Genome-wide characterization of lncRNAs in acute myeloid leukemia

Lijun Lei*, Siyu Xia*, Dan Liu*, Xiaoqing Li*, Jing Feng, Yaqi Zhu, Jun Hu, Linjian Xia, Lieping Guo, Fei Chen, Hui Cheng, Ke Chen, Hanyang Hu, Xiaohua Chen, Feng Li, Shan Zhong, Nupur Mittal, Guohua Yang, Zhijian Qian, Leng Han and Chunjiang He

Corresponding authors: Chunjiang He, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China. Tel.: +86-27-68759702; E-mail: che@whu.edu.cn; Leng Han, Department of Biochemistry and Molecular Biology, The University of Texas Health Science Center at Houston McGovern Medical School, Houston, TX 77030, USA. Tel.: 713-500-6039; Fax: 713-500-0652; E-mail: leng.han@uth.tmc.edu

*These authors contributed equally to this work.

Abstract

Long noncoding RNAs (lncRNAs) are a large family of noncoding RNAs that play a critical role in various normal bioprocesses as well as tumorigenesis. However, the expression patterns and biological functions of lncRNAs in acute leukemia have not been well studied. Here, we performed transcriptome-wide lncRNA expression profiling of acute myeloid leukemia (AML) patient samples, along with non-leukemia control hematopoietic samples. We found that lncRNAs were differentially expressed in AML samples relative to control samples. Notably, we identified that lncRNAs upregulated in AML (relative to the control samples) are associated with a lower degree of DNA methylation and a higher ratio of being bound by transcription factors such as SP1, STAT4, ATF-2 and ELK-1 compared with those downregulated in AML. Moreover, an enrichment of H3K4me3 and a depletion of H3K27me3 were observed in upregulated lncRNAs in AML. Expression patterns of three types of lncRNAs (antisense, enhancer and intergenic lncRNAs) have previously been characterized. Of the identified lncRNAs, we

Lijun Lei is a Master Student in School of Basic Medical Sciences, Wuhan University.

Siyu Xia is a PhD Student in School of Basic Medical Sciences, Wuhan University.

Dan Liu is a Postdoc in School of Basic Medical Sciences, Wuhan University.

Xiaoqing Li is an Associate Professor in Union Hospital, Tongji Medical College, Huazhong University of Science and Technology.

Jing Feng is an Associate Professor in International School of Software, Wuhan University.

Yaqi Zhu is a PhD Student in School of Basic Medical Sciences, Wuhan University.

Jun Hu is a Postdoc in School of Basic Medical Sciences, Wuhan University.

Lanjian Xia is a Master Student in School of Basic Medical Sciences, Wuhan University.

Lieping Guo is a Physician in Eastern Hepatobiliary Hospital, Third Affiliated Hospital of Second Military Medical University.

Fei Chen is an Associate Professor in Zhongnan hospital of Wuhan University.

Hui Cheng is an Associate Professor in Changhai Hospital, Second Military Medical University.

Ke Chen is a Postdoc in Department of Medicine and Cancer Research Center, University of Illinois at Chicago.

Hanyang Hu is a Postdoc in School of Basic Medical Sciences, Wuhan University.

Xiaohua Chen is a PhD Student in School of Basic Medical Sciences, Wuhan University.

Feng Li is a Professor in School of Basic Medical Sciences, Wuhan University.

Shan Zhong is a Professor in School of Basic Medical Sciences, Wuhan University.

Nupur Mittal is a Postdoc in Department of Medicine and Cancer Research Center, University of Illinois at Chicago.

Guohua Yang is an Associate Professor in School of Basic Medical Sciences, Wuhan University.

Zhijian Qian is an Associate Professor in Department of Medicine and Cancer Research Center, University of Illinois at Chicago.

Leng Han is an Assistant Professor and CPRIT Scholar at Department of Biochemistry and Molecular Biology, the University of Texas Health Science Center at Houston McGovern Medical School. He is an expert for high-throughput data mining, especially for RNA sequencing.

Chunjiang He is an Associate Professor at School of Basic Medical Sciences, Wuhan University. He is an expert for identification and functional characterization of noncoding RNAs.

