Supplemental Data

Clinical and functional significance of circular RNAs in cytogenetically normal AML

Appendices

Participating institutions

The following Cancer and Leukemia Group B (CALGB)/Alliance for Clinical Trials in Oncology (Alliance) institutions participated in this study and contributed at least two patients. For each of these institutions, the current or last principal investigators are listed as follows:

The Ohio State University Medical Center, Columbus, OH: Claire F. Verschraegen; Wake Forest University School of Medicine, Winston-Salem, NC: Heidi D. Klepin; Washington University School of Medicine, St. Louis, MO: Nancy L. Bartlett; North Shore University Hospital, Manhasset, NY: Daniel R. Budman; Dana Farber Cancer Institute, Boston, MA: Harold J. Burstein; Roswell Park Cancer Institute, Buffalo, NY: Ellis G. Levine; University of Chicago Medical Center, Chicago, IL: Hedy L. Kindler; University of Iowa Hospitals, Iowa City, IA: Daniel A. Vaena; University of North Carolina, Chapel Hill, NC: Thomas C. Shea; Ft. Wayne Medical Oncology/Hematology, Ft. Wayne, IN: Sreenivasa Nattam; University of Maryland Cancer Center, Baltimore, MD: Heather D. Mannuel; Dartmouth Medical School, Lebanon, NH: Konstantin H. Dragnev; Christiana Care Health Services, Inc., Newark, DE: Gregory Masters; University of Vermont Cancer Center, Burlington, VT: Steven Ades; Duke University Medical Center, Durham, NC: Jeffrey Crawford; Mount Sinai School of Medicine,

New York, NY: Lewis R. Silverman; Weill Medical College of Cornell University, New York, NY: Scott Tagawa; University of Massachusetts Medical Center, Worcester, MA: William V. Walsh; Eastern Maine Medical Center, Bangor, ME: Thomas H. Openshaw; University of California at San Francisco: Charalambos Andreadis; Western Pennsylvania Hospital, Pittsburgh, PA: John Lister; University of Puerto Rico School of Medicine, San Juan, PR: Eileen I. Pacheco; SUNY Upstate Medical University, Syracuse, NY: Stephen L. Graziano; University of Alabama at Birmingham: Robert Diasio; Rhode Island Hospital, Providence, RI: Howard Safran; University of Illinois, Chicago, IL: Arkadiusz Z. Dudek; Moores University of California San Diego Cancer Center, San Diego, CA: Barbara A. Parker; Walter Reed National Military Medical Center, Bethesda, MD: Mary Kwok; Massachusetts General Hospital, Boston, MA: David Ryan, Justin Gainor; Virginia Commonwealth University, Richmond, VA: Steven Grossman; University of Minnesota, Minneapolis, MN: Bruce A. Peterson; Long Island Jewish Medical Center, Lake Success, NY: Daniel R. Budman; University of Missouri/Ellis Fischel Cancer Center, Columbia, MO: Clint Kingsley; Nevada Cancer Research Foundation NCORP, Las Vegas, NV: John Ellerton; University of Tennessee Cancer Center, Memphis, TN: Harvey B. Niell; University of Nebraska Medical Center, Omaha, NE: Apar Ganti.

Patients and methods

Treatment protocols

All patients included in our study were treated on CALGB/Alliance first-line protocols for patients with acute myeloid leukemia (AML), and received standard cytarabine/daunorubicin-based induction chemotherapy.[1] Per protocol, all patients were scheduled to receive at least one induction cycle. For patients with residual leukemia present in a bone marrow (BM) biopsy after one induction cycle, a second cycle of induction was administered. None of the protocols included allogeneic stem cell transplantation (SCT) in first complete remission (CR). Patients enrolled on the treatment protocols also provided written informed consent to participate in the companion protocols CALGB 20202 (molecular studies in AML), CALGB 8461 (prospective cytogenetic companion), and CALGB 9665 (leukemia tissue bank), which involved collection of pretreatment BM aspirates and blood samples.

Patients were enrolled on the following treatment protocols: CALGB 19808, CALGB 10503, CALGB 9621, CALGB 10603, CALGB 9222, CALGB 8525, CALGB 9022 and CALGB 8721, CALGB 8821 and CALGB 9120. Patients enrolled onto CALGB 19808 (n=111) were randomly assigned to receive induction chemotherapy with cytarabine/daunorubicin, and etoposide with or without the multidrug resistance protein inhibitor PSC-833 (valspodar).[1] Upon achievement of CR, patients were assigned to intensification with high-dose cytarabine and etoposide for stem-cell mobilization followed by myeloablative treatment with busulfan and etoposide supported by autologous peripheral blood SCT. Patients on CALGB 10503 (n=109) received cytarabine/daunorubicin-based induction chemotherapy and those who achieved CR further received a two-step consolidation with chemo-mobilization and autologous SCT. For patients not eligible for autologous SCT, high-dose cytarabine-based consolidation was given. Maintenance with decitabine began as soon as possible after recovery

from consolidation.[2] Patients enrolled onto CALGB 9621 (n=59) were treated similarly to those on CALGB 19808, as previously reported.[3] Patients on CALGB 10603 (n=39) were stratified by *FLT3* mutation subtype [*FLT3*-TKD vs *FLT3*-ITD-high allelic ratio (>0.7) vs low allelic ratio (0.05-0.7)], and were randomized to receive cytarabine/daunorubicin-based induction chemotherapy and high-dose cytarabine consolidation in combination with either the multi-kinase inhibitor midostaurin or placebo. One-year midostaurin or placebo maintenance was administered after the last cycle of consolidation therapy.[4] Patients enrolled on CALGB 9222 (n=25) received cytarabine/daunorubicin-based induction chemotherapy, and those who achieved CR received either three cycles of high-dose cytarabine or three cycles of a so-called non cross-resistant regimen (the first cycle of this regimen was high-dose cytarabine, the second was cyclophosphamide plus etoposide, and the third was mitoxantrone plus diaziquone).[5] Patients enrolled onto CALGB 8525 (n=17) who achieved CR after cytarabine/daunorubicin-based induction chemotherapy were randomly assigned to consolidation with different doses of cytarabine followed by maintenance treatment.[6] Patients who participated in CALGB 9022 (n=2) and achieved CR after cytarabine/daunorubicin-based induction chemotherapy received one course of high-dose cytarabine consolidation, followed by one course of cyclophosphamide and etoposide, followed by one course of mitoxantrone and diaziquone (AZQ).[7]

Transcriptome analysis: library generation, sequencing and data analysis

Extracted total RNA was assessed for quality on an Agilent 2100 Bioanalyzer (BioA) using the RNA 6000 Nanochip and for quantity on a Qubit 2.0 Fluorometer (Agilent Technologies, Santa Clara, CA) using the RNA HS Assay Kit. Samples with a RNA Integrity Number (RIN) larger than four, with no visible sign of genomic DNA (gDNA) contamination and a concentration of >40 ng/μL were used for total RNA library generation. RNA-seq libraries were prepared using the Illumina TruSeq Stranded Total RNA Sample Prep Kit with RiboZero

Gold (#RS1222201) according to the manufacturer's instructions. Paired-end sequencing was performed with the Illumina HiSeq 2500 system using the HiSeq version 3 sequencing reagents to an approximate cluster density of 800,000/mm^2. Image analysis, base calling, error estimation, and quality thresholds were performed using the HiSeq Controller Software (version 2.2.38) and the Real Time Analyzer (RTA) software (version 1.18.64).

Cutadapt and FastQC were used to apply quality control and adapter trimming to FastQ files. The HISAT2[8] aligner was used to align the short reads to the human genome (hg38) using GENCODE annotation (ver. 22)[9] to quantify and annotate RNAs transcripts. Raw data were transformed into reads per million (RPM) and FPKM (fragments per kilobase per million) prior to statistical analysis. To minimize noise, each RNA species was evaluated as present in a sample only when at least nine reads were present in a total of 40 million reads. Information on RNA-seq characteristics and parameters of each patient sample are provided in Supplemental Table S4.

MScircRNA: A novel algorithm for the identification of exonic circular RNAs

To identify and quantify exonic *bona-fid*e circRNAs from next generation total RNA-seq datasets, we developed a novel computational pipeline called Massive Scan for circRNA (MScircRNA). We first used the STAR aligner[10] and generated a list of all potential circularized transcripts that could be detected in our dataset of younger adults with CN-AML. Specifically, we selected all reads that could represent head-to-tail fusion products of different parts of the genome. We filtered these results and retained the reads that included at least 15 nucleotides of each of the identified fused sequences. The candidate fusion reads were further filtered to those that represented non-canonical fusions of exons of the same gene. The generated list was further enriched with exonic circRNAs previously reported and included in the circBase database (http://www.circbase.org)[11] (Supplemental Table S1). Each candidate

circRNA sequence was concatenated with itself to produce a head-to-tail linear tandem sequences in a FASTA file format. To quantify the expression levels of circRNAs, we first removed all linear reads from the RNA-seq profiles. Next, we aligned all remaining reads to the pool of candidate circRNA sequences in the FASTA file described above. The alignment of the non-linear reads to the custom FASTA file was performed using the STAR aligner. The length of each fragment mapped to the custom circRNA reference sequences was determined using the CIGAR field feature in the SAM files. For quantification of each candidate circRNA, reads containing the back-splicing junction as well as reads that did not directly support, but aligned against both sides of the junction were counted. This step was included to boost sensitivity and dynamic range, since for circRNAs longer than 100 nucleotides, most reads would not contain the back-splicing junction. Finally, a robust candidate selection was performed to retain well-defined and expressed circRNAs, filtering by the number of reads spanning the junction with at least 6 or more nucleotides in at least 75% of the AML samples (Supplemental Figure S1).

Definition of clinical endpoints

Clinical endpoints were defined according to generally accepted criteria.[12] CR required a BM aspirate with cellularity >20% with maturation of all cell lines, <5% blasts and undetectable Auer rods; in peripheral blood, an absolute neutrophil count of ≥1.5 x 10^9/L, platelet count of >100 x 10^9/L, and leukemic blasts absent; and no evidence of extramedullary leukemia, all of which had to persist for ≥4 weeks.[12] Relapse was defined by the presence of ≥5% BM blasts, or circulating leukemic blasts, or the development of extramedullary leukemia. Disease-free survival (DFS) was measured from the date of CR until the date of relapse or death (from any cause); patients alive and in continuous first CR were censored at last follow-up. Overall survival (OS) was measured from the date of study entry until the date of death (from any

cause); patients alive at last follow-up were censored. Event-free survival (EFS) was measured from the date of study entry until the date of failure to achieve CR, relapse or death. Patients alive and in CR at last follow-up were censored.

