Supplemental Methods

Cells and cell culture

All human cell lines used in this publication (CMK, K562, HEL and HEK293T) were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig). Culturing and maintenance were performed according to the supplier's instructions. Murine fetal liver cells were isolated from E12.5-E13.5 *Cas9* heterozygous C57BL/6J (Jackson Laboratory) embryos upon depletion of erythroid cells using anti-Ter119 immunomagnetic microbeads (StemCell Technologies). Neonatal HSPCs were obtained from the cord blood of healthy donors and fetal HSPCs from the MRC-Wellcome Trust Human Developmental Biology Resource (HDBR). $CD34^+$ cells were enriched using anti-CD34 immunomagnetic microbeads (Miltenyi Biotech and StemCell Technologies). Pediatric AML samples were collected from patients enrolled in the AML Berlin-Frankfurt-Münster treatment protocols for children and adolescents. The study was approved by the institutional review boards of all participating centers. All investigations were performed in accordance to the Declaration of Helsinki and local laws and regulations, and informed consent was obtained from all patients and custodians. Murine FLCs were stimulated before transduction in DMEM (Gibco, Life Technologies) with 10% FCS (Capricorn Scientific), 1% streptomycin/penicillin (Millipore), 1% L-glutamine (Millipore), 1% sodium pyruvate (Gibco, Life Technologies), 1% non-essential amino acids (Gibco, Life Technologies), 25ng/mL Scf and 25ng/mL Tpo, and were further expanded in the same medium except with 2ng/mL Scf. The same medium with low Scf conditions was used in megakaryocytic differentiation assays. For erythroid differentiation, the same cells were expanded in IMDM (Lonza Bioscience) with 15% FCS, 2% streptomycin/penicillin, 1% L-glutamine, 2U/mL Epo, 10μM dexamethasone (Sigma-Aldrich), 5ng/mL Scf and 1ng/mL Il3. For myeloid differentiation, IMDM with 10% FCS, 2% streptomycin/penicillin, 1% sodium pyruvate, 20ng/mL Scf, 40ng/mL G-CSF, 40ng/mL M-

CSF, 10ng/mL Il3 and 10 ng/mL Il6 was used. Human HSPCs were expanded in StemSpan SFEM (StemCell Technologies) with 1% streptomycin/penicillin, 50ng/mL SCF, 50ng/mL FLT3-ligand, 20ng/mL TPO, 10ng/mL IL3 and 10ng/mL IL6. For megakaryocytic differentiation, StemSpan SFEM with 1% streptomycin/penicillin, 0.25% CD-lipid concentrate (Gibco, Life Technologies), 30ng/mL TPO, 1ng/mL SCF, 7.5ng/mL IL6 and 13.5ng/mL IL9 was used. For combined megakaryocytic/erythroid differentiation, cells were cultured in StemSpan SFEM with 1% streptomycin/penicillin, 1% L-glutamine, 0.25% CD-lipid concentrate, 50ng/mL SCF, 0.08U/mL EPO, 100ng/mL TPO, 10ng/mL IL3 and 10ng/mL IL6 for 7 days. On day 7, the concentrations of SCF and EPO were increased to 100ng/mL and 0.5U/mL, respectively, while the TPO concentration was reduced to 50ng/mL. For myeloid differentiation, the culture medium consisted of RPMI (Lonza Bioscience), with 10% FCS, 1% streptomycin/penicillin, 1% L-glutamine, 5ng/mL SCF, 5ng/mL GM-CSF, 10ng/mL G-CSF and 5ng/mL IL3. ML-DS blasts from PDX models were cultured *in vitro* in StemSpan SFEM with 1% streptomycin/penicillin, 0.25% CD-lipid concentrate, 50ng/mL SCF, 50ng/mL FLT3, 20ng/mL TPO, 2.5ng/mL IL3, 10ng/mL IL6, 0.75μM Stemregenin1 (SR1) (APExBIO) and 35nM UM171 (APExBIO). Leukemic bone marrow cells obtained from primary recipients were expanded *in vitro* in StemSpan SFEM with 1% streptomycin/penicillin, 10ng/mL Scf, 20ng/mL Tpo and 10ng/mL Il3. All human and murine cytokines were purchased from Peprotech, except for EPO, which was purchased from GoldBio. For drug response curves using MYCi361 (Seleckchem), cell viability was assessed using the CellTiter-Glo® Cell Viability Assay (Promega) according to the manufacturer's instructions.

