Research Article

Identification of the 7-lncRNA Signature as a Prognostic Biomarker for Acute Myeloid Leukemia

Chun-yu Liu,[1] Hui-hui Guo,[2] Hai-xia Li,[1] Ying Liang,[1] Cong Tang,[1] and Nan-nan Chen [1]

[1]Department of Hematology, The No. 967 Hospital of PLA Joint Logistics Support Force, Dalian, Liaoning, China
[2]Department of Obstetrics and Gynecology, The No.967 Hospital of PLA Joint Logistics Support Force, Dalian, Liaoning, China

Correspondence should be addressed to Nan-nan Chen; chnngax@hotmail.com

A lot of evidence has emphasized the function of long noncoding RNAs (lncRNAs) in tumors' development and progression. Nevertheless, there is still a lack of lncRNA biomarkers that can predict the prognosis of acute myeloid leukemia (AML). Our goal was to develop a lncRNA marker with prognostic value for the survival of AML. AML patients' RNA sequencing data as well as clinical characteristics were obtained from the public TARGET database. Then, differentially expressed lncRNAs were identified in female and male AML samples. By adopting univariate and multivariate Cox regression analyses, AML patients' survival was predicted by a seven-lncRNA signature. It was found that 95 abnormal expressed lncRNAs existed in AML. Then, the analysis of multivariate Cox regression showed that, among them, 7 (LINC00461, RP11-309M23.1, AC016735.2, RP11-61I13.3, KIAA0087, RORB-AS1, and AC012354.6) had an obvious prognostic value, and according to their cumulative risk scores, these 7 lncRNA signatures could independently predict the AML patients' overall survival. Overall, the prognosis of AML patients could be predicted by a reliable tool, that is, seven-lncRNA prognostic signature.

1. Introduction

As a malignant and aggressive disease, acute myeloid leukemia (AML) has a lot of advantages, such as abnormal expansion of myeloid blasts, and it appears especially in the elderly [1, 2]. The cure rate of AML patients with an age of less than 60 is 20-35% higher than that of those over 60 years old [3]. Although AML patients have been treated with potential therapies (such as intensive chemotherapy and allogeneic stem cell transplantation), only a few are eligible for these treatments (because of treatment intolerance and lack of matched donors); for AML patients who receive transplants, more than 50% end up relapsing [4, 5]. Therefore, it is of great significance to explore effective biomarkers for the early detection and improvement of the prognosis of AML.

Nowadays, long noncoding RNAs (lncRNAs) have been paid more and more attention to in cancer research [6]. Meanwhile, lncRNAs are a class of nonprotein coding transcripts with a length of generally greater than 200 nucleotides [7]. In recent years, some researchers have proved that lncRNAs are indispensable through comprehensive mechanisms in a number of biological events, including cell differentiation, cell cycle, and apoptosis [8, 9]. Especially in the field of tumors, a larger number of researches have confirmed that lncRNAs are involved in tumor genesis and metastasis by multiple mechanisms, including sponging miRNAs, epigenetic regulation, translation regulation, cell differentiation regulation, and therapy resistance [10, 11]. It is noteworthy that these lncRNAs can be biomarkers as well as therapeutic targets for tumors. In this study, a seven-lncRNA model was established by using the Cox regression method, aiming to independently access prognosis and precisely predict survival probability in AML patients from the TARGET database.

2. Materials and Methods

2.1. Data Sources. The RNAseq data as well as their respective clinical follow-up material were downloaded from the

public TARGET database [12], including 186 male AML samples and 172 female AML samples.

2.2. Differentially Expressed lncRNA Extraction and Analysis. The differentially expressed lncRNAs were estimated by using R statistical software package "limma." |log (FC)| ≥ 0.5 and p value < 0.05 were used as cutoffs to screen dysregulated lncRNAs. Then the "ggpubr" as well as "pheatmap" packages produced boxplots and heat maps, respectively.

