Supplemental Material

Deregulated Expression of Circular RNAs in Acute Myeloid Leukemia

Susanne Lux,[1] Tamara J. Blätte,[1,2] Bernhard Gillissen,[2] Antje Richter,[2] Sibylle Cocciardi,[1] Sabrina Skambraks,[1] Klaus Schwarz,[3,4] Hartmut Döhner,[1] Konstanze Döhner,[1] Anna Dolnik,[1,2] and Lars Bullinger[1,2,*]

[1]Internal Medicine III, University Hospital Ulm, Ulm, Germany
[2]Department of Hematology, Oncology and Tumorimmunology, Charité University Medicine, Berlin, Germany
[3]Institute for Transfusion Medicine, University of Ulm, Ulm, Germany
[4]Institute for Clinical Transfusion Medicine and Immunogenetics Ulm, German Red Cross Blood Service, Ulm, Germany

***Corresponding author:**

Lars Bullinger, Department of Hematology, Oncology and Tumorimmunology, Charité University Medicine Campus Virchow-Klinikum, Augustenburger Platz 1, 13353 Berlin, Germany. Phone: +49 30 450 553 111 ; Fax: +49 30 450 553 914 ; E-Mail: lars.bullinger@charite.de

SUPPLEMENTAL INFORMATION

Supplemental Methods
Supplemental Tables
Supplemental Figures
Supplemental References

Supplemental Methods

Circular to linear proportion (CLP) calculations

To quantify the ratio of circRNA versus linear mRNA expression for each circRNA isoform, reads mapping to the two linear junctions flanking the backsplice junction of a circRNA were counted. To this end, reads mapping to the linear junctions involving either the backsplice donor or backsplice acceptor were taken into account, and the mean of the two linear junction counts was used for the calculation of the CLP. The Circ.filter function of the CircTest[1] R package was used to filter for circRNAs with at least 2 read counts in at least 5 samples.

CLP = (circular junction counts) / ((reads mapping to linear junction A + reads mapping to linear junction B)/2)

$$CLP = \frac{circular\ junction\ counts}{(reads\ mapping\ to\ linear\ junction\ A\ +\ reads\ mapping\ to\ linear\ junction\ B)/2}$$

Cell lines

AML cell lines KASUMI-1, ME-1, NB-4, OCI-AML3, OCI-AML5 and the CML cell line K-562 and 293T cells were purchased from the German cell line repository (DSMZ, Braunschweig, Germany) and MV-4-11 was purchased from ATCC (Manassas, VA, USA). All cells were grown in accordance with the standard protocols. Cell lines were grown in RPMI 1640 supplemented with 15% fetal calf serum (FCS).

ShRNA-mediated functional evaluation of circRNAs

Oligos used for the scrambled and circ*BCL11B* shRNA constructs are listed in Suppl. Table 1. The constructs were ligated into pLKO.1 vectors (adapted from #8453, Addgene, Watertown, MA, USA) carrying either a puromycin resistance or GFP. 293T cells were used for lentivirus production. $0.2x10^6$ OCI-AML5, NB-4 and MV4-11 cells were lentivirally transduced in the presence of 2 µg/ml polybrene followed by puromycin selection with 2 µg/ml puromycin for 5 days.

Immunoblotting

Cells were washed twice with ice-cold PBS and lysed in 10 mM Tris-HCl pH 7.5, 137 mM NaCl, 1% Triton X-100, 2 mM EDTA, 1 µM pepstatin, 1 µM leupeptin, and 100 µM phenylmethyl sulfonylfluoride (PMSF). Protein concentration was determined using the Thermo Scientific Pierce BCA Protein Assay Kit (Life Technologies GmbH, Darmstadt, Germany). Equal amounts of protein were separated by SDS-PAGE, electroblotted onto nitrocellulose membrane. The membrane was blocked for 1 h in PBST (PBS, 0.05% Tween-20) containing

3% nonfat dry milk and incubated with primary antibodies for 1 h. After the membrane had been washed three times in PBST, the secondary antibody in PBST was applied for 1 h. Finally, the membrane was washed again in PBST and protein bands were detected using the ECL enhanced chemiluminescence system (Amersham Buchler, Braunschweig, Germany). BCL11B was detected using anti-BCL11B (rabbit anti-human, ab187668, 1:10000, Abcam, Cambridge, UK). Actin served as a loading control using anti-actin mAb (AC-74, Sigma-Aldrich, Taufkirchen, Germany). Secondary anti-rabbit and anti-mouse horseradish peroxidase (HRP)-conjugated antibodies were from Southern Biotechnology Associates Inc. (Birmingham, AL, USA).

