

PRIMARY RESEARCH

Open Access



Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia

Xuejiao Yin¹, Sui Huang¹, Ruiqi Zhu¹, Fengjuan Fan¹, Chunyan Sun^{1,2*} and Yu Hu^{1,2*} 

Abstract

Background: LncRNAs can regulate miRNAs and mRNAs by sequestering and binding them. Indeed, many researchers have reported lncRNA mediated-competing endogenous RNAs (ceRNAs) could regulate the progression of solid tumors. However, the roles of ceRNA in acute myeloid leukemia (AML), especially in pediatric and adolescent AML, were not completely expounded.

Materials and methods: 27 cytogenetically normal acute myeloid leukemia (CN-AML) patients under 18 years old with corresponding clinical data were selected from the cancer genome atlas (TCGA), which was a large sample sequencing database of RNA sequencing. We constructed a survival specific ceRNA network, and investigated its associations with patients' clinical information by analyzing the data from TCGA.

Results: We identified survival specific lncRNAs, miRNAs and mRNAs, and constructed a survival specific ceRNA network of CN-AML patients and a weighted correlation network. Furthermore, we identified 4 biological pathways associated with OS and selected the most enriched pathway 'Transcriptional misregulation in cancer' to verify that it could accurately predict younger CN-AML patients' prognosis to guide treatment.

Conclusions: We successfully constructed a survival specific ceRNA network which could provide a new approach to lncRNA research in younger CN-AML. Importantly, we constructed a weighted correlation network to overcome the difficulty in biological interpretation of individual genes.

Keywords: Pediatric and adolescent, CN-AML, LncRNA, ceRNA network, WGCNA, Biological pathways, Prognosis

Background

Long non-coding RNAs (lncRNA) were a group of noncoding RNAs with more than 200 bp, which has intrinsic advantages over the use of protein-coding RNAs in diagnostics [1]. LncRNAs play critical roles in diverse biological functions, including nuclear architecture, regulation of gene expression, immune

surveillance, cancer development and maintenance of tumorigenesis. LncRNAs can regulate miRNAs and mRNAs by sequestering and binding them, and competing endogenous RNA (ceRNA) mechanism attracted more and more attention since it was firstly proposed by Salmena et al. [2]. CeRNA is a complex post-transcriptional regulatory network using miRNA response elements (MREs) to compete for the binding of miRNAs thereby implementing mutual control between mRNAs, lncRNAs and miRNA [3]. A lot of studies have reported that ceRNA mechanism

*Correspondence: suncy0618@163.com; dr_huyu@126.com

¹ Institute of Hematology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China
Full list of author information is available at the end of the article



plays critical roles in solid tumor progression [4–6]. Junge et al. [7], Ding et al. [8] and Chen et al. [9] also reported highly-upregulated RUNX1T1, C-Myc and CCAT1 acts as competing endogenous RNA (ceRNA) in adult acute myeloid leukemia. However, little has been done in pediatric and adolescent cytogenetically normal acute myeloid leukemia (CN-AML) patients. Although board overlap lies between AML adults and pediatric patients in pathogenesis, diagnosis, treatment and prognosis of the diseases, differences still exist.

CN-AML, a most common AML type, is characterized by the absence of microscopically detectable chromosome abnormalities, but mutations, epigenetic changes and dysregulated expression signatures exists, respectively [10]. Some reports have shown CN-AML patients harboring mutations including NPM1, CEBPA, FLT3-ITD and WT1 were associated with an disparate prognosis [11]. With the advent of high-throughput technologies such as microarray and next-generation sequencing, some prognostic gene expression signatures have been proposed. These findings helped focus targeted therapies for CN-AML with significantly heterogeneous outcomes [12]. The genes in a prognostic gene signature were considered as independent individuals, which may result in ignoring potential relationships between the genes [13]. To overcome these challenges and to reduce the difficulty in biological interpretation, we constructed a weighted correlation network and the significant modules associated with prognosis were selected [14]. Based on these modules information, we identified 4 biological pathways associated with overall survival (OS). Next the most enriched pathway was selected to verify whether it could accurately predict pediatric and adolescent CN-AML patients' prognosis and guide treatment. To our knowledge, our study is the first to construct survival specific ceRNAs co-expression network and identify survival specific biological pathways in pediatric and adolescent CN-AML patients. This approach can help to clarify the functions of lncRNAs in younger CN-AML patients.

Materials and methods

TCGA AML dataset

A total of 27 patients with AML were collected from the cancer genome atlas (TCGA) database. The criteria of exclusion were set as follows: (i) patients were not cytogenetically normal or cytogenetically information unknown; (ii) patients didn't have lncRNA, mRNA, miRNA information; (iii) patients were alive and the last contact days were unavailable. Overall, 27 CN-AML patients who were under 18 at diagnosis were included

in our study. This study was fully compliance with the publication guidelines provided by TCGA. The data were obtained from TCGA, so the approval of ethics committee was not needed.

Data processing

TCGA database provided the normalized count data of RNA sequencing including lncRNA and mRNA expression profiles by RNASeqV2 system. The STAD level 3 microRNA sequencing (miRNAseq) data, downloaded from TCGA database, were collected by Illumina HiSeq 2000 microRNA sequencing platforms. TCGA database have already normalized these RNA expression profile data, so no further normalization was required. In the next step, each lncRNA, miRNA and mRNA were respectively put into univariate cox's model. P value < 0.05 as screening criteria were used to select lncRNA, miRNA and mRNA which were significantly associated with OS. Clinical information including age at diagnosis, gender, wbc at diagnosis, bone marrow blasts, peripheral blasts and molecular mutations (FLT3-ITD, NPM1, CEBPA) were also used to build univariate cox's model under the same standard. To explore the relationship between prognosis and the key lncRNAs, miRNAs and mRNAs involved in the ceRNA network, Kaplan–Meier curve were carried out at a P value < 0.05. Each significant mRNA identified by univariate cox was further chosen to construct a weighted correlation network and the significant modules associated with prognosis were selected for further exploration. The mRNAs in the specific modules associated with OS were further implemented by functional enrichment analysis. The most enriched pathway was selected by carrying out multivariate cox regression formula. Prognostic models were reconstructed based on the relative contributions of each of the genes in the cox analysis, as described in the following equation: The risk score = $Z1G1 + Z2G2 + Z3G3 + \dots + ZnGn$ (z is z scores obtained from the Cox analysis and G is the expression value of each gene). Then, the patients were classified into 2 groups (high risk vs low risk) based on the median value of the predictor score. Survival times were compared between the 2 groups by using the Kaplan–Meier analysis at a P value < 0.05.

Construction of the ceRNA network

We used miRanda (<http://www.microrna.org/microrna/home.do>) to find the lncRNA–miRNA interactions, and using miRBase targets (<http://mirdb.org/miRDB/>), miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/>) and Targetscan (<http://www.targetscan.org/>) to predict target genes. Predicted targets are ranked according to the predicted efficacy of targeting as calculated using cumulative weighted context ++ scores of the sites, and scores < -0.4 were selected as target genes of each miRNA. Then,

according to the theory of ceRNA, we chose the miRNA negatively regulated intersection expression of lncRNAs and mRNAs to construct the ceRNA network. The ceRNA networks were constructed and visualized using Cytoscape v3.0.