2.3. Risk Signature Construction and Validation. The definition of risk signature scores is Risk score = expression of LINC 00461 × 0.097559 + expression of RP11 − 309 M23.1 × 0.087081 + expression of AC016735.2 × 0.107281 + expression of RP11 − 61I13.3 × 0.100018 + expression of KIAA0087 × 0.072299 + expression of RORB − AS1 × 0.081708 + expression of AC012354.6 × −0.24187. Then, the score was calculated in overall cohort, and the sample was divided into groups with high risk and low risk by using the median risk score. Through assessing the area under the curve (AUC), the "timeROC" R package was adopted to verify this immune-related risk signal's prognostic utility. Besides, the overall survival (OS) of patients with high and low risk was contrasted by the Kaplan-Meier curves as well as the "survival" R package [13].

2.4. Survival Analysis. lncRNAs whose expressions were correlated with AML patients' overall survival were recognized by multivariate Cox regression analysis. Meanwhile, genes whose p values are smaller than 0.05 were taken as candidate ones that are correlated with patients' survival. Moreover, the general survival was explored by the Kaplan-Meier survival curve, so as to evaluate the differences in patients' survival. More importantly, the receiver operating characteristic (ROC) curve (AUC) was used to prove the specificity for diagnostic accuracy, and then the cBioPortal tool was adopted to determine the alterations of the survival-associated lncRNAs in AML [14].

2.5. Statistical Analysis. According to the critical conditions of |log (FC)| ≥0.5 as well as an adjusted p value < 0.05, DELs were recognized by adopting the "edgeR" package [15] in the R software. Then, univariate and multivariate Cox regression analyses were implemented in order to recognize the lncRNAs with a prognostic value. The "timeROC" R package [16] was adopted to calculate the zone under the ROC curve (AUC), and Kaplan-Meier curves as well as log-rank tests "survival" R package were used to analyze patients' survival outcomes. p < 0.05 was considered to have a statistical significance, and all tests were two-sided. We used R (v.3.6.1, R Core Team, Boston, MA, USA) for the above data analysis.

3. Results

3.1. The Identification of Differentially Expressed lncRNAs in AML. For the identification of dysregulated lncRNAs in AML patients, we downloaded microarray data from TARGET database and performed "R studio" software for statistical analysis. We found 95 abnormal expressed lncRNAs in

AML, which were shown using volcano plot and heat map (Figures 1(a) and 1(b)).

3.2. The Identification of Prognostic lncRNAs in AML Patients. Next, a multivariate Cox regression analysis was conducted to recognize the differentially expressed lncRNAs that were significantly related the AML patients' overall survival. It was found that a total of 7 of these differentially expressed lncRNAs were greatly related to the survival of AML patients (Figure 2).

3.3. Construction and Validation of Prognostic lncRNA Signals. The following prognostic model score was constructed by using these IRGs as well as their corresponding regression coefficient values: Risk score = the expression of LINC00461 × 0.097559 + that of RP11 − 309 M23.1 × 0.087081 + that of AC016735.2 × 0.107281 + that of RP11 − 61I13.3 × 0.100018 + that of KIAA0087 × 0.072299 + that of RORB − AS1 × 0.081708 + that of AC012354.6 × −0.24187. Six prognostic lncRNAs were associated with elevated risk (LINC00461, RP11-309M23.1, AC016735.2, RP11-61I13.3, KIAA0087, and RORB-AS1; Coef > 0), and one is a protective gene that was correlated with decreased risk (AC012354.6; Coef < 0) (Table 1). Then, the risk score method was adopted to provide scores for all samples, and the median risk score values were applied to divide patients into two groups, including low-risk (n = 148) as well as high-risk (n = 147) patients. Later, this risk scoring method is used to score all samples. The general survival of the low-risk group was obviously longer than that of the high-risk group (p = 5.406e − 07, Figure 3(a)), with a 3-year AUC value of 0.721 (Figure 3(b)). Besides, patients who had continuous risk scores had different clinical outcomes in different groups (Figures 4(a) and 4(b)). The expressing pattern of LINC00461, RP11-309M23.1, AC016735.2, RP11-61I13.3, KIAA0087, RORB-AS1, and AC012354.6 was shown in Figure 4(c).

