Prognostic and biologic significance of long non-coding RNA profiling in younger adults with cytogenetically normal acute myeloid leukemia

Geoffrey L. Uy
Washington University School of Medicine in St. Louis
et al.

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/6208

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Long non-coding ribonucleic acids (RNAs) are a novel class of RNA molecules, which are increasingly recognized as important molecular players in solid and hematologic malignancies. Herein we investigated whether long non-coding RNA expression is associated with clinical and molecular features, as well as outcome of younger adults (aged <60 years) with cytogenetically normal acute myeloid leukemia. Whole transcriptome profiling was performed in a training (n=263) and a validation set (n=114). Using the training set, we identified 24 long non-coding RNAs associated with event-free survival. Linear combination of the weighted expression values of these transcripts yielded a prognostic score. In the validation set, patients with high scores had shorter disease-free (P<0.001), overall (P=0.002) and event-free survival (P<0.001) than patients with low scores. In multivariable analyses, long non-coding RNA score status was an independent prognostic marker for disease-free (P=0.01) and event-free survival (P=0.002), and showed a trend for overall survival (P=0.06). Among multiple molecular alterations tested, which are prognostic in cytogenetically normal acute myeloid leukemia, only double CEBPA mutations, NPM1 mutations and FLT3-ITD associated with distinct long non-coding RNA signatures. Correlation of the long non-coding RNA scores with messenger RNA and microRNA expression identified enrichment of genes involved in lymphocyte/leukocyte activation, inflammation and apoptosis in patients with high scores. We conclude that long non-coding RNA profiling provides meaningful prognostic information in younger adults with cytogenetically normal acute myeloid leukemia. In addition, expression of prognostic long non-coding RNAs associates with oncogenic molecular pathways in this disease.
poor. Only 40% of younger adult (aged <60 years) and 10% of older (aged ≥60 years) AML patients achieve long-term survival.1 Currently, chromosomal aberrations2-4 and recurrent gene mutations5-8 are considered the most reliable and reproducible prognostic markers in AML, and are used in the clinic to identify patients at high risk of death and to guide treatment. Aberrant levels of messenger RNA (mRNA)9-11 and microRNA (miR) transcripts12,13 also have prognostic significance, and efforts have been made to incorporate gene-expression profiling into prognostic algorithms.14-16

Long non-coding RNAs (lncRNAs) are a novel class of RNA molecules that are longer than 200 nucleotides, with no protein coding potential and are either located within the intergenic stretches of the genome or overlap (in sense or antisense direction) protein coding genes.17-20 These transcripts regulate key cellular functions, such as chromosomal dosage compensation,21 imprinting,22 cell cycle progression,23 and differentiation.24 In cancer, individual lncRNAs have been shown to play an important role in malignant transformation.25-27 Despite the growing understanding of the biologic significance of deregulated lncRNA expression in malignant diseases, the value of these molecules as potential biomarkers in the clinical setting has not been extensively studied.26,27 With regard to cytogenetically normal AML (CN-AML), the prognostic and biologic significance of lncRNAs in younger adult patients remains unknown. Therefore, we analyzed, using whole transcriptome sequencing (RNA-seq), the lncRNA profiles of younger adults with de novo CN-AML, who were comprehensively characterized with regard to molecular abnormalities and outcome. Herein, we show that lncRNA profiling provides independent prognostic information in these patients. We also show that expression levels of prognostic lncRNAs correlate with distinct mRNA and miRNA signatures, and provide insights into the leukemogenic pathways that these lncRNAs potentially regulate.

Methods

Patients and treatment

Pretreatment bone marrow (BM) or blood samples were obtained from a training (n=263) and a validation set (n=114) of younger adult patients (aged 17-59 years) with de novo CN-AML, who received intensive cytarabine/anthracycline-based first-line therapy on Cancer and Leukemia Group B (CALGB)/Alliance for Clinical Trials in Oncology (Alliance) trials and were alive 30 days after initiation of treatment. Per protocol, no patient received allogeneic stem cell transplantation in first complete remission (CR). Details regarding treatment protocols are provided in the Online Supplementary Appendix. All patients provided written informed consent, and all study protocols were in accordance with the Declaration of Helsinki and approved by institutional review boards at each center.