Weighted correlation network construction

Gene correlation is described as a network in which the relationship between the connected genes is represented by the weight. Each gene is described as a node. The edge weight between the connected nodes is the pairwise Pearson coefficient. In gene correlation network construction, an adjacency matrix and an adjacency function are defined. Using adjacency function the co-expression similarity between genes can be described as connection strength. The node dissimilarity is input to hierarchical clustering to define network modules. From the clustering tree many gene co-expression modules are discovered. In the construction of hierarchical clustering tree, a dynamic shear algorithm based on tree branch shape is used. R Bioconductor was performed to construct a weighted correlation network.

GO, pathway analysis, PPI network establishment and leave-one-out cross validation (LOO-CV)

The Database for Annotation, Visualization and Integrated Discovery (DAVID, <http://david.abcc.ncifcrf.gov/>) was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Protein–protein interactions (PPI) network of survival specific genes was constructed using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <http://string.embl.de/>). In the network, genes represent nodes and the interactions between the nodes represent edges. The important nodes with high degree in the network were obtained, namely hub nodes. The co-expresses value >0.4 was used as the cutoff criterion. R Bioconductor was used to perform leave-one-out cross validation.

Results

Survival specific lncRNAs, miRNAs and mRNAs in cytogenetically normal AML and ceRNA network construction

After univariate cox’s model filtering of the data extracted from TCGA database, we identified 1402 lncRNAs, 56 miRNA, and 2232 mRNA negatively associated with overall survival, as well as 34 lncRNAs, 46 miRNA, and 54 mRNA positively associated with overall survival (P<0.05). Next, based on the theory that lncRNA can regulate miRNA abundance by sequestering and binding them, acting as so-called miRNA sponges, we intended to establish the lncRNA–miRNA–mRNA ceRNAs network. 7 miRNAs targeted 47 key lncRNAs were predicted through miRcode, and 56 target genes of these 7 miRNAs were predicted by three prediction tools—TargetScan, miRanda and miRTarBase. Considering false positive rate of prediction tools, predicted targets are ranked according to the predicted efficiency of targeting as calculated using cumulative weighted context ++ scores of the sites. According to the theory of ceRNA, we then chose the miRNA that negatively regulated intersection expression of lncRNAs and mRNAs to construct the ceRNA network using Cytoscape 3.0 (Table 1).

Finally, 35 lncRNAs, 7 miRNAs and 44 mRNAs were involved in the ceRNA network (Fig. 1). Furthermore, to identify their association with prognostic characteristics, the Kaplan–Meier curve was applied. 6 lncRNAs, 1 miRNA and 14 mRNAs were found significantly associated with cytogenetically normal AML patients’ overall survival (log-rank P<0.05) (Additional file 1: Figure S1; Additional file 2: Figure S2; Additional file 3: Figure S3). Among these significant RNAs, 4 lncRNAs (AL356475.1, CRNDE, LINC00158 and LINC00504), 1 miRNA (hsa-mir-363) and all 14 mRNAs (BRWD1, CNIH1, FMR1, GOLGA8J, GPBP1L1, HMBOX1, KIF5B, NUFIP2, PTAR1, RRN3, SNX4, TMF1, ZFC3H1 and ZKSCAN8) were negatively associated with overall survival

Table 1 miRNAs targeted specific intersection key lncRNAs and mRNAs in ceRNA network

| miRNAs | lncRNAs | mRNAs |
|--------------|---|--|
| hsa-mir-25 | TTTY9B, AC068020.1, C11orf44, DLEU1, CASC2, AC024563.1, AL356475.1, AC083799.1, AC006305.1, POLR2J4, AC009495.1, LINC00365, LINC00502, LINC00392, MIR4500HG, AC104445.1, CRNDE, LINC00504, C8orf49, VENTXP1 | RSBN1, CNIH1, TWLF1, MFF, CCSER2, BAZ2B, SMU1, PTEN, SLC25A32, FOXN2, RRN3, DDX3X, EDEM1, NUFIP2, ZFC3H1, TMF1, KLHL15, VPS4B, CNEP1R1, DNAJB9, SLC25A36, GOLGA8J, SOX11, KIF5B, KIAA1109, PTAR1, SSFA2, GPBP1L1 |
| hsa-mir-363 | AC011498.1, AC092811.1, AC108134.2 | |
| hsa-mir-506 | AC108134.2 | |
| hsa-mir-551a | LINC00470, DLEU1, AC079949.1 | |
| hsa-mir-221 | C15orf54, CASC2, LINC00158, AC011467.1, AC006305.1, LINC00326, BPESC1, LINC00460, AL356489.1, CRNDE, C5orf17 | PCDHA4, CHORDC1, RAB1A, HMBOX1, PPP6C, BRWD1, FBXO28, SNX4, FMR1, UBE2J1, PCDHA6, PCDHA5, ZKSCAN8, PAIP2 |
| hsa-mir-100 | DLEU1, AC105206.1, POLR2J4, LINC00365, C8orf49 | KBTBD8, SMARCA5 |
| hsa-mir-204 | AC092811.1, AC068831.1 | |

($P < 0.05$), while the remaining 2 lncRNA (AC011498.1 and AC092811.1) were positively associated with overall survival ($P < 0.05$).

Correlation of lncRNAs and miRNAs signature with clinical or laboratory features and gene alterations

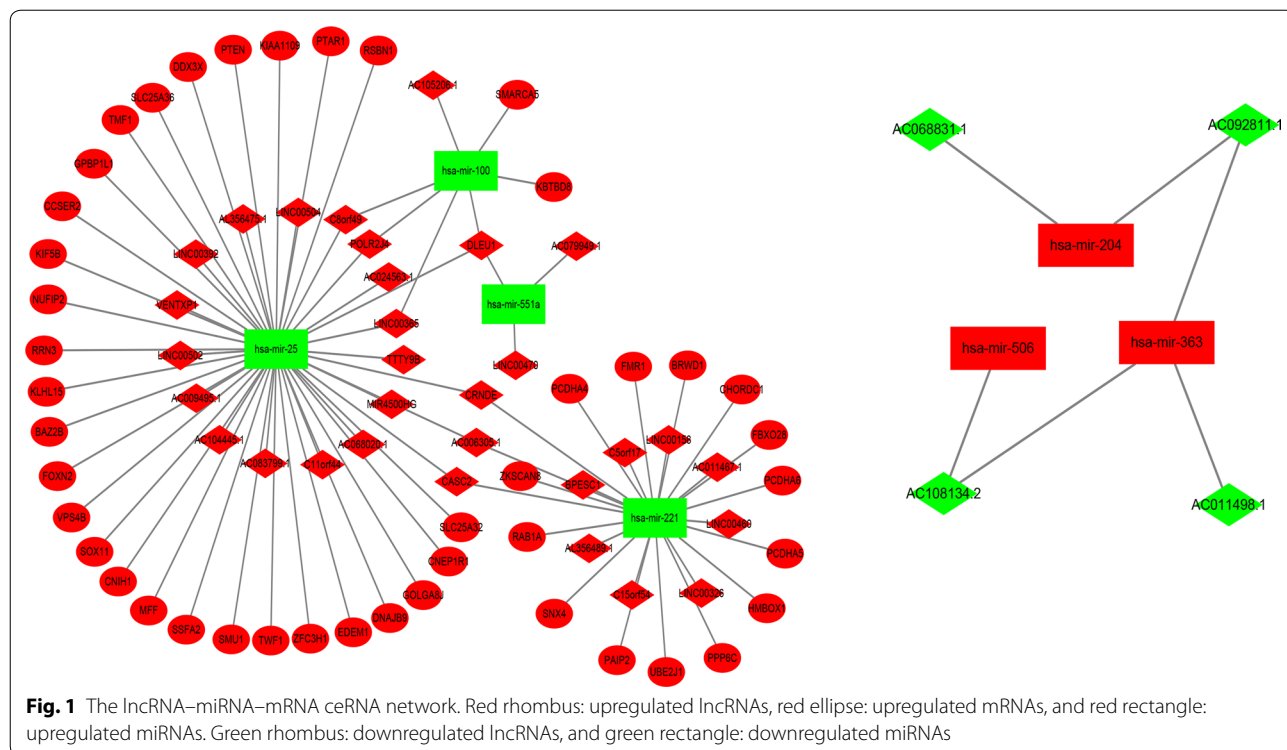
The 35 key lncRNAs and 7 miRNAs in the ceRNA network were further analyzed for the association with the clinical features, such as age at diagnosis, gender, WBC at diagnosis, bone marrow blasts, peripheral blasts and molecular mutations (FLT3-ITD, NPM1, CEBPA) (Additional file 4: Table S1). We found that CRNDE, LINC00504 and hsa-mir-363 were associated with peripheral blasts, CRNDE were associated with bone marrow blasts, CASC2 and CRNDE were associated with WBC at diagnosis. However, there was no association of other lncRNA, miRNA with any clinical features. In addition, we found patients with high level of AC011498.1 were more likely to carry NPM1 mutation and CEBPA mutation compared with patients with low level of the lncRNA. Moreover, patients with high levels of CRNDE, LINC00504 and hsa-mir-363 were carried CEBPA mutation. Association between miRNAs, race and ethnicity were also assessed (Additional file 5: Table S3). However, no links were found among miRNA, race and and ethnicity.