4. Discussion

In the past few years, many nonprotein-coding transcripts of the genome, including lncRNAs, have been treated as irrelevant transcriptional junk [17]. Due to the success of ENCODE as well as the implementation of the Cancer Genome Atlas (TCGA) project, lncRNAs have attracted attention because of their important function in cancer occurrence and development [18, 19]. lncRNAs participate in a series of basic biological processes, such as cell cycle regulation, apoptosis, and DNA damage response, and their roles in some human diseases have been reported more and more [20, 21]. Recently, a number of researches have indicated that altered lncRNA expression levels were correlated with disease progression, but their prognostic value based on multiple models has rarely been studied [22–24].

In this study, we identified 95 dysregulated lncRNAs between 186 male AML patients and 172 female AML patients. Using multivariate Cox regression analysis, we identified 7 prognosis-related lncRNAs, including LINC00461, RP11-309M23.1, AC016735.2, RP11-61I13.3,
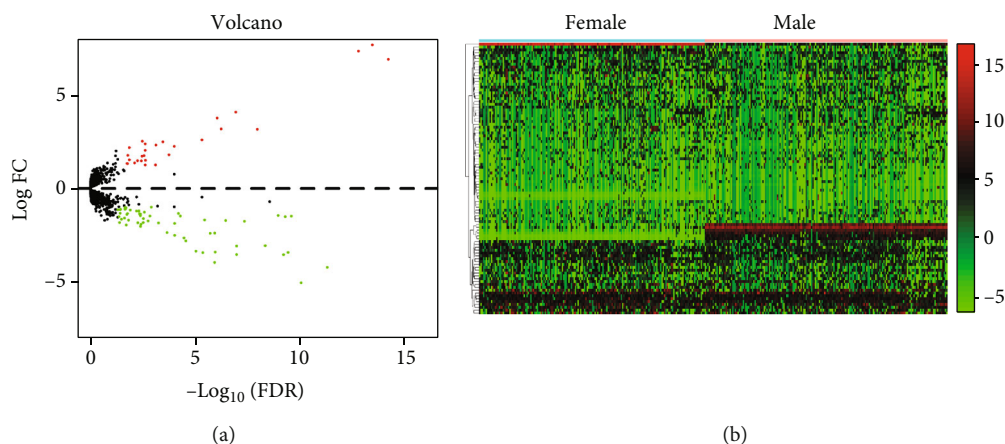
(a)

(b)

FIGURE 1: The expressing pattern of lncRNAs in AML. (a) Volcano plot of the differentially expressed lncRNAs in AML based on TARGET database. (b) Heat map of the differentially expressed lncRNAs in AML.
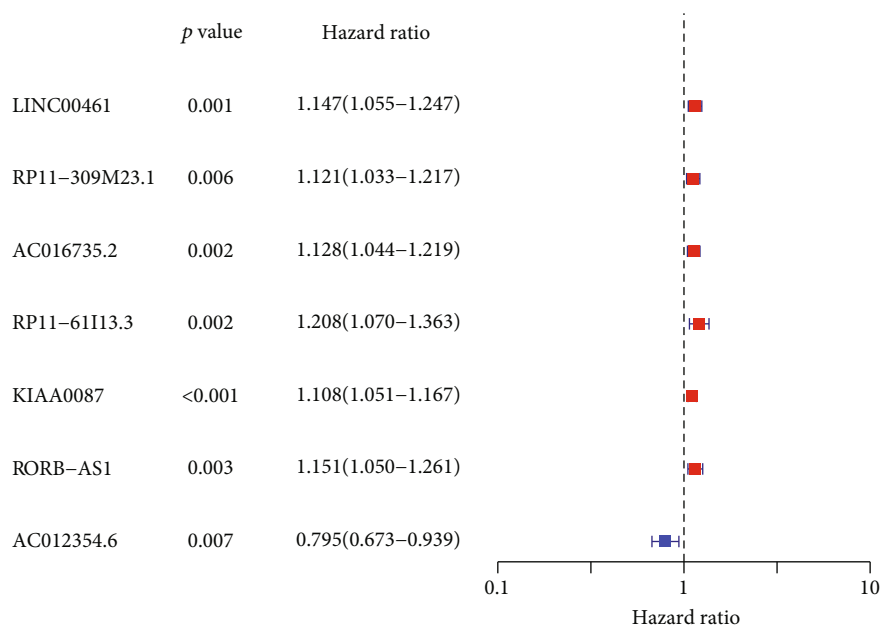