Cytogenetic and molecular analyses

Cytogenetic analyses were performed in CALGB/Alliance-approved institutional laboratories and results confirmed by central karyotype review.25 The diagnosis of normal karyotype was based on at least 20 metaphase cells analyzed in BM specimens subjected to short-term (24- or 48-hour) unstimulated cultures.

Targeted amplicon sequencing using the MiSeq platform (Illumina) was used to analyze DNA samples for presence of gene mutations that have been reported to associate with clinical outcome of CN-AML patients (i.e., mutations in the ASXL1, DNMT3A [R882 and non-R882], IDH1, IDH2 [R140 and R172], NPM1, RUNX1, TET2 or WT1 genes, and FLT3-tyrosine kinase domain [FLT3-TKD] mutations), as described previously.26,27 A variant allele frequency of ≥10% was used as the cutoff to distinguish between mutated versus wild-type alleles of these genes. The presence of mutations in the CEBPA gene and FLT3-internal tandem duplications (FLT3-ITD) were evaluated using Sanger sequencing28 and fragment analysis,29 as described previously. Since only double CEBPA mutations are favorable prognostic markers in CN-AML,30 we considered only this genotype as mutated.

Transcriptome analyses

RNA samples of all studied patients (n=377) were analyzed with total RNA sequencing (after depletion of ribosomal and mitochondrial RNA) using the Illumina HiSeq 2500 platform. Due to RNA quality restrictions, a subset of 300 patients could be additionally analyzed with small RNA sequencing for profiling of miRNA expression. Further details are provided in the Online Supplementary Appendix. To determine the expression status of patients (i.e., high versus low expressers) with regard to prognostic expression markers (e.g., BAALC), the median values of normalized RNA sequencing reads were used as the cutoff.

Statistical analyses

Clinical endpoint definitions are given in the Online Supplementary Appendix. Baseline demographic, clinical, and molecular features were compared between patients with low and those with high lncRNA scores (later on referred to as favorable and unfavorable, see below), and between the training and validation sets using the Wilcoxon rank-sum and Fisher’s exact tests for continuous and categorical variables, respectively.31 The estimated probabilities of disease-free (DFS), overall (OS) and event-free survival (EFS) were calculated using the Kaplan–Meier method, and the log-rank test evaluated differences between survival distributions.32 Cox proportional hazard models were used to calculate hazard ratios (HR) for DFS, OS and EFS.33 Multivariable proportional hazards models were constructed using a backward selection procedure. All statistical analyses were performed by The Alliance Statistics and Data Center.

Results

Global expression of lncRNAs

To investigate the role of lncRNAs in AML, we first identified all known lncRNAs which were present in the transcriptomes of the younger CN-AML patients who were studied (n=377). After exclusion of contaminating ribosomal RNA molecules, we identified 22,166 non-coding RNA transcripts. According to the GENCODE v22 database,34 23% of these transcripts were categorized as processed pseudogenes, 21% as intergenic/intervening lncRNAs, 21% as antisense lncRNAs, 4% as sense intronic/overlapping lncRNAs and 31% were classified as other transcripts (e.g., as unitary pseudogenes, unprocessed pseudogenes etc.; Figure 1).

Generation of a prognostic lncRNA score in the training set

To assess the prognostic significance of lncRNA expression in younger adults with CN-AML, we performed exploratory analysis in a training set (n=263) of younger CN-AML patients and used a separate patient cohort to
validate our findings (validation set, n=114). Comparison of clinical and molecular characteristics at diagnosis between the training and validation sets showed that they were relatively similar, with the exceptions that patients in the training set had higher percentages of blood blasts \( (P=0.03) \), were more frequently \( \text{FLT3-TKD} \)-positive \( (P=0.02) \), and had higher \( \text{ERG} \) \( (P=0.01) \) and \( \text{BAALC} \) \( (P=0.002) \) expression levels (Online Supplementary Table S1).

