Deregulated expression of circular RNAs in acute myeloid leukemia

Susanne Lux,1 Tamara J. Blätte,1,2 Bernhard Gillissen,2 Antje Richter,2 Sibyle Cocciardi,1 Sabrina Skambraks,1 Klaus Schwarz,3,4 Hubert Schrezenmeier,3,4 Hartmut Döhner,1 Konstanze Döhner,1 Anna Dolnik,1,2 and Lars Bullinger1,2

1Internal Medicine III, University Hospital Ulm, Ulm, Germany; 2Department of Hematology, Oncology, and Tumorimmunology, Charité University Medicine, Berlin, Germany; 3Institute for Transfusion Medicine, University of Ulm, Ulm, Germany; and 4Institute for Clinical Transfusion Medicine and Immunogenetics Ulm, German Red Cross Blood Donor Service Baden-Württemberg–Hessen, Ulm, Germany

Key Points
• Compared with healthy HSPCs, circRNA expression is deregulated in AML.
• Deregulation mainly affects leukemia-relevant genes, highlights key signaling pathways, and is tightly linked to leukemia dedifferentiation.

Circular RNAs (circRNAs) are dynamically regulated during differentiation and show cell type–specific expression, which is altered in cancer and can have a direct impact on its various hallmarks. We hypothesized that circRNA expression is deregulated in acute myeloid leukemia (AML) and that circRNA candidates might contribute to the pathogenesis of the disease. To identify leukemia-associated and differentiation-independent changes in circRNA expression, we determined the circular RNAome of 61 AML patients and 16 healthy hematopoietic stem and progenitor cell (HSPC) samples using ribosomal RNA–depleted RNA sequencing. We found hundreds of circRNAs that were differentially expressed between AML and healthy HSPCs. Gene set analysis found that many of these circRNAs were transcribed from genes implicated in leukemia biology. We discovered a circRNA derived from the T-cell transcription factor gene B cell CLL/lymphoma 11B, circBCL11B, which was exclusively expressed in AML patients, but not detected in healthy HSPCs, and associated with a T-cell–like gene expression signature. We were able to validate this finding in an independent cohort of 332 AML patients. Knockdown of circBCL11B had a negative effect on leukemic cell proliferation and resulted in increased cell death of leukemic cells, thereby suggesting circBCL11B as a novel functionally relevant candidate in AML pathogenesis. In summary, our study enables comprehensive insights into circRNA expression changes upon leukemic transformation and provides valuable information on the biology of leukemic cells and potential novel pathway dependencies that are relevant for AML therapy.

Introduction

Circular RNAs (circRNAs) are a class of abundant and stable RNAs that result from ligation of a downstream splice donor to an upstream splice acceptor.1 Some circRNAs have been shown to serve as microRNA (miRNA) sponges in the cytoplasm,2 whereas others in the nucleus can enhance the expression of their parental gene through association with RNA polymerase II.3 Recently, circRNAs that can be translated into novel peptides have been discovered.4-6 This challenges the classification of circRNAs as “noncoding” and adds another level of complexity to this RNA species that only recently attracted the attention of the scientific community. Many groups have now shown that circRNAs are deregulated in cancer and leukemia7-9 and can contribute to the pathogenesis of the disease by influencing the hallmarks of cancer.10

circRNA deregulation might also have an impact in acute myeloid leukemia (AML), a hematologic malignancy that is characterized by a block in the differentiation of myeloid progenitors and an increased self-renewal capacity of the respective immature cells. Some transforming events are strongly
associated with distinct gene expression signatures, including cases with mutated nucleophosmin (NPM1) or core binding factor (CBF) leukemias, whereas ~10% to 20% of AML cases also carry mutations in splicing factors, such as SF3B1, SRSF2, and U2AF1. Noncoding RNA species like miRNAs and long noncoding RNAs have been shown to contribute to the pathogenesis and global deregulation of gene expression in AML.11-14

In accordance, in 1 of our previous studies we described dynamically regulated circRNA expression throughout myeloid differentiation of healthy hematopoiesis, as well as an alteration in circRNA expression patterns in 10 AML cases.9 Because these AML-associated circRNA expression changes were also dependent on the underlying leukemic driver mutation, we now wanted to comprehensively characterize the overall deregulation of circRNA expression in AML and elaborate the differences in distinct AML subgroups. Therefore, we determined the circular RNAome of a large AML cohort (n = 61), including samples from patients with mutated NPM1 (NPM1mut), CBF leukemias with a translocation t(8;21) or an inversion inv(16), and AML patients with mutations in the splicing factors SF3B1, SRSF2, or U2AF1. To define the leukemia-associated circular expression patterns in these cases, we compared the circRNA expression of primary AML cells with the 1 of the hematopoietic stem and progenitor cells (HSPCs) derived from healthy controls (n = 16). Findings were validated by sequencing full-length circRNAs using ultra-long read Oxford Nanopore technology. We further functionally characterized a circRNA that was exclusively expressed in AML and not detected in healthy HSPCs. Its expression was associated with a distinct gene expression signature in AML patients, which we could validate using a previously published independent cohort of 332 AML patients.15

Methods

Patients and healthy controls

Diagnostic samples were taken from 61 AML patients who were chosen from larger cohorts of patients based on sample availability and the availability of exome or targeted resequencing data. For 18 of these patients an additional matched sample was collected at the time of relapse. The patients were enrolled in the AMLSG BiO Registry study (NCT01252485). The median age of the patients was 50 years (range, 21-63 years). Healthy CD34+ HSPCs mobilized from the bone marrow to the peripheral blood via granulocyte colony-stimulating factor were used as a control. Informed patient consent was obtained for the study in accordance with the principles of the Declaration of Helsinki, and ethical approval was obtained from the local ethics committee.