Lentiviral vector construction and transduction

For the CRISPR/Cas9 screen, a library of 1090 sgRNAs targeting 218 genes on Hsa21 was cloned into the SGL40C.EFS.dTomato vector (Addgene, #69147 and #89395) [1]. Individual sgRNAs targeting *RUNX1 (1-5)* were also cloned into the SGL40C.EFS.dTomato vector,

while the sgRNA targeting *Gata1 (sgGata1.4)* was cloned into the SGL40C.EFS.dTomato and the pLeGO-iG2 vectors (Addgene, #27341) [2]. A detailed list of all protospacer and shRNA sequences is provided in **Supplemental Table 1 and 3**. shRNAs were designed using the miR-N online tool [3] and cloned into the SIN40C.SFFV.tBFP2.miR30n backbone (Addgene, #169278). The *RUNX1A* cDNA was cloned into the pLeGO.iG2 vector. Codon-optimized 5'HA-tagged RUNX1A and RUNX1C cDNAs carrying a single HA-tag at the 5´-end were synthesized at Twist Bioscience and introduced into the pRRL.PPT.SFFV.MCS.IRES.eGFP vector (Addgene, #171173) and into the doxycycline-inducible SIN40C.TRE.MCS.IRES.idTomato.PGK.sfGFP.P2A.Tet3G vector system (Addgene, #169283). Additionally, Cre recombinase cDNA sequences were cloned into the pLKO5d.SFFV.IRES.dTomato backbone. All sequences were verified by Sanger sequencing.

Lentiviral particles were generated by co-transfection of HEK293T cells with the vector of interest, pMD2.G and psPAX2 (Addgene, #12259 and #12260) using the polyethylenimine (PEI) method [4,5]. CMK cells, murine FLCs, $CD34^+$ HSPCs, PDX-derived ML-DS blasts and leukemic BM cells were lentivirally transduced as previously described [6].

CRISPR/Cas9 library screening

6×10^6 CMK and K562 cells were lentivirally transduced with the library of sgRNAs, in order to achieve a 1000-fold coverage of the library. Both cell lines were previously transduced with the pLKO5d.SFFV.Cas9.P2A.BSD vector (Addgene, #181981) and selected with Blasticidin, to establish stable expression of the Cas9 nuclease. The cells were maintained in culture and samples were harvested for DNA extraction at an early (day 4 post-transduction) and a late time point (after 18 population doublings). DNA was isolated using the QIAmp DNA Blood Mini Kit (Qiagen). Adapters including sample specific barcodes were added by PCR. Libraries were subjected to next generation sequencing (NGS) for sgRNA quantification. NGS data

were mapped to the sgRNA library and quantified using the MAGeCK count command. Beta-Scores and p-values were calculated using the MAGeCK maximum-likelihood estimation (mle)[7]. The fluorescent reporter assay to assess cutting efficacy was performed in HEL cells as previously described[8].

In vitro cell growth assays

In all growth assays, the same standard protocol was followed. Every 2-3 days, a sample of cells was taken to measure cell counts or percentage of fluorescent cells on the CytoFLEX platform (Beckman Coulter, BC) or the FACS Canto Flow Cytometer (BD Biosciences). Competition assays were performed by using a starting transfection efficiency of approximately 50%.

In vitro differentiation assays

Upon transduction, 5×10^4 murine FLCs were cultured in the megakaryocytic differentiation-supporting medium, as described above. Similarly, 3×10^4 cells were cultured in the erythroid medium and 2.5×10^4 cells in the myeloid medium. On days 4 and 7 of differentiation, cell surface marker expression was analyzed by flow cytometry. The following fluorochrome-coupled antibodies were used for flow cytometry: CD41-PECy7 (BioLegend) clone MWReg30, CD42d-APC (eBioscience) clone 1C2, Ter119-Pacific Blue (BioLegend) clone TER-119, CD71-PerCP-Cy5.5 (BioLegend) clone RI7217, CD117-APC-H7 (BD) clone 2B8, CD11b-V500 (BD) clone M1/70 and Gr1-PerCP-Cy5.5 (BD) clone RB6-8C5. In the case of $CD34^+$ HSPCs, 8×10^4 transduced cells were seeded for megakaryocytic differentiation assays, while 5×10^4 cells were used for combined megakaryocytic/erythroid and myeloid assays. These cells were also stained and evaluated for the expression of cell surface markers on days 7 and 10 of differentiation. The following fluorochrome-coupled antibodies were used for these assays: CD41-PECy7 (BC) clone P2, CD41-PE (BC) clone P2, CD61-

PECy7 (BC) clone SZ21, CD42b-APC (BD) clone HIP1, CD235a-APC (BD) clone GA-R2, CD71-APC-AF750 (BC) clone YDJ1.2.2, CD117-PE (BC) clone 104D2D1, CD11b-APC-AF750 (BC) clone Bear1, CD14-APC (BC) clone RMO52, CD15-V500 (BD) clone HI98, CD16-PE (BC) clone 3G8, CD34-BV605 (BD) clone 8G12 and CD66b-PECy7 (eBioscience) clone G10F5. All measurements were performed on a FACS Canto Flow Cytometer (BD Bioscience) or a CytoFLEX Flow Cytometer (Beckman Coulter). Data were subsequently analyzed with FlowJo v10 (FlowJo, LLC) or Kaluza 2.1 (Beckman Coulter). Fluorescence-based cell sorting was conducted on a FACSAria II at the FACS core facility of the medical faculty (Martin-Luther-University Halle-Wittenberg).

