The long non-coding RNA landscape in juvenile myelomonocytic leukemia

Juvenile myelomonocytic leukemia (JMML) is a rare and aggressive myelodysplastic and myeloproliferative disorder of early childhood. It is characterized by proliferation of granulocytic and monocytic cells. Currently, hematopoietic stem cell transplantation (HSCT) is the standard of care and results in long-term overall survival (OS) of only 50-60%. In approximately 90-95% of patients, hyperactivation of the RAS/MAPK pathway can be observed through typical mutations in NRAS and KRAS (20-25%), PTPN11 (35%), NFI (10-15%) or CBL (10-15%) genes. This underlying molecular defect can in part be used to guide patient management (e.g. no HSCT in patients with mutations in CBL). Additional somatic mutations, including mutations in SETBP1, JAK3 and members of the polycomb repressive complex 2 (PRC2), are associated with worse outcome. In addition, LIN28B overexpression and DNA hypermethylation profiles can identify JMML patients with a more aggressive clinical course.

The non-coding transcriptome consists of a variety of different RNA types such as tRNAs, snoRNAs, microRNAs, circRNAs and long non-coding RNAs (lncRNAs). LncRNAs are a class of RNA genes with a minimum length of 200 nucleotides. Although only a fraction of known lncRNAs have been functionally characterized, there is growing evidence of their involvement in a variety of biological processes, human diseases and malignancies.

To explore lncRNA expression in JMML, RNA from isolated mononuclear cell preparations from 44 previously untreated JMML patients (median age: 2.02 years) and 7 bone marrow samples from healthy controls (siblings screened for transplantation; median age 6.83 years) was profiled on a custom designed Agilent (8x60K) microarray platform described by Volders et al. The dataset (GEO Series accession n. GSE71449) had been previously used in an mRNA expression analysis, leading to the identification of overexpression of LIN28B in a fetal-like subgroup of JMML patients. Details of patients’ characteristics can be found in the Online Supplementary Table S1.

Only probes (n=23,728) representing known lncRNAs were withheld for further analyses. Therefore, the probes were mapped against the most up-dated version of different lncRNA databases, including Incpiedia (www.lncipedia.org), Hugo Gene Nomenclature Committee (HGNC), and Long Noncoding RNA Database v.2.0 (www.lncrnadb.org). The Incpiedia 3.1 annotations for lncRNAs have been used throughout this manuscript. Differential gene expression analysis, with samr package (R BioConductor) and BRB-Array tools, revealed a total of 300 differentially expressed lncRNAs (adjusted P<0.05; FDR<0.10) between JMML patients and controls, of which 15 up-regulated and 285 down-regulated in JMML patients (Figure 1A and B and Online Supplementary Table S2). The 5 most up-regulated

Figure 1. Long non-coding RNA transcriptome in juvenile myelomonocytic leukemia (JMML). (A) Diagonal plot representing the average expression versus the log2 fold change added to the average expression of lncRNAs when comparing JMML (n=44) to healthy donors (n=7). lncRNAs significantly up-regulated in patients are represented in red and lncRNAs significantly down-regulated in blue. The names of the top 5 up- and down-regulated lncRNAs based on log2 fold change are indicated. (B) Circos plot showing lncRNAs and mRNAs on human chromosomes. From the outside in, the first layer of the Circos plot represents a chromosome map of the human genome. Black and white bars are chromosome cytobands, and red bars represent centromeres. The second circle shows all differentially expressed coding genes when comparing JMML patients with healthy donors, and the inner circle shows all differentially expressed lncRNAs. Up-regulated lncRNA and mRNA genes are represented in red and down-regulated genes in blue. The height of the bars is representative for the fold change between JMML and healthy donors. Names of lncRNA genes with fold change of more than 2 between JMML and healthy donors are indicated between the circles with mRNA and lncRNA genes for the down-regulated lncRNAs and in the middle of the plot for the up-regulated lncRNAs. (C) Density plots representing the average of the normalized expression of all lncRNAs and mRNAs detected by the array in all samples (left) and in JMML samples only (right).
lncRNAs based on log2 fold change were Inc-THADA-4, Inc-SYF2-2, Inc-BASP1-3, Inc-ACSL1-1 and Inc-CROCC-1, whereas Inc-SUPT3H-1, Inc-SOX6-1, Inc-HMCN1-2, Inc-ACVR2A-2 and Inc-RAN-1 were the top 5 most down-regulated lncRNAs. As shown in other disease entities, the mean expression level of lncRNAs was nearly 2-fold lower as compared to mRNA (P<0.0001) (Figure 1C). To exclude the impact of differences in cellular composition between JMML and healthy controls, we compared our dataset with an IncRNA expression atlas of flow cytometry sorted blood cell populations (Online Supplementary Table S3). This analysis revealed 2 of 15 differentially up-regulated IncRNAs in JMML to be specific for granulocytic lineage (Inc-ACSL1-1 and Inc-BASP1-3) and none for monocytic differentiation. Conversely, one monocytic specific IncRNA showed significant downregulation in JMML patients (Inc-AL713998.1-2).