Statistical analyses

Multivariable proportional hazards models were constructed for DFS, OS and EFS, using a limited backwards elimination procedure. Variables considered for model inclusion were: *circKLHL8* (high vs low), *circFCHO2* (high vs low), *circCFLAR* (high vs low), *circSMC1A* (high vs low), age (as a continuous variable, in 10-year increments), sex (male vs female), race (white vs non-white), white blood cell count [(WBC) 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 vs absent), *ASXL1* mutations (mutated vs wild-type), *CEBPA* mutations (double-mutated vs single-mutated or wild-type), *DNMT3A* mutations (mutated vs wild-type), *FLT3*-ITD (present vs absent), *FLT3*-TKD (present vs absent), *IDH1* mutations (mutated vs wild-type), *IDH2* mutations (mutated vs wild-type), *NPM1* mutations (mutated vs wild-type), *RUNX1* mutations (mutated vs wild-type), *SF1* (mutated vs wild-type), *SF3A1* (mutated vs wild-type), *SF3B1* (mutated vs wild-type), *SRSF2* (mutated vs wild-type), *TET2* mutations (mutated vs wild-type), *U2AF1* (mutated vs wild-type), U2AF2 (mutated vs wild-type), *WT1* mutations (mutated vs wild-type), *ZRSR2* (mutated vs wild-type), *ERG* expression levels (high vs low), *BAALC* expression levels (high vs low), *MN1* expression levels (high vs low), miR-181a expression levels (high vs low), miR-3151 (expressed vs not expressed), and miR-155 expression levels (high vs. low). For *ERG*, *BAALC*, *MN1*, miR-181a and miR-155, the median expression value was used as the cut point to divide patients into high and low expressers. Variables significant at α=0.2 from the univariable analyses were considered for multivariable

analyses. For the time-to-event endpoints, the proportional hazards assumption was checked for each variable individually.

RNase R digestion and library generation or cDNA transcription and real time PCR

For RNA sequencing of RNAse-R treated versus mock-treated samples, 10 µg of total RNA were subjected to rRNA depletion step using the RiboZero rRNA Removal Kit (Epicentre) followed by an AMPure bead (Beckman Coulter) clean-up with a 1.8-fold volume of beads, according to the manufacturer's instructions. Isolated RNA was treated with 20 Units of RNAse R or water in RNAse-R buffer (Epicentre, Illumina, San Diego, CA, USA) by incubation at 37ºC for 15 minutes. A 10-fold volume of QIAzol Lysis Reagent (QIAGEN) and 3-fold volume of chloroform were added to the RNase R reaction (50 µl). After mixture and phase separation (30 min, 15,000 rpm, 4ºC), the aqueous phase was re-extracted with 600 µl chloroform. RNA was precipitated with addition of 1-fold volume of isopropanol and glycogen followed by centrifugation. The pellet was washed once with 70% ethanol and re-suspended in DEPC water. A second round of RNase R digestion, QIAzol/chloroform extraction and isopropanol precipitation was followed by library generation as described above. For cDNA synthesis the rRNA depletion step was omitted. The RNA was reversed transcribed using random hexamer primers and the first strand cDNA synthesis kit by Invitrogen.

Custom designed qPCR primers that targeted the back-splicing regions of circRNAs were purchased from Integrated DNA Technologies. qRT-PCR assays were performed according to standard protocols.

CircRNA knock-downs

For the knock-down (KD) of circRNAs in OCI-AML3 and KG1-a cells as well as samples from AML patients we used custom-designed locked nuclei acid-modified RNase-H recruiting oligonucleotides (gapmers, Exiqon A/S). The sequences of the gapmers are provided in Table

S7. For *in vitro* delivery of the gapmers, the Nucleofector device (Lonza) and the corresponding reagents (Solution T for OCI-AML3 cells, solution L for KG1-a cells and human Monocyte Solution for AML blasts, Lonza) were used according to the instructions of the manufacturer (programs X-001 for OCI-AML3 cells, T-001 for KG1-a cells and Y-001 for patient blasts). CircRNA targeting or non-targeting control gapmers were delivered at a final concentration of 500 nM.

Flow cytometry analyses

Cell cycle analyses were performed with the BrdU staining kits of BD Pharmingen. Experiments were analyzed with flow cytometry on an LSRII instrument, according to the instructions of the manufacturer.

***In vitro* colony forming assays**

For colony formation unit assays with AML cell lines, the cells were electroporated, cultured for 24 hours and then mixed with basic pre-warmed methylcellulose-based medium without cytokines (MethoCult H4100, STEMCELL Technologies Inc), supplemented with 10% FBS. AML cells were seeded at a concentration of 1000 cells/ml, and colonies were counted on day 14.

For CFU assays with human AML patient blasts, the cells were electroporated, cultured for 24 hours (as described above) and then mixed with pre-warmed methylcellulose-based medium (Methocult H4034 Optimum, STEMCELL Technologies Inc.) according to the instructions of the manufacturer. Leukemic blasts were seeded at a concentration of 50.000 blasts/ml of medium (on day 0) and colonies were counted on day 14.

Differential gene expression analyses

To conduct differential gene expression analysis we divided the patients in our cohort in high versus low expressers of prognostic or functionally relevant circRNAs, based on the median expression value of each circular transcript. We then performed 2-class comparison analyses of the protein-coding transcripts and identified those that were differentially expressed among the two groups. We used a stringent threshold of false discovery rate $\leq 10^{-7}$ and $P<0.001$ to filter the identified mRNA transcripts. For clustering analysis we used the Eisen Cluster and Java Treeview. To identify biological processes that are associated with circRNA expression status we performed Gene Ontology analysis using the Panther Database (http://pantherdb.org/).

Material availability

All reagents used or generated in order to conduct the experiments described in this work are available upon request. The Massive Scan for circular RNAs pipeline is also available for analysis of total RNA sequencing datasets upon request to the corresponding author.

References

1. Kolitz JE, George SL, Marcucci G, et al. P-glycoprotein inhibition using valspodar (PSC-833) does not improve outcomes for patients younger than age 60 years with newly diagnosed acute myeloid leukemia: Cancer and Leukemia Group B study 19808. *Blood*. 2010;116(9):1413-1421.

2. Blum W, Sanford BL, Klisovic R, et al. Maintenance therapy with decitabine in younger adults with acute myeloid leukemia in first remission: a phase 2 Cancer and Leukemia Group B study (CALGB 10503). *Leukemia*. 2017;31(1):34-39.

3. Kolitz JE, George SL, Dodge RK, et al. Dose escalation studies of cytarabine, daunorubicin, and etoposide with and without multidrug resistance modulation with PSC-833 in untreated adults with acute myeloid leukemia younger than 60 years: final induction results of Cancer and Leukemia Group B study 9621. *J Clin Oncol*. 2004;22(21):4290-4301.

4. Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus chemotherapy for acute myeloid leukemia with a *FLT3* mutation. *N Engl J Med*. 2017;377(5):454-464.

5. Moore JO, George SL, Dodge RK, et al. Sequential multiagent chemotherapy is not superior to high-dose cytarabine alone as postremission intensification therapy for acute myeloid leukemia in adults under 60 years of age: Cancer and Leukemia Group B study 9222. *Blood*. 2005;105(9):3420-3427.

6. Mayer RJ, Davis RB, Schiffer CA, et al. Intensive postremission chemotherapy in adults with acute myeloid leukemia. *N Engl J Med*. 1994;331(14):896-903.

7. Moore JO, Dodge RK, Amrein PC, et al. Granulocyte-colony stimulating factor (filgrastim) accelerates granulocyte recovery after intensive postremission chemotherapy for acute myeloid leukemia with aziridinyl benzoquinone and mitoxantrone: Cancer and Leukemia Group B study 9022. *Blood*. 1997;89(3):780-788.

8. Kim D, Paggi JM, Park C, et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37(8):907-915.

9. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22(9):1760-1774.

10. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2012;29(1):15-21.

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

12. Cheson BD, Cassileth PA, Head DR, et al. Report of the National Cancer Institute-sponsored workshop on definitions of diagnosis and response in acute myeloid leukemia. *J Clin Oncol*. 1990;8(5):813-819.

Supplemental Table S1. List of back-splicing junctions of candidate circular RNAs that were included in the analysis of our cohort of younger adults with CN-AML with the Massive Scan for circular RNAs (MScircRNA) pipeline (provided as Excel Table).

Supplemental Table S2. List of circular RNAs that are detected in the RNAse R versus the mock-treated samples of AML patients and cell lines (provided as Excel Table).

Supplemental Table S3. Circular fraction (i.e., circular to linear ratio) of circular to linear RNAs that are detected in the RNAse R versus the mock-treated samples of AML patients and cell lines (provided as Excel Table).

Supplemental Table S4. Technical features and characteristics of the total RNA sequencing data of the cohort of younger adults with CN-AML by sample (provided as Excel Table).

Supplemental Table S5. List of the 180 circRNAs used for outcome analysis

UniqueID	NAME
chr1:113829591-113834439	*PTPN22*
chr1:117402185-117414831	*MAN1A2*
chr1:117402185-117420649	*MAN1A2*
chr1:117402185-117442325	*MAN1A2*
chr1:1223243-1223968	*SDF4*
chr1:149655698-149669310	*LINC00869*
chr1:155438326-155459898	*ASH1L*
chr1:155675009-155679512	*YY1AP1*
chr1:155725381-155726019	*DAP3*
chr1:155853275-155853806	*GON4L*
chr1:16564806-16567351	*NBPF1*
chr1:23030468-23050520	*KDM1A*
chr1:231795241-231818517	*DISC1*
chr1:247159005-247159813	*ZNF124*
chr1:44411980-44412722	*RNF220*
chr1:45640209-45642499	*GPBP1L1*
chr1:51827795-51834170	*NRD1*
chr1:58506059-58539310	*OMA1*
chr1:7777159-7778169	*VAMP3*
chr1:92333390-92380873	*RPAP2*
chr1:9931890-9934860	*LZIC*
chr10:103438014-103438808	*PDCD11*
chr10:110596397-110598290	*SMC3*
chr10:110964124-110985765	*SHOC2*
chr10:7276891-7285954	*SFMBT2*
chr10:7797046-7802854	*ATP5C1*
chr11:108267170-108267342	*ATM*
chr11:130260855-130261929	*ZBTB44*
chr11:14771936-14789242	*PDE3B*
chr11:18291441-18292976	*HPS5*
chr11:33286412-33287511	*HIPK3*
chr11:77619605-77625818	*CLNS1A*
chr11:77624962-77625818	*CLNS1A*
chr12:108652271-108654410	*CORO1C*
chr12:116096668-116111512	*MED13L*
chr12:120782654-120784593	*SPPL3*
chr12:26996741-26999676	*TM7SF3*
chr12:28225794-28259442	*CCDC91*
chr12:32598496-32611283	*FGD4*
chr12:46229152-46243314	*SLC38A1*
chr12:56666203-56670364	*PTGES3*
chr12:66203710-66228370	*IRAK3*
chr12:69251128-69262562	*CPSF6*
chr12:69800208-69801721	*RAB3IP*