Colony-forming assays

$1x10^5$ sorted $CD34^+$ HSPCs were seeded in for collagen-based colony-forming assays (MegaCult™- C medium with cytokines, StemCell Technologies), as previously described [6]. Similarly, 10^4 sorted $CD34^+$ HSPCs were plated for methylcellulose-based colony forming assays [Human Methylcellulose Base Medium HSC002 (R&D Systems), which were supplemented with 2% streptomycin/penicillin, 10ng/mL SCF, 10ng/mL FLT3, 4U/mL EPO, 50ng/mL TPO, 50ng/mL GM-CSF, 20ng/mL IL3 and 10ng/mL IL6] according to the manufacturer's instructions and incubated for 7 days in a 37°C incubator with 5% CO_2 and ≥95% humidity. Serial replatings were also performed. Counting and classification of the colonies were performed with the Axiovert 40CFL microscope (Zeiss).

Immunophenotyping

ML-DS blasts were analyzed for the expression of surface cell markers 5 and 14 days after lentiviral transduction. The following fluorochrome-coupled antibodies were used for the staining: CD33-PerCP-Cy5.5 (BC) clone D3HL60.251, CD117-PE (BC) clone 104D2D1, CD235a-APC(BD) clone GA-R2, CD41-PECy7 (BC) clone P2, CD42b-APC (BD) clone HIP1

and CD15-V500 (BD) clone HI98. Measurements were performed on a FACS Canto Flow Cytometer (BD Biosciences) and the acquired data were analyzed with FlowJo v10 (FlowJo, LLC). Apoptosis was measured using APC or FITC-conjugated Annexin V (BD Biosciences) using a CytoFLEX Flow Cytometer (Beckman Coulter).

Drug responsive assay

MYC:MAX dimerization was inhibited in CMK cells, murine Gata1s-RUNX1A-FLCs, PDX blasts and CD34+ HSPCs by incubation with MYCi361 (Seleckchem) for 24h. Cell viability was assessed using the CellTiter-Glo® Cell Viability Assay (Promega) according to the manufacturer´s instructions.

Animal studies

All animal experiments were performed according to protocols approved by the local authorities (Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit, Landesverwaltungsamt Sachsen-Anhalt and Regierungspräsidium Darmstadt). The animals were maintained under pathogen-free conditions.

PDX transplantation into MISTRG mice

For the generation of patient-derived xenografts (PDX), mononuclear patient cells were injected intravenously into irradiated (2.5Gy) 8-16 week old, female MISTRG mice (Regeneron Pharmaceuticals) [9]. For xenotransplantation experiments, cells from tertiary or quaternary recipients were transduced with lentiviral vectors for constitutive expression of either RUNX1A, RUNX1C and EV cDNA, sgRUNX1.1 and sgCtrl sgRNAs or shMAX or shCtrl. 24 hours post transduction, cells were mixed in a 1:1 ratio (GFP:dTomato or GFP:E2Crimson), washed, and injected via tail vein into irradiated (2.5Gy), 8-16 week old, female MISTRG mice, with groups of five mice per condition. A total of $1x10^6$ cells were

injected per mouse. Animals were sacrificed before leukemia onset and the hematopoietic compartments were analyzed. Engraftment of the transplanted cells was evaluated via tail-vein bleeding of the recipients every 4 weeks. When signs of disease occurred, bone marrow cells and splenocytes were harvested and stained with fluorochrome-coupled antibodies against human CD117 and human CD45. All measurements were performed on a FACS Canto Flow Cytometer (BD Biosciences) or a Cytoflex Flow Cytometer (BD Biosciences) and the acquired data were analyzed with FlowJo v10 (FlowJo, LLC).

FLC transplantation into C57BL-6J mice

Cas9 heterozygous murine FLCs were obtained by timed breeding of C57BL/6J (Jackson Laboratory) females with B6J.129(Cg)-Gt(ROSA)26$^{Sortm1.1(CAG-cas9*,-EGFP)}$Fezh/J (Cas9 knock-in mice on C57BL/6J background) males. On embryonic day E12.5/E13.5, the fetal livers were mechanically detached from the embryos and further processed as described above. For the purposes of bone marrow transplantation, Ter119-depleted FLCs were lentivirally transduced with the pLeGO.sgGata1.4.iG2 vector and expanded *in vitro*. After at least 20 days of culture, these cells were further transduced with the pLeGO-iG2 cDNA vectors. 24 hours later, the FLCs were transplanted intravenously into sublethally irradiated (7.5 Gy) female C57BL/6JRj mice (aged 6-15 weeks). Secondary transplantations were performed by intravenously injecting 5×10^6 primary bone marrow cells into sublethally irradiated (7 Gy) female C57BL/6J mice. Engraftment of the transplanted cells was evaluated by retro-orbital bleeding of the recipients every 4 weeks. When signs of disease occurred, the mice were sacrificed, and harvested bone marrow cells and splenocytes were stained with the following fluorochrome-coupled antibodies: CD3e-Pacific Blue (BioLegend) clone 17A2, B220-Pacific Blue (BioLegend) clone RA3-6B2, Lineage cocktail-Pacific Blue (BioLegend), Gr1-PerCP-Cy5.5 (BD) clone RB6-8C5, CD11b-PE (BD) clone M1/70, CD117-APC-Cy7 (BD) clone 2B8, CD117-PerCP-Cy5.5 (BD) clone 2B8, CD41-PE (BD) clone MWReg30, Sca1-PECy7 (BD)