Submitted: 14 December 2016; Received (in revised form): 22 December 2016

© The Author 2017. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

found that high expression level lncRNA LOC285758 is associated with the poor prognosis in AML patients. Furthermore, we found that LOC285758 regulates proliferation of AML cell lines by enhancing the expression of HDAC2, a key factor in carcinogenesis. Collectively, our study depicts a landscape of important lncRNAs in AML and provides novel potential therapeutic targets and prognostic markers for AML treatment.

Key words: lncRNA; acute myeloid leukemia

Introduction

Acute myeloid leukemia (AML) is the second most common type of leukemia in adults and children. In total, 60–80% of AMLs are associated with a chromosome translocation in somatic cells [1]. Prognosis of adult AML is dismal, and ongoing research in the field is focused on identifying new targets for treatment of AML. Previous studies have mostly focused on messenger RNA (mRNA) and microRNA and the interactions between them in the regulation of AML [2–4]. The long noncoding RNAs (lncRNAs) are a group of RNAs >200bp, which have been increasingly discovered by high-throughput assays. Thousands of lncRNAs have been identified in various human tissues [5]. The function of the majority of these long noncoding transcripts is still not clear. The role of disease- and development-related lncRNAs has been established in variant pathological and physiological processes [6–8]. In certain types of leukemia, such as childhood MLL-rearranged acute lymphoblastic leukemia, lncRNAs have been systematically identified and correlated with several target protein-coding genes [9]. In acute megakaryoblastic leukemia, two lncRNAs were detected as regulators of erythro-megakaryocytic development and contributors to the maintenance of leukemic growth [10]. One intergenic lncRNA, HOTAIRM1, was shown to regulate cell proliferation in acute promyelocytic leukemia [11]. There are also a few lncRNAs implicated in the pathogenesis of AML. lncRNA CCDC26 has been abnormally amplified in childhood AML and regulates KIT expression, which can control the growth of AML cells [12]. Also, methylation of lncRNA MEG3 has been shown to confer an overall worse prognosis for AML patients [13]. Furthermore, it was demonstrated that intragenic lncRNAs, IRAIN and RUNXOR, interact with chromatin DNA and are involved in the formation of the chromatin loop, which has been detected in AML cells [14, 15]. Despite the discovery of all the potential functions of lncRNAs in AML samples, there has been a lack of global exploration of lncRNA-mRNA networks and understanding of the mechanism of coordination of lncRNAs and epigenetic modification in pathogenesis of AML. Garzon et al. [16] identified several lncRNAs and established their correlation with the prognosis of cytogenetically normal AML. As the lncRNA may be correlated with chromatin modification and possible interaction with other elements in the genome, it is necessary to perform a systematic study to inspect the global pattern of lncRNA regulation in AML.

Through a comprehensive analysis on our in-house patient samples, we detected the expression of whole transcriptome in cytogenetically abnormal AML patients. We integrated bioinformatic and experimental methods to reveal the network of long noncoding transcripts and the role of transcriptome changes in leukemogenesis. Our results of histone modification and the DNA methylation within the promoter regions of AML-related lncRNAs conclusively revealed the interactions between lncRNAs and epigenetic modifications. Moreover, we examined the transcription factors binding within the lncRNA regions and showed the lncRNAs LOC285758 targeted HDAC2, which can inhibit the cell proliferation of AML cell lines. Our results

suggested that the lncRNAs are a good repository to further research the mechanism of pathogenesis of AML.

Materials and methods

Patient samples

Bone marrow samples from six AML patients, and two non-leukemia controls were collected from Zhongnan Hospital, Wuhan University. The six AML patients had different morphologic types: AML-1: unknown; AML-2: m1/m4; AML-3: m1; AML-4: m2; AML-5: m2; and AML-6: m4. The study was approved by the institutional review board and consents were obtained. The mononuclear cell layer of bone marrow was extracted by Ficoll-Urografin solution (Sigma, USA) for RNA extraction.

Cell lines and culture condition

Human acute myelocytic leukemia cell line KG-1a and acute monocytic leukemia cell line THP-1 were purchased from American Type Cell Culture (Manassas, VA, USA). All cells were grown in RPMI (Roswell Park Memorial Institute) 1640 medium supplemented with 10% fetal bovine serum under the conditions of 5% CO₂ in incubator at 37 °C.