chr13:32517856-32527532	*N4BP2L2*
chr13:50927406-50949505	*RNASEH2B*
chr14:102040235-102040673	*DYNC1H1*
chr14:21392149-21392253	*SNORD9*
chr14:21503172-21503881	*METTL3*
chr14:22909482-22911403	*RBM23*
chr14:34862043-34862322	*BAZ1A*
chr14:39179090-39179462	*PNN*
chr14:39276933-39279537	*RP11-407N17.3*
chr14:49586578-49586878	*RN7SL1*
chr14:50150007-50150230	*SOS2*
chr14:73147794-73148094	*PSEN1*
chr14:73147794-73148106	*PSEN1*
chr14:99458278-99465813	*SETD3*
chr15:101235081-101235577	*CHSY1*
chr15:25405460-25411971	*UBE3A*
chr15:41668827-41669958	*MGA*
chr15:49235850-49239367	*GALK2*
chr15:50038767-50047464	*ATP8B4*
chr15:58912562-58916999	*SLTM*
chr15:62007307-62013992	*VPS13C*
chr15:75859877-75873568	*UBE2Q2*
chr15:80120327-80122800	*ZFAND6*
chr15:89113724-89116521	*ABHD2*
chr16:30483826-30484263	*ITGAL*
chr16:68121986-68123121	*NFATC3*
chr16:68121986-68126610	*NFATC3*
chr16:69370482-69372355	*TERF2*
chr16:85633913-85634132	*GSE1*
chr17:20204332-20205912	*SPECC1*
chr17:59353214-59353526	*YPEL2*
chr17:64886085-64889752	*LRRC37A3*
chr17:67945408-67975958	*BPTF*
chr17:83084937-83085323	*METRNL*
chr18:12999420-13019206	*CEP192*
chr18:24064139-24069271	*TTC39C*
chr18:2890560-2892486	*EMILIN2*
chr18:32111753-32113860	*RNF138*
chr18:79695224-79704917	*CTDP1*
chr18:9182381-9221999	*ANKRD12*
chr19:12928341-12928847	*FARSA*
chr19:21033455-21034184	*ZNF430*
chr19:29985222-29986417	*URI1*
chr19:34430575-34434968	*UBA2*
chr19:40583397-40583717	*SHKBP1*
chr19:52282110-52283413	*ZNF766*
chr19:5604582-5604936	*SAFB2*

chr19:57455652-57456182	*AC004076.9*
chr19:8548035-8548317	*MYO1F*
chr2:106158057-106166083	*UXS1*
chr2:112299848-112300029	*ZC3H6*
chr2:147972738-147975975	*ORC4*
chr2:201145377-201149835	*CFLAR*
chr2:24135118-24147086	*FAM228B*
chr2:31917925-31932135	*MEMO1*
chr2:40428472-40430304	*SLC8A1*
chr2:45546731-45553730	*SRBD1*
chr2:58221941-58232112	*FANCL*
chr2:61522610-61526521	*XPO1*
chr2:61522610-61533903	*XPO1*
chr2:99169549-99171429	*MITD1*
chr20:32366383-32369123	*ASXL1*
chr20:58438944-58441083	*VAPB*
chr21:15014343-15043574	*NRIP1*
chr21:15762890-15766141	*USP25*
chr21:36247516-36248568	*DOPEY2*
chr3:138570317-138571356	*CEP70*
chr3:138570317-138572932	*CEP70*
chr3:142093060-142101841	*TFDP2*
chr3:149846010-149921227	*RNF13*
chr3:158122102-158123991	*RSRC1*
chr3:170136418-170149244	*PHC3*
chr3:170145422-170149244	*PHC3*
chr3:172247532-172251541	*FNDC3B*
chr3:18378169-18420991	*SATB1*
chr3:195381008-195392147	*ACAP2*
chr3:196391812-196403019	*UBXN7*
chr3:31576395-31580096	*STT3B*
chr3:50108069-50108304	*RBM5*
chr3:56592969-56594028	*CCDC66*
chr4:102714437-102726683	*MANBA*
chr4:105424195-105456745	*PPA2*
chr4:128992166-129003876	*SCLT1*
chr4:139125163-139139497	*ELF2*
chr4:143543508-143543972	*SMARCA5*
chr4:152411302-152412529	*FBXW7*
chr4:177353307-177360677	*NEIL3*
chr4:36228581-36229645	*ARAP2*
chr4:37631384-37638504	*RELL1*
chr4:38089931-38103157	*TBC1D1*
chr4:39913610-39925933	*PDS5A*
chr4:87195323-87195690	*KLHL8*
chr5:123545416-123557564	*CSNK1G3*
chr5:137985256-137988315	*FAM13B*

chr5:168488601-168494650	*RARS*
chr5:32135571-32143880	*GOLPH3*
chr5:50402285-50411383	*EMB*
chr5:56864733-56865977	*MAP3K1*
chr5:57246299-57251141	*GPBP1*
chr5:65988634-65994864	*ERBB2IP*
chr5:73074741-73077493	*FCHO2*
chr5:77463094-77464809	*WDR41*
chr5:83537006-83542268	*VCAN*
chr5:95755395-95763620	*RHOBTB3*
chr6:138943512-138944622	*REPS1*
chr6:18236451-18258405	*DEK*
chr6:4891712-4892379	*CDYL*
chr7:100023418-100024307	*ZKSCAN1*
chr7:155672866-155680908	*RBM33*
chr7:158759485-158764853	*ESYT2*
chr7:158788003-158799072	*ESYT2*
chr7:16258389-16278226	*ISPD*
chr7:23611170-23611553	*CCDC126*
chr7:55982209-55983868	*GBAS*
chr7:64543706-64544432	*ZNF680*
chr7:66127703-66134374	*AC068533.7*
chr7:74219743-74220734	*LAT2*
chr8:37765525-37766355	*PROSC*
chr8:431340-435707	*FBXO25*
chr8:51860844-51861246	*PCMTD1*
chr8:61680967-61684188	*ASPH*
chr8:70213902-70216764	*NCOA2*
chr9:111386376-111391824	*KIAA0368*
chr9:121161855-121166180	*CNTRL*
chr9:131506113-131506453	*POMT1*
chr9:135881632-135883078	*CAMSAP1*
chr9:33953284-33963791	*UBAP2*
chr9:33971650-33973237	*UBAP2*
chr9:37126311-37126942	*ZCCHC7*
chr9:6880011-6893232	*KDM4C*
chr9:85618982-85633374	*AGTPBP1*
chr9:86303085-86304947	*ZCCHC6*
chr9:93471140-93498886	*FAM120A*
chrX:131749305-131794466	*FIRRE*
chrX:1593443-1595532	*AKAP17A*
chrX:53403565-53403893	*SMC1A*

Supplemental Table S6. Clinical and molecular features in patients in the training and validation sets that were used for the evaluation of the prognostic significance of circRNA expression in CN-AML

Characteristic	Training set (n=254)	Validation set (n=111)	*P*
Age, y			.52
Median	47	45	
Range	17-59	18-59	
Sex, n (%)			.82
Male	132 (52)	56 (50)	
Female	122 (48)	55 (50)	
Race, n (%)			1.00
White	225 (90)	99 (91)	
Non-white	24 (10)	10 (9)	
Hemoglobin, g/dL			.27
Median	9.3	8.8	
Range	4.6-14.4	4.2-25.1	
Platelet count, $x10^9$/L			.99
Median	60	54	
Range	8-445	8-433	
WBC count, $x10^9$/L			.93
Median	27.3	30.4	
Range	0.6-308.8	0.9-475.0	
% Blood blasts			.02
Median	63	50	
Range	0-97	0-97	
% Bone marrow blasts			.22
Median	70	65	
Range	18-96	18-95	
Extramedullary involvement, n (%)			1.00
Present	74 (30)	32 (30)	
Absent	175 (70)	75 (70)	
ASXL1, n (%)			1.00
Mutated	9 (4)	3 (3)	
Wild-type	232 (96)	102 (97)	
CEBPA, n (%)			.63
Double Mutated	40 (16)	14 (14)	
Wild-type	206 (84)	88 (86)	
DNMT3A, n (%)			.91
Mutated	96 (39)	43 (40)	
R882	71	34	
Non-R882	25	9	
Wild-type	149 (61)	64 (60)	
FLT3-ITD, n (%)			.72
Present	94 (38)	43 (40)	
Absent	156 (62)	65 (60)	
FLT3-TKD, n (%)			.007
Present	32 (13)	4 (4)	
Absent	212 (87)	102 (96)	
IDH1, n (%)			.67
Mutated	20 (8)	7 (7)	
Wild-type	225 (92)	100 (93)	
IDH2, n (%)			.56

Characteristic	Training set (n=254)	Validation set (n=111)	*P*
Mutated	22 (9)	12 (11)	
Wild-type	223 (91)	95 (89)	
NPM1, n (%)			.29
Mutated	135 (56)	66 (63)	
Wild-type	104 (44)	39 (37)	
RUNX1, n (%)			.80
Mutated	14 (6)	5 (5)	
Wild-type	231 (94)	102 (95)	
SF1, n (%)			1.00
Mutated	2 (1)	0 (0)	
Wild-type	243 (99)	107 (100)	
SF3A1, n (%)			.30
Mutated	0 (0)	1 (1)	
Wild-type	245 (100)	106 (99)	
SF3B1, n (%)			.50
Mutated	6 (2)	4 (4)	
Wild-type	239 (98)	103 (96)	
SRSF2, n (%)			1.00
Mutated	8 (3)	3 (3)	
Wild-type	235 (97)	104 (97)	
TET2, n (%)			.28
Mutated	31 (13)	9 (8)	
Wild-type	214 (87)	98 (92)	
U2AF1, n (%)			1.00
Mutated	3 (1)	1 (1)	
Wild-type	242 (99)	106 (99)	
U2AF2, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	245 (100)	107 (100)	
WT1, n (%)			.71
Mutated	26 (11)	13 (12)	
Wild-type	219 (89)	94 (88)	
ZRSR2, n (%)			1.00
Mutated	11 (4)	4 (4)	
Wild-type	234 (96)	103 (96)	
ELN Genetic Group*, n (%)			.77
Favorable	139 (58)	60 (60)	
Intermediate	61 (26)	27 (27)	
Adverse	39 (16)	13 (13)	
NPM1 and *DNMT3A*			.90
Either or neither mutated	165 (69)	74 (70)	
Both mutated	74 (31)	31 (30)	
NPM1 and *FLT3*-ITD			.35
Either or neither mutated	179 (76)	73 (70)	
Both mutated	58 (24)	31 (30)	
ERG expression group†, n (%)			.008
High	140 (56)	44 (40)	
Low	112 (44)	66 (60)	
BAALC expression group†, n (%)			.002
High	129 (56)	39 (38)	
Low	101 (44)	65 (63)	
MN1 expression group†, n (%)			.10

Characteristic	Training set (n=254)	Validation set (n=111)	*P*
High	129 (54)	47 (44)	
Low	112 (46)	61 (56)	
miR-181a expression group†, n (%)			.32
High	106 (52)	42 (46)	
Low	97 (48)	50 (54)	
miR-3151, n (%)			.24
Expressed	39 (19)	12 (13)	
Not expressed	164 (81)	80 (87)	
miR-155 expression group†, n (%)			1.00
High	101 (50)	46 (50)	
Low	102 (50)	46 (50)	

n, number; WBC, white blood cell; BM, bone marrow; ELN, European LeukemiaNet; *FLT3*-ITD, internal tandem duplication of the *FLT3* gene; *FLT3*-TKD, tyrosine kinase domain mutation in the *FLT3* gene, NA, *P*-value could not be calculated because of 0 cell counts.

* Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN Favorable-risk category comprises patients with double-mutated *CEBPA*, patients with mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}. The ELN Intermediate-risk category includes patients with wild-type or single mutated *CEBPA*, patients with wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}, and/or patients with mutated *NPM1* and *FLT3*-ITD^{high}. The ELN Adverse-risk category comprises patients with wild-type or single mutated *CEBPA* and wild-type *NPM1* with *FLT3*-ITD^{high}, and/or patients with mutated *RUNX1* and/or *ASXL1* (if these mutations do not co-occur with Favorable-risk subtypes), and/or patients with mutated *TP53*. *FLT3*-ITD^{low} is defined by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of less than 0.5 and *FLT3*-ITD^{high} is defined as by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of equal to or more than 0.5.

† The median expression value was used as a cut point.

Supplemental Table S7. CircRNAs associated with EFS in the training set.

CircRNA	EFS (*P*-value)
CCDC91	.02
CFLAR	.03
DAP3	<.01
FCHO2	<.01
ISPD	.02
KLHL8	<.01
PCMTD1	<.01
PDCD11	<.01
RAB3IP	<.01
SMARCA5	.01
SMC1A	<.01
ZCCHC6	<.01

Supplemental Table S8. Clinical and molecular features of CN-AML patients with high and those with low *circKLHL8* expression in the validation set

Characteristic	Low *circKLHL8** (n=56)	High *circKLHL8** (n=55)	*P*
Age, y			.42
Median	46	44	
Range	18-59	18-59	
Sex, n (%)			.71
Male	27 (48)	29 (53)	
Female	29 (52)	26 (47)	
Race, n (%)			1.00
White	51 (91)	48 (91)	
Non-white	5 (9)	5 (9)	
Hemoglobin, g/dL			.60
Median	8.8	9.1	
Range	5.5-13.4	4.2-25.1	
Platelet count, $x10^9$/L			.05
Median	46	60	
Range	8-270	10-433	
WBC count, $x10^9$/L			.47
Median	30.3	31.6	
Range	2.2-475.0	0.9-208.8	
% Blood blasts			.002
Median	66	40	
Range	0-95	0-97	
% Bone marrow blasts			.05
Median	67	63	
Range	21-95	18-94	
Extramedullary involvement, n (%)			.83
Present	17 (31)	15 (28)	
Absent	37 (69)	38 (72)	
ELN Genetic Group*, n (%)			<.001
Favorable	22 (44)	38 (76)	
Intermediate	22 (44)	5 (10)	
Adverse	6 (12)	7 (14)	
ASXL1, n (%)			1.00
Mutated	2 (4)	1 (2)	
Wild-type	54 (96)	48 (98)	
CEBPA, n (%)			.25
Double Mutated	5 (9)	9 (18)	
Wild-type	48 (91)	40 (82)	
DNMT3A, n (%)			.56
Mutated	21 (38)	22 (43)	
R882	16	18	
Non-R882	5	4	
Wild-type	35 (63)	29 (57)	
FLT3-ITD, n (%)			<.001
Present	32 (58)	11 (21)	
Absent	23 (42)	42 (79)	
FLT3-TKD, n (%)			.05
Present	0 (0)	4 (8)	
Absent	55 (100)	47 (92)	
IDH1, n (%)			.12

Characteristic	Low *circKLHL8** (n=56)	High *circKLHL8** (n=55)	*P*
Mutated	6 (11)	1 (2)	
Wild-type	50 (89)	50 (98)	
IDH2, n (%)			.54
Mutated	5 (9)	7 (14)	
Wild-type	51 (91)	44 (86)	
NPM1, n (%)			.69
Mutated	36 (65)	30 (60)	
Wild-type	19 (35)	20 (40)	
RUNX1, n (%)			.67
Mutated	2 (4)	3 (6)	
Wild-type	54 (96)	48 (94)	
SF1, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	56 (100)	51 (100)	
SF3A1, n (%)			1.00
Mutated	1 (2)	0 (0)	
Wild-type	55 (98)	51 (100)	
SF3B1, n (%)			.35
Mutated	1 (2)	3 (6)	
Wild-type	55 (98)	48 (94)	
SRSF2, n (%)			.60
Mutated	1 (2)	2 (4)	
Wild-type	55 (98)	49 (96)	
TET2, n (%)			.73
Mutated	4 (7)	5 (10)	
Wild-type	52 (93)	46 (90)	
U2AF1, n (%)			1.00
Mutated	1 (2)	0 (0)	
Wild-type	55 (98)	51 (100)	
U2AF2, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	56 (100)	51 (100)	
WT1, n (%)			.24
Mutated	9 (16)	4 (8)	
Wild-type	47 (84)	47 (92)	
ZRSR2, n (%)			.05
Mutated	0 (0)	4 (8)	
Wild-type	56 (100)	47 (92)	
ERG expression group†, n (%)			<.001
High	31 (56)	13 (24)	
Low	24 (44)	42 (76)	
BAALC expression group†, n (%)			.55
High	22 (41)	17 (34)	
Low	32 (59)	33 (66)	
MN1 expression group†, n (%)			.33
High	26 (49)	21 (38)	
Low	27 (51)	34 (62)	
miR-181a expression group†, n (%)			.30
High	24 (52)	18 (39)	
Low	22 (48)	28 (61)	
miR-3151, n (%)			.76
Expressed	5 (11)	7 (15)	

Characteristic	Low *circKLHL8** (n=56)	High *circKLHL8** (n=55)	*P*
Not expressed	41 (89)	39 (85)	
miR-155 expression group†, n (%)			<.001
High	34 (74)	12 (26)	
Low	12 (26)	34 (74)	

n, number; WBC, white blood cell; BM, bone marrow; ELN, European LeukemiaNet; *FLT3*-ITD, internal tandem duplication of the *FLT3* gene; and *FLT3*-TKD, tyrosine kinase domain mutation in the *FLT3* gene, NA, p-value cannot be calculated because of 0 cell counts.

* Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN Favorable-risk category comprises patients with double-mutated *CEBPA*, patients with mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}. The ELN Intermediate-risk category includes patients with wild-type or single mutated *CEBPA*, patients with wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}, and/or patients with mutated *NPM1* and *FLT3*-ITD^{high}. The ELN Adverse-risk category comprises patients with wild-type or single mutated *CEBPA* and wild-type *NPM1* with *FLT3*-ITD^{high}, and/or patients with mutated *RUNX1* and/or *ASXL1* (if these mutations do not co-occur with Favorable-risk subtypes), and/or patients with mutated *TP53*. *FLT3*-ITD^{low} is defined by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of less than 0.5 and *FLT3*-ITD^{high} is defined as by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of equal to or more than 0.5.

† The median expression value was used as a cut point.

Supplemental Table S9. Clinical and molecular features in CN-AML patients with high and those with low *circSMC1A* in the validation set of the studied cohort

n=111	Low *circSMC1A* (n=43)	High *circSMC1A* (n=68)	*P*
Age, y			.63
Median	49	44	
Range	18-59	18-59	
Sex, n (%)			.002
Male	30 (70)	26 (38)	
Female	13 (30)	42 (62)	
Race, n (%)			.74
White	40 (93)	59 (89)	
Non-white	3 (7)	7 (11)	
Hemoglobin, g/dL			.68
Median	8.7	9.0	
Range	5.5-13.4	4.2-25.1	
Platelet count, $x10^9$/L			.25
Median	48	55	
Range	12-266	8-433	
WBC count, $x10^9$/L			.06
Median	32.2	30.0	
Range	2.3-303.6	0.9-475.0	
% Blood blasts			.29
Median	54	47	
Range	0-97	2-91	
% Bone marrow blasts			.64
Median	66	65	
Range	21-95	18-94	
Extramedullary involvement, n (%)			1.00
Present	12 (30)	20 (30)	
Absent	28 (70)	47 (70)	
ELN Genetic Group*, n (%)			.12
Favorable	20 (51)	40 (66)	
Intermediate	15 (38)	12 (20)	
Adverse	4 (10)	9 (15)	
ASXL1, n (%)			.29
Mutated	0 (0)	3 (5)	
Wild-type	40 (100)	62 (95)	
CEBPA, n (%)			.39
Double Mutated	4 (10)	10 (16)	
Wild-type	37 (90)	51 (84)	
DNMT3A, n (%)			.55
Mutated	18 (44)	25 (38)	
R882	15	19	
Non-R882	3	6	
Wild-type	23 (56)	41 (62)	
FLT3-ITD, n (%)			.04
Present	22 (52)	21 (32)	
Absent	20 (48)	45 (68)	
FLT3-TKD, n (%)			1.00
Present	1 (2)	3 (5)	
Absent	40 (98)	62 (95)	
IDH1, n (%)			1.00

n=111	Low *circSMC1A* (n=43)	High *circSMC1A* (n=68)	*P*
Mutated	3 (7)	4 (6)	
Wild-type	38 (93)	62 (94)	
IDH2, n (%)			1.00
Mutated	5 (12)	7 (11)	
Wild-type	36 (88)	59 (89)	
NPM1, n (%)			.53
Mutated	27 (68)	39 (60)	
Wild-type	13 (33)	26 (40)	
RUNX1, n (%)			.15
Mutated	0 (0)	5 (8)	
Wild-type	41 (100)	61 (92)	
SF1, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	41 (100)	66 (100)	
SF3A1, n (%)			1.00
Mutated	0 (0)	1 (2)	
Wild-type	41 (100)	65 (98)	
SF3B1, n (%)			1.00
Mutated	1 (2)	3 (5)	
Wild-type	40 (98)	63 (95)	
SRSF2, n (%)			1.00
Mutated	1 (2)	2 (3)	
Wild-type	40 (98)	64 (97)	
TET2, n (%)			.48
Mutated	2 (5)	7 (11)	
Wild-type	39 (95)	59 (89)	
U2AF1, n (%)			1.00
Mutated	0 (0)	1 (2)	
Wild-type	41 (100)	65 (98)	
U2AF2, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	41 (100)	66 (100)	
WT1, n (%)			.24
Mutated	7 (17)	6 (9)	
Wild-type	34 (83)	60 (91)	
ZRSR2, n (%)			.30
Mutated	0 (0)	4 (6)	
Wild-type	41 (100)	62 (94)	
ERG expression group†, n (%)			.16
High	21 (49)	23 (34)	
Low	22 (51)	44 (66)	
BAALC expression group†, n (%)			1.00
High	15 (38)	24 (37)	
Low	24 (62)	41 (63)	
MN1 expression group†, n (%)			1.00
High	18 (43)	29 (44)	
Low	24 (57)	37 (56)	
miR-181a expression group†, n (%)			.14
High	12 (35)	30 (52)	
Low	22 (65)	28 (48)	
miR-3151, n (%)			.52
Expressed	3 (9)	9 (16)	
Not expressed	31 (91)	49 (84)	

n=111	Low *circSMC1A* (n=43)	High *circSMC1A* (n=68)	*P*
miR-155 expression group†, n (%)			.13
High	21 (62)	25 (43)	
Low	13 (38)	33 (57)	

n, number; WBC, white blood cell; BM, bone marrow; ELN, European LeukemiaNet; *FLT3*-ITD, internal tandem duplication of the *FLT3* gene; and *FLT3*-TKD, tyrosine kinase domain mutation in the *FLT3* gene, NA, p-value cannot be calculated because of 0 cell counts.

* Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN Favorable-risk category comprises patients with double-mutated *CEBPA*, patients with mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}. The ELN Intermediate-risk category includes patients with wild-type or single mutated *CEBPA*, patients with wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}, and/or patients with mutated *NPM1* and *FLT3*-ITD^{high}. The ELN Adverse-risk category comprises patients with wild-type or single mutated *CEBPA* and wild-type *NPM1* with *FLT3*-ITD^{high}, and/or patients with mutated *RUNX1* and/or *ASXL1* (if these mutations do not co-occur with Favorable-risk subtypes), and/or patients with mutated *TP53*. *FLT3*-ITD^{low} is defined by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of less than 0.5 and *FLT3*-ITD^{high} is defined as by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of equal to or more than 0.5.

† The median expression value was used as a cut point.

Supplemental Table S10. Clinical and molecular features in CN-AML patients with high and those with low *circCFLAR* expression in the validation set of the studied cohort

n=111	Low *circCFLAR* (n=56)	High *circCFLAR*(n=55)	*P*
Age, y			.20
Median	47	43	
Range	18-59	18-59	
Sex, n (%)			.71
Male	27 (48)	29 (53)	
Female	29 (52)	26 (47)	
Race, n (%)			.20
White	52 (95)	47 (87)	
Non-white	3 (5)	7 (13)	
Hemoglobin, g/dL			.42
Median	8.9	8.8	
Range	5.5-25.1	4.2-13.2	
Platelet count, $x10^9$/L			.07
Median	49	59	
Range	8-270	10-433	
WBC count, $x10^9$/L			.26
Median	30.7	30.4	
Range	0.9-475.0	1.4-141.0	
% Blood blasts			.12
Median	58	39	
Range	0-91	0-97	
%Bone marrow blasts			.17
Median	66	64	
Range	21-95	18-94	
Extramedullary involvement, n (%)			.83
Present	17 (31)	15 (28)	
Absent	37 (69)	38 (72)	
ELN Genetic Group*, n (%)			.03
Favorable	24 (48)	36 (72)	
Intermediate	19 (38)	8 (16)	
Adverse	7 (14)	6 (12)	
ASXL1, n (%)			.11
Mutated	0 (0)	3 (6)	
Wild-type	54 (100)	48 (94)	
CEBPA, n (%)			.08
Double Mutated	4 (8)	10 (20)	
Wild-type	49 (92)	39 (80)	
DNMT3A, n (%)			.56
Mutated	20 (37)	23 (43)	
Wild-type	34 (63)	30 (57)	
FLT3-ITD, n (%)			< .001
Present	34 (62)	9 (17)	
Absent	21 (38)	44 (83)	
FLT3-TKD, n (%)			1.00
Present	2 (4)	2 (4)	
Absent	52 (96)	50 (96)	
IDH1, n (%)			.72
Mutated	3 (6)	4 (8)	
Wild-type	51 (94)	49 (92)	

n=111	Low *circCFLAR* (n=56)	High *circCFLAR*(n=55)	*P*
IDH2, n (%)			.76
Mutated	7 (13)	5 (9)	
Wild-type	47 (87)	48 (91)	
NPM1, n (%)			.32
Mutated	36 (68)	30 (58)	
Wild-type	17 (32)	22 (42)	
RUNX1, n (%)			.68
Mutated	2 (4)	3 (6)	
Wild-type	52 (96)	50 (94)	
SF1, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	54 (100)	53 (100)	
SF3A1, n (%)			1.00
Mutated	1 (2)	0 (0)	
Wild-type	53 (98)	53 (100)	
SF3B1, n (%)			.36
Mutated	1 (2)	3 (6)	
Wild-type	53 (98)	50 (94)	
SRSF2, n (%)			1.00
Mutated	2 (4)	1 (2)	
Wild-type	52 (96)	52 (98)	
TET2, n (%)			.74
Mutated	4 (7)	5 (9)	
Wild-type	50 (93)	48 (91)	
U2AF1, n (%)			.50
Mutated	0 (0)	1 (2)	
Wild-type	54 (100)	52 (98)	
U2AF2, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	54 (100)	53 (100)	
WT1, n (%)			**.02**
Mutated	11 (20)	2 (4)	
Wild-type	43 (80)	51 (96)	
ZRSR2, n (%)	2 *unknown*	2 *unknown*	.06
Mutated	0 (0)	4 (8)	
Wild-type	54 (100)	49 (92)	
ERG expression group†, n (%)			.08
High	27 (48)	17 (31)	
Low	29 (52)	37 (69)	
BAALC expression group†, n (%)			1.00
High	20 (37)	19 (38)	
Low	34 (63)	31 (62)	
MN1 expression group†, n (%)			.25
High	27 (49)	20 (38)	
Low	28 (51)	33 (62)	
miR-181a expression group†, n (%)			1.00
High	22 (46)	20 (45)	
Low	26 (54)	24 (55)	
miR-3151, n (%)			1.00
Expressed	6 (13)	6 (14)	
Not expressed	42 (88)	38 (86)	
miR-155 expression group†, n (%)			**.002**
High	32 (67)	14 (32)	

n=111	Low *circCFLAR* (n=56)	High *circCFLAR*(n=55)	*P*
Low	16 (33)	30 (68)	

n, number; WBC, white blood cell; BM, bone marrow; ELN, European LeukemiaNet; *FLT3*-ITD, internal tandem duplication of the *FLT3* gene; and *FLT3*-TKD, tyrosine kinase domain mutation in the *FLT3* gene, NA, p-value cannot be calculated because of 0 cell counts.

* Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN Favorable-risk category comprises patients with double-mutated *CEBPA*, patients with mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}. The ELN Intermediate-risk category includes patients with wild-type or single mutated *CEBPA*, patients with wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}, and/or patients with mutated *NPM1* and *FLT3*-ITD^{high}. The ELN Adverse-risk category comprises patients with wild-type or single mutated *CEBPA* and wild-type *NPM1* with *FLT3*-ITD^{high}, and/or patients with mutated *RUNX1* and/or *ASXL1* (if these mutations do not co-occur with Favorable-risk subtypes), and/or patients with mutated *TP53*. *FLT3*-ITD^{low} is defined by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of less than 0.5 and *FLT3*-ITD^{high} is defined as by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of equal to or more than 0.5.

† The median expression value was used as a cut point.

Supplemental Table S11. Clinical and molecular features in CN-AML patients with high and those with low *circFCHO* expression in the validation set of the studied cohort.

n=111	Low *circFCHO2* (n=56)	High *circFCHO2* (n=55)	*P*
Age, y			.56
Median	47	44	
Range	18-59	18-59	
Sex, n (%)			.71
Male	27 (48)	29 (53)	
Female	29 (52)	26 (47)	
Race, n (%)			.20
White	52 (95)	47 (87)	
Non-white	3 (5)	7 (13)	
Hemoglobin, g/dL			.82
Median	8.8	9.0	
Range	5.5-12.9	4.2-25.1	
Platelet count, x10^9/L			.38
Median	53	55	
Range	8-270	10-433	
WBC count, x10^9/L			.004
Median	36.2	25.0	
Range	1.6-303.6	0.9-475.0	
% Blood blasts			.88
Median	50	50	
Range	0-97	1-91	
% Bone marrow blasts			.14
Median	69	63	
Range	18-95	20-94	
Extramedullary involvement, n (%)			.09
Present	12 (22)	20 (38)	
Absent	43 (78)	32 (62)	
ELN Genetic Group*, n (%)			.03
Favorable	26 (50)	34 (71)	
Intermediate	20 (38)	7 (15)	
Adverse	6 (12)	7 (15)	
ASXL1, n (%)			.61
Mutated	1 (2)	2 (4)	
Wild-type	53 (98)	49 (96)	
CEBPA, n (%)			.02
Double Mutated	3 (6)	11 (22)	
Wild-type	49 (94)	39 (78)	
DNMT3A, n (%)			.33
Mutated	19 (35)	24 (45)	
R882	16	18	
Non-R882	3	6	
Wild-type	35 (65)	29 (55)	
FLT3-ITD, n (%)			< .001
Present	31 (57)	12 (22)	
Absent	23 (43)	42 (78)	
FLT3-TKD, n (%)			1.00
Present	2 (4)	2 (4)	
Absent	51 (96)	51 (96)	
IDH1, n (%)			1.00
Mutated	4 (7)	3 (6)	

n=111	Low *circFCHO2* (n=56)	High *circFCHO2* (n=55)	*P*
Wild-type	50 (93)	50 (94)	
IDH2, n (%)			.56
Mutated	5 (9)	7 (13)	
Wild-type	49 (91)	46 (87)	
NPM1, n (%)			.005
Mutated	41 (76)	25 (49)	
Wild-type	13 (24)	26 (51)	
RUNX1, n (%)			.21
Mutated	1 (2)	4 (8)	
Wild-type	53 (98)	49 (92)	
SF1, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	54 (100)	53 (100)	
SF3A1, n (%)			.50
Mutated	0 (0)	1 (2)	
Wild-type	54 (100)	52 (98)	
SF3B1, n (%)			.36
Mutated	1 (2)	3 (6)	
Wild-type	53 (98)	50 (94)	
SRSF2, n (%)			.62
Mutated	1 (2)	2 (4)	
Wild-type	53 (98)	51 (96)	
TET2, n (%)			1.00
Mutated	5 (9)	4 (8)	
Wild-type	49 (91)	49 (92)	
U2AF1, n (%)			1.00
Mutated	1 (2)	0 (0)	
Wild-type	53 (98)	53 (100)	
U2AF2, n (%)			NA
Mutated	0 (0)	0 (0)	
Wild-type	54 (100)	53 (100)	
WT1, n (%)			.24
Mutated	9 (17)	4 (8)	
Wild-type	45 (83)	49 (92)	
ZRSR2, n (%)			.36
Mutated	1 (2)	3 (6)	
Wild-type	53 (98)	50 (94)	
ERG expression group†, n (%)			1.00
High	22 (39)	22 (41)	
Low	34 (61)	32 (59)	
BAALC expression group†, n (%)			.01
High	13 (25)	26 (50)	
Low	39 (75)	26 (50)	
MN1 expression group†, n (%)			.70
High	22 (41)	25 (46)	
Low	32 (59)	29 (54)	
miR-181a expression group†, n (%)			.04
High	16 (34)	26 (58)	
Low	31 (66)	19 (42)	
miR-3151 expression group†, n (%)			.01
Expressed	2 (4)	10 (22)	
Not expressed	45 (96)	35 (78)	
miR-155 expression group†, n (%)			.40

n=111	Low *circFCHO2* (n=56)	High *circFCHO2* (n=55)	*P*
High	26 (55)	20 (44)	
Low	21 (45)	25 (56)	

n, number; WBC, white blood cell; BM, bone marrow; ELN, European LeukemiaNet; *FLT3*-ITD, internal tandem duplication of the *FLT3* gene; and *FLT3*-TKD, tyrosine kinase domain mutation in the *FLT3* gene, NA, p-value cannot be calculated because of 0 cell counts.