An independent test data set of 332 AML patients was downloaded from the Gene Expression Omnibus (GEO; accession number GSE137851).15

RNA-sequencing and identification of circRNAs

Primary AML bone marrow (n = 50) and peripheral blood (n = 29) samples were enriched for >85% mononuclear cells by Ficoll-Hypaque density gradient centrifugation; healthy controls were already enriched for CD34+ HSPCs. RNA was isolated and RNA-sequencing (RNA-seq) libraries were prepared as previously described using a protocol depleting ribosomal RNA.16 Subsequently, sequencing on a HiSeq 2000 (Illumina, San Diego, CA) was used to obtain 100-bp paired-end reads yielding an average coverage of 76 (±14) × 106 reads per sample. The average ribosomal RNA content was 5.6%. Moreover, we used an independent test data set (GEO accession number GSE137851) with 50-bp paired-end reads.

Reads were aligned and quantified using STAR,17 and reads mapping to circular junctions were identified using an in-house analysis pipeline, as previously described.9 In short, chimeric junctions with the splice donor lying downstream of the splice acceptor were extracted. Moreover, only backsplice junctions that overlapped known exon-exon borders of the same gene, with a minimum of 15 bp in each exon of the junction (chimSegmentMin and chimJunctionOverhangMin = 15), were considered; the paired mate had to map between the acceptor and donor position. Downstream analyses followed 2 approaches: (1) for the gene-based approach, all circular-derived reads were summed for each parental gene, regardless of the specific circRNA isoform(s) expressed, and (2) for the circRNA-based approach, reads were summed per backsplice junction. circRNA-derived reads were normalized and rlog transformed using DESeq2.18

Data analysis

Principal component analysis (PCA) was performed based on the 500 circRNA-expressing genes with the highest variance across all samples using DESeq2. Heat maps were generated using the R package pheatmap.19 Unsupervised clustering was performed based on Euclidean distance and the 5000 circRNA-expressing genes with the highest coefficient of variation across all samples.

Thresholds for differential expression were set at log2 fold change ≥ 0.6 and false discovery rate (FDR) ≤ 0.1. Gene set overlap computation and gene ontology analyses were performed using the Molecular Signature Database20,21 (Broad Institute, http://software.broadinstitute.org/gsea/msigdb/compute_overlaps.jsp, gene set collections H, C2, and C7). Enrichment with an FDR q < 0.1 was considered significant. Gene Set Enrichment Analysis (GSEA) was performed using a gene list preranked by fold change estimates from DESeq2, and FDR q < 0.25 and a high normalized enrichment score were considered informative.

Oxford Nanopore sequencing validation of circRNA expression

For selected circRNAs, polymerase chain reaction (PCR) products were generated from cell line (supplemental Methods) and primary AML patient–derived RNA via the use of divergent primers and purification using a QiAquick PCR purification kit (Qiagen), according to the manufacturer’s instructions. Primer sequences are listed in supplemental Table 1. Libraries were prepared from 1 μg of pooled purified PCR products using a SQK-MAP006 or SQK-NSK007 Nanopore Sequencing Kit (Oxford Nanopore Technologies Ltd, Oxford, United Kingdom), following the “Amplicon sequencing for the MinION device” protocol. After sequencing, FAST5 files were converted to FASTQ using poretools.22 Target reads containing primer sequences were extracted using the bash grep command, and alignment of these reads to the target genes was performed using the National Center for Biotechnology Information’s Basic Local Alignment Search Tool.23
circRNA abundance in AML patients is comparable to that of healthy HSPCs

Across all 77 ribosomal RNA–depleted diagnostic AML and control HSPC RNA-seq samples, we detected circRNAs from a total of 9683 genes. In most cases (79%), >1 circular isoform was detected per gene, with a median of 3 and a range of 1 to 126 different circular isoforms per gene (Figure 1A). Among the highly expressed genes (mean read count across all AML samples ≥ 20), we observed a tendency toward more circRNA isoforms for genes with a higher exon number (Figure 1B).

An overview of the average number of circRNAs that were detected in the control and AML sample groups, as well as the number of genes from which they were transcribed, is given in Table 1. Total number of circRNAs and the average number of circRNAs per gene that we detected did not differ between healthy HSPCs and diagnostic AML samples, whereas the number of different circRNAs detected was slightly higher in AML relapse samples. Moreover, upon closer inspection of the defined AML subgroups (NPM1mut AML, CBF leukemia, and AML with splicing factor mutations), the presence of mutations in splicing factors did not alter the number of detected circRNAs compared with the other AML patients (Table 1).

To quantify the overall abundance of circRNAs in comparison to parental messenger RNA (mRNA) abundance, we calculated the ratio of all circular-derived (backsplice-supporting) reads/total reads per sample. To distinguish circRNAs expressed in AML patients from those in healthy HSPCs, we detected 30 circRNAs derived from 26 genes that were differentially expressed in all AML patients at diagnosis with the one in healthy HSPCs, we detected 30 circRNAs derived from 26 genes that were differentially expressed in all AML cases (Figure 2E; Table 1). With regard to distinct AML subgroups, we detected 124 differentially expressed circRNAs derived from 93 genes when comparing NPM1mut AML and CBF leukemia (Figure 2F; Table 1), among which 18 patients, in addition to the diagnostic sample, RNA-seq was conducted at the time of relapse. Furthermore, we used CD34+ HSPCs (n = 16) as a healthy control to compare with AML blasts to determine leukemia-associated changes in the circular RNAome.

circRNA expression patterns are different in AML compared with healthy HSPCs

Although the overall abundance of circular transcripts was not altered in AML compared with healthy HSPCs, we next investigated whether, for well-defined AML subgroups (ie, NPM1mut patients and CBF leukemias), we could also observe defined changes in the circular RNAome and whether the pattern of circRNA expression is altered by mutations in splicing factors.