Cell counting

The respective cell lines were cultured in the presence and absence of FCS and counted by using the CASY 1 Cell-Counter and Analyzer System (Schärfe System/Innovatis) at the indicated time points.

Detection of apoptotic cell death by flow cytometry

Cell death was determined by staining cells with Annexin V-fluorescein isothiocyanate (FITC) and propidium iodide (PI). Briefly, cells were washed twice with cold PBS and resuspended in 10 mM Hepes (pH7.4), 140 mM NaCl, 2.5mM CaCl2 at 1x106 cells/ml. To 100 µl of the solution (1 x 105 cells), 5 µl of Annexin-V-FITC (BD Biosciences, Heidelberg, Germany) and 10 µl PI (20 mg/ml, Sigma-Aldrich) were added. The samples were analyzed using a BD FACSCalibur equipped with the CELLQuest software (Becton Dickinson, Heidelberg, Germany). Cells positive for annexin V-FITC (phosphatidylserine exposure) and negative for PI staining represent early apoptotic cells, whereas late apoptotic or necrotic cells are positive for both annexin V-FITC and PI. Results are represented as % of cells.

Cytospin

Cytospins were performed with $1x10^5$ OCI-AML5 cells transduced with scrambled shRNA control or the circ*BCL11B* shRNA. Cells were stained with May-Grünwald and Giemsa solution. Images were taken at an original magnification ×60 using an Eclipse Ti-S light microscope (Nikon Precision, Belmont, CA, USA) and the INFINITY software (Lumenera, Ottawa, ON, CA).

Supplemental Tables

Supplemental Table 1. Primer sequences used in this study.

Gene	Sequence	Exon	Usage
AFF2	5' TTTTGGGGAGCCATACAAGGTAGCTGAA 3' se 5' CTTGATTCCTCCGCTCCCATTCCCT 3' as	E.2 E.2	Oxford Nanopore sequencing
AFF2	5' CCAGGACCAGGCACCGGACATCT 3' se 5' CCAGTTTGGCTTGAGACTGGTCCTGTGT 3' as	E.3 E.3	Oxford Nanopore sequencing
ARID1B	5' CCACAGCGGTATCCAATTGG 3' se 5' GCCCATGCCATACAACTGAG 3' as	E.2 E.2	Oxford Nanopore sequencing
ARID1B	5' CCCCGGAAAACCTAACCATG 3' se 5' CTGTCCATACTGAGGTGGCA 3' as	E.5 E.4	Oxford Nanopore sequencing
ARID1B	5' GGAGCCAGTCAGAATCCAGT 3' se 5' ACTGCTGAGCTCAAAGTTGC 3' as	E.7 E.6	Oxford Nanopore sequencing
BCL11B	5' CTATGACAAGGCCCTGGACA 3' se 5' GCTCCTCTATCTCCAGACCC 3' as	E.2 E.2	Oxford Nanopore sequencing
BCL11B	5' GGAGAACATTGCAGCAGAGG 3' se 5'CCACAGGTGAGCAGGTCA 3' as	E.2 E.2	Validation of shRNA knockdown
BCL11B	5'CCGGGAACATTGCAGCAGAGGCTGACTCGAGTCAGCCTCTGCTGCAATGTTCTTTTTG 3' 5'AATTCAAAAAGAACATTGCAGCAGAGGCTGACTCGAGTCAGCCTCTGCTGCAATGTTC 3'	E.2 E.2	circ*BCL11B* shRNA oligo
CBL	5' AGCAAACCATAAGCCTCTTCA 3' se 5' CCAGCTTTGGGTTCTGACAC 3' as	E.2 E.2	Oxford Nanopore sequencing
CBL	5' TGCTGTAACTCATCCTGGCT 3' se 5' TGTAGAGCCTGTCGAAAGCT 3' as	E.5 E.4	Oxford Nanopore sequencing
CBL	5' CCCAAGTTCCAGTGATCCCT 3' se 5' CAGAAGGTCAAGTCGTGGTG 3' as	E.11 E.10	Oxford Nanopore sequencing
CBL	5' TCTATCACCGAGAGCAGCAC 3' se 5' CCGTAGAGGCCTGGAAGAG 3' as	E.14 E.13	Oxford Nanopore sequencing
GSE1	5' CGCAAGCTCGCCAAACAG 3' se 5' GTCGCGGTGGAAAGCATC 3' as	E.2 E.2	Oxford Nanopore sequencing
GSE1	5' AATGTAGACGACTGGAGGCC 3' se 5' ATGTTGTAGTGGACGGACGT 3' as	E.15 E.14	Oxford Nanopore sequencing