***** Among patients with cytogenetically normal acute myeloid leukemia (CN-AML), the ELN Favorable-risk category comprises patients with double-mutated *CEBPA*, patients with mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}. The ELN Intermediate-risk category includes patients with wild-type or single mutated *CEBPA*, patients with wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low}, and/or patients with mutated *NPM1* and *FLT3*-ITD^{high}. The ELN Adverse-risk category comprises patients with wild-type or single mutated *CEBPA* and wild-type *NPM1* with *FLT3*-ITD^{high}, and/or patients with mutated *RUNX1* and/or *ASXL1* (if these mutations do not co-occur with Favorable-risk subtypes), and/or patients with mutated *TP53*. *FLT3*-ITD^{low} is defined by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of less than 0.5 and *FLT3*-ITD^{high} is defined as by a *FLT3*-ITD/*FLT3* wild-type allelic ratio of equal to or more than 0.5.

† The median expression value was used as a cut point.

Supplemental Table S12. List of genes, whose differential expression associates with *circCFLAR* expression status (provided as Excel Table).

Supplemental Table S13. List of genes, whose differential expression associates with *circKLHL8* expression status (provided as Excel Table).

Supplemental Table S14. List of biologic processes that associate with high *circCFLAR* expression as evaluated by Gene Ontology Pathway Analysis. The top-5 pathways are listed. Pathways are ranked according to false discovery rate.

Molecular pathway	Number of genes	Fold enrichment	False discovery rate
Regulation of Immune Response	22	4.74	1.98E-06
Leukocyte migration	10	6.51	1.23E-03
Positive regulation of cytokine production	21	4.84	1.85E-03
Leukocyte chemotaxis	7	11.87	9.53E-04
Immune response-regulating signaling pathway	12	4.95	1.86E-03

Supplemental Table S15. List of biologic processes that associate with high *circKLHL8* expression as evaluated by Gene Ontology Pathway Analysis. The top-5 pathways are listed. Pathways are ranked according to false discovery rate.

Molecular pathway	Number of genes	Fold enrichment	False discovery rate
Regulation of cell differentiation	83	8.11	1.69E-27
Regulation of cytokine secretion	71	7.68	2.94E-23
Regulation of apoptotic process	85	5.53	2.29E-19
Regulation of leukocyte activation	54	4.19	1.82E-15
Regulation of immune response	15	2.46	3.85E-04

Supplemental Table S16. List of primers used for linear and circular RNA profiling

	RT-qPCR Assays
	circCFLAR
Fw_Primer	5'-ACAAGATGAAGAGCAAGCCC -3'
Rev_Primer	5'-CCAAGCTGTTCCTTAAGTCTTTG -3'
Probe	5'-/56-FAM/TCCCATTAT/ZEN/GGAGCTGTCTCATTGCC/3IABkFQ/-3'
	linCFLAR
Fw_Primer	5'-AGAACATCCACAGAATAGACCTG -3'
Rev_Primer	5'-CTTGAGACTCTTTTGGATTGCTG -3'
Probe	5'-/56-FAM/CCTGTAACT/ZEN/TGTCCCTGCTCCTTGA/3IABkFQ/
	circPCMTD1
Fw_Primer	5'-AACTTCAACCAGGATTGTCTTTTC
Rv_Primer	5'-ATGGCTTCCAATATTGCACTTG
Probe	5'-/56-FAM/ACTTAAATA/ZEN/TCCGGTTCCACTTCCCAGG/3IABkFQ/
	linPCMTD1
Fw_Primer	5'- GTGCAGAAAGACCATGAAAACTAC-3'
Rev_Primer	5'- GGAGCAAATGAAACAGCAAGG-3'
Probe	5'-/56-FAM/TATGCGAAC/ZEN/TGGACAGAACACTTGGG/3IABkFQ/
	circKLHL8
Fw_Primer	5'- TTATGAAAATGGAGAACTCTGTGATG -3'
Rev_Primer	5'- GAGAAGGCAGAAGGTAGATGTAG -3'
Probe	5'-/56-FAM/ACTGTACAA/ZEN/TGGCAGTTGAGTGGGT/3IABkFQ/
	linKLHL8
Fw_Primer	5'-GGATGAAACACTTGCACAGG
Rv_Primer	5'-CTGCTACTCAAGTGAAGGTGG
Probe	5'-/56-FAM/CGGTTGATT/ZEN/TTCTTATGGGTGTTGTGGC/3IABkFQ/
	cirFBXW7
Fw_Primer	5'- CTAACAGTGTCACGAACTCCAG -3'
Rev_Primer	5'- CTCCACTTCTGGCAATCTATCC -3'
Probe	5'-/56-FAM/CGTTCACCA/ZEN/ACTCTCCTCCCCATT/3IABkFQ/

	RT-qPCR Primers
	linKLHL8
Fw_Primer	GTCTCCCAAACCTGACTGTC
Rv_Primer	GTTGGTTCCCTTCCTCCTTC
Probe	/56-FAM/TCCTCTTTT/ZEN/CCTCTTCCTGGGTCTTTTC/3IABkFQ/
	circFCH02
Fw_Primer	TCACCAGCAATCCAACTCC
Rv_Primer	AGGTATCCTGATTTCCAAGGC
Probe	56-FAM/TTTCACCCT/ZEN/GAAGCACAACACAGC/3IABkFQ
	circSMC1A
Fw_Primer	GGTGTTTGAAGAGTTTTGTCGG
Rv_Primer	CATTAATGCGAGGCCCAAAG
Probe	56-FAM/AGGTGAAAC/ZEN/GGCAGAATGAAATCGC/3IABkFQ/
	Cloning Primers
	circCFLAR
Fw_Primer	5'-CTAGGAATCTGCCTGATAATCGATTGC-3'
Rv_Primer	5'-GCTCTTCATCTTGTATCTCTCTTCAGG-3'
	circKLHL8
Fw_Primer	5'-CTGAATTGTAGTGTCTACATCTACCTTCTG-3'
Rv_Primer	5'-CTGCCATTGTACAGTTTGCTGTG-3'
	circPCMTD1
Fw_Primer	5'-CTATCATGGGAGGAGCTGTGAG-3'
Rv_Primer	5'-GTAGAAATGGCTTCCAATATTGCACTTG-3'
	cirFBXW7
Fw_Primer	5'-CAAGGTCCAAGAAGTAGCAAGCTG-3'
Rv_Primer	5'-CTAGACTATCAGAAGATGCAAAGGTTTTCAC-3'

Abbreviations: Fw, forward; Rev: reverse; FAM, 6-carboxyfluorescein; ZEN, ZEN internal quencher; 3IABkFQ, 3' end Iowa black quencher

Supplemental Table S17. Gapmers sequences

	Gapmers
Negative Control	5'-+C*+G*+A*A*T*A*G*T*T*A*G*T*A*+G*+C*+G-3'
Anti-*circCFLAR*	5'-+A*+T*+T*A*T*G*G*A*G*C*T*G*T*C*T*+A*+T-3'
Anti-*circPCMTD1*	5'-+A*+T**C*T*A*A*A*A*T*T*A*A*G*C*+C*+C*+A-3'
Anti-*circKLHL8*	5'-+T*+T*G*C*T*G*T*G*A*C*T*T*G*A*G*T*+G*+T-3'
Anti-*circFBXW7*	5'-+T*+C*+C*+T*A*A*G*G*A*A*G*T*A*A*T*C*T*+T*+T-3'

Abbreviations/Symbols: +, locked-nucleic acid-modified DNA base; *, phosphorothioated

Supplemental Table S18. List of genes, whose differential expression associates with *circFBXW7* expression status (provided as Excel Table).

Supplemental Table S19. List of biologic processes that associate with high *circFBXW7* expression as evaluated by Gene Ontology Pathway Analysis. The top-3 pathways are listed. Pathways are ranked according to false discovery rate.

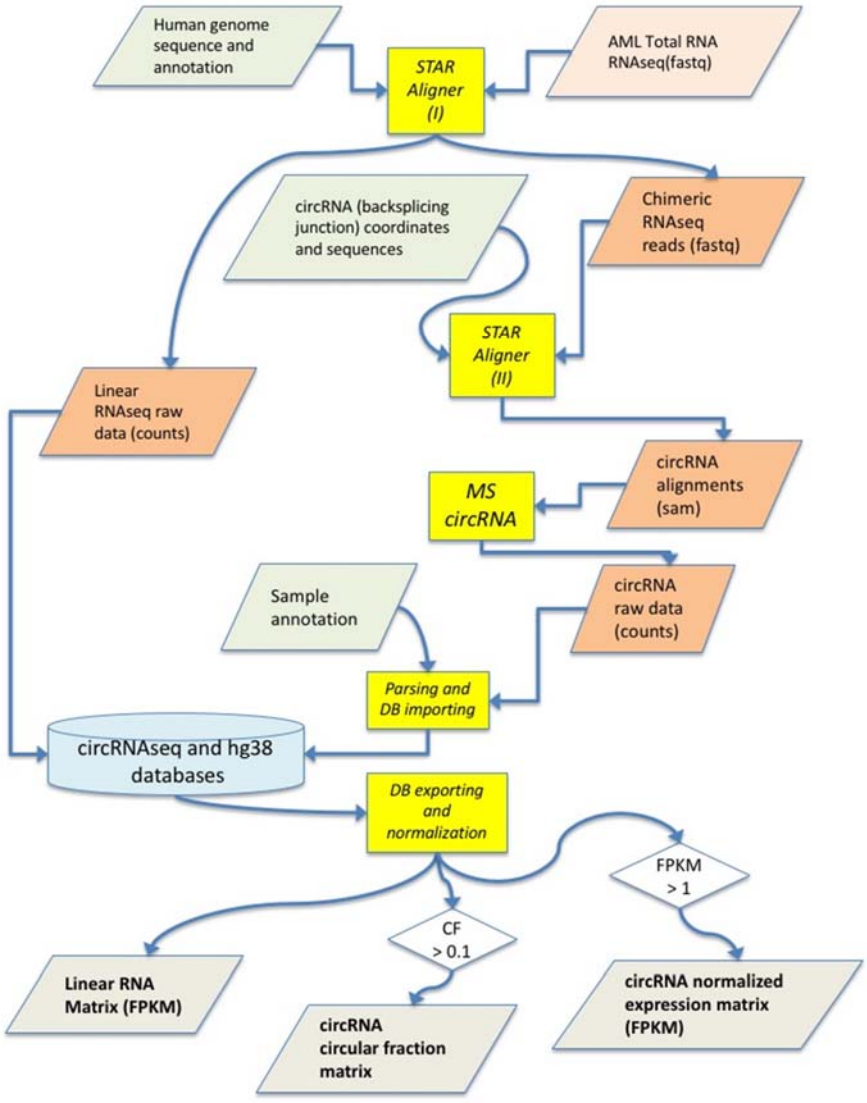
Molecular pathway	Number of genes	Fold enrichment	False discovery rate
Regulation of signal transduction	41	8.18	3.64E-10
Positive regulation of Rho protein signal transduction	4	22.22	5.97E-08
Leukocyte differentiation	18	7.89	4.91E-07

Supplemental Table S20. List of biologic processes that associate with low *circFBXW7* expression as evaluated by Gene Ontology Pathway Analysis. The top-3 pathways are listed. Pathways are ranked according to false discovery rate.

