within the entire pathway) was used in a previous study [19]. Compared with the entire pathway, the subpathway provided more precise information for functional enrichment analyses. In addition, studies showed that abnormalities of subpathways, and not the entire pathway, were associated with disease etiology. The subpathway-level analysis method was also applied to explore other biological mechanisms, for example, the actions of drugs [28]. The subpathway concept could therefore be used for identifying AML survival signatures, and the roles of miRNAs should also be considered.

Using the training set, we performed an analysis to identify the most robust survival subpathways, and three representative subpathways were identified, including path: 05200\_10 from Pathways in cancer, path: 04110\_20 from Cell cycle, and path: 04510\_8 from Focal adhesion. Pathways in cancer is a tumor pathways which was associated with multiple of tumor types. Also, the cell cycle arrest was involved in the molecular mechanism of AML cells, which correlated with known responses to therapy [29, 30]. Moreover, the study of Carter *et al.* found that the kinase FAK within the focal adhesion pathway regulated leukemia-stromal interactions and supported leukemia cell survival, thus could be a potential therapeutic target in AML [31]. Based on these robust subpathways, we further calculated single and combined risk scores, which combined the Cox coefficient and expression values of all miRNAs and gene components involved in corresponding subpathway graphs. Notably, survival subpathways were identified by considering both survival-related miRNAs and survival-related genes using a hypergeometric test; however, all miRNA and gene components (including significant and insignificant) within the subpathway were included in the calculation of the risk score. The risk score we calculated therefore reflected the overall risk condition at the subpathway level.

We integrated high-throughput miRNAs, mRNA expressions, and pathway structures to systematically identify robust subpathways, and further calculated subpathway-based risk scores for predicting AML patient prognoses. With more sample-matched miRNA and mRNA expression data sets, we further validated the predictive performance of the subpathway-based signatures. Notably, the subpathway signatures we identified contained smaller set of miRNA and gene components than corresponding entire pathways. And the key components within these subpathways also had the potential to be prognostic marker for clinical applications. In the current study, our findings were potentially useful for understanding AML biological events and identifying functional signatures for clinical use.

## Supporting information

**S1 Fig. The distribution of count values for significant subpathways.**

(TIF)

**S2 Fig. (A)** Subpathway\_10 from Pathways in cancer (path: 05200). **(B)** The 05200\_10 in the training set. **(C)** The 05200\_10 in the testing set.

(TIF)

**S1 Table. Detailed information of subpathway graphs.**

(XLS)

**S2 Table. The clinical features of acute myeloid leukemia patients in The Cancer Genome Atlas training and testing sets.**

(DOC)

**S3 Table. Detailed information on the top 50 subpathways.**

(XLS)

**S4 Table. Top 30 significant Gene Ontology terms enriched by the gene components within three subpathways.**

(XLS)