We first identified all lncRNAs that were highly associated with EFS \( (P<10^{-6}) \) in the training set by univariable
Cox analysis (Figure 2). EFS was used because it comprehensively evaluates the lncRNAs that are associated with response to chemotherapy, probability of relapse and probability of survival. We detected 24 lncRNAs associated with EFS ($P < 10^{-6}$; Online Supplementary Table S2). Next, we derived a prognostic lncRNA score by linear combination of the weighted expression values of these 24 lncRNAs. The median value of the lncRNA score was used to dichotomize the training set of patients. Patients with low lncRNA scores (n=132) had longer DFS ($P < 0.001$), OS ($P < 0.001$) and EFS ($P < 0.001$) than patients with high lncRNA scores (n=131). We therefore classified low lncRNA scores as “favorable” and high as “unfavorable” (Online Supplementary Table S3 and Online Supplementary Figure S1).

### Association of lncRNA score with patient characteristics and clinical outcome in the training set

With regard to clinical and molecular characteristics, patients with favorable lncRNA scores in the training set were more likely to present with higher hemoglobin levels ($P=0.02$), lower white blood cell (WBC) counts ($P<0.001$), and lower percentages of BM blasts ($P=0.02$). They were also less likely to harbor FLT3-ITD ($P<0.001$), DNMT3A ($P=0.01$) and RUNX1 ($P=0.009$) mutations and more likely to harbor double CEBPA mutations ($P<0.001$). Patients with favorable lncRNA scores in the training set differed with regard to their distribution in the Risk Categories of the European LeukemiaNet (ELN) classification of AML, when compared with patients with unfavorable lncRNA scores ($P<0.001$); patients with favorable lncRNA scores were more frequently classified as favorable and less frequently as intermediate or adverse risk than those with unfavorable lncRNA scores (Online Supplementary Table S4). Favorable lncRNA score status also associated with high expression of miR-181a ($P<0.001$) and low expression of miR-155 ($P=0.03$, Online Supplementary Table S4). Association of a favorable lncRNA score with longer DFS, OS and EFS remained significant in multivariable analyses ($P<0.001$ for all 3 end points, Online Supplementary Table S5), after adjusting for other co-variates.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Favorable lncRNA Score (n=57)</th>
<th>Unfavorable lncRNA Score (n=57)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNMT3A, n. (%)</td>
<td>Mutated 20 (36) R882 14 Non-R882 6 Wild-type 36 (64)</td>
<td>23 (43) 20 3 31 (57)</td>
<td>0.56</td>
</tr>
<tr>
<td>RUNX1, n. (%)</td>
<td>Mutated 3 (5) Wild-type 53 (95)</td>
<td>2 (4) 52 (96)</td>
<td>1.00</td>
</tr>
<tr>
<td>ELN Risk Category,* n. (%)</td>
<td>Favorable 37 (71) Intermediate 11 (21) Adverse 4 (8)</td>
<td>23 (43) 20 (38) 10 (19)</td>
<td>0.02</td>
</tr>
<tr>
<td>ERG expression group,' n. (%)</td>
<td>High 22 (39) Low 35 (61)</td>
<td>23 (41) 33 (59)</td>
<td>0.85</td>
</tr>
<tr>
<td>BAALC expression group,* n. (%)</td>
<td>High 19 (36) Low 34 (64)</td>
<td>21 (39) 33 (61)</td>
<td>0.64</td>
</tr>
<tr>
<td>MVI expression group,* n. (%)</td>
<td>High 18 (33) Low 37 (67)</td>
<td>29 (52) 27 (48)</td>
<td>0.06</td>
</tr>
<tr>
<td>miR-181a expression group,* n. (%)</td>
<td>High 24 (40) Low 24 (40)</td>
<td>18 (30) 27 (60)</td>
<td>0.41</td>
</tr>
<tr>
<td>miR-3151, n. (%)</td>
<td>Expressed 8 (17) Not expressed 40 (83)</td>
<td>4 (9) 41 (91)</td>
<td>0.56</td>
</tr>
<tr>
<td>miR-155 expression group,* n. (%)</td>
<td>High 16 (33) Low 32 (67)</td>
<td>31 (69) 14 (31)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN favorable risk category comprises patients with double-mutated CEBPA and patients with mutated NPM1 without FLT3-ITD or with FLT3-ITD". The ELN intermediate risk category includes patients with wild-type NPM1 without FLT3-ITD or wild-type NPM1 and FLT3-ITD-- or mutated NPM1 and FLT3-ITD--. The ELN adverse risk category comprises patients with wild-type NPM1 and FLT3-ITD-- and/or mutated RUNX1 (if it does not co-occur with a favorable AML subtype) and/or mutated ASXL1 (it does not cooccur with a favorable AML subtype) and/or mutated TP53. FLT3-ITD-- is defined by a FLT3-ITD/FLT3 wild-type allelic ratio of less than 0.5 and FLT3-ITD++ is defined by a FLT3-ITD/FLT3 wild-type allelic ratio of equal to or more than 0.5." The median expression value was used as the cut point. WBC: white blood cell; HCT: hematopoietic cell transplant; CR: complete remission; ELN: European LeukemiaNet; FLT3-ITD: internal tandem duplication of the FLT3 gene; FLT3-TKD: tyrosine kinase domain mutation in the FLT3 gene; lncRNA: long non-coding ribonucleic acid; miR: microRNA.
Association of lncRNA score with patient characteristics and clinical outcome in the validation set