FIGURE 2: Multivariate Cox regression analysis confirmed the prognostic lncRNAs in AML.

TABLE 1: Seven lncRNAs associated with the survivals of AML patients.

id	coef	exp (coef)	se (coef)	z	Pr (>|z|)
LINC00461	0.097559	1.102477	0.045227	2.157084	0.030999
RP11-309M23.1	0.087081	1.090985	0.04358	1.998193	0.045696
AC016735.2	0.107281	1.113247	0.039321	2.728368	0.006365
RP11-61I13.3	0.100018	1.105191	0.06338	1.578059	0.114552
KIAA0087	0.072299	1.074976	0.028702	2.518928	0.011771
RORB-AS1	0.081708	1.085139	0.04816	1.696594	0.089774
AC012354.6	-0.24187	0.785156	0.088975	-2.71844	0.006559

KIAA0087, RORB-AS1, and AC012354.6. In recent years, several studies have reported the tumor-related function of LINC00461 in various tumors. For instance, LINC00461 was shown to strengthen the colorectal cancer cells' prolifer-ation and invasion ability by miRNA-323b-3p/NFIB Axis [25]. LINC00461 was reported to be highly expressed in lung tumor, and its knockdown suppressed the tumor cells' pro-liferation and metastasis via miR-4478/E2F1 [26]. However,
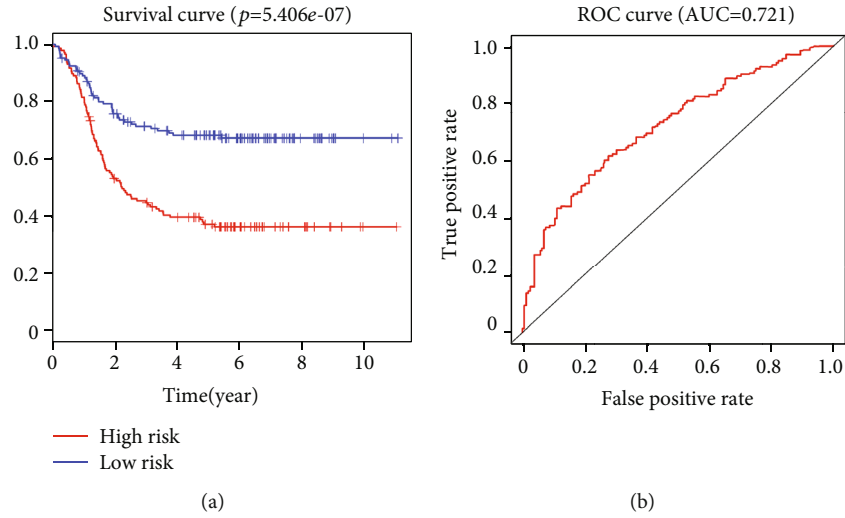
(a)

(b)