Quantitative real-time polymerase chain reaction analysis

Total RNA was extracted from the patient samples and cell lines with the TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions, and stored at –80 °C. In total, 1 µg RNA was reverse transcribed to complementary DNA (cDNA) using RevertAid™ First Strand cDNA Synthesis Kit (Fermentas). Quantitative polymerase chain reaction (qPCR) was performed on iQ5 real-time PCR detection system (Bio-Rad, USA) using SYBR Green Realtime PCR Master Mix (TOYOBO, Japan). All primers used in this study are listed in [Supplementary Table S1](#).

RNA transfection

Transfection of small interfering RNAs (siRNAs) into cells (THP-1, KG-1a) was performed using the PepMute siRNA transfection reagent (SignaGen) according to the manufacturer's instructions. The siRNAs for LOC285758 were purchased from RIBOBIO (Guangzhou, China). si-HDAC2 was purchased from GenePharma (Shanghai, China). The Stealth RNAi siRNA sequences were as follows: si-LOC285758: 5'-GCCTATTGACCAACATGTT-3'; si-HDAC2: 5'-GGAUUACAUAUGCU AAGATT-3'; and si-Control: 5'-UUCUCCGAACGUGUCACGUTT-3'.

Cell proliferation CCK-8 assays

THP-1 and KG-1a cells were transfected with siRNAs and then seeded in 96-well plates with 3 × 10⁴ cells per well in 100 µl of cell culture medium. Absorbance was measured at wavelengths of 450 nm consecutively at 0, 24, 48, 72 and 96 h. At each time

point, 10 μ l of CCK-8 solution (Biosharp, Japan) was added to each well.

Western blot

Cells (THP-1 and KG-1a) were washed in phosphate buffer saline (PBS) and lysed on ice for 30 min in RIPA buffer (add 1% (v/v) protease inhibitor phenylmethanesulfonyl fluoride (PMSF)). Equal amounts of proteins from each sample were separated in sodium dodecyl sulfate polyacrylamide gel electrophoresis gel. The proteins were then transferred to polyvinylidene fluoride (PVDF) membranes and probed with anti-HDAC2 (ABclonal, CHINA) and anti- β -actin (CWBIO, CHINA) antibodies.

Statistical analysis

Quantitative data were expressed as means $\pm$ SD for at least three independent experiments. Statistical significance was determined by Student's t-test. A P-value <0.05 was considered significant (*P <0.05 , **P <0.01 , ***P <0.001). Statistical test for distribution of histone modification and CpG methylation level were performed by WinSTAT. We downloaded the latest clinical data from The Cancer Genome Atlas (TCGA) official Web site (<https://tcga-data.nci.nih.gov>). We used log rank test to access whether the lncRNA expression is associated with the overall survival. We further used multivariate Cox model to determine whether the prognostic impact of expression profile of LOC285758 is dependent or independent from other clinical factors such as age, white blood count, cytogenetic and molecular abnormalities.

lncRNA microarray profiling

Arraystar Human lncRNA Microarray V3.0 is designed for the global profiling of human lncRNAs and protein-coding transcripts, which can detect 30 586 lncRNAs and 26 109 coding transcripts. Quantile normalization was performed using the GeneSpring GX v12.1 software package, and differentially expressed lncRNAs and mRNAs were identified with statistical significance [fold change >2.0 , P-value <0.05 , false discovery rate (FDR) <0.05]. All array data were uploaded to NCBI GEO database as GSE85030.

ChIP-seq data analysis

ChIP-seq data for targeting H3K4me3 and H3K27me3 in AML patient samples were downloaded from the GEO database (GSE54580). Reads were mapped to hg19 using Bowtie [17], and peaks in differentially expressed lncRNA promoters (± 2 kb to transcription start site (TSS)) were called by MACS [18], and the heatmaps were generated by deepTools [19].

Pathway analysis

Neighboring coding genes [20] were analyzed based on their relative genomic locations to the differentially expressed lncRNAs for GO biological processes by DAVID Bioinformatics Resources v6.7 [21].

Bisulfite sequencing analysis

Reduced representation bisulfite sequencing (RRBS) data in AML patients and normal bone marrow samples were downloaded from GEO databases (GSE37454). RRBS data analysis was performed with Bismark [22]. Reads were aligned to human genome (hg19) with 2 bp mismatch in a directional manner for

calling CpG methylation score in the differentially expressed lncRNA and mRNA promoters (± 2 kb to TSS). Heatmaps were plotted by deepTools [19].