## Author Contributions

**Data curation:** Wei Ding, Xueqin Qing.

**Formal analysis:** Huijuan Chang.

**Investigation:** Huijuan Chang.

**Methodology:** Xueqin Qing.

**Validation:** Huijuan Chang.

**Visualization:** Qiuying Gao, Wei Ding.

**Writing – original draft:** Xueqin Qing.

**Writing – review & editing:** Qiuying Gao, Xueqin Qing.

## References

1. Lowenberg B, Downing JR, Burnett A. Acute myeloid leukemia. *N Engl J Med.* 1999; 341(14):1051–62. <https://doi.org/10.1056/NEJM199909303411407> PMID: 10502596
2. de Lima M, Anagnostopoulos A, Munsell M, Shahjahan M, Ueno N, Ippoliti C, et al. Nonablative versus reduced-intensity conditioning regimens in the treatment of acute myeloid leukemia and high-risk myelodysplastic syndrome: dose is relevant for long-term disease control after allogeneic hematopoietic stem cell transplantation. *Blood.* 2004; 104(3):865–72. <https://doi.org/10.1182/blood-2003-11-3750> PMID: 15090449
3. Garzon R, Volinia S, Liu CG, Fernandez-Cymering C, Palumbo T, Pichiorri F, et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood.* 2008; 111(6):3183–9. <https://doi.org/10.1182/blood-2007-07-098749> PMID: 18187662
4. Chen HY, Yu SL, Li KC, Yang PC. Biomarkers and transcriptome profiling of lung cancer. *Respirology.* 2012; 17(4):620–6. <https://doi.org/10.1111/j.1440-1843.2012.02154.x> PMID: 22372638
5. Costa AFO, Menezes DL, Pinheiro LHS, Sandes AF, Nunes MAP, Lyra Junior DP, et al. Role of new Immunophenotypic Markers on Prognostic and Overall Survival of Acute Myeloid Leukemia: a Systematic Review and Meta-Analysis. *Sci Rep.* 2017; 7(1):4138. <https://doi.org/10.1038/s41598-017-00816-2> PMID: 28646224

6. Huang R, Liao X, Li Q. Identification and validation of potential prognostic gene biomarkers for predicting survival in patients with acute myeloid leukemia. *Onco Targets Ther.* 2017; 10:5243–54. <https://doi.org/10.2147/OTT.S147717> PMID: 29138577
7. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, 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):15545–50. <https://doi.org/10.1073/pnas.0506580102> PMID: 16199517
8. Hung JH. Gene Set/Pathway enrichment analysis. *Methods Mol Biol.* 2013; 939:201–13. Epub 2012/11/30. [https://doi.org/10.1007/978-1-62703-107-3\\_13](https://doi.org/10.1007/978-1-62703-107-3_13) PMID: 23192548
9. Yang X, Regan K, Huang Y, Zhang Q, Li J, Seiwert TY, et al. Single sample expression-anchored mechanisms predict survival in head and neck cancer. *PLoS Comput Biol.* 2012; 8(1):e1002350. <https://doi.org/10.1371/journal.pcbi.1002350> PMID: 22291585
10. Ooi CH, Ivanova T, Wu J, Lee M, Tan IB, Tao J, et al. Oncogenic pathway combinations predict clinical prognosis in gastric cancer. *PLoS Genet.* 2009; 5(10):e1000676. <https://doi.org/10.1371/journal.pgen.1000676> PMID: 19798449
11. Huang S, Yee C, Ching T, Yu H, Garmire LX. A novel model to combine clinical and pathway-based transcriptomic information for the prognosis prediction of breast cancer. *PLoS Comput Biol.* 2014; 10(9):e1003851. <https://doi.org/10.1371/journal.pcbi.1003851> PMID: 25233347
12. Rottiers V, Naar AM. MicroRNAs in metabolism and metabolic disorders. *Nat Rev Mol Cell Biol.* 2012; 13(4):239–50. <https://doi.org/10.1038/nrm3313> PMID: 22436747
13. Tomasetti M, Santarelli L, Neuzil J, Dong L. MicroRNA regulation of cancer metabolism: role in tumour suppression. *Mitochondrion.* 2014; 19 Pt A:29–38.
14. Zhang C, Li C, Li J, Han J, Shang D, Zhang Y, et al. Identification of miRNA-mediated core gene module for glioma patient prediction by integrating high-throughput miRNA, mRNA expression and pathway structure. *PLoS One.* 2014; 9(5):e96908. <https://doi.org/10.1371/journal.pone.0096908> PMID: 24809850
15. Wan YW, Beer DG, Guo NL. Signaling pathway-based identification of extensive prognostic gene signatures for lung adenocarcinoma. *Lung Cancer.* 2012; 76(1):98–105. <https://doi.org/10.1016/j.lungcan.2011.09.016> PMID: 22047960
16. Lu TP, Chuang EY, Chen JJ. Identification of reproducible gene expression signatures in lung adenocarcinoma. *BMC Bioinformatics.* 2013; 14:371. <https://doi.org/10.1186/1471-2105-14-371> PMID: 24369726
17. Shah NM, Zaitseva L, Bowles KM, MacEwan DJ, Rushworth SA. NRF2-driven miR-125B1 and miR-29B1 transcriptional regulation controls a novel anti-apoptotic miRNA regulatory network for AML survival. *Cell Death Differ.* 2015; 22(4):654–64. <https://doi.org/10.1038/cdd.2014.152> PMID: 25323587
18. Zhang C, Bai G, Zhu W, Bai D, Bi G. Identification of miRNA-mRNA Network Associated with Acute Myeloid Leukemia Survival. *Med Sci Monit.* 2017; 23:4705–14. <https://doi.org/10.12659/MSM.903989> PMID: 28965123
19. Li C, Li X, Miao Y, Wang Q, Jiang W, Xu C, et al. SubpathwayMiner: a software package for flexible identification of pathways. *Nucleic Acids Res.* 2009; 37(19):e131. <https://doi.org/10.1093/nar/gkp667> PMID: 19706733
20. Li C, Han J, Yao Q, Zou C, Xu Y, Zhang C, et al. Subpathway-GM: identification of metabolic subpathways via joint power of interesting genes and metabolites and their topologies within pathways. *Nucleic Acids Res.* 2013; 41(9):e101. <https://doi.org/10.1093/nar/gkt161> PMID: 23482392
21. Li X, Li C, Shang D, Li J, Han J, Miao Y, et al. The implications of relationships between human diseases and metabolic subpathways. *PLoS One.* 2011; 6(6):e21131. <https://doi.org/10.1371/journal.pone.0021131> PMID: 21695054
22. Srinivasan S, Patric IR, Somasundaram K. A ten-microRNA expression signature predicts survival in glioblastoma. *PLoS One.* 2011; 6(3):e17438. <https://doi.org/10.1371/journal.pone.0017438> PMID: 21483847
23. Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, et al. miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2011; 39(Database issue):D163–9. <https://doi.org/10.1093/nar/gkq1107> PMID: 21071411
24. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, et al. miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res.* 2009; 37(Database issue):D98–104. <https://doi.org/10.1093/nar/gkn714> PMID: 18927107
25. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.* 2009; 37(Database issue):D105–10. <https://doi.org/10.1093/nar/gkn851> PMID: 18996891

26. Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, et al. TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res.* 2012; 40 (Database issue):D222–9. <https://doi.org/10.1093/nar/gkr1161> PMID: 22135297
27. Feng L, Xu Y, Zhang Y, Sun Z, Han J, Zhang C, et al. Subpathway-GMir: identifying miRNA-mediated metabolic subpathways by integrating condition-specific genes, microRNAs, and pathway topologies. *Oncotarget.* 2015; 6(36):39151–64. <https://doi.org/10.18632/oncotarget.5341> PMID: 26472186
28. Li C, Shang D, Wang Y, Li J, Han J, Wang S, et al. Characterizing the network of drugs and their affected metabolic subpathways. *PLoS One.* 2012; 7(10):e47326. <https://doi.org/10.1371/journal.pone.0047326> PMID: 23112813
29. Li P, Yan S, Dong X, Li Z, Qiu Y, Ji C, et al. Cell Cycle Arrest and Apoptosis Induction Activity of Nitidine Chloride on Acute Myeloid Leukemia Cells. *Med Chem.* 2018; 14(1):60–6. <https://doi.org/10.2174/1573406413666170620091543> PMID: 28629311
30. Behbehani GK, Samusik N, Bjornson ZB, Fantl WJ, Medeiros BC, Nolan GP. Mass Cytometric Functional Profiling of Acute Myeloid Leukemia Defines Cell-Cycle and Immunophenotypic Properties That Correlate with Known Responses to Therapy. *Cancer Discov.* 2015; 5(9):988–1003. <https://doi.org/10.1158/2159-8290.CD-15-0298> PMID: 26091827
31. Carter BZ, Mak PY, Wang X, Yang H, Garcia-Manero G, Mak DH, et al. Focal Adhesion Kinase as a Potential Target in AML and MDS. *Mol Cancer Ther.* 2017; 16(6):1133–44. <https://doi.org/10.1158/1535-7163.MCT-16-0719> PMID: 28270436