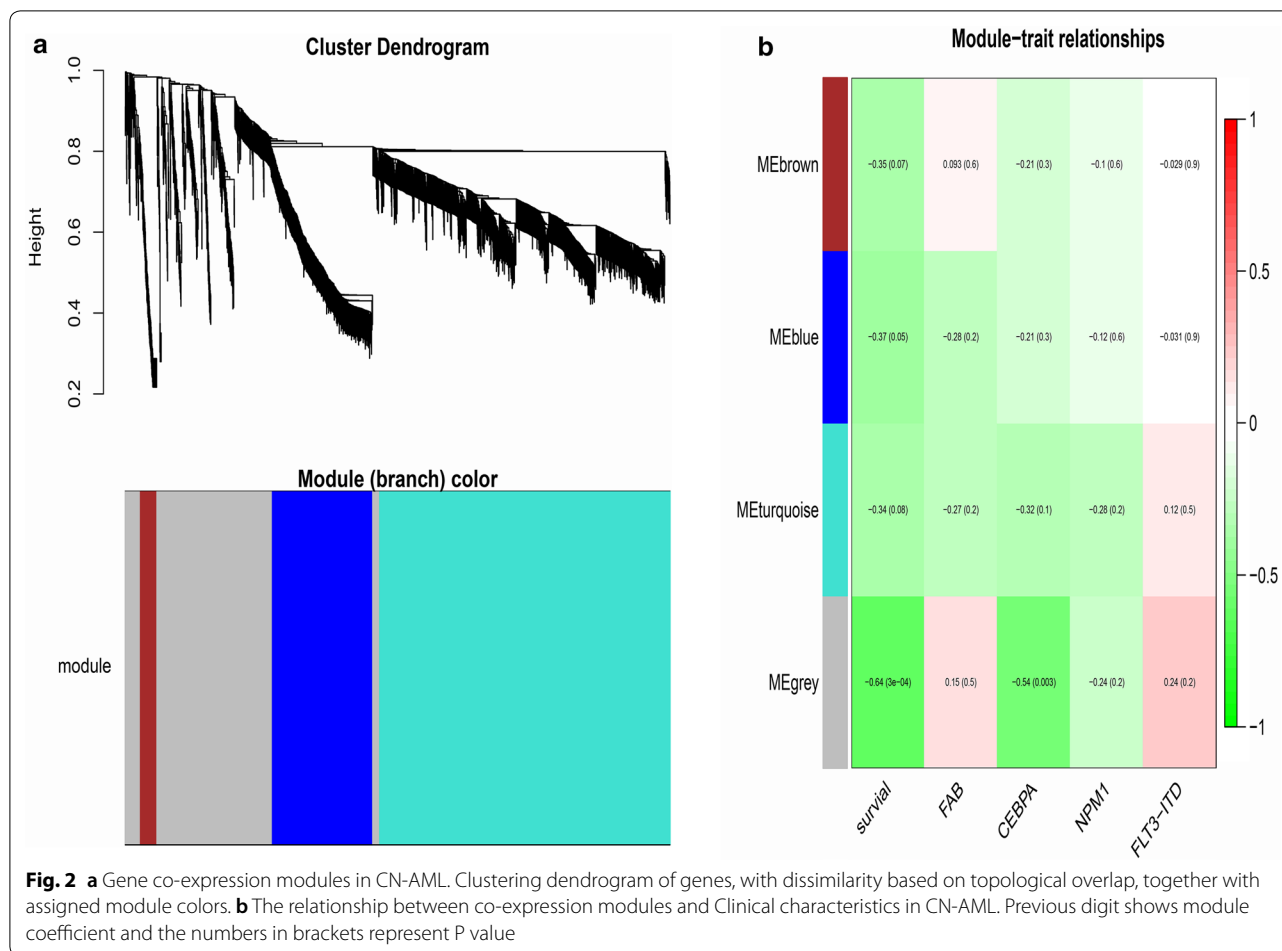
Co-expression module establishment

The 2286 survival specific mRNAs were further chosen for co-expression exploring and a weighted correlation network was constructed. Weighted gene co-expression network analysis is a systems biology method for describing the correlation patterns among genes across micro-array samples and exploring the hiding and biologic patterns Fig. 2a showed hierarchical clustering of these 2286 mRNAs and the corresponding gene co-expression module. The color bars correspond to four gene co-expression modules including blue, brown, grey and turquoise modules. The gene number in these modules is 421, 69, 577 and 1218 respectively. To further identify co-expression modules associated with disease progression, clinical characteristics of the patients including survival time, FAB category and molecular alteration associated with prognosis (FLT3-ITD, NPM1, CEBPA) were explored Fig. 2b showed the module significance. Grey module were positively associated with the cytogenetically normal AML patients' overall survival ($cor = 0.64$, $P < 0.05$) and CEBPA mutation ($cor = 0.54$, $P < 0.05$).

Functional enrichment analysis and establishment PPI network of survival specific genes

To provide more information for predicting prognosis and guiding treatment in younger CN-AML patients, mRNAs in grey module were applied to select biological pathways associated with OS using gene ontology (GO)





enrichment analysis and KEGG pathway analysis. GO analysis showed that the most significantly enriched GO terms were “transcription, DNA-templated” (ontology: BP), “integral component of membrane” (ontology: CC) and “DNA binding” (ontology: MF) (Fig. 3). Enrichment provides a measure of the significance of the function, and as the enrichment increases, the corresponding function is more specific.