Transcription factor binding analysis

Transcription factor binding analysis was performed with bedtools [23], which mapped human 81 conserved transcription factor binding sites (TFBSs) collected from UCSC Genome Browser (<http://genome.ucsc.edu>) to differentially expressed lncRNA promoters (-2 kb to TSS).

Results

lncRNAs are differentially expressed between AML and control samples

Using ArrayStar lncRNA Array v3.0 platform, we acquired expression of lncRNAs and mRNAs in six AML patients and two non-leukemia controls (NC-1 and NC-2). A total of 9134 lncRNAs and 13 542 mRNAs were detected in at least six of the eight samples. Principle component analysis (PCA) established clustering of different subtypes based on lncRNA expression profiles: two non-leukemia controls and six AML patients were clustered separately; AML-1, 5, 6 were clustered together, and AML-2, 3, 4 were clustered together, according to the subtypes of AML (Figure 1A). Similarly, mRNA profile can also distinguish the eight samples (Supplementary Figure S1A).

After performing an unpaired t-test on lncRNA profile, we discovered 2616 differentially expressed lncRNAs between AML and NC samples [fold change >2 and Benjamin-Hochberg adjusted P value (FDR) <0.05]. Among these, 1657 lncRNAs were upregulated and 959 lncRNAs were downregulated in AML as compared with NC samples (Supplementary Table S2, Figure 1B). Similarly, we observed 3518 differentially expressed mRNAs between AML and NC [fold change >2 and Benjamin-Hochberg adjusted P value (FDR) <0.05], of which 2013 mRNAs were upregulated and 1505 mRNAs were downregulated in AML (Supplementary Table S3 and Figure S1B). We deposited all lncRNA data into database lncAMLDdb (<http://gb.whu.edu.cn/lncAMLDdb>).

With the DAVID Web tool, we performed pathway enrichment analysis for upregulated and downregulated lncRNAs in AML based on their neighboring coding genes. Based on biological processes, the 1657 upregulated lncRNAs were mainly enriched in protein localization, transport and cell cycle (P <0.01) (Figure 1C). Also, the 959 downregulated lncRNAs were mainly enriched in cell apoptosis and migration (P <0.01) (Figure 1D).

Differentially expressed lncRNAs are associated with DNA methylation and transcription factor binding in the promoter regions in AML

With publicly available data sets from GEO, we further examined the DNA methylation levels of the differentially expressed lncRNAs in AML patients. We set a 4 kb window (± 2 kb) surrounding the TSS region of lncRNAs and observed the changes in DNA methylation in lncRNA promoters. In all three samples we checked, DNA methylation level in the TSS region of the upregulated lncRNAs was lower (CpG methylation score <30) than that of downregulated lncRNAs (CpG methylation score >30) (Figure 2A). Wilcoxon test indicated that the CpG methylation levels in TSS regions were significantly different between upregulated and downregulated lncRNAs (Supplementary Figure S2), revealing a correlation between differentially expressed lncRNAs and DNA methylation.

Figure 1. Expression pattern of lncRNAs in leukemia and non-leukemia controls. (A) PCA of lncRNAs expressed in samples. (B) Heatmap of differentially expressed lncRNAs in AML and non-leukemia controls (NC). (C) Pathways enriched in upregulated lncRNAs in AML. (D) Pathways enriched in downregulated lncRNAs in AML. P-values were calculated by DAVID Web server.

Figure 2. Cis- or trans-regulation of differentially expressed lncRNAs in AML. (A) DNA methylation distribution in TSS regions of differentially expressed lncRNAs in AML and NC. (B) Enrichment of TFBSs in upregulated (left panel) and downregulated (right panel) lncRNAs in AML. Color scale represents the mean centered binding ratio (percentages) of those transcription factors.