FIGURE 3: The construction of a prognostic risk signature related to outcomes of AML patients. (a) Low and high-risk AML patients' overall survivals were determined by the use of Kaplan-Meier curves. (b) Time-dependent ROC assays of the identified 7-lncRNA risk signature in AML patients.
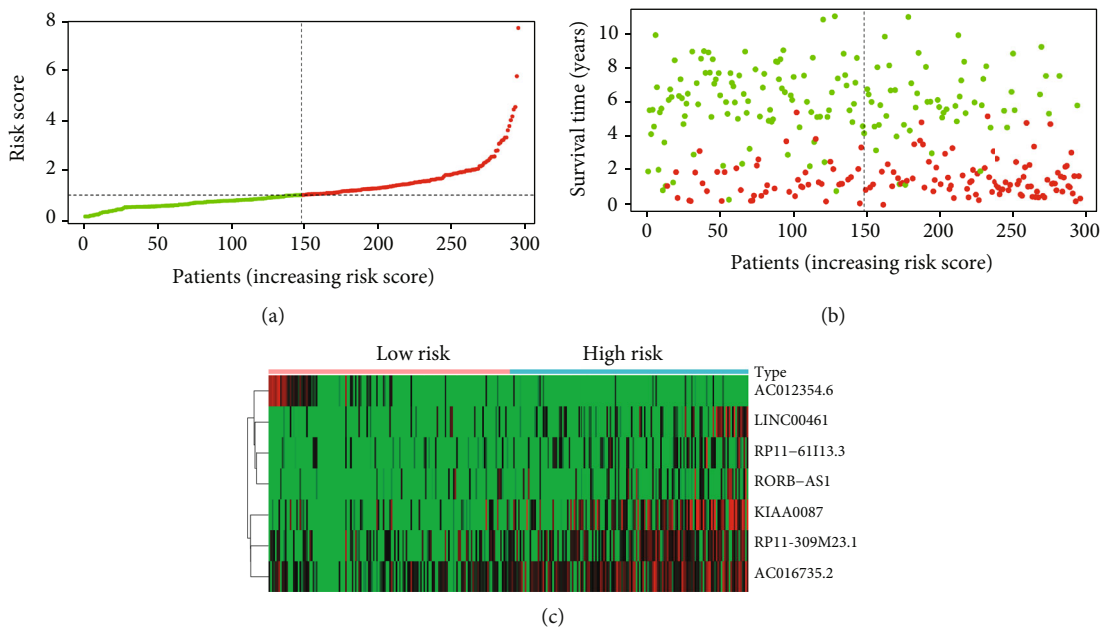

(a)

(b)

(c)

FIGURE 4: The characteristic of 7-lncRNA risk signature. (a) and (b) The distribution of risk scores, patient survival time, and status for overall survival. (c) The heat map of the seven prognostic lncRNAs in different risk groups.

the expression and function of RP11-309 M23.1, AC016735.2, RP11-61I13.3, KIAA0087, RORB-AS1, and AC012354.6 were rarely reported. More experiments were needed to confirm their functions in various types of tumors.

In recent years, many studies have reported that lncRNAs have some potentials to be used as novel biomarkers for AML patients [27, 28]. For instance, it was reported that lncRNA ANRIL could be used to regulate the development of AML by regulating the AdipoR1/AMPK/SIRT1's glucose metabolism pathway [29]. Tao and his group showed that downregulating CD27-AS1 could inhibit proliferation and induce apoptosis of AML cells through the miR-

224-5p/PBX3 [30]. In recent years, several prognostic models based on lncRNAs were developed in different types of tumors [31, 32]. However, at present, there are no reliable models based on critical lncRNAs that can predict prognosis in AML patients. In this study, the following prognostic model score was established by adopting the above 7 prognosis-related lncRNAs as well as their corresponding regression coefficient values, and it was found that high-risk patients displayed a generally shorter survival than low-risk ones, which suggested that this innovative lncRNA expression signature might be a powerful biomarker for the AML patients' prognosis.

However, there are several limitations to the present analysis. First, this study included 358 AML patients from TARGET database, which had a limited sample size. Second, our results' reliability is restrained because of a lack of in vitro or in vivo experiments. Third, as this research is retrospective, so a prospective study is needed to verify the findings of this study.