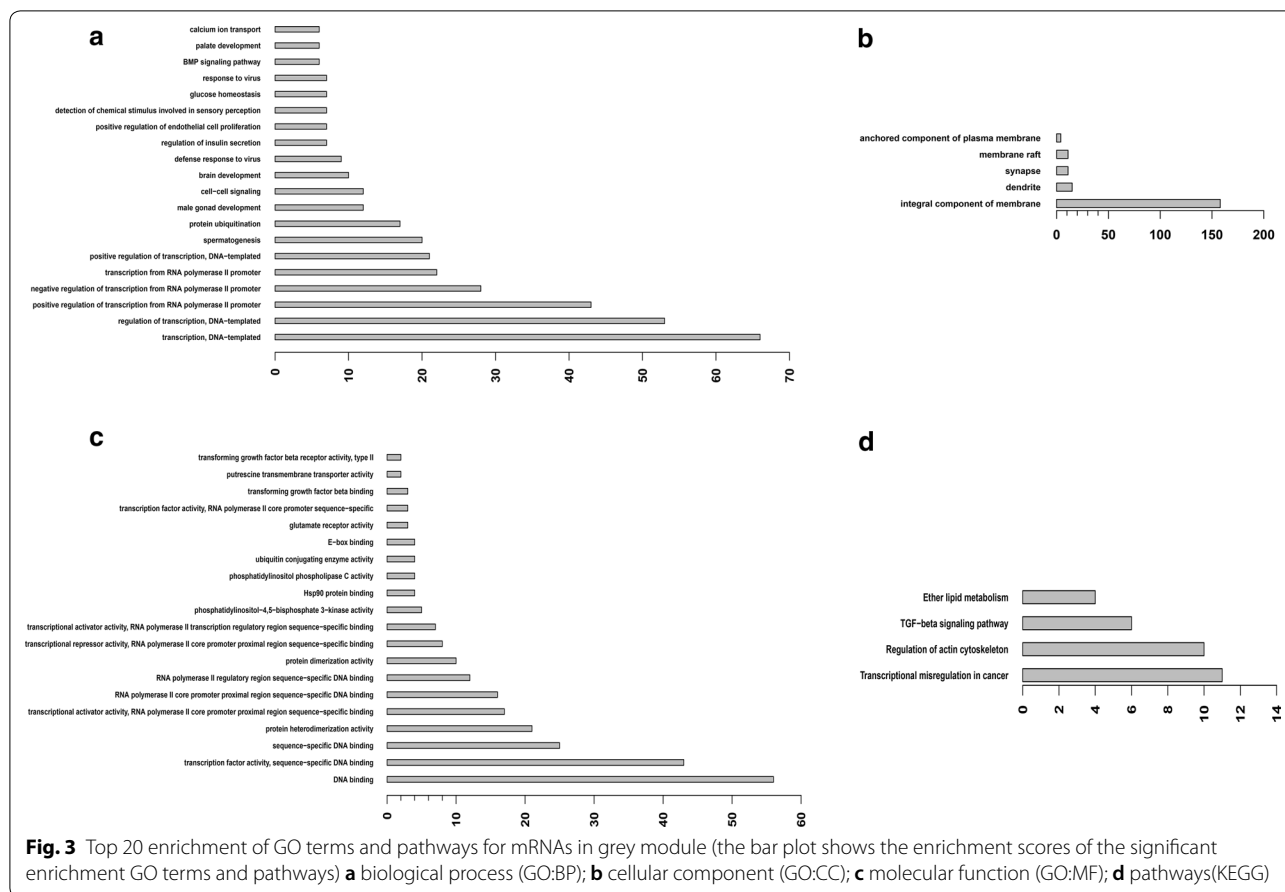
Furthermore, KEGG pathway analysis showed that 4 pathways associated with OS and the most enriched pathway was ‘Transcriptional misregulation in cancer’ (Fig. 3). Among these pathways, the ‘Transcriptional misregulation in cancer’ pathway was associated with myeloid leukemogenesis and AML cell functions [15], the ‘TGF-beta signaling pathway’ was involved in the carcinogenesis [16], the ‘Regulation of actin cytoskeleton’ has been investigated as a cause of tumor invasion and metastasis [17] and ‘Ether lipid metabolism’ was involved in the development of cancer [18].

The protein–protein interactions (PPI) network of survival specific genes in grey module was next constructed

by STRING, which was composed of 312 nodes and 601 edges (Fig. 4). Moreover, 30 genes in the PPI network were identified as hub genes in cytogenetically normal AML, including BCL2 and HIF1A genes, when “Degrees ≥ 9 ” was set as the cut-off criterion.

Construction and performance of prognostic pathway-based signatures

Next, the most enriched network ‘Transcriptional misregulation in cancer’ was carried out multivariate cox to verify whether it could accurately predict pediatric and adolescent CN-AML patients’ prognosis. After multivariate cox’s model filtering, we identified 4 mRNAs (MLLT3, CEBPA, HIST2H3C and AFF1) in ‘Transcriptional misregulation in cancer’ pathway were independently associated with prognosis. Only CEBPA were positively associated with OS, the remaining 3 mRNAs were all negatively associated with OS. The risk score was calculated through the four mRNAs status and their weight on OS, which is represented by the β coefficient in multivariate cox model. The risk score = (0.670 * status



of MLLT3) – (0.749*status of CEBPA) + (0.962*status of HIST2H3C) + (0.824*status of AFF1). The patients were then classified into 2 groups (high risk vs low risk) based on the median value of the predictor score and the Kaplan–Meier curve was applied. As shown in Fig. 5, the survival time between the 2 groups (high risk vs low risk), created based on the prognostic pathways, was significantly different ($P < 0.05$) (Fig. 5). The 3-year-Area Under Curve (AUC) of signature was 0.978 (Fig. 5). The heatmap were also created to evaluate the signature, and the results distinctly demonstrated that most dead patients were in high risk group and had worse OS (Fig. 5). Therefore, our 4 mRNA signatures (MLLT3, CEBPA, HIST2H3C and AFF1) may offer an approach for risk assessment and predicting the prognosis in pediatric and adolescent CN-AML patients.

Comparison of ‘Transcriptional misregulation in cancer’ pathway with other prognostic parameters and leave-one-out cross validation signature

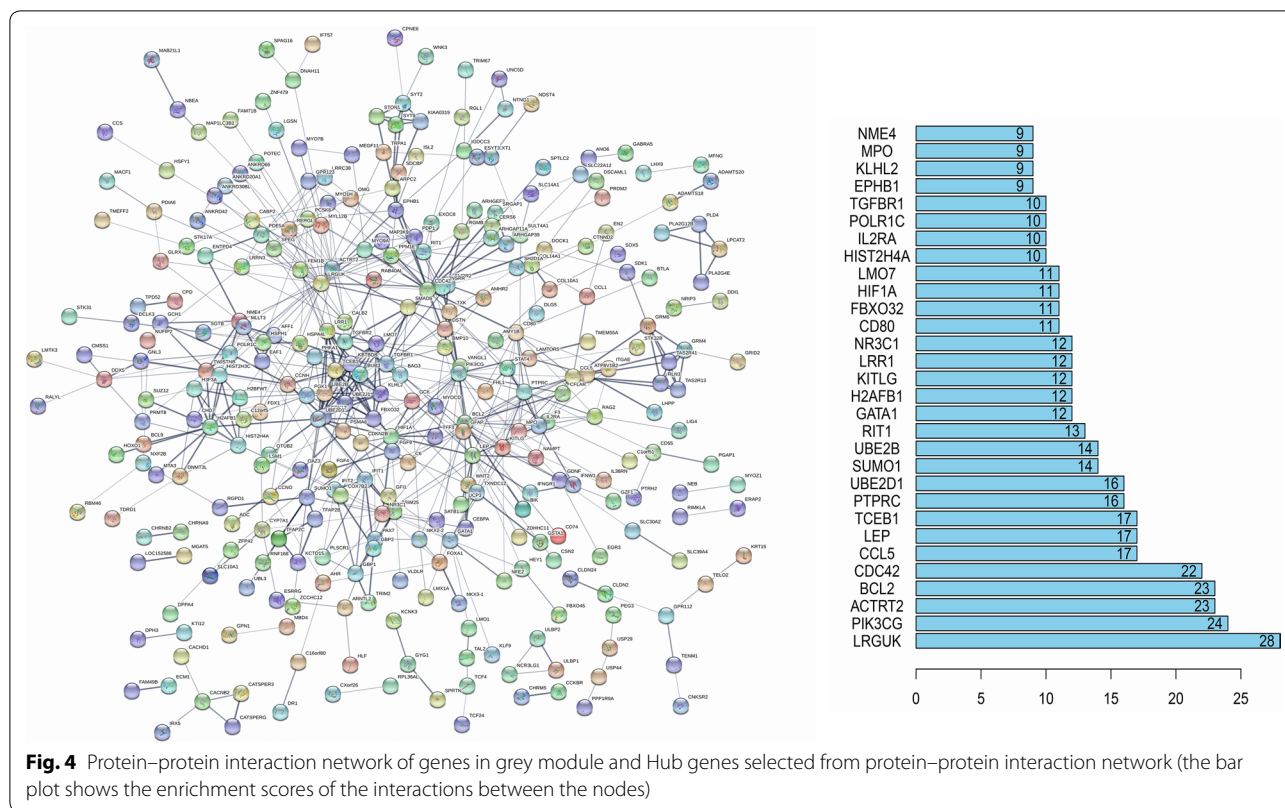
Univariate Cox analysis was used to test clinical parameters with the prognosis. The results showed gender, bone marrow blasts and peripheral blasts could not predict