To explore whether transcription factor binding is involved in the variation of expression of the lncRNAs, we downloaded 81 TFBSs from the UCSC genome browser (<http://genome.ucsc.edu>) and examined the distribution of TFBS in our differentially expressed lncRNAs. To better compare the TFBS variance in lncRNA promoters, we selected the region from -2 kb to TSS of lncRNAs and inspected the distribution of TFBS. Our results revealed different transcription factor binding events in upregulated and

downregulated lncRNAs in AML. For example, sp1 has the highest binding ratio in upregulated lncRNAs (4.57%) in all binding transcription factors, while AREB6 has the highest binding ratio in downregulated lncRNAs (3.28%). STAT4, ATF-2 and Elk-1 have more than two times binding ratio in upregulated (0.42, 0.7 and 3.1%) than in downregulated lncRNAs (0.17, 0.29 and 1.42%), while Chx10 has more than two times binding ratio in downregulated (1.22%) than in upregulated lncRNAs (0.46%)

(Supplementary Table S4; Figure 2B). These results indicated that transcription factor binding is also involved in lncRNA regulation in AML.

Differentially expressed lncRNA are associated with histone modifications in AML

Previous studies revealed that lncRNA might be involved in histone modification in cancer and developmental processes [9, 15, 24]. Particularly, previously studies revealed that H3k4me3 and H3k27me3 may define the promoter activities in genome [25, 26]. To further test the potential interaction between epigenetic changes in AML with lncRNAs, we collected ChIP-Seq data of H3k4me3 and H3k27me3 in five AML patients from the GEO database and inspected the global distribution of those two modifications in the 4 kb window surrounding the lncRNA TSS region. We observed that the distribution of H3k4me3 and H3k27me3 establishes differences between upregulated and downregulated lncRNAs. In the TSS region of upregulated lncRNAs, H3k4me3 has more abundance than in downregulated lncRNAs (Figure 3A). In contrast, H3k27me3 has greater distribution in the TSS region of downregulated lncRNAs than in the upregulated group (Figure 3C). To quantify the difference, we performed a statistical test (Mann-Whitney U-test) for the distribution of histone modification coverage (tags/peak length) between upregulated and downregulated lncRNAs. We found the two types of histone modifications, and both showed significant differences between two groups of lncRNAs. The distribution of H3k4me3 is significantly higher in upregulated lncRNAs ($P = 5.84 \times 10^{-14}$, 0.014, 0.0055 , 4.13×10^{-5} and 0.0001 in five samples, Figure 3B), and H3k27me3 is significantly higher in the downregulated lncRNAs ($P = 0.0$, 0.01, 1.53×10^{-14} , 6.38×10^{-5} and 4.39×10^{-8} , Figure 3D). These results are consistent with previous studies in cells [24], and further confirmed that differentially expressed lncRNAs in AML are also correlated with histone modifications.

lncRNA interact with AML coding genes in different patterns

lncRNAs might affect the coding gene expression through genomic position effect [27]. lncRNAs can be classified as three types based on their positions with the coding genes: antisense lncRNA (lncRNA located in antisense strand of coding gene) [28, 29], enhancer lncRNA (lncRNA located in enhancer region of coding gene) [25, 30] and lincRNA (intergenic lncRNA) [31]. To explore whether these three types of lncRNAs are involved in AML coding genes, we separated the differentially expressed lncRNAs into three groups. We identified 120 antisense lncRNAs, which are differentially expressed between AML and NC, including 94 (78.3%) upregulated and 26 (21.7%) downregulated. Corresponding to 94 upregulated antisense lncRNAs, 70 sense genes were upregulated and 18 were downregulated in AML. Similarly, corresponding to 26 downregulated antisense lncRNAs, 14 sense genes were upregulated and 13 were downregulated in AML (Supplementary Table S5; Figure 4A). These results indicated that antisense lncRNAs are correlated with their host genes in both positive and negative ways.

Using the enhancer lncRNAs repository identified by previous work [30], we detected that 96 enhancer lncRNAs have different expression between AML and NC, including 55 (57.3%) upregulated lncRNAs and 41 (42.7%) downregulated lncRNAs. Among them, 51 upregulated and 37 downregulated coding genes corresponded to the 55 upregulated lncRNAs. In total, 36 upregulated and 30 downregulated coding genes corresponded to the 41 downregulated lncRNAs (Supplementary Table S6; Figure 4B).