We made use of unsupervised PCAs, based on the 500 circRNA-expressing genes with the highest coefficient of variation across all samples, to visualize global circRNA expression in all AML patients at diagnosis (Figure 2A), as well as in distinct AML subgroups (ie, NPM1mut patients (Figure 2B), CBF leukemia patients (Figure 2C) and AML patients with mutations in splicing factors (Figure 2D; supplemental Figure 1E) in contrast to healthy HSPCs. In these plots, based on PC1 and PC2, healthy HSPCs were clearly separated from leukemic samples. Unsupervised hierarchical clustering, based on the top 5000 circRNA-expressing genes with the highest coefficient of variation, also resulted in a clear separation of healthy HSPCs and all AML subgroups (supplemental Figure 1A-D). For NPM1mut AML and CBF leukemia, the intersample variation was higher than in healthy cells; nevertheless, they exhibited a distinct group-specific circular expression pattern. As for the CBF leukemias, inv(16) and t(8;21) cases could clearly be separated based only on their circRNA expression (Figure 2C,G; supplemental Figure 1C). Although circRNA expression in patients with mutations in 1 of the 3 splicing factors was more heterogeneous, circRNA expression in these AMLs was still distinguishable from healthy HSPCs.

In a supervised analysis, we tested for differentially expressed circRNAs in all AML subgroups. A detailed list of differentially expressed circRNAs and linear mRNAs for all comparisons is provided in supplemental Table 4. By comparing circRNA expression in all AML patients at diagnosis with the one in healthy HSPCs, we detected 30 circRNAs derived from 26 genes that were differentially expressed in all AML cases (Figure 2E; Table 1). With regard to distinct AML subgroups, we detected 124 differentially expressed circRNAs derived from 93 genes when comparing NPM1mut patients and healthy HSPCs (Figure 2F; Table 1), 138 differentially expressed circRNAs derived from 98 genes in CBF leukemias vs healthy HSPCs (Figure 2G; Table 1), and 42 differentially expressed circRNAs derived from 39 genes when comparing AML cases with mutations in splicing factors and healthy HSPCs (Figure 2H; Table 1). These differentially expressed circRNAs were sufficient to clearly separate each cohort from healthy HSPCs by hierarchical clustering. Although some of the subgroup-associated differentially expressed circRNAs were found to be deregulated in all of the AML subgroups (n = 20; supplemental Figure 1F), many circRNA isoforms were deregulated in only 1 of the AML subgroups (ie, 40%, 51%, and 24% of the differentially expressed circRNAs in NPM1mut, in CBF leukemia, and in AMLs with splicing factor mutations, respectively) (supplemental Figure 1F).

AML-associated circRNA expression patterns are enriched for leukemia-relevant genes

Because the mRNA expression of genes is also known to be disturbed in AML, our aim was to investigate whether the same
genes that are differentially expressed at the linear mRNA level when comparing AML and healthy HSPC samples also differentially express circRNAs. The numbers of genes differentially expressed at the linear mRNA level (DE_linear; Figure 2I-L), differentially expressing circRNAs (DE_circ; Figure 2I-L), and differentially expressing mRNA and circRNAs (DE_both; Figure 2I-L) in the different groups are shown in Table 1. Of note, in more than half of the genes with differential circRNA expression, this was independent of parental gene regulation. These DE_circ genes were significantly enriched for cancer-relevant genes (KEGG_PATHWAYS_IN_CANCER gene set (Figure 2J-L; Table 2)), whereas this enrichment was not observed among genes with differential mRNA expression (DE_linear).

Genes that showed differential circRNA and mRNA expression were significantly enriched for hematopoietic stem cell signature genes...
In NPM1mut AML, the previously described NPM1 signature genes [VERHAAK_AML_WITH_NPM1_MUTATED gene set] were enriched in the differentially expressed circRNAs (Figure 2J; Table 2); in CBF leukemia patients, a respective CBF signature [ROSS_AML_CBF gene set] (Figure 2K; Table 2) was detected. Comparing AML with splicing mutations vs healthy HSPCs revealed an enrichment of other known leukemia-relevant genes among the genes with differential circRNA expression, including gene sets previously defined in AML with rearrangements involving the mixed-lineage leukemia (MLL) gene (Figure 2L; Table 2).