the outcome, while age at diagnosis, FLT3-ITD and WT1 mutations were significantly associated with survival (Additional file 6: Table S2). After adjusting for age at diagnosis, FLT3-ITD and WT1 mutations, our prognostic signatures remained as independent prognostic factors in the multivariate analysis ($P < 0.05$).

Leave-one-out cross validation test (LOO-CV) is powerful in estimating a model’s performance. To verify whether the 4-mRNA signature could predict other pediatric and adolescent CN-AML patients’ prognosis, LOO-CV was applied. The results showed the AUC of signature was 0.604 which validated the 4-mRNA signature performed well and could successfully predict the prognosis of pediatric and adolescent CN-AML patients.

Discussion

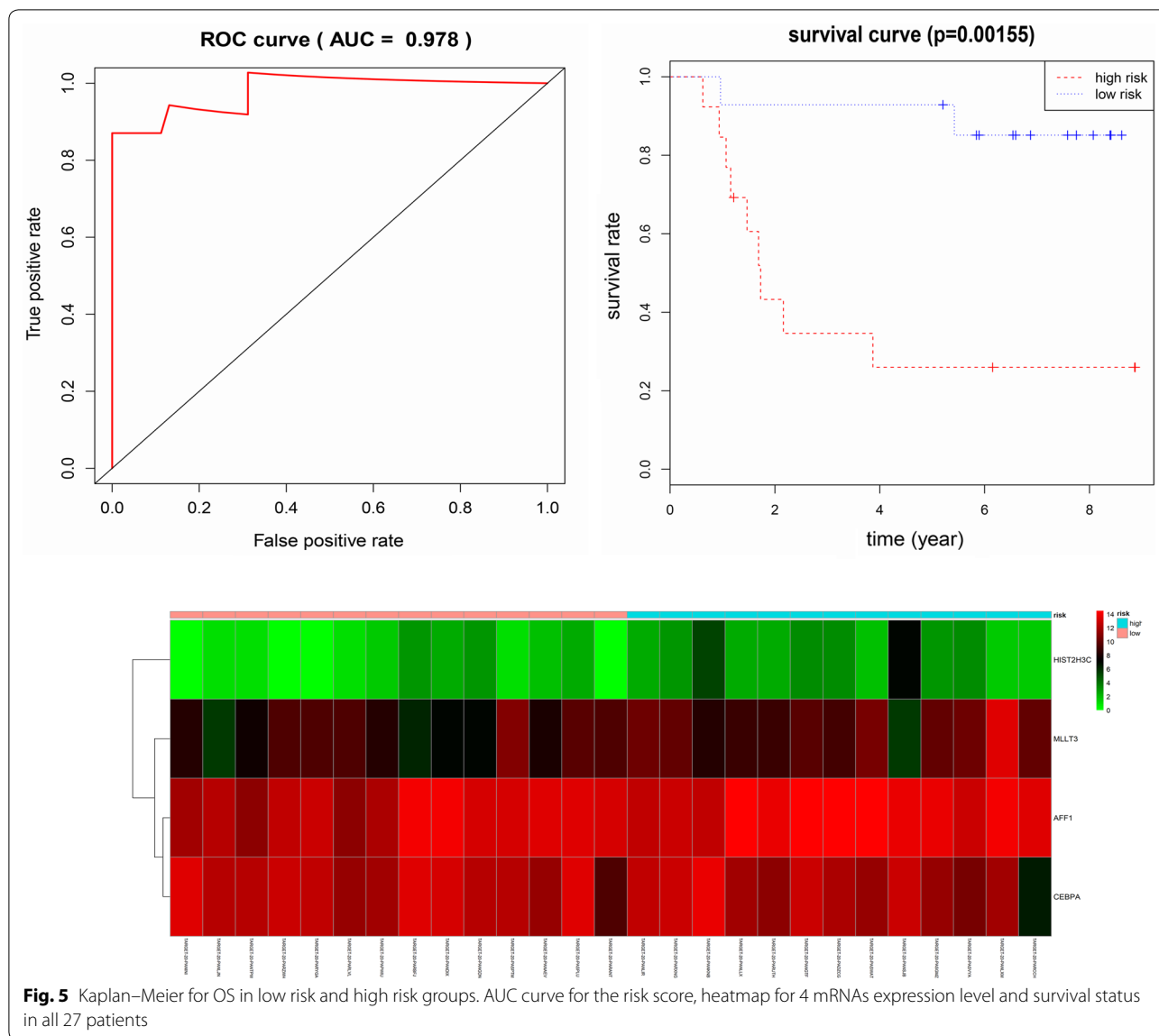
Pediatric and adolescent acute myeloid leukemia (AML) with incidence of approximately 7 occurrences per 1 million children annually is a rare type of childhood cancer [19]. Because high incidence of severe and dose-limiting short- and long-term toxicities happened, younger AML therapy was a big challenge for patients, their families, and care providers [20]. Although multiple national and



international cooperative groups have contributed to an evolving treatment strategy including four to five kinds of intensive myelosuppressive chemotherapy and stem cell transplantation over several decades, there was only a mild decline in younger AML mortality. Döhner et al. [21] has been reported State-of-the-art recommendations in adult AML. But little has been done in pediatric and adolescent AML patients. Although there were broad overlap in diagnosis, treatment and prognosis for AML, differences still exist between adolescent and adult patients [11]. Hence, to improve younger AML's prognosis, the pivotal genes and regulatory mechanism in AML's development and progress need to be identified. Studies revealed that lncRNAs play important roles in diverse gene expression and cellular processes regulation. Dysregulation of lncRNAs has been reported to contribute to oncogenesis and tumor metastasis, including AML [22, 23]. Therefore, the investigation of lncRNAs' expression and function could help us to understand leukemogenesis and identify novel therapeutic targets. However, to date, only a few studies have reported mechanical and functional characterization of AML-associated aberrant gene networks [24–27].