We identified 194 differentially expressed lincRNAs based on lincRNA annotations from previous research [31]. Among them, 79 (40.7%) were upregulated and 115 (59.3%) were downregulated in AML. The genomic neighboring coding genes were used as potential target genes of these lincRNAs. Corresponding to 79 upregulated lincRNAs, 104 coding genes were also upregulated

Figure 3. Association between differentially expressed lncRNAs and histone modifications in AML. ChIP-Seq data sets from five AML samples were analyzed. (A) Heatmaps of H3k4me3 distribution in TSS regions (± 2 kb) of differentially expressed lncRNAs. (B) Differential distribution of H3k4me3 in upregulated and downregulated lncRNA. Mann-Whitney U-test was used to calculate the P-value. Y-axis represents the geometric average of tags/peak length, and error bar represents 95% of confidence interval (CI). (C) Heatmaps of H3k27me3 distribution in TSS regions (± 2 kb) of differentially expressed lncRNAs. (D) Differential distribution of H3k27me3 in upregulated and downregulated lncRNAs. Mann-Whitney U-test was used to calculate the P-value. Y-axis represents the geometric average of tags/peak length, and error bar represents 95% of confidence interval (CI).

Figure 4. Co-expression between three types of lncRNAs and coding genes. Upregulated and downregulated lncRNAs in AML are clustered. (A) Co-expression between antisense lncRNAs and their sense mRNAs. (B) Co-expression between enhancer lncRNAs and neighboring genes. (C) Co-expression between lincRNAs and their neighboring genes. (D) Expression pattern of four lncRNAs (AC005562.2, RP11-363G2.4, SPTY2D1-AS1 and LOC285758) and HDAC2 in AML and NC by array. Numbers in bar represent the log2 normalized expression. (E) qPCR validation of four lncRNAs and HDAC2 in AML and NC. * $P < 0.05$, ** $P < 0.01$.

and 59 were downregulated in AML. Corresponding to 115 downregulated lincRNAs, 100 coding genes were upregulated and 117 were downregulated in AML (Supplementary Table S7; Figure 4C). Pearson correlation analysis revealed expression of these three types of lncRNA were significantly correlated with above corresponding differentially expressed coding genes. Among that, antisense lncRNAs were more positively than negatively correlated with neighboring coding genes (Supplementary Figure S3). Pathway analysis of those three groups of lncRNAs indicated different biological process enrichment (Supplementary Figure S4).

For further function analysis, we performed real-time PCR for four differentially expressed lncRNAs, which were randomly selected: antisense lncRNA SPTY2D1-AS1, enhancer lncRNA LOC285758, lincRNA AC005562.2 and RP11-363G2.4, in AML and NC samples (Figure 4D). Consistent with the array data, qPCR analysis showed that they were significantly upregulated in AML (Figure 4E).

lncRNA might affect AML cell proliferation by regulating the expression of HDAC2

Among these lncRNAs, enhancer lncRNA LOC285758 showed the significant association with the poor overall survival based on 197 AML samples from TCGA (<https://tcga-data.nci.nih.gov>) (Log rank test $P = 0.023$; Figure 6A). To further consider other clinical factors such as age, white blood count, cytogenetic and molecular abnormalities, we built a univariate Cox model for all these clinical

factors, and observed that age ($P = 2.9 \times 10^{-4}$) and white blood count ($P = 0.018$) are significantly associated with the overall survival (OS), while cytogenetic and molecular abnormalities are marginally associated with OS ($P = 0.058$). We then built a multivariate Cox model to include these clinical factors, and observed that the expression level of LOC285758 is no longer associated with OS ($P = 0.287$), suggesting that the expression level of LOC285758 is dependent from other clinical factors. To further explore how LOC285758 exerts functions in AML development, we performed function analysis of LOC285758 in AML cell lines. LOC285758 located in ~60 kb upstream of HDAC2 (histone deacetylase2). Inhibition of HDAC2 could increase cell apoptosis and inhibit cell proliferation [32]. We hypothesized that LOC285758 might function as an enhancer for HDAC2 and regulated the expression of HDAC2. We observed that HDAC2 is significantly increased in AML samples, which is consistent with LOC285758. We designed siRNA against HDAC2 and expressed HDAC2 specific-siRNA in two AML cell lines including KG-1a and THP-1. We observed that the proliferation of both cell lines was inhibited after knockdown of HDAC2 (Figure 5A and B). Next, we determined whether LOC285758 has a similar function as HDAC1 does, we knocked down expression of LOC285758 by siRNA in KG-1a and THP-1. Our results showed that the proliferation of those two cell lines were inhibited after knockdown of LOC285758 (Figure 5C and D). We further showed that inhibition of LOC285758 leads to downregulation of HDAC2 (Figure 6B–E), suggested that inhibition of cell proliferation of AML cell lines by LOC285758 is mediated by HDAC2 in AML cells.