GSE1	5'AGGAGGATGATGAAGATGGAGAAGATGAGG AGGA 3' se 5' ATTGTGCTGGGGTGGAGGGGCGC 3' as	E.15 E.15	Oxford Nanopore sequencing
FLT3	5' AAGGCTGGAAGAAGAGGAGG 3' se 5' ATTGGTCCTGACAGTGTGCA 3' as	E.19 E.16/17	Oxford Nanopore sequencing
MEIS1	5' CAGCTCATACCAACGCCATG 3' se 5' ATGCCTACTCCATCCATGCC 3' as	E.2 E.2	Oxford Nanopore sequencing
MEIS1	5' CACAAGACACGGGACTCAC 3' se 5' GTCATCATCGTCACCTGTGC 3' as	E.9 E.8	Oxford Nanopore sequencing
RUNX1	5' CTACCGCAGCCATGAAGAAC 3' se 5' CAACGCCTCGCTCATCTTG 3' as	E.4 E.3	Oxford Nanopore sequencing
RUNX1	5' CACCTACCACAGAGCCATCA 3' se 5' CATCATTGCCAGCCATCACA 3' as	E.5 E.4	Oxford Nanopore sequencing
RUNX1	5' GAGCTTGTCCTTTTCCGAGC 3' se 5' TGATGGCTCTGTGGTAGGTG 3' as	E.6 E.5	Oxford Nanopore sequencing
SPI1	5' CGCACGAGTATTACCCCTATCTCAGCAGT 3' se 5' TTGGCGTTGGTATAGATCCGTGTCATAGG 3' as	E.2 E.2	Oxford Nanopore sequencing
SPI1	5' GGTGCCACCCCATCCCAGTCT 3' se 5' TGGCTCTCCCCATCACTGCTGAGA 3' as	E.3 E.2	Oxford Nanopore sequencing
SPI1	5' ACCTGCCCCGGATGTGCCTCCA 3' se 5' TGGCTCTCCCCATCACTGCTGAGA 3' as	E.4 E.2	Oxford Nanopore sequencing
SPI1	5' TGGAGGTGTCTGACGGCGAGGCG 3' se 5' GGGGACAGGGATGGGTACTGGAGGCA 3' as	E.4 E.4	Oxford Nanopore sequencing
Scrambled control	5'CCGGCCTAAGGTTAAGTCGCCCTCGCTCGAG CGAGGGCGACTTAACCTTAGGTTTTTG 3' se 5'AATTCAAAAACCTAAGGTTAAGTCGCCCTCG CTCGAGCGAGGGCGACTTAACCTTAGG 3' as		Scrambled shRNA oligo

as = antisense; se = sense; E. = exon; KD = knockdown.

Supplemental Table 2. Gene ontology analysis of highly expressed genes (mean normalized read count across all samples ≥20) also expressing circular RNAs.

GO term	FDR
GO_CELL_CYCLE	< 0.0001
GO_CELL_CYCLE_PROCESS	< 0.0001
GO_REGULATION_OF_CELL_CYCLE	< 0.0001
GO_MITOTIC_CELL_CYCLE	< 0.0001
GO_REGULATION_OF_RESPONSE_TO_STRESS	< 0.0001
GO_CELLULAR_RESPONSE_TO_STRESS	< 0.0001

Supplemental Table 3. Gene ontology analysis of highly expressed genes (mean normalized read count across all samples ≥20) not expressing circular RNAs.

GO term	FDR
GO_RIBOSOMAL_SUBUNIT	< 0.0001
GO_RIBOSOME_BIOGENESIS	< 0.0001
GO_RRNA_METABOLIC_PROCESS	< 0.0001
GO_STRUCTURAL_CONSTITUENT_OF_RIBOSOME	< 0.0001
GO_LARGE_RIBOSOMAL_SUBUNIT	< 0.0001
GO_RIBOSOME	< 0.0001

Supplemental Table 4. List of genes differentially expressing circRNAs, differentially expressed circRNAs and genes differentially expressing linear mRNA for all AML versus healthy HSPC comparisons.

deregulated_circRNAs_AML_vs_healthy.xlsx

Supplemental Table 5. Circular to linear proportions (CLP). Circular junction counts and linear junction counts for 61 AML patients and 16 healthy HSPC controls.