Recently, many studies have reported that endogenous lncRNA had miRNA responsive elements (MRE) and modulated miRNAs via acting as miRNA sponges and

binding with them [28]. These lncRNAs acting as competing endogenous RNAs (ceRNAs) participate in post-transcriptional regulation by interfering with the miRNA pathways [29]. ceRNAs have been shown to play critical roles in diverse biological functions and the disruption of the equilibrium of ceRNAs were implicated in tumorigenesis [3, 28]. For example, ATB lncRNA blocked miR-200 family by binding to its targets and upregulated ZEB1 and ZEB2, thereby inducing EMT and invasion in hepatocellular carcinoma [30]. Hence, understanding the intricate interplay among protein-coding messenger RNAs, miRNAs and lncRNA would help to identify gene regulatory networks which played critical roles in the progress and development of younger CN-AML. In the present study, we identified prognosis related specific lncRNAs, miRNAs and mRNAs of CN-AML from TCGA database. Moreover, we constructed a ceRNA network which could provide an integrated biological view based on the bioinformatics differential analysis. Some of these network have been reported to be solid tumor-associated gene-network, such as lncRNA-miRNA (CASC2-miR-221 [31]), miRNA-mRNA (miR-221-RAB1A [32], miR-25-PTEN [33], and miR-221-FMR1 [34]). The ceRNA network that we built revealed a previously unknown ceRNA regulatory network in pediatric and adolescent CN-AML. However, we failed to find



target mRNAs when we constructed a ceRNA network with lncRNAs positively associated with overall survival and miRNAs negatively associated with prognosis.

In order to identify the function of the key lncRNAs, miRNAs and mRNAs in the ceRNA network, we analyzed their associations with OS by Kaplan–Meier curve. Our results suggested that 6 lncRNAs, 1 miRNA and 14 mRNAs were associated with CN-AML overall survival. Among them, has-mir-363 has been reported to be associated with prognosis of AML [35]; lncRNA CRNDE played critical roles in promoting cell proliferation, invasion and migration of solid tumor [36]; mRNA HMBOX1 and KIF5B was involved in the carcinogenesis [37, 38]. The other lncRNAs (AL356475.1, AC011498.1, AC092811.1, LINC00158 and LINC00504) and mRNAs

(BRWD1, CNIH1, FMR1, GOLGA8J, GPBP1L1, NUFIP2, PTAR1, RRN3, SNX4, TMF1, ZFC3H1 and ZKSCAN8) were not reported previously. We analyzed the relationships of the above 35 key lncRNAs, 7 miRNAs with clinical features including age at diagnosis, gender, WBC at diagnosis, bone marrow blasts, peripheral blasts and molecular mutations (FLT3-ITD, NPM1, CEBPA). The results showed that AC011498.1, CRNDE, LINC00504 and hsa-mir-363 were the indicators of CEBPA mutation of AML. Furthermore, AC011498.1 was also associated with NPM1 mutation. LncRNA CRNDE and miRNA hsa-mir-363 have been previously reported to be associated with clinical features of leukemia [35, 39]. However, the other lncRNAs we found here have not been reported to be the indicators of relevant features

previously. The imprint of genetic ancestry and population structure carried in the genome of each individual and groups has led to the remarkable racial and ethnic diversity, which integrate biological, geopolitical, linguistic and cultural factors and are widely applied in population study [40]. Besides, Bonham et al. [41] have reported that the assessment of race and ethnicity at the individual level will make us closer to more individualized genetic-based medicine. Hence, we have estimated the association between miRNAs, race and ethnicity, but we failed to find any relation among them.

To date, the prognostic gene expression signatures have been extensively proposed for targeted therapies in cancer patients with significantly heterogeneous outcomes. However, the potential relationships of the genes have been ignored [13]. Genes were considered as independent individuals, which may result in undermining potential relationship between genes [13]. To overcome these challenges and to reduce the difficulty in biological interpretation, we constructed a weighted correlation network [14]. Then, we analyzed the relationships among the modules, OS, FAB category and the molecular mutations associated with prognosis (FLT3-ITD, NPM1, CEBPA). To identify biological pathways that were indicators of AML's prognosis, significant modules associated with OS and clinical features were selected to carry out functional enrichment analysis. The GO analysis showed that the functions of significant differences in the aspects of "transcription, DNA-templated" (ontology: BP), "integral component of membrane" (ontology: CC) and "DNA binding" (ontology: MF). Furthermore, by KEGG pathway analysis, we identified 4 biological pathways associated with OS which may predict prognosis and guide treatment in younger CN-AML patients, including 'Transcriptional misregulation in cancer, Regulation of actin cytoskeleton, TGF-beta signaling pathway and Ether lipid metabolism'. Recently, many studies have shown epigenetic abnormalities caused by 'Transcriptional misregulation in cancer' play important roles in tumor biology, such as DNA methylation, histone modifications and noncoding RNAs [42–44]. Therefore, we selected the most enriched network 'Transcriptional misregulation in cancer' to verify the relationship between epigenetic abnormalities and CN-AML patients' prognosis. The result showed that the 'Transcriptional misregulation in cancer' pathway could strongly predict younger CN-AML patients' survival. In our study, we only showed FLT3-ITD mutations could predict unfavorable outcome of pediatric CN-AML patients ($P=0.003$), but we failed to find the relationship between NPM1 and CEBPA mutation and patients' prognosis. One reason may be the sample size in our study was small and it couldn't come to a significant conclusion. Another reason may be that

patients simultaneously possess FLT3-ITD, NPM1 or CEBPA mutations in our study. Some reports have shown FLT3-ITD could implement a negative effect on OS irrespective of NPM1 or CEBPA mutations [45–47].

In PPI network analysis, we identified 30 hub genes and selected the top ten genes to analysis in detail. Among them, the PIK3CG gene, which encodes the catalytic subunit of phosphoinositide 3-OH-kinase- γ (PI3K γ) namely p110 γ , is located in chromosome band 7q22 and is often missing in myeloid malignancies [48]. Grimwade et al. [48] has implied PIK3CG was evaluated as a candidate suppressor gene of myeloid tumor. Dysregulation of apoptosis, which leads to the accumulation of tumor cells by slowing the rate of cell turnover, is a hallmark of cancer [49]. BCL-2 proteins encoded by BCL-2 gene can be classified into two families, anti-apoptotic and pro-apoptotic proteins. As a central regulator of cell polarity, loss of CDC42 suppresses AML cell polarity and division asymmetry [50]. Primary AML cells show constitutive release of a wide range of chemokines (including CCL5) involved in leucocyte chemotaxis, differentiation and angiogenesis [51]. LEP acting as a growth factor promotes cellular proliferation and may affect leukemic hematopoiesis [52]. Hence, the genes PIK3CG, BCL-2, CDC42, CCL5 and LEP in the treatment of AML are now being widely and actively taken into consideration. In previous studies, there were no reports showing the other top ten hub genes were associated with AML.

Conclusion

Taken together, we filtered survival related specific lncRNAs, miRNAs and mRNAs after analyzing by univariate cox's model from TCGA database. We successfully constructed a survival specific ceRNA network which could provide a new approach to lncRNA research in younger CN-AML. Importantly, we constructed a weighted correlation network to overcome the difficulty in biological interpretation of independently individual genes. We showed that 4 biological pathways associated with OS may be considered as potential specificity biomarkers in the diagnosis, prognosis and classification of pediatric and adolescent CN-AML. In addition, we select the most enriched network 'Transcriptional misregulation in cancer' to verify that it could accurately predict pediatric and adolescent CN-AML patients' prognosis.

Additional files

Additional file 1: Figure S1. Kaplan–Meier survival curves for 6 lncRNAs associated with overall survival from ceRNA network.