Figure 5. Function analysis of AML cell lines after knockdown of HDAC2 and LOC285758. (A) The inhibition of cell proliferation in KG-1a after transfection with si-HDAC2. (B) The inhibition of cell proliferation after transfection with si-HDAC2 in THP-1. (C) The inhibition of cell proliferation after transfection with si-LOC285758 in KG-1a. (D) The inhibition of cell proliferation after transfection with si-LOC285758 in THP-1. The error bar represents the SDs in replicated experiments. * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Figure 6. Regulation between lncRNAs and HDAC2 in AML cell lines. (A) Kaplan-Meier curves for overall survival in AML patients based on the expression level of LOC285758. (B) qPCR of LOC285758 and HDAC2 after knockdown of LOC285758 in KG-1a. (C) qPCR of LOC285758 and HDAC2 after knockdown of LOC285758 in THP-1. (D) Western blot of HDAC2 after knockdown of LOC285758 in KG-1a. (E) Western blot of HDAC2 after knockdown of LOC285758 in THP-1.

Interestingly, knockdown of HDAC2 could also inhibit the expression of LOC285758 (Supplementary Figure S5A and B), suggested the interactions between LOC285758 and HDAC2.

Conclusion

With lncRNA array, we identified thousands of differentially expressed lncRNAs in AML compared with control samples, which may involve in AML pathology and physiology. We performed the pathway enrichment analysis to understand the potential biological processes of AML. Our results showed that lncRNAs are regulated by cis- or trans-factors such as DNA methylation or transcription factors, which is similar to coding genes [33, 34]. Interestingly, the distributions of TFBSs were different between upregulated and downregulated lncRNAs. We also indicated that lncRNAs are involved in regulation of promoter activities, which are represented by H3k4me3 and H3k27me3. Despite the fact that we obtain the DNA methylation and histone modification data from different AML samples, our integrative analysis showed that these lncRNAs could interact with their host or neighboring coding genes.

Importantly, we identified an enhancer lncRNA LOC285758, which may play a key role in the pathogenesis of AML. Analysis of TCGA AML data sets showed that high expression level of LOC285758 is associated with poor prognosis in AML patients, suggesting that LOC285758 is a potential prognostic marker. Furthermore, we analyzed the function and expression regulation of lncRNA LOC285758 and its neighboring coding gene HDAC2, suggesting that LOC285758 might play an important role in AML by regulating HDAC2 expression.

Key Points

- We provided a novel repository of potential functional lncRNAs for AML research.
- We found the expression pattern of lncRNAs was associated with DNA methylation and transcription factor binding in AML.
- We revealed the potential regulation of lncRNAs with histone modifications in AML.
- We revealed three potential interactions between lncRNAs and coding genes in AML.
- We verified the function and potential regulation between lncRNA and HDAC2 in AML cell lines.

Supplementary Data

Supplementary data are available online at <http://bib.oxfordjournals.org/>.

Acknowledgements

The authors thank Ms Jennifer Strong for English editing. We also appreciate Wuhan University for financial support to this research.

Funding

National Natural Science Foundation of China (grant number 81500140 to C.H.); Natural Science Foundation of Hubei Province, China (grant number 2015CFB170 to C.H.); the Fundamental Research Funds for the Central Universities of China (grant number 2042015kf0018 to C.H.); Cancer Prevention and Research Institute of Texas (grant number

RR150085 to L.H.); and China Scholarship Council (grant number 201606175095 to J.F.).