### Circular to linear proportion of selected target circRNAs

We selected 9 circRNA-expressing genes of potential pathogenic relevance for further investigation: AF4/FMR2 family member 2 (AFF2), AT-rich interaction domain 1B (ARID1B), B cell CLL/lymphoma 11B (BCL11B), casitas B-lineage lymphoma protooncogene (CBL), Fms-related tyrosine kinase 3 (FLT3), genetic suppressor element 1 (GSE1), myeloid ecotropic viral integration site 1 homolog (MEIS1), runt-related transcription factor 1 (RUNX1), and spleen focus forming virus proviral integration oncogene (SPI1) coding for the transcription factor PU.1. Circular target gene expression in the 16 control HSPC samples and 61 AML samples ranged from low (CBL, MEIS1, RUNX1; supplemental Figure 2C,F,G), to intermediate (ARID1B, SPI1; supplemental Figure 2B,H), to high (BCL11B, Figure 3B; AFF2, FLT3, GSE1, supplemental Figure 2A,D,E) based on normalized circular-derived read counts. In addition to the circRNA expression in general, we were interested in the ratio of circular/linear RNAs that are expressed by a gene, the so-called “circular to linear proportion” (CLP). Linear gene expression was calculated by taking the mean read count of reads mapping to the 2 linear junctions flanking the exons that take part in the circularization, while circular expression was again defined by the number of reads mapping to the backsplice junction (supplemental Methods). The CLP was highest for the AFF2 gene, with a median CLP of 0.85 in all AML patients, indicating that the abundance of circular transcripts is comparable to that of the linear mRNA isoforms, whereas circAFF2 is even more abundant in AML patients with mutations in splicing factors (supplemental Figure 2A). Circular transcripts are also very abundantly transcribed from the GSE1 gene, with a median CLP of 0.35 in NPM1mut patients, whereas circGSE1 is less common in all AML patients with mutations in splicing factors (median CLP, 0.007; supplemental Figure 2E). For 12,888 circRNAs, raw read counts mapping to circular and linear junctions and the respective CLPs are given for all 61 AML samples at diagnosis and the 16 healthy HSPC control samples (supplemental Table 5).

### Ultralong read full-length circRNA sequencing using Oxford Nanopore technology

To validate our RNA-seq findings and to also gain full-length sequence information about the target gene circRNAs, we made use of the ultralong read Oxford Nanopore sequencing technology. Using divergent primers, we produced circular-derived PCR products that were then pooled and sequenced on the Oxford Nanopore Technologies MinION (Figure 3D; supplemental Figure 3). Exemplary circRNA sequencing results are shown for all target genes. In addition to 7 AML cell lines (supplemental Methods), we validated circRNAs detected in the RNA-seq data from 4 AML patients. In brief, we were able to validate 15 of 17 target gene circRNAs previously detected by RNA-seq for 7 genes (ARID1B, BCL11B, CBL, FLT3, GSE1, MEIS1, and RUNX1) in those patients (supplemental Table 6). Ten of the respective circRNAs were expressed at very low levels and only supported by a single read in the RNA-seq data. Of these, we nevertheless successfully validated 8 circRNAs, and we detected 18 additional circRNA isoforms for ARID1B, CBL, FLT3, MEIS1, and RUNX1 by Oxford Nanopore sequencing (supplemental Table 6). We did not detect intron retention within the circles, but exon skipping events were detected for circRNAs derived from ARID1B, where exon 3 was absent in the circRNA isoform of the gene (supplemental Figure 3B).

### CircBCL11B is associated with AML samples and is not detected in healthy HSPCs

The BCL11B gene was the only gene in our dataset for which circRNAs were exclusively expressed in AML and not in healthy
HSPCs. BCL11B is known as a key transcription factor for T-cell development and function. In our study, the main circular transcript isoform of the BCL11B gene, circBCL11B, is formed through circularization of exon 2 (Figure 3A-B) and could be validated using Oxford Nanopore sequencing technology, with no detectable alternative splicing events within the circle (Figure 3D). Although circBCL11B expression was restricted to AML samples, not all AML patients expressed this circRNA (Figure 3B). Of note, circBCL11B was not restricted to any of the subgroups that we included in
the study. The CLP of circBCL11B compared with BCL11B mRNA was comparably high in all AML patients, with a median CLP of 0.38 (Figure 3C), meaning that only around three quarters of all RNA molecules derived from the BCL11B gene are linear mRNAs that are translated into protein, whereas around one quarter of the RNA molecules are circularized and

Table 2. Gene sets enriched among genes differentially expressing circRNAs between AML samples and healthy HSPCs

| MSigDB gene set                        | NPM1<sup>mut</sup> vs healthy HSPCs | CBF vs healthy HSPCs | Splice factor mutation vs healthy HSPCs |
|----------------------------------------|--------------------------------------|----------------------|----------------------------------------|
|                                        | DE list | FDR | DE list | FDR | DE list | FDR |
| JAATINEN_HEMATO-POIETIC STEM_CELL_UP   | DE_both | <1 E-04 | DE_both | <1 E-04 | DE_both | 2.0 E-02 |
|                                        | DE_circ | 1.1 E-03 | DE_circ | <1 E-04 | DE_circ | <1 E-04 |
| KEGG_PATHWAYS_IN_CANCER                | DE_circ | 2.0 E-04 | DE_circ | 3.0 E-04 | DE_circ | 5.0 E-02 |
| VERHAAK_AML_WITH_NPM1_MUTATED          | DE_both | 1.2 E-03 |         |       |         |       |
| ROSS_AML_CBF                           |         |       |         |       | DE_circ | 2.1 E-03 |
| MULLIGHAN_MLL_SIGNATURE_1_UP           |         |       |         |       |         |       |
| MULLIGHAN_MLL_SIGNATURE_2_UP           |         |       |         |       |         |       |
| YAGI_AML_WITH_REARRANGED               |         |       |         |       |         |       |

MSigDB, Molecular Signature Database.

Figure 3. CircBCL11B is exclusively expressed in AML. (A) Schematic diagram of the BCL11B gene consisting of 4 exons and circBCL11B resulting from a circularization of exon 2. The backsplice site is indicated by a red line. (B) Normalized circular read counts of circBCL11B in 61 AML patients and 16 healthy HSPC samples, as detected in the ribosomal RNA–depleted RNA-seq data. (C) CLP of BCL11B circRNA junction counts to BCL11B linear mRNA junction counts. (D) CircBCL11B-derived PCR products were generated using different AML cell lines and the CML cell line K562 and divergent primers (red arrows) specific for exon 2 of BCL11B. PCR products were pooled and sequenced on an Oxford Nanopore MinION. Nanopore reads were aligned to BCL11B using the National Center for Biotechnology Information’s Basic Local Alignment Search Tool. A sequencing result of circBCL11B is shown; the fusion site is indicated by a red line. NTC, no template control.
fulfill different functions. With a median value of 1.1, the CLP of BCL11B was highest in SRSF2mut patients; the proportion of circular transcripts was higher than the linear mRNA counterparts and was lowest in U2AF1mut AML patients (median CLP 0.1; Figure 3C).