Additional file 2: Figure S2. Kaplan–Meier survival curves for 1 miRNA associated with overall survival from ceRNA network.

Additional file 3: Figure S3. Kaplan–Meier survival curves for 14 mRNAs associated with overall survival from ceRNA network.

Additional file 4: Table S1. The correlation among cancer specific lncRNAs, miRNAs and clinical features.

Additional file 5: Table S3. The correlation among cancer specific miRNAs, race and ethnicity.

Additional file 6: Table S2. Univariate Cox analysis of clinical parameters with the prognosis.

Abbreviations

AML: acute myeloid leukemia; CN-AML: cytogenetically normal acute myeloid leukemia; TCGA: the cancer genome atlas; OS: overall survival; AUC: the area under receiver operating characteristic curve; LOO-CV: leave-one-out cross validation; GO: gene ontology; DAVID: The Database for Annotation, Visualization and Integrated Discovery; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: the protein–protein interactions; CeRNA: competitive endogenous RNA; lncRNA: long noncoding RNA; miRNA: microRNA; mRNA: messenger RNA.

Authors' contributions

XY collected and analyzed the data, wrote the paper; SH research literature and edit the paper; FJF and RZ revise the paper; CYS and YH conceived and designed this study, analyzed the data, wrote the paper; and all authors reviewed the paper, and approved the final manuscript. All authors read and approved the final manuscript.

Author details

¹ Institute of Hematology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China. ² Collaborative Innovation Center of Hematology, Huazhong University of Science and Technology, Wuhan 430022, China.

Acknowledgements

We would like to thank the researchers and study participants for their contributions.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Please contact author for data requests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (Research Ethics Committees of Wuhan Union Hospital, Tongji Medical College, Huazhong University of Science and Technology) and with the Helsinki Declaration of 1964 and later versions.

Funding

This work was supported by grants of the National Natural Science Foundation of China (No. 81670197 for Chunyan Sun); and the Integrated Innovative Team for Major Human Diseases Program of Tongji Medical College, HUST; and the Clinical Research Physician Program of Tongji Medical College, HUST.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 27 April 2018 Accepted: 16 August 2018

Published online: 28 August 2018

References

- Kunej T, Obsteter J, Pogacar Z, Horvat S, Calin GA. The decalog of long non-coding RNA involvement in cancer diagnosis and monitoring. *Crit Rev Clin Lab Sci*. 2014;51(6):344–57. <https://doi.org/10.3109/10408363.2014.944299>.
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–8. <https://doi.org/10.1016/j.cell.2011.07.014>.
- Song X, Cao G, Jing L, et al. Analysing the relationship between lncRNA and protein-coding gene and the role of lncRNA as ceRNA in pulmonary fibrosis. *J Cell Mol Med*. 2014;18(6):991–1003. <https://doi.org/10.1111/jcmm.12243>.
- Pan W, Liu L, Wei J, et al. A functional lncRNA HOTAIR genetic variant contributes to gastric cancer susceptibility. *Mol Carcinog*. 2016;55(1):90–6. <https://doi.org/10.1002/mc.22261>.
- Huang C, Cao L, Qiu L, et al. Upregulation of H19 promotes invasion and induces epithelial-to-mesenchymal transition in esophageal cancer. *Oncol Lett*. 2015;10(1):291–6. <https://doi.org/10.3892/ol.2015.3165>.
- Loewen G, Jayawickramarajah J, Zhuo Y, Shan B. Functions of lncRNA HOTAIR in lung cancer. *J Hematol Oncol*. 2014;7:90. <https://doi.org/10.1186/s13045-014-0090-4>.
- Junge A, Zandi R, Havgaard JH, Gorodkin J, Cowland JB. Assessing the miRNA sponge potential of RUNX1T1 in t (8, 21) acute myeloid leukemia. *Gene*. 2017;615:35–40. <https://doi.org/10.1016/j.gene.2017.03.015>.
- Ding Y, Wang ZC, Zheng Y, et al. C-Myc functions as a competing endogenous RNA in acute promyelocytic leukemia. *Oncotarget*. 2016;7(35):56422–30. <https://doi.org/10.18632/oncotarget.10896>.
- Chen L, Wang W, Cao L, Li Z, Wang X. Long non-coding RNA CCAT1 acts as a competing endogenous RNA to regulate cell growth and differentiation in acute myeloid leukemia. *Mol Cells*. 2016;39(4):330–6. <https://doi.org/10.14348/molcells.2016.2308>.
- Mrozek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. *Blood Rev*. 2004;18(2):115–36. [https://doi.org/10.1016/s0268-960x\(03\)00040-7](https://doi.org/10.1016/s0268-960x(03)00040-7).
- Creutzig U, van den Heuvel-Eibrink MM, Gibson B, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood*. 2012;120(16):3187–205. <https://doi.org/10.1182/blood-2012-03-362608>.
- Moreaux J, Klein B, Bataille R, et al. A high-risk signature for patients with multiple myeloma established from the molecular classification of human myeloma cell lines. *Haematologica*. 2011;96(4):574–82. <https://doi.org/10.3324/haematol.2010.033456>.
- 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):15545–50. <https://doi.org/10.1073/pnas.0506580102>.
- Ivliev AE, t Hoen PA, Borisevich D, Nikolsky Y, Sergeeva MG. Drug repositioning through systematic Mining of Gene Coexpression Networks in cancer. *PLoS ONE*. 2016;11(11):e0165059. <https://doi.org/10.1371/journal.pone.0165059>.
- Jasielec J, Saloura V, Godley LA. The mechanistic role of DNA methylation in myeloid leukemogenesis. *Leukemia*. 2014;28(9):1765–73. <https://doi.org/10.1038/leu.2014.163>.
- Matveeva A, Kovalevska L, Kholodnyuk I, Ivanivskaya T, Kashuba E. The TGF-beta—SMAD pathway is inactivated in chronic lymphocytic leukemia cells. *Exp Oncol*. 2017;39(4):286–90.
- Peng JM, Bera R, Chiou CY, et al. Actin cytoskeleton remodeling drives epithelial-mesenchymal transition for hepatoma invasion and metastasis. *Hepatology*. 2017. <https://doi.org/10.1002/hep.29678>.
- Mollinedo F, de la Iglesia-Vicente J, Gajate C, et al. In vitro and in vivo selective antitumor activity of Edelfosine against mantle cell lymphoma and chronic lymphocytic leukemia involving lipid rafts. *Clin Cancer Res*. 2010;16(7):2046–54. <https://doi.org/10.1158/1078-0432.ccr-09-2456>.
- Zwaan CM, Kolb EA, Reinhardt D, et al. Collaborative efforts driving progress in pediatric acute myeloid leukemia. *J Clin Oncol*. 2015;33(27):2949–62. <https://doi.org/10.1200/jco.2015.62.8289>.
- Hasle H, Niemeyer CM, Chessells JM, et al. A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. *Leukemia*. 2003;17(2):277–82. <https://doi.org/10.1038/sj.leu.2402765>.