References

1. Haferlach T, Bacher U, Kern W, et al. Diagnostic pathways in acute leukemias: a proposal for a multimodal approach. *Ann Hematol* 2007;**86**:311–27.
2. Li Z, Huang H, Li Y, et al. Up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML. *Blood* 2012;**119**:2314–24.
3. Rathert P, Roth M, Neumann T, et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* 2015;**525**:543–7.
4. Sun XJ, Wang Z, Wang L, et al. A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* 2013;**500**:93–7.
5. Nielsen MM, Tehler D, Vang S, et al. Identification of expressed and conserved human noncoding RNAs. *RNA* 2014;**20**:236–51.
6. Klattenhoff CA, Scheuermann JC, Surface LE, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* 2013;**152**:570–83.
7. Ounzain S, Micheletti R, Beckmann T, et al. Genome-wide profiling of the cardiac transcriptome after myocardial infarction identifies novel heart-specific long non-coding RNAs. *Eur Heart J* 2015;**36**:353–68a.
8. Brunner AL, Beck AH, Edris B, et al. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. *Genome Biol* 2012;**13**:R75.
9. Fang K, Han BW, Chen ZH, et al. A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target. *Hum Mol Genet* 2014;**23**:3278–88.
10. Emmrich S, Streltsov A, Schmidt F, et al. LincRNAs MONC and MIR100HG act as oncogenes in acute megakaryoblastic leukemia. *Mol Cancer* 2014;**13**:171.
11. Zhang X, Weissman SM, Newburger PE. Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. *RNA Biol* 2014;**11**:777–87.
12. Hirano T, Yoshikawa R, Harada H, et al. Long noncoding RNA, CCDC26, controls myeloid leukemia cell growth through regulation of KIT expression. *Mol Cancer* 2015;**14**:90.
13. Benetatos L, Hatzimichael E, Dasoula A, et al. CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes. *Leuk Res* 2010;**34**:148–53.
14. Sun J, Li W, Sun Y, et al. A novel antisense long noncoding RNA within the IGF1R gene locus is imprinted in hematopoietic malignancies. *Nucleic Acids Res* 2014;**42**:9588–601.
15. Wang H, Li W, Guo R, et al. An intragenic long noncoding RNA interacts epigenetically with the RUNX1 promoter and enhancer chromatin DNA in hematopoietic malignancies. *Int J Cancer* 2014;**135**:2783–94.
16. Garzon R, Volinia S, Papaioannou D, et al. Expression and prognostic impact of lncRNAs in acute myeloid leukemia. *Proc Natl Acad Sci USA* 2014;**111**:18679–84.
17. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;**10**:R25.
18. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;**9**:R137.

19. Ramirez F, Dundar F, Diehl S, et al. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 2014;**42**:W187–91.
20. He C, Hu H, Wilson KD, et al. Systematic characterization of long noncoding RNAs reveals the contrasting coordination of Cis- and trans-molecular regulation in human fetal and adult hearts. *Circ Cardiovasc Genet* 2016;**9**:110–18.
21. Dennis G, Jr, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;**4**:P3.
22. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 2011;**27**:1571–2.
23. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–2.
24. Wu SC, Kallin EM, Zhang Y. Role of H3K27 methylation in the regulation of lncRNA expression. *Cell Res* 2010;**20**:1109–16.
25. Ounzain S, Pezzuto I, Micheletti R, et al. Functional importance of cardiac enhancer-associated noncoding RNAs in heart development and disease. *J Mol Cell Cardiol* 2014;**76**:55–70.
26. Marques AC, Hughes J, Graham B, et al. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. *Genome Biol* 2013;**14**:R131.
27. Morlando M, Ballarino M, Fatica A. Long non-coding RNAs: new players in hematopoiesis and leukemia. *Front Med (Lausanne)* 2015;**2**:23.
28. Bhan A, Hussain I, Ansari KI, et al. Antisense transcript long noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by estradiol. *J Mol Biol* 2013;**425**:3707–22.
29. Punzi G, Ursini G, Shin JH, et al. Increased expression of MARCKS in post-mortem brain of violent suicide completers is related to transcription of a long, noncoding, antisense RNA. *Mol Psychiatry* 2014;**19**:1057–9.
30. Orom UA, Derrien T, Beringer M, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell* 2010;**143**:46–58.
31. Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011;**25**:1915–27.
32. Lei WW, Zhang KH, Pan XC, et al. Histone deacetylase 1 and 2 differentially regulate apoptosis by opposing effects on extracellular signal-regulated kinase 1/2. *Cell Death Dis* 2010;**1**:e44.
33. Kumegawa K, Maruyama R, Yamamoto E, et al. A genomic screen for long noncoding RNA genes epigenetically silenced by aberrant DNA methylation in colorectal cancer. *Sci Rep* 2016;**6**:26699.
34. Li H, Li H, Hao Y, et al. Differential long noncoding RNA and mRNA expression in differentiated human glioblastoma stem cells. *Mol Med Rep* 2016;**14**:2067–76.