To investigate in more detail what distinguishes patients with high circBCL11B expression from those with no circBCL11B expression, irrespective of the leukemic driver, we looked for genes that were differentially expressed between these 2 groups. In this analysis, 29 leukemia cases were included that were allocated to a circBCL11B_high group (top 25% of AML cases with highest normalized circBCL11B read counts; n = 12) or a circBCL11B_low group (all AMLs with normalized circBCL11B read count = 0; n = 17). GSEA revealed that genes that were more highly expressed in circBCL11B_high were associated with genes that were highly expressed in T cells compared with myeloid cell types (Figure 4A-B). Moreover, genes that are normally downregulated in hematopoietic stem cells compared with T cells are highly expressed in the circBCL11B_high leukemic blasts (Figure 4A). This indicates a more T-cell like, less myeloid, and less stem cell–like gene expression pattern in circBCL11B_high AML cases.

Vice versa, genes that were more highly expressed in the circBCL11B_low AML group were associated with genes with low expression in T cells compared with myeloid cell types (Figure 4A-B). Instead, highly expressed genes in circBCL11B_low

---

**Figure 4. CircBCL11B expression is associated with a T-cell–like gene expression signature.** Gene set enrichment analysis of genes differentially expressed between circBCL11B_high and circBCL11B_low AML patients. (A-B) Results of the original cohort (29 patients included in this analysis; with n = 12 circBCL11B_high and n = 17 circBCL11B_low). (C-D) Results of the independent test cohort (of 332 patients, 19 were assigned to circBCL11B_high, and 155 were assigned to circBCL11B_low). In panels A and C, bars show the normalized enrichment score (NES) of gene sets enriched in circBCL11B_low patients (left, negative NES) and in circBCL11B_high patients (right, positive NES). FDR is color coded as indicated. (B,D) Exemplary GSEA plots of gene sets enriched in circBCL11B_low (left panels) and in circBCL11B_high patients (right panels).
patients were associated with genes that were upregulated in hematopoietic stem cells compared with T cells (Figure 4A). This indicates a more myeloid and stem cell–like gene expression signature in circBCL11B_low AML cases. The same observations were also made when comparing circBCL11B_high/low groups within the well-defined NPM1mut and CBF cohorts and were less pronounced within the splicing factor–mutated AML cohort (supplemental Table 7).

Figure 5. shRNA-mediated circBCL11B knockdown has a negative effect on AML cell proliferation in the absence of FCS, whereas BCL11B protein levels remain unchanged. (A) Schematic diagram of the shRNA design specific for the circBCL11B backsplice sequence and confirmation of the shRNA-mediated knockdown of circBCL11B in OCI-AML5 cells via PCR. A scrambled shRNA served as a control. (B) Western blot analysis of BCL11B protein expression in OCI-AML5 (top), NB-4 (middle) and MV4-11 (bottom) cells generated by lentiviral transduction with a vector containing a scrambled shRNA or an shRNA against circBCL11B. Actin served as a loading control. (C) Proliferation of AML cell lines OCI-AML5 (top panel), NB-4 (middle panel), and MV4-11 (bottom panel) transduced with scrambled shRNA or circBCL11B shRNA. Cells were cultured in the presence or absence of FCS. *P < .05, Student t test.
Independent validation cohort confirms association of circBCL11B expression with a T-cell–like gene expression signature

To further validate our findings for circBCL11B, we used a previously published RNA-seq data set.15 Of 332 AML patients from this study, we found 76 cases with detectable circBCL11B expression. Of these, in accordance with our study, we assigned 19 cases to the circBCL11B_high group (top 25% of patients with highest normalized circBCL11B read counts). All patients with normalized circBCL11B and total BCL11B read count = 0 were assigned to the circBCL11B_low group (n = 155). In agreement with our findings, this large independent validation cohort also showed an enrichment of genes expressed by T cells in the genes that were more highly expressed in circBCL11B_high patients (Figure 4C-D), whereas genes that were highly expressed in

Figure 6. shRNA-mediated circBCL11B knockdown increases apoptosis induced by FCS withdrawal. The respective circBCL11B shRNA or scrambled shRNA-transduced OCI-AML5 (A), NB-4 (B) and MV4-11 (C) cell lines were cultured in the presence or absence of FCS. After 72 hours, cells were harvested, and induction of apoptosis was analyzed by Annexin V–fluorescein isothiocyanate/PI staining and flow cytometry. Representative flow cytometry dot plots (left panels) and mean ± standard deviation of triplicates (right panels). *P < .05, Student t test.
circBCL11B_low patients were associated more with cells of the myeloid lineage and a hematopoietic stem cell–like signature (Figure 4C-D).

Functional evaluation of circBCL11B

To elucidate a possible pathogenic function of circBCL11B in leukemic cells, we performed short hairpin RNA (shRNA)-mediated knockdown of this circRNA in the AML cell lines OCI-AML5, NB-4, and MV-4-11, which showed high circBCL11B expression (Figure 3D). The successful knockdown of BLC11B circRNA was confirmed by PCR (Figure 5A). Knockdown of circBCL11B did not change BCL11B protein levels compared with a scrambled shRNA control in any of the 3 cell lines, as assessed by western blotting (Figure 5B).