We used the median value of the lncRNA score, as calculated in the training set to divide the validation set into favorable and unfavorable lncRNA score groups (Figure 2). Patients with favorable lncRNA scores (n=57) were less likely to present with higher WBC counts at the time of diagnosis (P=0.009) or to harbor FLT3-ITD (P=0.007). lncRNA score status also associated with significantly different distribution of the patients in the Risk Categories of the ELN guidelines (P=0.02). Patients with favorable lncRNA scores were more likely to belong to the favorable and less likely to belong to the intermediate or adverse risk category. Patients with favorable lncRNA scores in the validation set were less likely to be miR-155 high-expressers (P<0.001) than patients with unfavorable lncRNA scores (n=57; Table 1).

Patients with favorable lncRNA scores had longer DFS than those with unfavorable lncRNA scores (P<0.001; Figure 3A). Five years after diagnosis, 51% of patients with favorable lncRNA scores remained alive and leukemia-free, in contrast to only 17% of those with unfavorable lncRNA scores. Favorable lncRNA score status also associated with longer OS (P=0.002, 5-year rates; 52% versus 26%; Figure 3B) and longer EFS (P=0.001, 5-year rates, 46% versus 16%; Figure 3C, Online Supplementary Table S6). The prognostic value of the lncRNA score in the validation set remained significant when it was analyzed as a continuous variable. Increasingly favorable lncRNA scores associated with longer DFS (P<0.001), OS (P=0.007) and EFS (P=0.002).

In multivariable analyses, favorable lncRNA score status was an independent marker for longer DFS (P=0.01), after adjusting for miR-155 expression status, and EFS (P=0.002), after adjusting for the presence of FLT3-ITD (Table 2). With regard to OS, patients with a favorable lncRNA score had a trend for longer survival (P=0.06), after adjustment for FLT3-ITD and MN1 expression status.

Associations of recurrent gene mutations with lncRNA expression

We evaluated if recurrent prognostic gene mutations in CN-AML associated with distinct expression patterns of lncRNAs in younger adults with CN-AML. For this purpose, mutation-related lncRNA signatures were derived in the training set using stringent criteria (for details see Methods and the Online Supplementary Appendix).

Double-mutated CEBPA showed the strongest association with lncRNA expression; 82 lncRNAs were upregulated and 186 lncRNAs were downregulated in patients who harbored double-mutated CEBPA (Figure 4A, Online Supplementary Table S7). Among the CEBPA mutation-related lncRNAs, NEAT1 was significantly underexpressed in the group of patients with CEBPA mutations. This lncRNA has been involved in myeloid differentiation of acute promyelocytic leukemia cells after all-trans retinoic acid treatment.