CLP_circular_to_linear_proportion_AML_HSPCs.xlsx

Supplemental Table 6. Validation of target circRNAs using Oxford Nanopore sequencing.

Patient	Gene	CircRNA	SuppReads	Backsplice	Validated
U2AF1^{mut}	*ARID1B*	Chr6 157150360 157431695	1	6→2	Yes
U2AF1^{mut}	*ARID1B*	Chr6 157405795 157431695	4	7→6	Yes
U2AF1^{mut}	*ARID1B*	Chr6 157150360 157431695	0	7→2	
U2AF1^{mut}	*ARID1B*	Chr6 157222509 157431695	0	7→4	
U2AF1^{mut}	*ARID1B*	Chr6 157357968 157431695	0	7→5	
U2AF1^{mut}	*ARID1B*	Chr6 157357968 157454341	0	8→5	
U2AF1^{mut}	*ARID1B*	Chr6 157405795 157454341	0	8→6	
U2AF1^{mut}	*ARID1B*	Chr6 157405795 157470085	0	9→6	
U2AF1^{mut}	*ARID1B*	Chr6 157405795 157488319	0	10→6	
AML_rl	*BCL11B*	Chr14 99723807 99724176	14	2→2	Yes
U2AF1^{mut}	*CBL*	Chr11 119155678 119156276	1	11→10	Yes
U2AF1^{mut}	*CBL*	Chr11 119167627 119168191	5	14→13	Yes
U2AF1^{mut}	*CBL*	Chr11 119155678 119158656	0	12→10	
NPM1^{mut}_1	*FLT3*	Chr13 28588588 28602425	1	23→16	Yes
NPM1^{mut}_1	*FLT3*	Chr13 28592603 28602425	4	20→16	Yes
NPM1^{mut}_1	*FLT3*	Chr13 28592603 28610180	1	20→11	Not found
NPM1^{mut}_1	*FLT3*	Chr13 28597486 28602425	5	19→16	Yes
NPM1^{mut}_1	*FLT3*	Chr13 28597486 28624359	1	19→6	Not found
NPM1^{mut}_1	*FLT3*	Chr13 28589293 28602425	0	22→16	
NPM1^{mut}_1	*FLT3*	Chr13 28588588 28624359	0	23→6	
NPM1^{mut}_1	*FLT3*	Chr13 28592603 28608544	0	20→13	
AML_rl	*GSE1*	Chr16 85667519 85667738	22	2→2	Yes
NPM1^{mut}_2	*MEIS1*	Chr2 66664868 66670180	1	6→2	Yes
NPM1^{mut}_2	*MEIS1*	Chr2 66691240 66739426	2	8→7	Yes
NPM1^{mut}_2	*MEIS1*	Chr2 66691240 66775151	1	9→7	Yes
NPM1^{mut}_2	*MEIS1*	Chr2 66664868 66667116	0	3→2	
NPM1^{mut}_2	*MEIS1*	Chr2 66664868 66691352	0	7→2	
NPM1^{mut}_2	*MEIS1*	Chr2 66664868 66665095	0	2→2	
NPM1^{mut}_2	*MEIS1*	Chr2 66691240 66691352	0	7→7	
NPM1^{mut}_2	*RUNX1*	Chr21 36202148 36231875	1	X→6	Yes
NPM1^{mut}_2	*RUNX1*	Chr21 36206706 36231875	8	7→6	Yes
NPM1^{mut}_2	*RUNX1*	Chr21 36231770 36259393	1	6→4	Yes
NPM1^{mut}_2	*RUNX1*	Chr21 36206706 36253010	0	7→5	
NPM1^{mut}_2	*RUNX1*	Chr21 36206706 36259393	0	7→4	
NPM1^{mut}_2	*RUNX1*	Chr21 36231770 36265260	0	6→3	