Molecular pathway	Number of genes	Fold enrichment	False discovery rate
Definitive hemopoiesis	11	85.18	7.94E-08
Negative regulation of myeloid cell differentiation	9	22.62	4.89E-07
Positive regulation of RNA biosynthetic process	16	9.75	6.79E-06

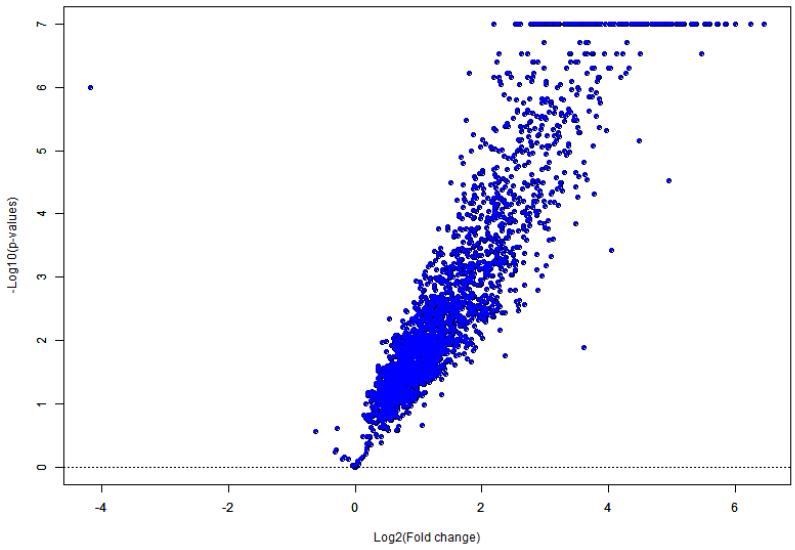
Supplemental Figures

Supplemental Figure S1. Schematic diagram outlining the sequential steps for analyzing circular RNA (circRNA) expression with the Massive Scan for circRNAs (MScircRNA) pipeline.

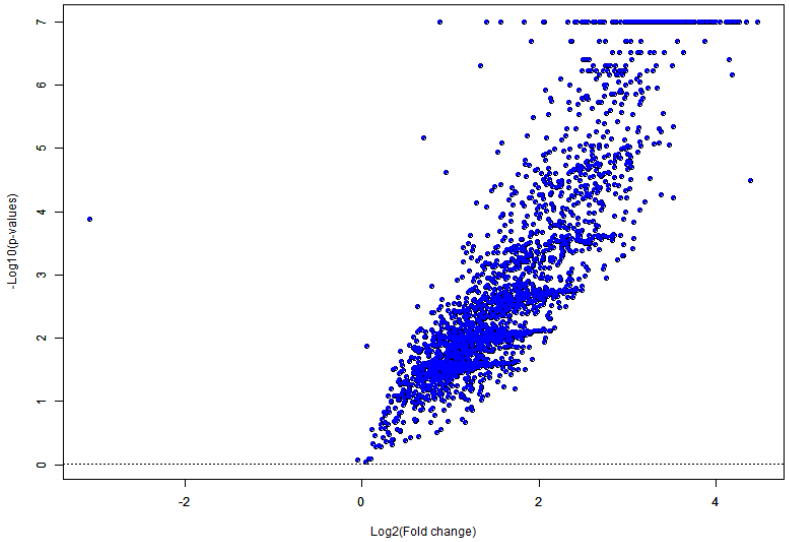


Supplemental Figure S2. Volcano plots showing the enrichment of circRNAs after treatment with RNAse R in mock versus RNAse R-treated samples of seven acute myeloid leukemia patients and three cell lines. (A) The FPKM expression levels are shown. (B) The calculated circular fraction of the predicted circRNAs is depicted. In both graphs, the negative log10 of the *P* value (Y axis) and the log2 of the fold change of expression (X axis) of candidate circRNAs in the RNAse R vs the scramble treated datasets are depicted.