Mutations in the NPM1 gene also strongly associated with a lncRNA signature, which comprised 35 transcripts upregulated and 37 transcripts downregulated in patients harboring NPM1 mutations (Figure 4B, Online Supplementary Table S8). Thirty-three of the 35 lncRNAs overexpressed in patients with NPM1 mutations, were downregulated in patients with CEBPA mutations. This finding is consistent with the observation that double CEBPA and NPM1 mutations rarely co-occur in CN-AML. NPM1 mutations were positively associated with lncRNAs embedded within the HOX gene loci (HOXAS3, HOXB-AS3) and other lncRNAs implicated in myelopoiesis (EGOT19) or carcinogenesis (e.g., PCAT16 and LUCAT120).

The FLT3-ITD-related lncRNA signature consisted of 26 transcripts, 19 of which were upregulated and 7 downregulated in patients with this mutation (Figure 4C, Online Supplementary Table S9). The host gene of miR-155 (MIR155HG) was among the lncRNAs overexpressed in FLT3-ITD-positive patients. High MIR155HG expression independently associates with poor outcome in CN-AML.
The WT1-AS lncRNA was also highly expressed among FLT3-ITD-positive patients; it has been reported to post-translationally regulate the protein levels of WT1. To assess the capacity of gene mutation-related lncRNA signatures to detect their corresponding molecular alterations in CN-AML patients, we applied these signatures to the validation set. The mutated CEBPA-related signature showed the highest level of accuracy (specificity and sensitivity of mutated CEBPA detection: ≥93% and ≥98%, respectively), followed by the mutated NPM1-related (sensitivity: ≥80%, specificity ≥73%) and the FLT3-ITD-related signatures (sensitivity ≥70%, specificity: ≥76%). The remaining prognostic gene mutations that were tested either did not associate with differential expression of lncRNAs (i.e., TET2 mutations) or generated signatures that failed to reliably detect the mutational status of patients in the validation set (e.g., DNMT3A, WT1 mutations).

**Biologic implications of the lncRNA score**

To gain biologic insights into the molecular pathways that may be affected by differences in the lncRNA score, we examined the correlation between the lncRNA score and the mRNA/miRNA expression in 300 younger CN-AML patients who had available mRNA and miRNA profiling data.

We identified 410 mRNA transcripts whose expression levels correlated with the lncRNA score, 172 of which correlated positively and 238 negatively with unfavorable lncRNA scores (Figure 5A, Online Supplementary Table S10). Among highly expressed genes in patients with unfavorable lncRNA scores, putative oncogenes and key mediators of the oncogenic AP-1 pathway such as ATF3, FOS, FOSB, JUN, and MAFF were identified. With regard to hematopoiesis, the AP-1 pathway has been shown to regulate proliferation of erythroblast cells, to mediate monocyte/macrophage differentiation of myeloid cells and to co-regulate miR-155 expression in stimulated macrophages. Genes that regulate immune responses (e.g., IL1B, IRF7, CD80) and genes that mediate immune evasion (e.g., IER3, ILIRB4) were also highly expressed in patients with unfavorable lncRNA profiles. Finally, oncomiRs promoting proliferation of malignant cells (e.g., RET, ETS2, PLK2, NEK6, PLK3 and SRC) were found to be overexpressed in patients with unfavorable lncRNA scores. Gene ontology analysis revealed that genes involved in lymphocyte/leukocyte activation, inflammation, response to wounding and regulation of apoptosis were enriched in the subset of patients with unfavorable lncRNA scores (Figure 5B, Online Supplementary Table S11).

Among mRNA molecules downregulated in patients with unfavorable lncRNA scores, we detected transcripts with reported tumor-suppressive function (APC, JADE1, BRM51L, and ING3). Gene ontology analysis showed that genes that participate in the regulation of transcription, the regulation of RNA metabolic processes and DNA binding were underexpressed in the group of patients with unfavorable lncRNA scores (Figure 5C, Online Supplementary Table S11).