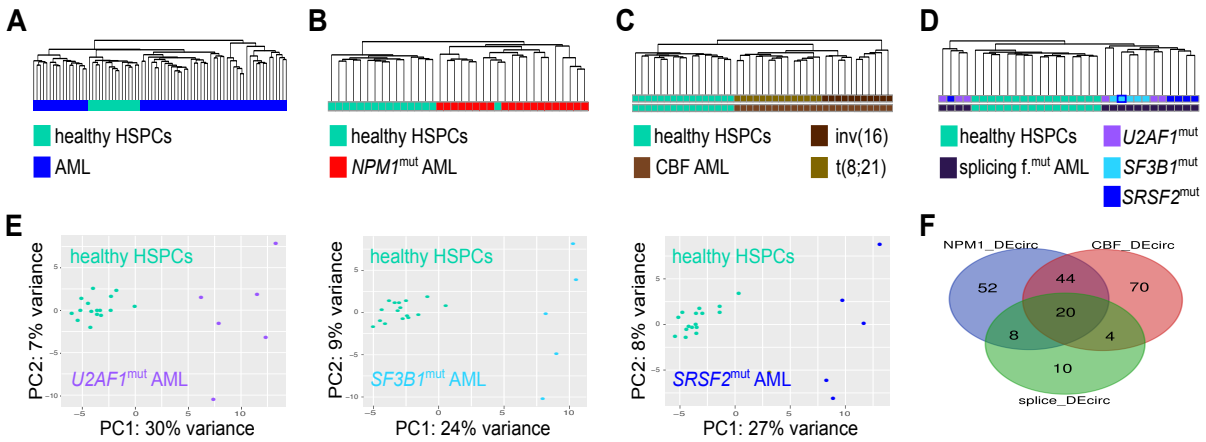
CircRNA isoforms that were not detected via RNA-seq but could be found in the Oxford Nanopore experiment are indicated in blue. Mut indicates mutated; rl, relapse; and SuppReads, supporting reads derived from RNA-seq data of the respective patient pre-normalization.

Supplemental Table 7. Gene set enrichment analysis of genes differentially expressed between circBCL11B_high and circBCL11B_low AML patients within *NPM1*mut (n=10), CBF leukemia (n=19), or within the splicing factor mutated cohort (n=9).

Patient group	***NPM1*mut**		**CBF**		**Splice factor mutation**	
Gene sets	**NES**	**FDR**	**NES**	**FDR**	**NES**	**FDR**
NAÏVE_TCELL_VS_ MONOCYTE_UP	1.81	4.15 E-03	0.88	1.00	0.88	1.00
NAÏVE_CD8_TCELL_VS_ MONOCYTE_UP	1.96	9.88 E-04	*N/A*	*N/A*	1.32	9.81 E-01
NAÏVE_CD4_TCELL_VS_ MONOCYTE_UP	1.94	4.94 E-04	0.89	1.00	1.03	1.00
CD4_TCELL_VS_ MYELOID_UP	1.78	6.03 E-03	0.79	1.00	1.01	1.00
HEMATOPOIETIC_STEM_ CELL_VS_CD4_TCELL _DN	1.50	2.51 E-01	1.73	1.27 E-01	1.25	1.00
NAÏVE_TCELL_VS_ MONOCYTE_DN	-1.59	2.12 E-02	-1.22	4.87 E-01	1.75	1.09 E-02
NAÏVE_CD8_TCELL_VS_ MONOCYTE_DN	-1.59	4.30 E-02	-1.51	1.38 E-01	1.81	8.94 E-03
NAÏVE_CD4_TCELL_VS_ MONOCYTE_DN	-1.66	1.09 E-02	-1.52	1.28 E-01	1.03	1.00
CD4_TCELL_VS_ MYELOID_DN	-1.81	3.90 E-03	-1.47	1.69 E-01	1.55	1.37 E-01
HEMATOPOIETIC_STEM_ CELL_VS_CD4_TCELL _UP	-0.91	1.00	-1.40	2.41 E-01	1.25	1.00