5. Conclusion

A 7-lncRNA risk signature correlated with AML prognosis is successfully constructed in the TARGET database. Our findings revealed that the signature is a potent predictive indicator for patients with AML.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

Chun-yu Liu and Hui-hui Guo contributed equally to this paper, and they are co-first authors. Chun-yu Liu and Hui-hui Guo contributed equally to this work.

Acknowledgments

References

[1] S. F. Cai and R. L. Levine, "Genetic and epigenetic determinants of AML pathogenesis," *Seminars in Hematology*, vol. 56, no. 2, pp. 84–89, 2019.

[2] A. Renneville, C. Roumier, V. Biggio et al., "Cooperating gene mutations in acute myeloid leukemia: a review of the literature," *Leukemia*, vol. 22, no. 5, pp. 915–931, 2008.

[3] S. A. Patel and J. M. Gerber, "A user's guide to novel therapies for acute myeloid leukemia," *Clinical Lymphoma, Myeloma & Leukemia*, vol. 20, no. 5, pp. 277–288, 2020.

[4] M. Yilmaz, T. Kadia, and F. Ravandi, "Identifying effective drug combinations for patients with acute myeloid leukemia," *Expert Review of Anticancer Therapy*, vol. 20, no. 7, pp. 591–601, 2020.

[5] M. Khan, A. E. R. Mansoor, and T. M. Kadia, "Future prospects of therapeutic clinical trials in acute myeloid leukemia," *Future Oncology*, vol. 13, no. 6, pp. 523–535, 2017.

[6] F. Ferrè, A. Colantoni, and M. Helmer-Citterich, "Revealing protein-lncRNA interaction," *Briefings in Bioinformatics*, vol. 17, no. 1, pp. 106–116, 2016.

[7] F. Kopp and J. T. Mendell, "Functional classification and experimental dissection of long noncoding RNAs," *Cell*, vol. 172, no. 3, pp. 393–407, 2018.

[8] Y. Ma, J. Zhang, L. Wen, and A. Lin, "Membrane-lipid associated lncRNA: a new regulator in cancer signaling," *Cancer Letters*, vol. 419, pp. 27–29, 2018.

[9] E. K. Robinson, S. Covarrubias, and S. Carpenter, "The how and why of lncRNA function: an innate immune perspective," *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1863, no. 4, article 194419, 2020.

[10] A. Bhan, M. Soleimani, and S. S. Mandal, "Long noncoding RNA and cancer: a new paradigm," *Cancer Research*, vol. 77, no. 15, pp. 3965–3981, 2017.

[11] T. Rajagopal, S. Talluri, R. L. Akshaya, and N. R. Dunna, "HOTAIR lncRNA: a novel oncogenic propellant in human cancer," *Clinica Chimica Acta*, vol. 503, pp. 1–18, 2020.

[12] H. Bolouri, J. E. Farrar, T. Triche Jr. et al., "The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions," *Nature Medicine*, vol. 24, no. 1, pp. 103–112, 2018.

[13] A. A. Rizvi, E. Karaesmen, M. Morgan et al., "gwasurvivr: an R package for genome-wide survival analysis," *Bioinformatics*, vol. 35, no. 11, pp. 1968–1970, 2019.

[14] E. Cerami, J. Gao, U. Dogrusoz et al., "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," *Cancer Discovery*, vol. 2, no. 5, pp. 401–404, 2012.

[15] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data," *Bioinformatics*, vol. 26, no. 1, pp. 139-140, 2010.

[16] P. Blanche, J. F. Dartigues, and H. Jacqmin-Gadda, "Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks," *Statistics in Medicine*, vol. 32, no. 30, pp. 5381–5397, 2013.

[17] S. Jathar, V. Kumar, J. Srivastava, and V. Tripathi, "Technological developments in lncRNA biology," *Advances in Experimental Medicine and Biology*, vol. 1008, pp. 283–323, 2017.