To explore the functional characteristics of differentially expressed IncRNAs, we used a guilt-by-association approach, based on the hypothesis that non-coding and protein coding genes belonging to the same biological pathways are co-ordinately regulated. Spearman’s rho values were calculated between IncRNAs of interest and all protein coding genes on the array. The resulting ranked gene lists served as inputs for pre-ranked gene-set enrichment analysis (GSEA), using the MSigDB v.6.0 C2 collection and FDR<0.1. This revealed an overlap between the 15 curated gene sets most positively correlated with the top 5 up-regulated IncRNAs (Figure 2 and Online Supplementary Table S4). Using the same approach for the top 5 down-regulated IncRNAs, a strong overlap both in gene sets positively correlated with these IncRNAs, as well as negatively correlated was observed (Figure 2 and Online Supplementary Table S4). Interestingly, several gene sets showed positive correlation with up-regulated IncRNAs as well as negative correlation with down-regulated IncRNAs, suggesting a common pathway. Notably, these gene sets are related to myeloid cell development, dendritic cell maturation, neutrophil function at skin wound, systemic lupus erythematosus, and different acute myeloid leukemia (AML) phenotypes.

Extensive cytogenetic and mutational data were available for our patient cohort and we evaluated whether the expression of IncRNAs was associated with specific JMML subgroups based on clinical, cytogenetic and mutational characteristics. First, patients with elevated fetal hemoglobin (HbF) levels for age were compared with patients with normal HbF levels. Only 8 IncRNAs were found differentially expressed between these two subgroups (Figure 2).
groups, all up-regulated in the group with normal HbF levels for age (adjusted \( p<0.05 \)) (Online Supplementary Table S5). Given the strong downregulation of lnc-C2orf54-1 and lnc-ATP6V0D2-2 in JMML patients with elevated HbF for age (log fold change -7.4 and -2.8, respectively), we predicted that these lncRNA would be anti-correlated with genes involved in HbF regulation. Indeed, HBG1 (Spearman rho value (\( \rho \)): -0.63 and -0.61), HBBP1 (Hemoglobin Subunit Beta Pseudogene 1; \( \rho \): -0.60 and -0.41), HBE1 (\( \rho \): -0.37 and -0.45) and LIN28B (\( \rho \): -0.58 and -0.70), belonged to the top anti-correlated protein coding genes, suggesting that both lnc-C2orf54-1 and lnc-ATP6V0D2-2 could be part of a molecular network controlling the formation of HbF. Moreover, lnc-ATP6V0D2-2 is associated with event-free survival (EFS) in cytogenetically normal AML and is down-regulated in patients with mutations in RUNX1. Subsequently, patients with monosomy 7 were compared to patients with a normal karyotype. To avoid confounding effects due to RAS gene mutations, a 2-class analysis (monosomy 7 vs. normal karyotype) blocking for RAS mutations through a randomized block design was applied. This analysis revealed a total of 36 lncRNAs significantly up-regulated in the monosomy 7 group and 5 lncRNAs down-regulated. Of note, 6 of 8 lncRNAs (lnc-C2orf54-1, PRINS, lnc-ATP6V0D2-2, lnc-CCT8-1, lnc-SH3BGRL2-1, lnc-FRMD4A-1) down-regulated in patients with elevated HbF levels for age were significantly up-regulated in JMML patients with monosomy 7. This observation corresponds with the established inverse relation between monosomy 7 and elevated HbF levels in JMML patients.