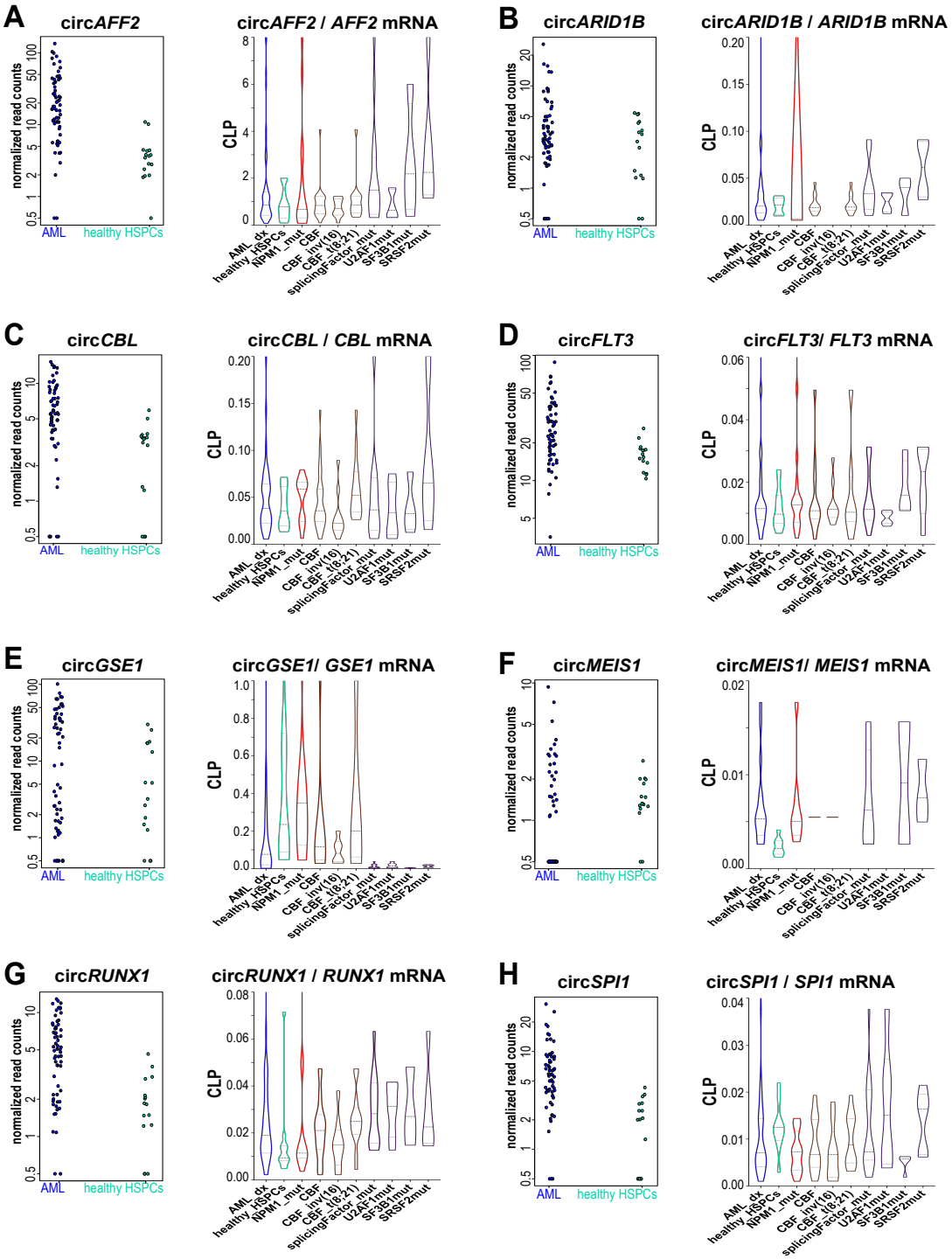
FDR = false discovery rate; NES = normalized enrichment score. A positive enrichment score indicates enrichment in circBCL11B_high patients, a negative enrichment score indicates enrichment in circBCL11B_low patients. CBF = core binding factor leukemia patients with t(8;21) or inv(16); *NPM1*^{mut} = AML patients with an *NPM1* mutation; splicing mut = AML patients with a mutation in splicing factor *U2AF1*, *SF3B1*, and/or *SRSF2*.