21. Dohner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010;115(3):453–74. <https://doi.org/10.1182/blood-2009-07-235358>.
22. Khandelwal A, Bacolla A, Vasquez KM, Jain A. Long non-coding RNA: a new paradigm for lung cancer. *Mol Carcinog*. 2015;54(11):1235–51. <https://doi.org/10.1002/mc.22362>.
23. Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: a new player in cancer. *J Hematol Oncol*. 2013;6:37. <https://doi.org/10.1186/1756-8722-6-37>.
24. Hughes JM, Legnini I, Salvatori B, et al. C/EBPalpha-p30 protein induces expression of the oncogenic long non-coding RNA UCA1 in acute myeloid leukemia. *Oncotarget*. 2015;6(21):18534–44. <https://doi.org/10.18632/oncotarget.4069>.
25. Wang Y, Zhou Q, Ma JJ. High expression of lnc-CRNDE presents as a biomarker for acute myeloid leukemia and promotes the malignant progression in acute myeloid leukemia cell line U937. *Eur Rev Med Pharmacol Sci*. 2018;22(3):763–70. https://doi.org/10.26355/eurev_201802_14310.
26. Fernando TR, Contreras JR, Zampini M, et al. The lncRNA CASC15 regulates SOX4 expression in RUNX1-rearranged acute leukemia. *Mol Cancer*. 2017;16(1):126. <https://doi.org/10.1186/s12943-017-0692-x>.
27. Cheng H, Huang CM, Wang Y, et al. Microarray profiling and co-expression network analysis of the lncRNAs and mRNAs associated with acute leukemia in adults. *Mol Biosyst*. 2017;13(6):1102–8. <https://doi.org/10.1039/c6mb00874g>.
28. Johnsson P, Ackley A, Vidarsdottir L, et al. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat Struct Mol Biol*. 2013;20(4):440–6. <https://doi.org/10.1038/nsmb.2516>.
29. Sen R, Ghosal S, Das S, Balti S, Chakrabarti J. Competing endogenous RNA: the key to posttranscriptional regulation. *Sci World J*. 2014;2014:896206. <https://doi.org/10.1155/2014/896206>.
30. Yuan JH, Yang F, Wang F, et al. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell*. 2014;25(5):666–81. <https://doi.org/10.1016/j.ccr.2014.03.010>.
31. Jin X, Cai L, Wang C, et al. CASC2/miR-24/miR-221 modulates the TRAIL resistance of hepatocellular carcinoma cell through caspase-8/caspase-3. *Cell Death Dis*. 2018;9(3):318. <https://doi.org/10.1038/s41419-018-0350-2>.
32. Sun T, Wang X, He HH, et al. miR-221 promotes the development of androgen independence in prostate cancer cells via downregulation of HECTD2 and RAB1A. *Oncogene*. 2014;33(21):2790–800. <https://doi.org/10.1038/ncr.2013.230>.
33. Yao Y, Sun F, Lei M. miRNA-25 inhibits sepsis-induced cardiomyocyte apoptosis by targeting PTEN. *Biosci Rep*. 2018. <https://doi.org/10.1042/bsr20171511>.
34. Zongaro S, Hukema R, D'Antoni S, et al. The 3' UTR of FMR1 mRNA is a target of miR-101, miR-129-5p and miR-221: implications for the molecular pathology of FXTAS at the synapse. *Hum Mol Genet*. 2013;22(10):1971–82. <https://doi.org/10.1093/hmg/ddt044>.
35. Mosakhani N, Raty R, Tyybakinoja A, Karjalainen-Lindsberg ML, Elonen E, Knuutila S. MicroRNA profiling in chemoresistant and chemosensitive acute myeloid leukemia. *Cytogenet Genome Res*. 2013;141(4):272–6. <https://doi.org/10.1159/000351219>.
36. Sun H, He L, Ma L, et al. lncRNA CRNDE promotes cell proliferation, invasion and migration by competitively binding miR-384 in papillary thyroid cancer. *Oncotarget*. 2017;8(66):110552–65. <https://doi.org/10.18632/oncotarget.22819>.
37. Gong J, Liu R, Zhuang R, et al. miR-30c-1* promotes natural killer cell cytotoxicity against human hepatoma cells by targeting the transcription factor HMBX1. *Cancer Sci*. 2012;103(4):645–52. <https://doi.org/10.1111/1/j.1349-7006.2012.02207.x>.
38. Viola D, Giani C, Mazzeo S, et al. KIF5B/RET rearrangement in a carcinoma of the thyroid gland: a case report of a fatal disease. *J Clin Endocrinol Metab*. 2017;102(9):3091–6. <https://doi.org/10.1210/jc.2017-00304>.
39. Subhash S, Andersson PO, Kosalai ST, Kanduri C, Kanduri M. Global DNA methylation profiling reveals new insights into epigenetically deregulated protein coding and long noncoding RNAs in CLL. *Clin Epigenet*. 2016;8:106. <https://doi.org/10.1186/s13148-016-0274-6>.
40. Sankar P, Cho MK. Genetics. Toward a new vocabulary of human genetic variation. *Science*. 2002;298(5597):1337–8. <https://doi.org/10.1126/science.1074447>.
41. Bonham VL, Knerr S. Social and ethical implications of genomics, race, ethnicity, and health inequities. *Semin Oncol Nurs*. 2008;24(4):254–61. <https://doi.org/10.1016/j.soncn.2008.08.005>.
42. Sharma A, Heuck CJ, Fazzari MJ, et al. DNA methylation alterations in multiple myeloma as a model for epigenetic changes in cancer. *Wiley Interdiscip Rev Syst Biol Med*. 2010;2(6):654–69. <https://doi.org/10.1002/wsbm.89>.
43. Dimopoulos K, Gimsing P, Gronbaek K. Aberrant microRNA expression in multiple myeloma. *Eur J Haematol*. 2013;91(2):95–105. <https://doi.org/10.1111/ejh.12124>.
44. Mithraprabhu S, Kalff A, Chow A, Khong T, Spencer A. Dysregulated Class I histone deacetylases are indicators of poor prognosis in multiple myeloma. *Epigenetics*. 2014;9(11):1511–20. <https://doi.org/10.4161/1559294.2014.983367>.
45. Taylor E, Morris K, Ellis M, et al. FLT3-ITD positive acute myeloid leukemia: a retrospective analysis of the role of allogeneic transplant and allelic ratio in patient management. *Asia Pac J Clin Oncol*. 2018. <https://doi.org/10.1111/ajco.12827>.
46. Weisberg E, Sattler M, Manley PW, Griffin JD. Spotlight on midostaurin in the treatment of FLT3-mutated acute myeloid leukemia and systemic mastocytosis: design, development, and potential place in therapy. *Oncotargets Ther*. 2018;11:175–82. <https://doi.org/10.2147/ott.s127679>.
47. Shimada A, Iijima-Yamashita Y, Tawa A, et al. Risk-stratified therapy for children with FLT3-ITD-positive acute myeloid leukemia: results from the JPLSG AML-05 study. *Int J Hematol*. 2018. <https://doi.org/10.1007/s12185-017-2395-x>.
48. Grimwade D, Walker H, Oliver F, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood*. 1998;92(7):2322–33.
49. Reed JC. Apoptosis-based therapies. *Nat Rev Drug Discov*. 2002;1(2):111–21. <https://doi.org/10.1038/nrd726>.
50. Mizukawa B, O'Brien E, Moreira DC, et al. The cell polarity determinant CDC42 controls division symmetry to block leukemia cell differentiation. *Blood*. 2017;130(11):1336–46. <https://doi.org/10.1182/blood-2016-12-758458>.
51. Bruserud O, Rynningen A, Olsnes AM, et al. Subclassification of patients with acute myelogenous leukemia based on chemokine responsiveness and constitutive chemokine release by their leukemic cells. *Haematologica*. 2007;92(3):332–41.
52. Yilmaz M, Kis C, Ceylan NO, et al. Serum leptin level in acute myeloid leukemia patients. *Hematology*. 2008;13(1):21–3. <https://doi.org/10.1179/102453308x315771>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