[18] G. L. Martin, "Encoder: a connectionist model of how learning to visually encode fixated text images improves reading fluency," *Psychological Review*, vol. 111, no. 3, pp. 617–639, 2004.

[19] A. Colaprico, T. C. Silva, C. Olsen et al., "TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data," *Nucleic Acids Research*, vol. 44, no. 8, article e71, 2016.

[20] K. Muret, C. Désert, L. Lagoutte et al., "Long noncoding RNAs in lipid metabolism: literature review and conservation analysis across species," *BMC Genomics*, vol. 20, no. 1, p. 882, 2019.

[21] F. Gao, Y. Cai, P. Kapranov, and D. Xu, "Reverse-genetics studies of lncRNAs-what we have learnt and paths forward," *Genome Biology*, vol. 21, no. 1, p. 93, 2020.

[22] C. Chen, P. Wang, W. Mo et al., "lncRNA-CCDC26, as a novel biomarker, predicts prognosis in acute myeloid leukemia," *Oncology Letters*, vol. 18, no. 3, pp. 2203–2211, 2019.

[23] L. Yang, J. D. Zhou, T. J. Zhang et al., "Overexpression of lncRNA *PANDAR* predicts adverse prognosis in acute myeloid leukemia," *Cancer Management and Research*, vol. 10, pp. 4999–5007, 2018.

[24] J. Zhang, H. Zhang, X. Wang, Y. Zhao, Y. Fu, and X. Liu, "PCAT18, as a novel differentially regulated long noncoding RNA in adult acute myeloid leukemia patients revealed by next-generation sequencing," *International Journal of Laboratory Hematology*, vol. 42, no. 6, pp. 858–865, 2020.

[25] H. Yu, J. Ma, J. Chen, Y. Yang, J. Liang, and Y. Liang, "lncRNA LINC00461 promotes colorectal cancer progression via miRNA-323b-3p/NFIB Axis," *Oncotargets and Therapy*, vol. 12, pp. 11119–11129, 2019.

[26] Q. Meng, M. Liu, and R. Cheng, "LINC00461/miR-4478/E2F1 feedback loop promotes non-small cell lung cancer cell proliferation and migration," *Bioscience Reports*, vol. 40, no. 2, 2020.

[27] A. A. Zimta, C. Tomuleasa, I. Sahnoune, G. A. Calin, and I. Berindan-Neagoe, "Long non-coding RNAs in myeloid malignancies," *Frontiers in Oncology*, vol. 9, p. 1048, 2019.

[28] S. Wei and K. Wang, "Long noncoding RNAs: pivotal regulators in acute myeloid leukemia," *Experimental Hematology & Oncology*, vol. 5, no. 1, p. 30, 2015.

[29] L. Y. Sun, X. J. Li, Y. M. Sun et al., "lncRNA ANRIL regulates AML development through modulating the glucose metabolism pathway of AdipoR1/AMPK/SIRT1," *Molecular Cancer*, vol. 17, no. 1, p. 127, 2018.

[30] Y. Tao, J. Zhang, L. Chen, X. Liu, M. Yao, and H. Zhang, "lncRNA CD27-AS1 promotes acute myeloid leukemia progression through the miR-224-5p/PBX3 signaling circuit," *Cell Death & Disease*, vol. 12, no. 6, p. 510, 2021.

[31] Y. Zhang, H. Li, W. Zhang, Y. Che, W. Bai, and G. Huang, "LASSO-based Cox-PH model identifies an 11-lncRNA signature for prognosis prediction in gastric cancer," *Molecular Medicine Reports*, vol. 18, no. 6, pp. 5579–5593, 2018.

[32] M. Zhou, Z. Zhang, H. Zhao, S. Bao, L. Cheng, and J. Sun, "An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme," *Molecular Neurobiology*, vol. 55, no. 5, pp. 3684–3697, 2018.