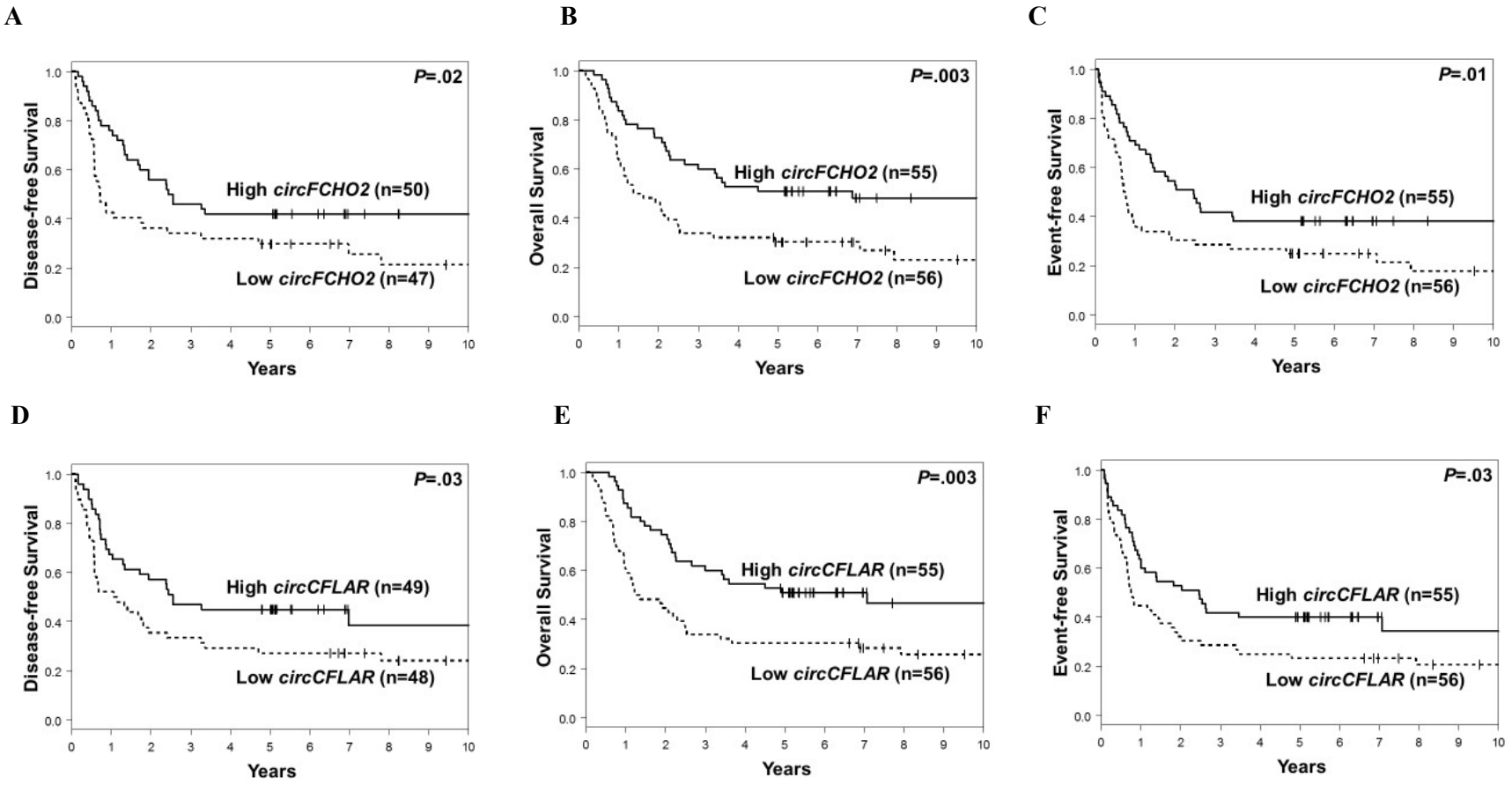
A



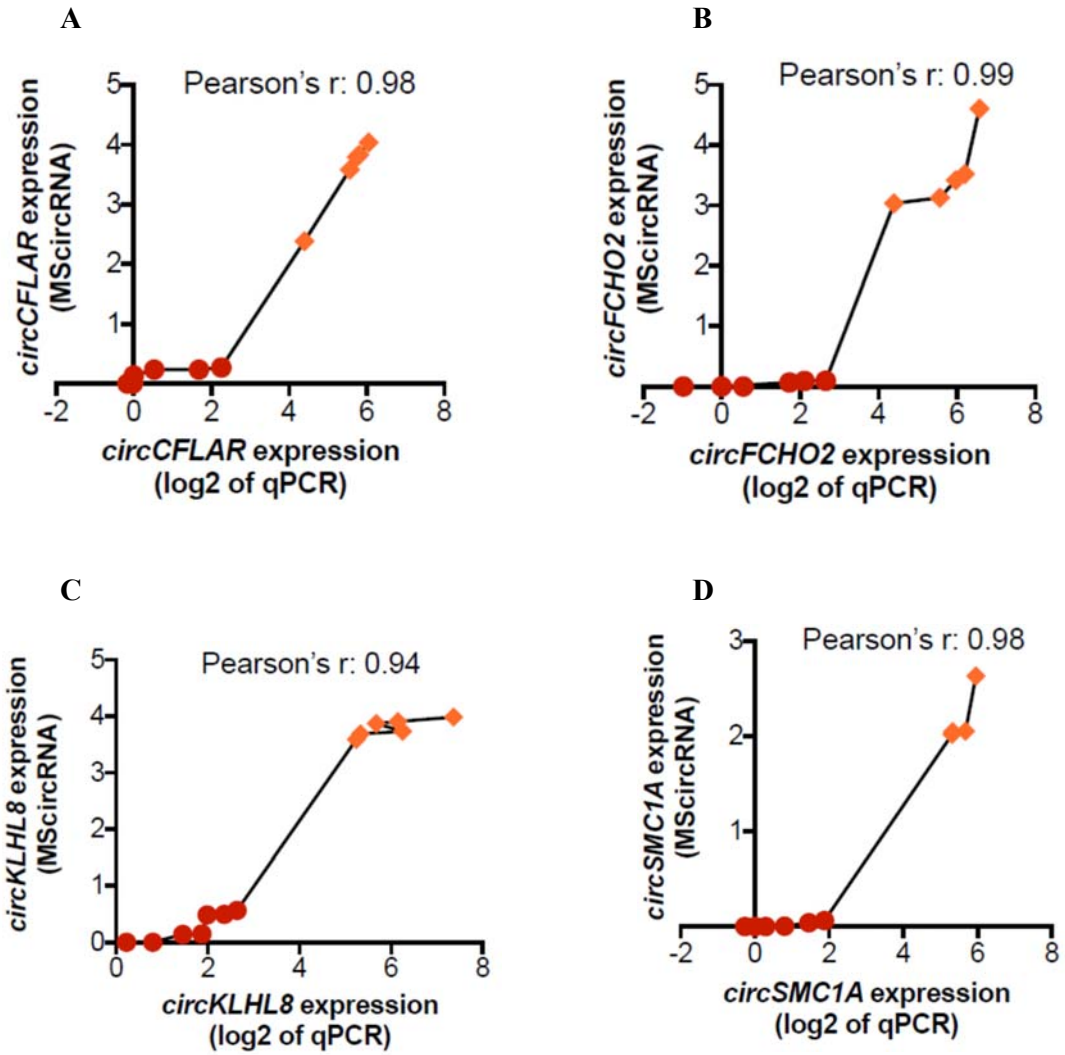
B



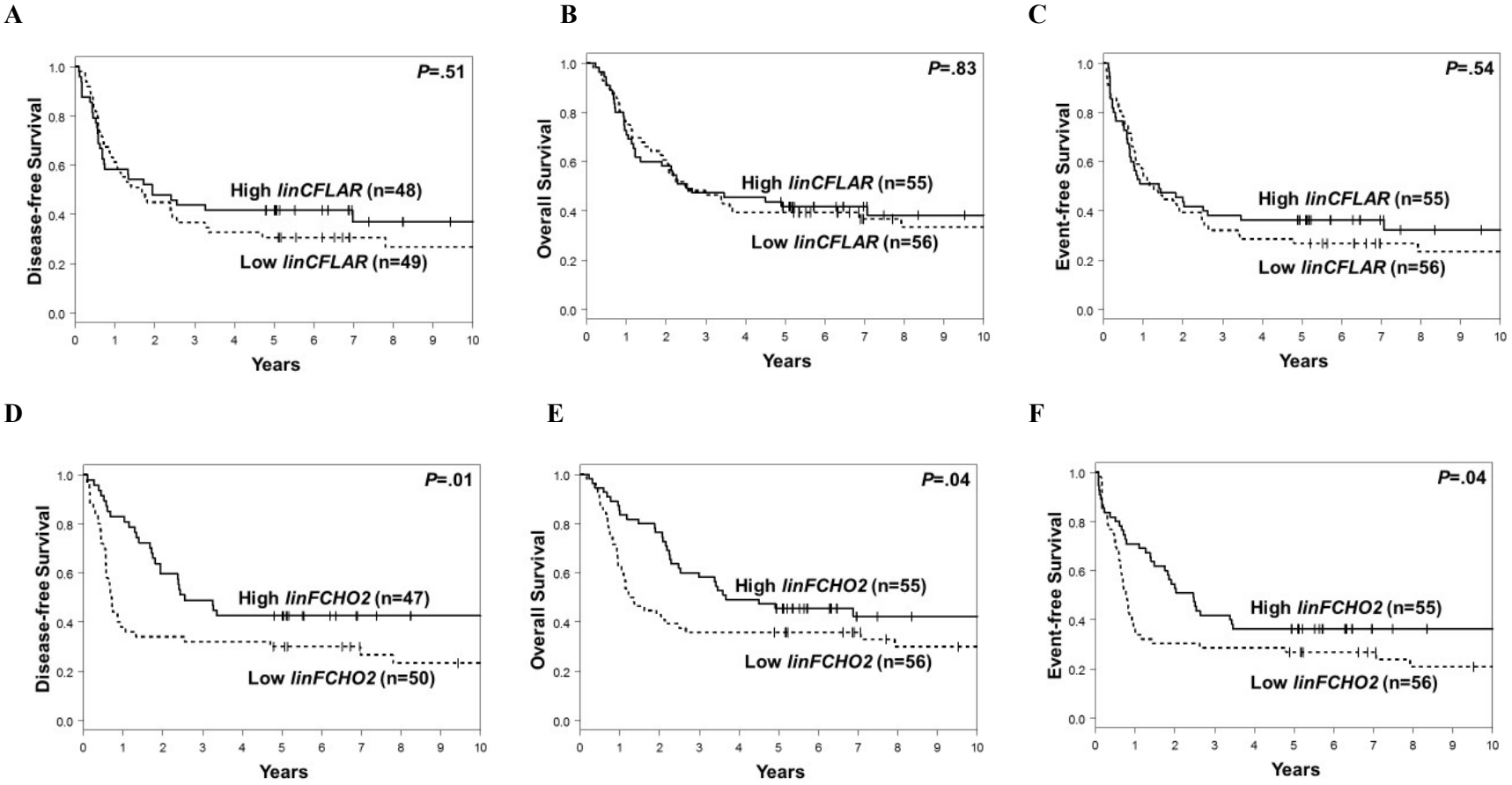
Supplemental Figure S3. Clinical outcome of 117 younger adult patients with cytogenetically normal acute myeloid in the validation set with either high or low expression of individual circRNAs. (A) Disease-free (DFS), (B) overall (OS) and (C) event free (EFS) survival of high versus low expressers of *circFCHO2*. (D) DFS, (E) OS and (F) EFS of high versus low expressers of circ*CFLAR*.

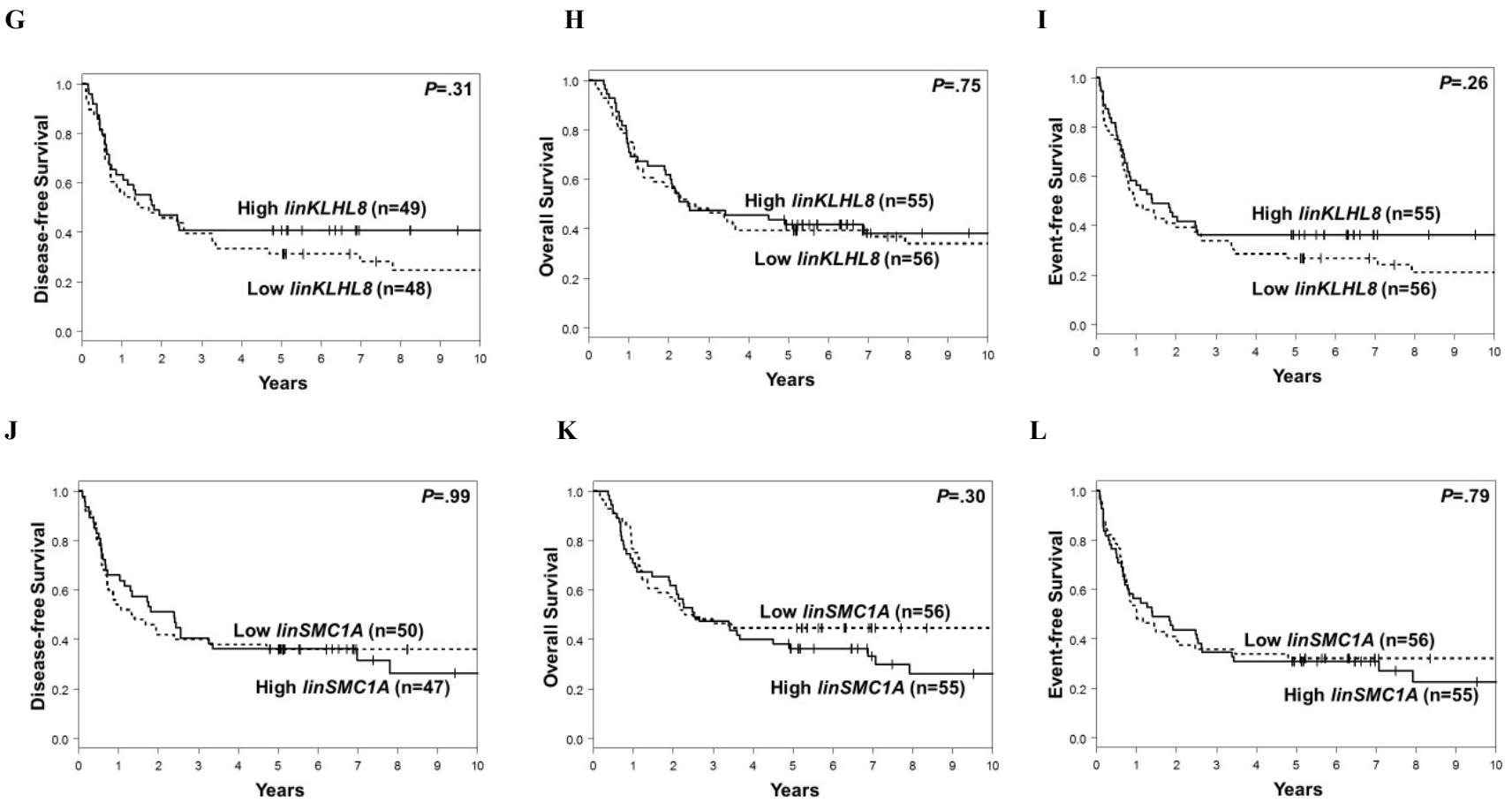


Supplemental Figure S4: Correlation between MScircRNA-based measurements and real time quantitative PCR results in younger adult CN-AML patients with high or low prognostic circRNA expression. Low expressers are annotated in red circles; high-expressers are annotated in yellow rhombi. (A) *circCFLAR*, (B) *circFCHO2*, (C) *circKLHL8*, (D) *circSMC1A*.



Supplemental Figure S5. Clinical outcome of 117 younger adult patients with cytogenetically normal acute myeloid in the validation set with either high or low expression of the linear mRNAs, corresponding to the four prognostic circRNAs. (A) Disease-free (DFS), (B) overall (OS) and (C) event free (EFS) survival of high versus low expressers of *linCFLAR*. (D) DFS, (E) OS and (F) EFS of high versus low expressers of *linFCHO2*. (G) DFS, (H) OS and (I) EFS of high versus low expressers of *linKLHL8*. (J) DFS, (K) OS and (L) EFS of high versus low expressers of *linSMC1A*.




G
P=.31
High linKLHL8 (n=49)
Low linKLHL8 (n=48)
Disease-free Survival
Years
H
P=.75
High linKLHL8 (n=55)
Low linKLHL8 (n=56)
Overall Survival
Years
I
P=.26
High linKLHL8 (n=55)
Low linKLHL8 (n=56)
Event-free Survival
Years
J
P=.99
Low linSMC1A (n=50)
High linSMC1A (n=47)
Disease-free Survival
Years
K
P=.30
Low linSMC1A (n=56)
High linSMC1A (n=55)
Overall Survival
Years
L
P=.79
Low linSMC1A (n=56)
High linSMC1A (n=55)
Event-free Survival
Years

Supplemental Figure S6. **Messenger RNA (mRNA) transcripts which associate with the expression status of *circFBXW7* in younger adults with CN-AML**. Heat map of the gene-expression signature, which associates with *circFBXW7* expression status. Rows represent protein-coding genes and columns represent patients. Patients are grouped by *circFBXW7* expression status. Expression values of the mRNA transcripts are represented by color: green: expression less than median value; red: expression greater than median value; grey: lack of detectable expression.