clone D7, CD34-AF647 (BD) clone RAM34, CD16/32-APC-Cy7 (BD) clone 2.4G2. All measurements were performed on a FACS Canto Flow Cytometer (BD Biosciences) and the acquired data were analyzed with FlowJo v10 (FlowJo, LLC).

Evaluation of RUNX1 isoforms expression by qRT-PCR

Total RNA was extracted from cell lines, adult CD34$^+$ HSPCs, trisomy 21 CD34+ HSPCs, erythroid cells, megakaryocytes, granulocytes, monocytes, and from AMKL, ML-DS, TAM, t(10;11) AML and t(9;11) AML patient samples, using the Quick-RNA Microprep Kit (Zymo Research). Complementary DNA (cDNA) was then synthesized using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher Scientific). Quantitative real-time PCR (qRT-PCR) was performed on the StepOnePlus™ Real-Time PCR System, using the TaqMan® Fast Advanced Master Mix (both from Applied Biosystems, Thermo Fisher Scientific).

Molecular studies

Oxford Nanopore sequencing of RUNX1 transcript isoforms

Total RNA from TAM and ML-DS patient-derived samples were isolated using the Quick-RNA Microprep Kit (Zymo Research). The MicroMACS mRNA isolation kit (Miltenyi Biotech) was used for isolation of polyA$^+$-enriched RNA. 1 ng polyA$^+$ RNA was processed for Oxford Nanopore full length transcript sequencing using the PCR-cDNA Sequencing Kit (SQK-PCS109; Oxford Nanopore Technologies) according to the manufacturer´s protocol. Sequencing was performed on an Oxford Nanopore MinION Mk1B. Base calling was done with guppy v3.0.6. Raw sequencing reads were first processed with filtlong (https://github.com/rrwick/Filtlong), removing reads shorter than 20 nt and associated with mean quality scores below 80. Subsequently, full-length cDNA reads were extracted using pychopper (https://github.com/nanoporetech/pychopper).

For gene expression analysis, full-length reads were mapped to the human genome (UCSC hg38) with minimap2 [10] using the long-read spliced alignment preset (-x splice) and without reporting secondary alignments (--secondary=no). Afterwards, sorting and indexing of mapped reads was performed using samtools [11]. Summarization of mapped reads was done with featureCounts [12] applying long read counting mode (-L) and using Ensembl GRCh38.89 (Aken et al., 2017) as the annotation base. Differential gene expression was determined using the R package edgeR [13], utilizing trimmed mean of M-values (TMM; [14]) for normalization. Normalized gene expression values were expressed as counts per million mapped reads (CPM). For determination of differential transcript usage, full-length reads were mapped to the human transcriptome (Gencode v29; [15] with minimap2 [10]using the Nanopore preset (-x map-ont) and without reporting secondary alignments (--secondary=no). Transcript expression was assessed with salmon (Patro et al., 2017) in alignment-based mode and activated options for correction of sequence-specific and GC-content bias. The R-package DRIMseq [16] was used to perform testing for differential transcript expression.

RNA sequencing data collection and analysis

RNA samples were generated from the BM cells of diseased mice, after sorting live fluorescent (GFP+) cells using the FACSAria cell sorter (BD Biosciences). For control samples, sorted murine HSPCs derived from E12.5-E13.5 FLCs were used (MEPs: $Lin^-Sca1^-cKit^+CD34^-CD16/32^-$, CMPs: $Lin^-Sca1^-cKit^+CD34^+CD16/32^{low}$, GMPs: $Lin^-Sca1^-cKit^+CD34^+CD16/32^+$, LSKs: $Lin^-Sca1^+cKit^+$; the fluorochrome-coupled antibodies used for staining are mentioned above).

In addition, RNA samples were generated from murine FLCs transduced with SGL40C.EFS.dTomato vectors carrying a sgRNA targeting *Gata1* or a control sgRNA (sgCtrl), and with pRRL.PPT.SFFV.iGFP or doxycycline-inducible SIN40C.TRE.sfGFP.idTom.Tet3G vectors expressing RUNX1A, RUNX1C or with an empty

vector control. Prior to sorting the double positive populations ($dTom^+GFP^+$) using a FACSAria cell sorter (BD Biosciences), the cells were cultured in our standard megakaryocytic differentiation medium for 72h, starting 24h post transduction. SIN40C.TRE.sfGFP.idTom.Tet3G transduced cells were expanded for up to 1 week prior to fluorescence sorting.