Upon knockdown, cytological examination did not reveal the induction of differentiation (supplemental Figure 4). However, under stress conditions (ie, in culture without supplementation of fetal calf serum [FCS]), we observed that knockdown of circBCL11B has a negative effect on the proliferation of OCI-AML5, NB-4, and MV-4-11 cells compared with cells transduced with scrambled shRNA control (Figure 5C; \( P < .05 \), Student t test). Under optimal cell culture conditions with 15% FCS supplemented, this difference was not observed (Figure 5C).

To investigate whether the knockdown of circBCL11B resulted in increased cell death, transduced OCI-AML5, NB-4, and MV-4-11 cells underwent Annexin V/propidium iodide (PI) staining, followed by flow cytometric analysis after 72 hours of culture, with or without FCS. In the absence of FCS, circBCL11B knockdown resulted in increased cell death in OCI-AML5, NB-4, and MV-4-11 cells compared with the respective scrambled shRNA control (\( P < .05 \), Student t test; Figure 6). Although this increase in cell death was dominated by an increase in late apoptotic and necrotic cells, represented by Annexin V+/PI− cells, in OCI-AML5 and MV-4-11 cells, the transduced NB-4 cells also showed an increase in Annexin V+/PI− early apoptotic cells upon circBCL11B knockdown.

Discussion

In this study, we determined circRNA expression in a large patient cohort covering several AML entities and investigated leukemia-associated changes in the circular RNAome by comparison with healthy hematopoietic cells of a comparable differentiation degree.

circRNAs have been overlooked for many decades, but the technological developments toward sequencing of longer reads of non-poly(A)–selected RNA libraries paved the way for studying circRNAs in an unbiased fashion. In addition to Illumina 100-bp paired-end sequencing, in which only backsplice-containing reads can be assigned to a circRNA, we made use of the third-generation ultralong read sequencing technology by Oxford Nanopore Technologies \(^{26}\) to acquire full-length circRNA sequences of 9 target genes, as we did previously for circNPM1 isoforms. \(^{6}\) Recently, the working group of Jorgen Kjems \(^{27}\) used Oxford Nanopore technology to sequence circRNAs of the brain in an untargeted approach by removing linear RNA through RNase R digestion and poly(A) depletion. Then, the remaining circRNAs were linearized through hydrolysis and sequenced using Oxford Nanopore technology. However, only 3.3% of the sequenced reads contained a backsplice site, because the circRNAs were cut multiple times. Although respective protocols can be further optimized in the future, our targeted approach has proven useful in validating the Illumina results, even for circRNAs that were only supported by a single read in the RNA-seq experiment. Furthermore, this approach revealed additional circRNAs that were not detected by RNA-seq, which, in turn, argues again for direct sequencing of circRNAs using an unbiased Oxford Nanopore long-read approach.

In healthy and leukemic samples, \( \sim 30\% \) of the expressed genes also expressed circRNAs. Although we detected \( \geq 1 \) circular isoform for most of these genes (79%), the number of distinct circRNA isoforms was substantially lower than theoretically possible, which supports a nonrandom circularization process. Moreover, we observed that, in most cases, only 1 dominant circRNA isoform showed differential expression, which further implies a parental gene expression–independent circRNA isoform-specific regulation. However, the overall circRNA abundance was comparable in AML and healthy HSPCs, whereas other cancer types often show lower circRNA expression in tumor compared with healthy tissue.\(^{28}\)

Although it is known that splicing factor depletion shifts the transcription process toward increased formation of circRNAs,\(^{29}\) mutations in the splicing factors U2AF1, SF3B1, and SRSF2 did not abrogate or enhance circRNA biogenesis. However, only few splicing factor–mutated patients were included in this study; in the future, larger cohorts will have to be investigated to get a more complete picture. Moreover, because the mutations might lead to the usage of alternative 3’ splice sites, noncanonical exon junctions taking part in the backsplice process will also have to be examined.

A promising candidate for functional characterization was circBCL11B, because it was exclusively expressed in AML samples in our cohort; no expression was detectable in healthy HSPCs, as well as in 20 other healthy tissues recently investigated.\(^{30}\) Moreover, it is derived from a gene whose expression in the hematopoietic system is normally restricted to T cells and their prethymic precursors.\(^{31}\) In accordance, circBCL11B was expressed in healthy T cells,\(^{32}\) whereas expression was negligible in healthy myeloid cells and in bulk peripheral blood cells.\(^{33}\) In addition, detection of circBCL11B in breast cancer\(^ {34}\) and in T-cell acute lymphoblastic leukemia\(^ {7}\) makes it a relevant candidate for further studies in cancer in general.

BCL11B is thought to be a haploinsufficient tumor suppressor. Deletions and missense mutations of BCL11B have been found in \( \sim 10\% \) of human T-cell acute lymphoblastic leukemia\(^ {35,37}\) and induced thymic lymphoma in mice.\(^ {38,39}\) Moreover, there are several case reports of translocations involving BCL11B in myeloid leukemias,\(^ {40,41}\) including a T/myeloid mixed-phenotype leukemia, and the ZEB2–BCL11B fusion was observed in T-cell precursor leukemia.\(^ {42}\) Riemke and colleagues\(^ {44,45}\) have shown that committed T-cell progenitors gave rise to leukemic cells of myeloid phenotype, while at the same time preserving a T-cell–associated gene expression signature, including the expression of Bcl11b. Knockdown of Bcl11b in these T-cell progenitor–derived myeloid leukemic blasts reduced proliferation in vitro and prevented leukemia development in transplanted mice. Furthermore, this lymphoid gene expression signature was also detected in myeloid cells from a small fraction of AML patients, leading to the hypothesis that these leukemias might have arisen from T-lymphoid progenitor cells. Interestingly enough, we found that high circBCL11B expression in AML patients was associated
with a naive T-cell–like gene expression signature. Consequently, it would be of interest to investigate whether, in some of these circBCL11B_high AML cases, the cell of origin might have also been a T-cell progenitor or whether, during malignant transformation, signaling cascades restricted to lymphoid cells can be reactivated in myeloid blasts by altering the circRNA expression pattern.