Supplemental Figures

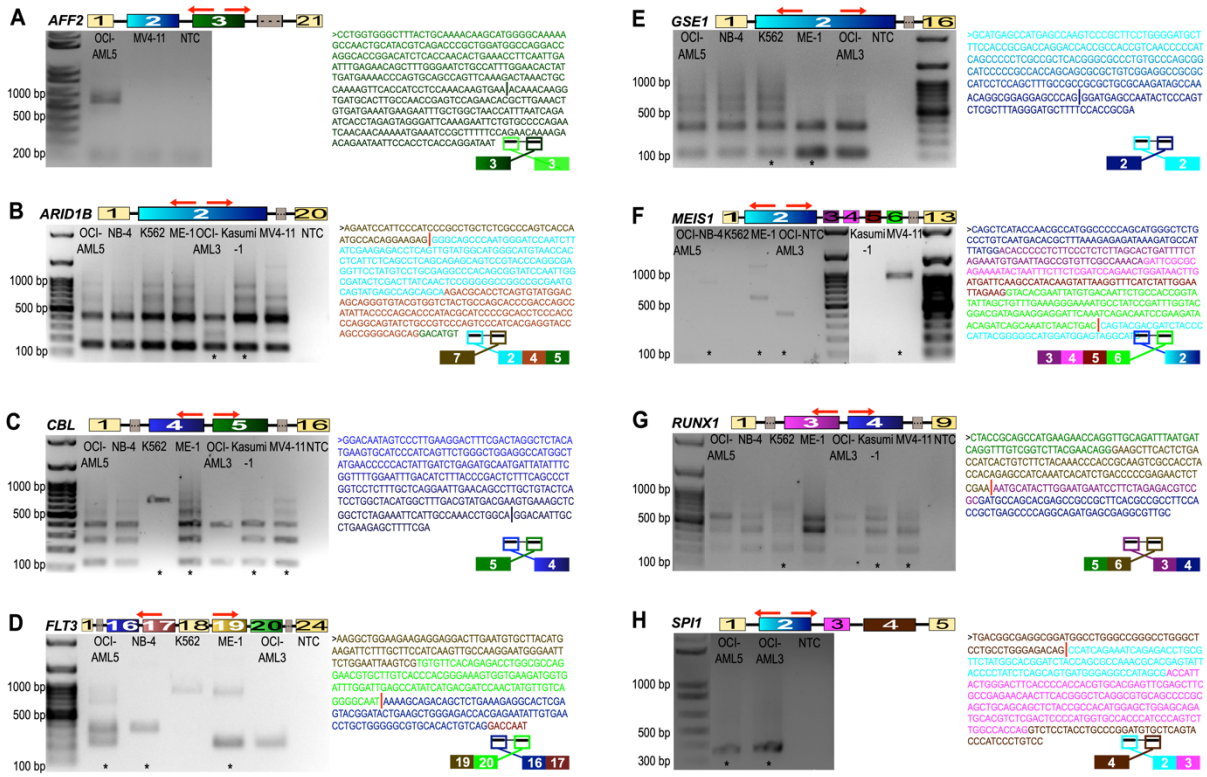


Supplemental Figure S1. Unsupervised hierarchical clustering and PCA illustrating circRNA expression in AML compared to healthy HSPCs.

(A)-(D) Unsupervised clustering was performed by Euclidean distance based on the 5,000 circRNA-expressing genes with the highest coefficient of variation across all respective samples. n=16 healthy HSPC (turquoise) and (A) n=61 AML samples at diagnosis (blue), (B) n=20 *NPM1*mut AML samples (red), (C) n=25 CBF leukemia samples, of which n=11 carry inv(16) (dark brown) and n=14 carry t(8;21) (beige), or (D) n=16 AML samples with mutations in splicing factors, including *U2AF1* (purple, n=6), *SF3B1* (light blue, n=5) and *SRSF2* (dark blue, n=6) (one case shares mutations in *SF3B1* and *SRSF2*), were included in the respective analyses. (E) PCA was performed based on the 500 circRNA-expressing genes with the highest variance across all respective samples using DESeq2. CircRNA expression is visualized for n=16 healthy HSPC samples (turquoise) compared to n=6 *U2AF1*mut (purple, left panel), n=5 *SF3B1*mut (light blue, middle panel) or n=6 *SRSF2*mut (dark blue, right panel) AML patients. PC = principal component. (F) Overlap of deregulated circRNA isoforms detected in the comparisons *NPM1*mut versus healthy HSPCs (blue), CBF leukemias versus healthy HSPCs (red) and AML patients with mutations in splicing factors versus healthy HSPCs (green).