In all cases, total RNA was extracted using the Quick-RNA Microprep Kit (Zymo Research). Paired-end libraries with 2 × 75 bp reads were prepared from the extracted RNA using the TruSeq Stranded total RNA LT Sample Prep (RiboZero Gold, Illumina) using Illumina methodology. RNA reads were aligned to GRCm38 using STAR (v2.6.0)[17]. Mapped reads were annotated using Ensembl v.91 [17,18]. Gene expression levels were quantified in fragments per kilobase of exon model per million mapped reads (FPKM). The FPKM calculation was performed using the R package edgeR including TMM normalization [14,19]. Functional enrichments were calculated using gene set enrichment analysis (GSEA; v3.0) on log2-transformed FPKM values [20]. This was done using previously described and curated gene sets and on ML-DS signatures [21]. Human gene symbols from the published leukemia gene sets were mapped to murine gene symbols using orthologue annotations provided by Ensembl considering only one-to-one orthologue relationships. Standard deviation of gene expression was calculated using log2-transformed FPKM values. Genes that showed a standard deviation >1 were selected for unsupervised hierarchical clustering.

Cleavage under targets and release using nuclease (CUT&RUN)

Gata1s-FLCs expressing doxycycline-inducible HA-RUNX1A/-RUNX1C or HA-GATA1/-GATA1s were harvested 36h and 48h after doxycline induction. Transgene expressing cells were sorted for GFP+ and dTomato+ expression. A minimum number of 100,000 cells per condition were used. CUT&RUN was performed [22] as previously described [22] using

antibodies recognizing endogenous Gata1s and full length Runx1 isoforms as well as an HA-specific antibody.

Sequencing was performed on an Illumina NextSeq machine at the sequencing core of the Max Planck Institute for Molecular Genetics (MPIMG, Berlin). Paired, raw reads were demultiplexed, resulting in reads 42 bases in length. The reads were quality- and adapter-trimmed using BBduk , then aligned to the mouse reference genome, (mm9), using BWA (Version: 0.7.17-r1188) in 'aln' mode . Mapped reads were normalized to the input cell number. For peak calling we used SEACR (version 1.1) with relaxed settings and 'norm' mode. Empty vector control (EV) was used as background [23]. Peaks were called using the homer software suite (v4.10) [24]. Enhancer regions were annotated by using EnhancerAtlas 2.0 [25] . We considered ESC hematopoietic progenitor, erythroblast fetal liver and erythroid fetal liver enhancers in mm9. A region was annotated as an enhancer region if it overlapped with any of the aforementioned enhancers. Overlapping peaks between transcription factors were determined with bedtools, and defined as overlapping to any degree. Heatmaps and cluster analyses were generated using deepTools v3.4.3. [26]. TFBS analysis was performed using the genomatix genome analyzer software suite v3.131123, TFBS Overrepresentation tool, with MatBase version 11.2, Mus musculus genome version GRCh38, ElDorado database 04-2020, and Matinspector version 8.3 for locating motif regions [27,28]. De novo motif discovery was performed using RSAT's peak-motif tool [29] with peaks trimmed to 800 bp from the center on either side, or otherwise default settings. The discovered motifs were subsequently searched in JASPAR's redundant vertebrate database [30].

Protein isolation and Western blotting

Total cellular protein was isolated using RIPA solution (Thermo Fisher Scientific) under standard conditions. A minimum of 100,000 cells were processed for Western blotting. Protein solutions were supplemented with 4x Laemmli loading buffer (Thermo Fisher

Scientific) and subsequently heated at 100°C for 10 min followed by the addition of 2-mercaptoethanol (Sigma-Aldrich). Proteins were separated by acrylamide gel electrophoresis and blotted according to standard protocols. Western blotting was performed under standard conditions using anti-RUNX1-RUNT, anti-GATA1, anti-HA, anti-MAX, anti-Vinculin and anti-GAPDH antibodies. Protein visualization was accomplished using anti-rabbit or anti-mouse HRP-conjugated secondary antibodies (Abcam) and Amersham™ ECL Prime Western Blotting Detection Reagent (Thermo Fisher Scientific) according to the manufacturer's instructions, followed by detection of chemiluminescent signals on a ChemiDoc station (BioRad).