Next, we performed pairwise analyses comparing patients with PTPN11 (n=18), KRAS (n=11), NRAS (n=8), NF1 (n=4) mutations or without any of the above mutations (quadruple negative; n=3) to all other patients (Online Supplementary Table S5). Patients harboring a KRAS mutation showed 26 differentially expressed lncRNAs (17 up-regulated and 9 down-regulated) compared to JMML patients without this mutation. NF1 mutated patients were characterized by 62 down-regulated lncRNAs whereas none were significantly up-regulated. PTPN11 mutated patients were marked by differential expression of 8 lncRNAs (6 up-regulated and 2 down-regulated) and 4 lncRNAs were differentially expressed in NRAS mutated JMML (all up-regulated). Remarkably, in quadruple negative patients, only one lncRNA, lnc-ATP10A-2, was differentially up-regulated with high fold change (FC=14.4). This lncRNA is anti-correlated with BAP1, a key regulator of many cancer-associated pathways.

Finally, we studied differentially expressed lncRNAs between JMML patients with and without LIN28B overexpression. Sixty-one lncRNAs were differentially expressed of which 60 were up-regulated in JMML patients with LIN28B overexpression (Online
Supplementary Table S5). Interestingly, both MEG3 and MEG8 were up-regulated in LIN28B overexpression and are known to be involved in regulation of pluripotency and fetal development.14

As the prognostic value of coding gene expression signatures in JMML has previously been reported by Bresolin et al.,15 we hypothesized that lncRNA expression might influence clinical outcome. Indeed, univariable Cox regression analysis identified 45 individual lncRNAs highly associated with OS (P<0.001) (Online Supplementary Table S6) of which 7 lncRNAs (UCKL1-AS1, Inc-GPR63-5, Inc-ZSWIM2-3, Inc-RORB-1, Inc-NPFFR2-1, Inc-EMB-3 and Inc-CCDC90A-2) remained significantly associated with OS after multivariable Cox regression analysis with age category (< or > 2 years old), HbF status, karyotype and molecular status (P<0.001). Interestingly, these lncRNAs did not associate with known prognostic characteristics and pre-ranked GSEA could not identify a common (anti-)correlated gene set. We devised an lncRNA score, derived as a linear combination of the expression of the significant lncRNAs (method described in Garzon et al.) and dichotomized patients based on lncRNA score (scores below the median were classified as favorable and scores above as unfavorable). Patients with unfavorable lncRNA score had a significantly lower OS and EFS after HSCT (log rank test P<0.001) (Figure 3A and C). Next, we compared the expression levels of 29,654 mRNA probes (21,659 unique genes) between patients with favorable and unfavorable lncRNA score. We identified 549 coding genes with differential expression between the two groups (absolute FC ≥ 2) (Figure 3B). Together, these results suggested that lncRNAs might be of prognostic significance in JMML. However, studies in independent patient cohorts are necessary to validate the prognostic value of lncRNA expression in JMML.

In conclusion, we documented for the first time the lncRNA landscape in 44 JMML patients and have associated lncRNA expression with clinical and molecular characteristics. We demonstrated that JMML patients exhibit a distinct lncRNA expression profile compared to healthy controls and different JMML subgroups have distinct lncRNA signatures contributing to disease heterogeneity. Follow-up studies are needed to confirm the prognostic value of lncRNAs and to perform an integrative analysis of lncRNA expression profiles with mutational patterns, DNA methylation profiles and clinical patients’ characteristics. This study paves the way for further functional research on the role of lncRNAs in JMML biology, and their diagnostic and therapeutic application.

References