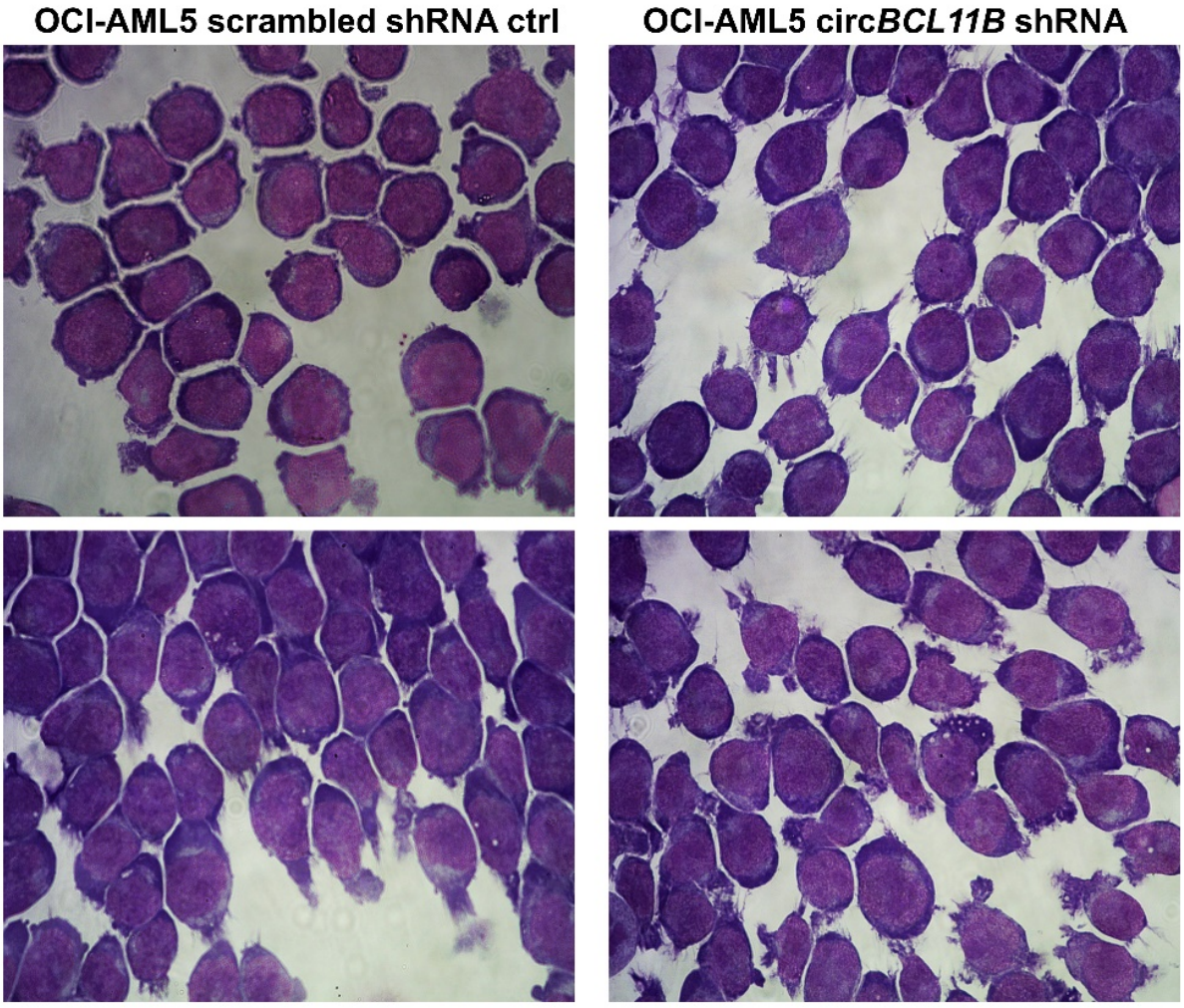


Supplemental Figure S2. CircRNA expression and circular to linear proportion (CLP) of target genes. Left panels: normalized circular read counts of target genes in n=61 AML patients (dark blue) and n=16 healthy HSPC samples (turquoise); right panels: circular to linear proportion (CLP) of target gene circular RNA counts to target gene mRNA counts. The horizontal solid line represents the median, dashed lines indicate the quartiles. Data is shown for (A) *AFF2*, (B) *ARID1B*, (C) *CBL*, (D) *FLT3*, (E) *GSE1*, (F) *MEIS1*, (G) *RUNX1* and (H) *SPI1*.



Supplemental Figure S3. Validation of target gene circRNAs using Oxford Nanopore sequencing.

(A)-(H) Circular-derived PCR products from target genes were generated using different AML cell lines and the CML cell line K562 and divergent primers (red arrows). PCR products marked with an asterisk were then pooled and sequenced on an Oxford Nanopore MinION. Nanopore reads were aligned to the respective genes using NCBI Blast. An exemplary sequencing result of each target gene PCR is shown. Exons are colored according to the target gene schemes. The fusion site of the downstream splice donor to the upstream splice acceptor is indicated by a red line. NTC=no template control.



Supplemental Figure S4. No increase in differentiation of OCI-AML5 cells upon knockdown of circ*BCL11B*.

OCI-AML5 cells were lentivirally transduced with a GFP-containing vector with either a scrambled shRNA control (ctrl) or an shRNA against circ*BCL11B*. Cytospin was performed and the cells were stained with May-Grünwald/Giemsa. Original magnification x60 using an Eclipse Ti-S light microscope (Nikon) and the INFINITY software (Lumenera).

Supplemental References

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