Co-immunoprecipitation

A minimum of $1x10^7$ K562 cells (Western blotting) or $1x10^8$ CMK cells (Mass spectrometry) stably transduced with doxycycline-inducible 1xHA-tagged RUNX1 or GATA1 isoforms as well as the empty control (EV) were induced for transgene expression using 0.25 - 1μg/ml doxycycline. 36-48h after induction, cells were gently lysed for 30 minutes on ice in 3 cell volumes of hypotonic lysis buffer (10 mM Tris/HCl (pH 7.4); 2 mM MgCl2; 0.5% TritonX) supplemented with PIC (Sigma-Aldrich) and PMSF protease inhibitor (Thermo Fisher Scientific), followed by solubilization of nuclei using a cell homogenizer. Nuclei were precipitated by centrifugation at 8500 rpm for 15 minutes at 4°C. Nuclear proteins were isolated by mechanistic disruption of the nuclear pellet via dounce-homogenization and incubation in nuclear extraction buffer (20 mM HEPES; 10 mM KCl; 1 mM EDTA; 350 mM NaCl; 20% glycerol) supplemented with PIC (Sigma-Aldrich), PMSF protease inhibitor (Thermo Fisher Scientific) and sodium-orthovanadate (Merckmillipore) for 30 minutes on ice. Insoluble cellular components were precipitated by centrifugation for 30 minutes at 13,000 rpm at 4°C. Protein complexes were immunoprecipitated using anti-HA-conjugated magnetic beads (Pierce™ anti-HA magnetic beads; Thermo Fischer Scientic) following the

manufacturer´s protocol. Protein complexes were eluted from the beads via boiling for 10 minutes (Western blotting) or incubation at 37°C in 8M urea solution supplemented with 10 mM DTT (Carl Roth) for 30 minutes, and processed for mass spectrometry.

Mass spectrometric analysis of protein complexes

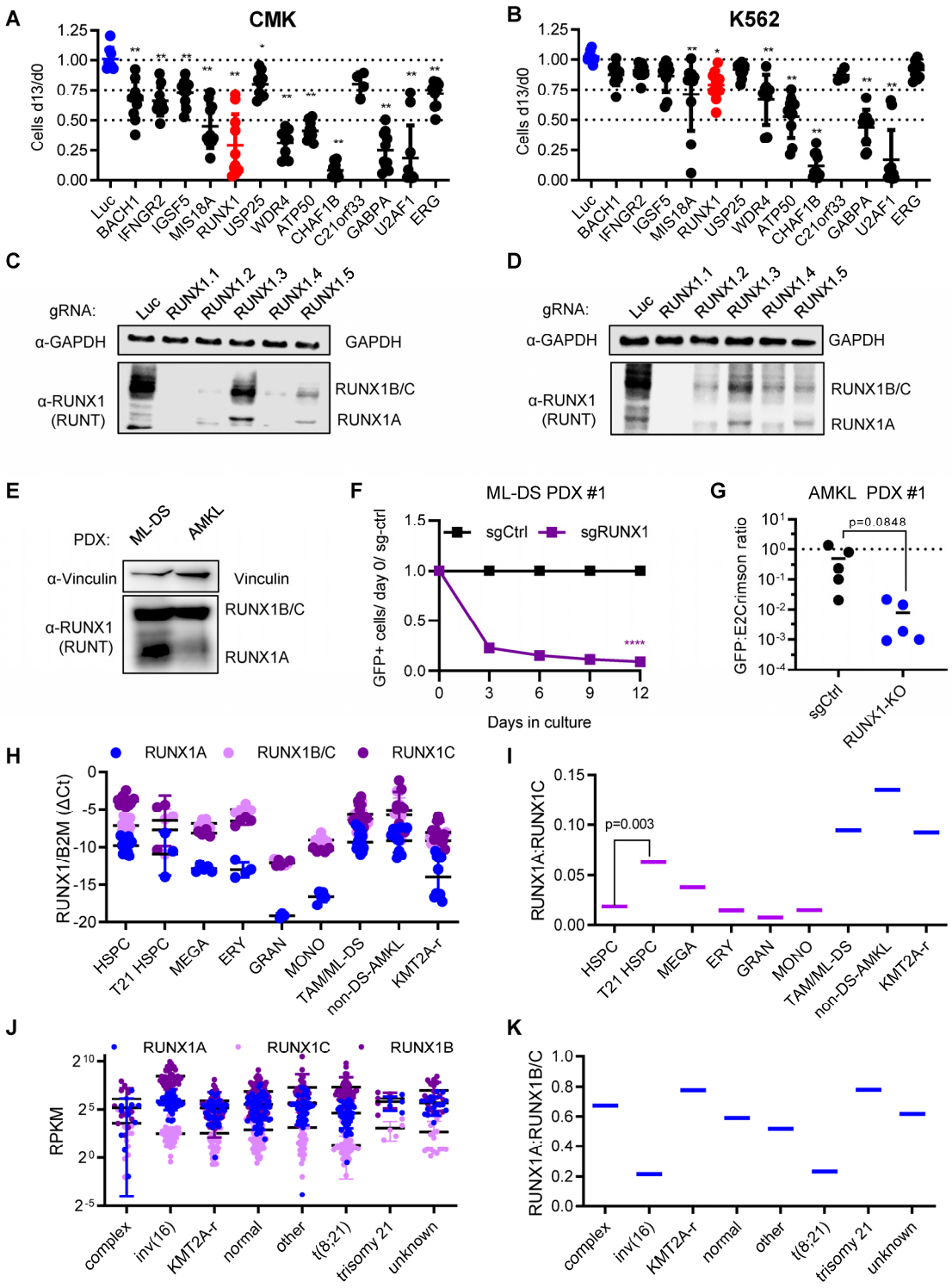