Moreover, we studied the effect of circBCL11B knockdown on leukemic cells representing different AML subtypes. Although effects of circBCL11B knockdown were not detectable under optimal cell culture conditions, we could show that knockdown of circBCL11B in 3 leukemia cell lines increased cell death and had a negative effect on leukemic cell proliferation under stress conditions (ie, serum deprivation). These findings further support that the circBCL11B isoform might be of pathogenic relevance in different AML subtypes. Interestingly, the circularized exon 2 of BCL11B contains experimentally validated binding sites for 2 miRNAs with antiproliferative activity, hsa-miR-32-3p and hsa-miR-92a-3p, both have been linked to leukemia. However, a potential role for circBCL11B in sponging and, thus, inactivating the respective miRNAs, will have to be investigated in more detail in future studies.

In summary, our data suggest that further determination of circRNA deregulation can improve the understanding of AML pathogenesis and can reveal pathway dependencies that might have been overlooked at the linear gene expression level.

Acknowledgments

The authors thank Tobias Jakobi for advice regarding the CircTest tool.

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 1074 project B3) (K.D. and L.B.) and the Studienstiftung des Deutschen Volkes (S.L.).

Authorship

Contribution: S.L. designed and performed the experiments, analyzed the data, and wrote the manuscript; T.J.B. created the pipeline and gave advice for all bioinformatic analyses; B.G., A.R., and S.S. performed experiments; S.C. assisted with bioinformatics analyses; K.S. and H.S. provided RNA from healthy control samples; H.D. and K.D. provided AML patients; A.D. and L.B. designed experiments and supervised the project; L.B. wrote the manuscript; and all authors read and approved the manuscript.

Conflict-of-interest disclosure: L.B. has served on advisory committees for AbbVie, Amgen, Astellas, Bristol Myers Squibb, Celgene, Daiichi Sankyo, Gilead, Hexal, Janssen Pharmaceuticals, Jazz Pharmaceuticals, Menarini, Novartis, Pfizer, Sanofi, and Seattle Genetics. K.D. has served on advisory committees for Novartis, Janssen Pharmaceuticals, Celgene, Bristol Myers Squibb, and Daiichi Sankyo. The remaining authors declare no competing financial interests.

ORCID: S.L., 0000-0003-8826-2542; T.J.B., 0000-0002-9859-1261.

Correspondence: Lars Bullinger, Department of Hematology, Oncology and Tumorimmunology, Charité University Medicine Campus Virchow-Klinikum, Augustenburger Platz 1, 13355 Berlin, Germany; e-mail: lars.bullinger@charite.de.

References

1. Memczak S, Jens M, Elefsinioti A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495(7441):333-338.
2. Hansen TB, Jensen TI, Clausen BH, et al. Natural RNA circles function as efficient microRNA sponges. Nature. 2013;495(7441):384-388.
3. Li Z, Huang C, Bao C, et al. Exon-intron circular RNAs regulate transcription in the nucleus [published correction appears in Nat Struct Mol Biol. 2017;24(2):194]. Nat Struct Mol Biol. 2015;22(3):256-264.
4. Legnini I, Di Timoteo G, Rossi F, et al. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. Mol Cell. 2017;66(1):22-37.e9.
5. Pamudurti NR, Bartok O, Jens M, et al. Translation of circRNAs. Mol Cell. 2017;66(1):9-21.e7.
6. Yang Y, Gao X, Zhang M, et al. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. Nat Struct Mol Biol. 2017;66(1):22-37.e9.
7. Buratin A, Paganin M, Gaffo E, et al. Large-scale circular RNA deregulation in T-ALL: unlocking unique ectopic expression of molecular subtypes. Blood Adv. 2020;4(23):5902-5914.
8. Guerriero J, Bezzi M, Jeong JC, et al. Oncogenic role of fusion-circRNAs derived from cancer-associated chromosomal translocations [published correction appears in Cell. 2016;166(4):1055-1056]. Cell. 2016;166(2):289-302.
9. Hirsch S, Blätte TJ, Grasedieck S, et al. Circular RNAs of the nucleophosmin (NPM1) gene in acute myeloid leukemia. Haematologica. 2017;102(12):2039-2047.
10. Lux S, Bullinger L. Circular RNAs in cancer. In: Xiao J, ed. Circular RNAs: Biogenesis and Functions, Singapore: Springer Nature Singapore Pte Ltd; 2018:215-230.
11. Russ AC, Sander S, Lück SC, et al. Integrative nucleophosmin mutation-associated microRNA and gene expression pattern analysis identifies novel microRNA - target gene interactions in acute myeloid leukemia. Haematologica. 2011;96(12):1783-1791.
12. Rücker FG, Russ AC, Cocciardi S, et al. Altered miRNA and gene expression in acute myeloid leukemia with complex karyotype identify networks of prognostic relevance. Leukemia. 2015;29(2):383-391.
13. Garzon R, Volinia S, Liu C-G, et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. Blood. 2008;111(6):3183-3189.
14. Bill M, Papaioannou D, Karunasiri M, et al. Expression and functional relevance of long non-coding RNAs in acute myeloid leukemia stem cells. Leukemia. 2019;33(9):2169-2182.
15. Papaioannou D, Volinia S, Nicolet D, et al. Clinical and functional significance of circular RNAs in cytogenetically normal AML [published correction appears in Blood Adv. 2020;4(12):2577]. Blood Adv. 2020;4(2):239-251.