With regard to miR expression, 10 miRs were found to correlate positively (miR-660, miR-502, miR-532-5p, miR-155, miR-500a-3p, miR-500a-5p, miR-532-3p, miR-362, miR-339 and miR-23a) and 4 miRs to correlate negatively (miR-192, miR-625, miR-100 and miR-194) with unfavorable lncRNA scores (Online Supplementary Table S12). Among the 10 miRs that positively correlated with unfavorable lncRNA scores, 7 were located in close proximity on chromosome X; miR-660, miR-502, miR-532-5p, miR-155, miR-500a-3p, miR-500a-5p, miR-532-3p, miR-362 are all imbedded in intron 5 of the CLCN5 gene. This miR cluster mediates an anti-apoptotic effect in chronic lymphocytic leukemia cells. miR-155, which also positively correlated...
with unfavorable lncRNA scores, is an established adverse prognosticator in CN-AML and has been implicated in leukemogenesis of FLT3-ITD-positive AML.

**Discussion**

Over the past 5 years, lncRNAs have emerged as new players in cancer biology and biomarker discovery. Our group has previously reported that distinctive lncRNA signatures are associated with prognostic gene mutations in older CN-AML patients, and that expression levels of a small group of lncRNAs have prognostic significance in these patients. Since CN-AML in younger adults differs with regard to clinical features, associated molecular abnormalities and outcome from that in older patients, we investigated the prognostic value and biologic implications of lncRNA expression in a total of 377 CN-AML adult patients younger than 60 years.

First, we identified 24 lncRNAs highly correlated with EFS. Similarly to the previously reported older CN-AML patients, only a small number of these prognostic lncRNAs associated with prognostic gene mutations: MIR155HG was upregulated in patients who harbor FLT3-ITD, AC006129.2 was upregulated in patients with double CEBPA mutations, whereas AL122127.25, RP11-946L16.2, and...

<table>
<thead>
<tr>
<th>Table 2. Multivariable analyses for outcome in the validation set of younger adults with cytogenetically normal acute myeloid leukemia.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables in final models</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>IncRNA score, favorable versus unfavorable</td>
</tr>
<tr>
<td>miR-155, high versus low*</td>
</tr>
<tr>
<td>FLT3-ITD, present versus absent</td>
</tr>
<tr>
<td>MN1, high versus low*</td>
</tr>
</tbody>
</table>

Hazard ratios greater than (less than) 1.0 indicate higher (lower) risk for relapse or death (disease-free survival) or death (overall survival) or for failure to achieve complete remission, relapse or death (event-free survival) for the first category listed. Variables considered for model inclusion were: IncRNA score status (favorable versus unfavorable), age (as a continuous variable, in 10-year increments), sex (male versus female), race (white versus non-white), white blood cell count (as a continuous variable, in 50-unit increments), hemoglobin (as a continuous variable, in 1-unit increments), platelet count (as a continuous variable, in 50-unit increments), extramedullary involvement (present versus absent), ASXL1 mutations (mutated versus wild-type), CEBPA mutations (double-mutated versus single-mutated or wild-type), DNMT3A mutations (mutated versus wild-type), FLT3-ITD (present versus absent), FLT3-TKD (present versus absent), IDH1 mutations (mutated versus wild-type), IDH2 mutations (mutated versus wild-type), NPM1 mutations (mutated versus wild-type), RUNX1 mutations (mutated versus wild-type), TET2 mutations (mutated versus wild-type), WT1 mutations (mutated versus wild-type), BAALC expression levels (high versus low), ERG expression levels (high versus low), MN1 expression levels (high versus low), miR-181a expression levels (high versus low), miR-3151 (expressed versus not expressed), and miR-155 expression levels (high versus low). * The median expression value was used as the cut point. HR: hazard ratio; CI: confidence interval; lncRNA: long non-coding RNA; FLT3-ITD: internal tandem duplication of the FLT3 gene.