Protein complexes were proteolyzed with trypsin following the FASP (filter aided sample preparation) protocol [31]. Samples were analyzed by LC/MS/MS using 180-min gradients on an U3000 nano-HPLC system coupled to a Q-Exactive Plus or an Orbitrap Fusion mass spectrometer (both Thermo Fisher Scientific). Peptides were separated on reversed phase C18 columns (trapping column: Acclaim PepMap 100, 300 µm × 5 mm, 5µm, 100 Å (Thermo Fisher Scientific); separation column: µPAC 50 cm C18 (Pharmafluidics) or self-packed RP C18 emitter column (PicoFrit, 75 µM × 250–500 mm, 15 µm tip diameter (New Objective), packed with *ReproSil-Pur* C18-AQ, 1.9 µm, 120 Å, Dr. Maisch, Germany). After desalting the samples on the trapping column, peptides were eluted and separated using linear gradients from 3% to 40% B (solvent A: 0.1% (v/v) formic acid in water, solvent B: 0.08% (v/v) formic acid in acetonitrile) with a constant flow rate of 300 nl/min over 180 or 240 min. For data acquisition with the Orbitrap Fusion MS, a data-dependent top 5s method was applied. FTMS survey scans were performed in the *m/z* range 350-1500 (R = 120,000 at *m/z* 200, target of the automatic gain control (AGC) 400,000, maximum injection time (IT) 50 ms). MS/MS scans of the most abundant signals of the survey scans were acquired in parallel by FTMS (HCD (higher energy collision-induced dissociation) at 28% NCE (normalized collision energy), AGC 50,000, IT 120 ms) and CID in linear ion trap (collision induced dissociation), 35% NCE, AGC 10,000, IT 35 ms), isolation window was 1.5 Th. For MS data acquisition with the Q-Exactive Plus mass spectrometer, high-resolution full scans (m/z 375 to 1799, R = 140,000 at m/z 200) in the orbitrap were followed by high-resolution product ion scans (R = 17,500) of the 10 most intense signals in the full-scan mass spectrum (isolation window 2 Th); AGC