16. Coccia SD, Dolnik A, Kapp-Schwoerer S, et al. Clinical and functional significance of circular RNAs in cytogenetically normal AML [published correction appears in Blood Adv. 2020;4(12):2577]. Blood Adv. 2020;4(2):239-251.

17. Bobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-Seq aligner. Bioinformatics. 2013;29(1):15-21.

18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

19. Kolde R, Kolde MR. Package “pheatmap”. R Package. 2015;17(7):790.

20. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003;34(3):267-273.

21. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-15550.

22. Loman NJ, Quinlan AR. Poretools: a toolkit for analyzing nanopore sequence data. Bioinformatics. 2014;30(23):3399-3401.

23. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res. 2008;36(Web Server issue):W5-9.

24. Li L, Leid M, Rothenberg EV. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. Science. 2010;329(9587):89-93.

25. Wakabayashi Y, Watanabe H, Inoue J, et al. Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. Nat Immunol. 2003;4(6):533-539.

26. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community [published correction appears in Genome Biol. 2016;17(1):256]. Genome Biol. 2016;17(1):239.

27. Rahimi K, Vene MT, Dupont DM, Kjems J. Nanopore sequencing of full-length circRNAs in human and mouse brains reveals circRNA-specific exon usage and intron retention. BioRxiv. 2019:567164.

28. Bachmayr-Heyda A, Reiner AT, Auer K, et al. Correlation of circular RNA abundance with proliferation--exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues. Sci Rep. 2015;5(1):8057.

29. Li L, Leid M, Rothenberg EV. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. Science. 2010;329(9587):89-93.

30. Maass PG, Glazier P, Memczak S, et al. A map of human circular RNAs in clinically relevant tissues. J Mol Med (Berl). 2017;95(11):1179-1189.

31. Li L, Zhang JA, Dose M, et al. A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. Blood. 2013;122(6):902-911.

32. Gaffo E, Boldrin E, Dal Molin A, et al. Circular RNA differential expression in blood cell populations and exploration of circRNA deregulation in pediatric acute lymphoblastic leukemia. Sci Rep. 2019;9(1):14670.

33. Haque S, Ames RM, Moore K, et al. circRNAs expressed in human peripheral blood are associated with human aging phenotypes, cellular senescence and mouse lifespan. Geroscience. 2020;42(1):183-199.

34. Yan N, Xu H, Zhang J, et al. Circular RNA profile indicates circular RNA VRK1 is negatively related with breast cancer stem cells. Oncotarget. 2017;8(56):95704-95718.

35. Neumann M, Vosberg S, Schlee C, et al. Mutational spectrum of adult T-ALL. Oncotarget. 2015;6(5):2754-2766.

36. Gutierrez A, Kentiss A, Sanda T, et al. The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. Blood. 2011;118(15):4169-4173.

37. De Keersmaecker K, Real PJ, Gatta GD, et al. The TLX1 oncogene drives aneuploidy in T cell transformation. Cell. 2004;149(1):72-76.

38. Yoshikai Y, Sato T, Morita S, et al. Effect of Bcl11b genotypes and T/cellular deletion on mouse peripheral blood. Blood Adv. 2018;2(12):550.

39. Goossens S, Radaelli E, Blanchet O, et al. ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. Nat Commun. 2015;6(1):5794.

40. Bezrookove V, van Zelder-Bhola SL, Brink A, et al. A novel t(6;14)(q25-q27;q32) in acute myelocytic leukemia involves the BCL11B gene. Blood. 2010;116(11):4169-4173.

41. De Keersmaecker K, Real PJ, Gatta GD, et al. The TLX1 oncogene drives aneuploidy in T cell transformation. Cell. 2004;149(1):72-76.

42. Goossens S, Radaelli E, Blanchet O, et al. ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. Nat Commun. 2015;6(1):5794.

43. Torkildsen S, Gorunova L, Beiske K, Tjønnfjord GE, Heim S, Panagopoulos I. Novel ZEB2-BCL11B fusion gene identified by RNA-sequencing in acute myeloid leukemia with t(2; 14)(q22; q32). PLoS One. 2015;10(7):e0132736.

44. Riemke P, Czeh M, Fischer J, et al. Myeloid leukemia with transdifferentiation plasticity developing from T-cell progenitors. EMBO J. 2016;35(22):2399-2416.

45. Bullinger L. T-lymphoid progenitors—we know what they are, but know not what they may be. EMBO J. 2016;35(22):2383-2385.

46. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. PLoS One. 2018;13(10):e0206239.
47. Chou C-H, Chang N-W, Shrestha S, et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* 2016;44(D1):D239-D247.

48. Gu Y, Si J, Xiao X, Tian Y, Yang S. miR-92a inhibits proliferation and induces apoptosis by regulating methylenetetrahydrofolate dehydrogenase 2 (MTHFD2) expression in acute myeloid leukemia. *Oncol Res.* 2017;25(7):1069-1079.

49. Cheng J, Metge F, Dieterich C. Specific identification and quantification of circular RNAs from sequencing data. *Bioinformatics.* 2016;32(7):1094-1096.