**Figure 3. Outcomes of younger adult patients with cytogenetically normal acute myeloid leukemia with favorable and unfavorable long non-coding RNA (lncRNA) scores in the validation set.** (A) Disease-free survival, (B) overall survival and (C) event-free survival. The IncRNA score of each individual patient was computed as a weighted score of 24 prognostic lncRNAs.
SDHAP3 and SENC3 were downregulated in patients with double CEBPA mutations. Of the 24 prognostic lncRNA genes, only MIR155HG has previously been associated with the clinical outcome of CN-AML patients.\(^4\) Linear combination of the weighted expression values of lncRNA transcripts yielded a prognostic score, which strongly associated with DFS, EFS and OS duration in younger adult CN-AML patients. Favorable lncRNA score status was an independent marker for longer DFS and EFS (and also showed a strong trend towards significance for longer OS). We were intrigued to find no overlap between the 48 prognostic lncRNAs that we previously identified in older CN-AML patients\(^26\) and the 24 transcripts reported herein in younger patients. This finding could be interpreted as an additional biologic difference between CN-AML of younger and that of older patients, similar to the age-dependent difference in frequency of some recurrent prognostic gene mutations (e.g., mutations in the ASXL1 and RUNX1 genes).\(^1\)

We also examined the associations between recurrent prognostic gene mutations and lncRNA expression, and found double CEBPA and NPM1 mutations and FLT3-ITD to associate with distinct lncRNA signatures. We identified several lncRNAs that were commonly associated with these gene mutations in both younger and older CN-AML patients\(^25\) (e.g., the HOX-loci embedded lncRNAs in the NPM1 mutation-related lncRNA signature, WT1-AS in the FLT3-ITD-related signature, etc.). On the other hand, such gene mutations as RUNX1 and ASXL1 that are more frequent in older CN-AML patients and were found to associate with differential expression of lncRNAs\(^26\) could not be tested in younger CN-AML patients, because too few younger patients harbored these mutations. Of note, mutations in the TET2 gene showed no correlation with differential expression of lncRNA molecules in either older\(^26\) or younger CN-AML patients, despite their impact on the epigenome\(^49\) and adequate numbers of patients in both studied cohorts.

To gain insights into biologic pathways affected by differences in the lncRNA score, we investigated correlations between mRNA and miR expression signatures and lncRNA scores. In concordance with the adverse outcome that unfavorable lncRNA scores bestow, a number of previously described oncogenes and oncomiRs were found over-expressed in patients with unfavorable lncRNA score status. Similarly, genes with reported tumor-suppressive activity were found downregulated in this patient group. Only a small fraction of these transcripts have been reported in gene mutation-related mRNA signatures or other prognostic gene-expression signatures.\(^14\) These findings indicate that, in addition to being independent of prognostic mutations, the differential expression of prognostic lncRNAs may regulate distinct molecular pathways in CN-AML. Notably, 5 important mediators of the AP-1 pathway (ATF3, FOS, FOSB, JUN, and MAFF) were found upregulated in patients with unfavorable lncRNA scores. The high number of cell cycle regulators and proliferation-inducing kinases that were also upregulated in this patient group is consistent with aberrant activation of the AP-1 pathway.

In this work, we used whole transcriptome sequencing techniques to identify and measure the expression of prognostic lncRNA molecules in younger adults with CN-AML. While this technology is becoming rapidly cheaper and widely available, it will most likely continue to serve...
as a research tool rather than a diagnostic test to guide patient treatment. Despite this, alternative techniques for measuring RNA transcripts in a clinically applicable manner are available and are used to risk stratify patients with certain solid malignancies. Similar assays could be developed in order to obtain targeted measurements of prognostic lncRNAs in a fast and clinically meaningful manner. The potential of such assays to improve risk stratification of AML patients should be evaluated in future prospective clinical studies.

Acknowledgments

The authors would like to thank: Donna Bucci and Wacharaphon Vongchucherd of The Alliance NCTN Biorepository and Biospecimen Resource for sample processing and storage services, Karl Kroll for technical support, and Lisa J. Sterling and Christine Finks of The Ohio State University, Comprehensive Cancer Center, Columbus, OH for data management.

Funding

This work was supported by the National Cancer Institute of the National Institutes of Health under Award Numbers CA180821 and CA180882 (to the Alliance for Clinical Trials in Oncology), CA077658, CA180840, CA150864, CA140158, CA16058, and CA197734. This work was also supported in part by the Leukemia Clinical Research Foundation, D Warren Brown Foundation, and the Pelotonia Fellowship Program. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

References