was set to 3 x 10^6 (MS) and 200,000 (MS/MS), IT were set to 100 ms (MS) and 150 ms (MS/MS). For acquisitions on both instruments, precursor ions with charge states <2+ and >6+ or were excluded from fragmentation. Dynamic exclusion was enabled (duration 60 seconds, window 3 ppm).

Raw data were processed using Proteome Discoverer 2.4 (Thermo Fisher Scientific). MS/MS data were searched against the UniProt database (version Nov. 2019, tax. *homo sapiens*, 73801 entries) using SequestHT. Label-free quantification of proteins was based on extracted peak areas of corresponding peptide precursor ions, quantitative data were normalized on total protein abundance.

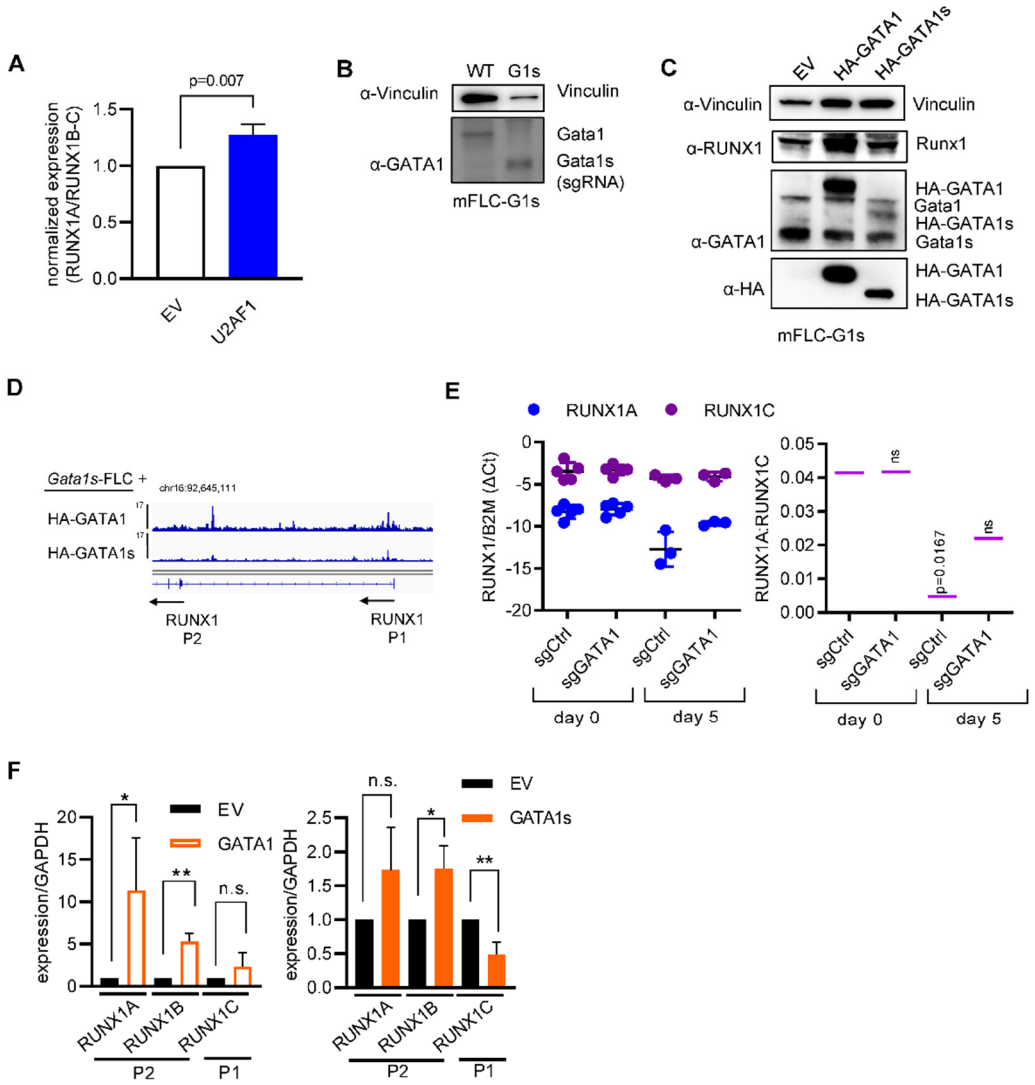
Proteins were selected on a minimum of 3 unique measured peptides. Common contaminants (MaxQuant contaminates.fasta, MPI Martinsried) were excluded. Abundance log2 ratios of either RUNX1A or RUNX1C versus empty control of greater than 1 (p-value < 0.05) were defined as significantly bound. Pathway enrichment analysis was done for either unique or common bound proteins for RUNX1A and RUNX1C using the STRING pathway enrichment analysis tool plugin of the Cytoscape software [32,33]. Cofactor core complexes were identified using the Corum database [34].

Supplemental Figures



Supplemental Figure 1. CRISPR-Cas9 screen reveals *RUNX1* dependency in ML-DS. (A-B) Dot plots showing the number of sgRNA-transduced (A) CMK and (B) K562 cells (with stable Cas9 expression) after 13 days of culture normalized to day 0 (n=3 per sgRNA, one-way ANOVA). $P_{ANOVA}<0.05$, $^{**}P_{ANOVA}<0.01$.

(C-D) Western blot showing RUNX1 protein levels in sorted sgRNA-transduced (sgRUNX1.1-5 and sgCtrl) CMK (C) and K562 (D) cells 5 days after transduction. Endogenous GAPDH was used as the loading control.
(E) Western blot showing RUNX1 protein levels in ML-DS- and AMKL-PDXs used for *in vitro* and *in vivo* experiments. GAPDH was used as the loading control.
(F) Percentage of GFP^+ sgRNA-transduced (sgCtrl and sgRUNX1.1) ML-DS PDX cells normalized to day 0 and sgCtrl (mean±s.d., n=4, two-way ANOVA, **** P_{ANOVA} <0.0001).
(G) Dot plot showing the number of sgRNA-transduced (sgCtrl and sgRUNX1.1) AMKL PDX cells (with stable Cas9 expression) after 12 days of culture normalized to day 0 (n=5, two-tailed unpaired t-test).
(H-I) Expression of *RUNX1A, RUNX1B/C* and *RUNX1C* normalized to the expression of β2-microglobulin (*B2M*) in $CD34^+$ HSPCs, erythrocytes, megakaryocytes, granulocytes and monocytes isolated from healthy donors, as well as in leukemic blasts from TAM/ML-DS, non-DS-AMKL and KMT2A-rearraged patients (H). Ratio of *RUNX1A* to *RUNX1C* expression (I).
(J-K) Expression of *RUNX1A, RUNX1B/C* and *RUNX1C* in subgroups of pediatric AML patients from the TARGET dataset[35] (J). Ratio of *RUNX1A* to *RUNX1B/C (K)*.



Supplemental Figure 2. Trisomy 21 and GATA1s cooperate deregulating RUNX1 isoform expression.

(A) Ratio of *RUNX1A* to *RUNX1B/C* expression in CMK cells 4 days after transduction with *U2AF1* or empty vector control (EV) (Student´s t-test, n=3).

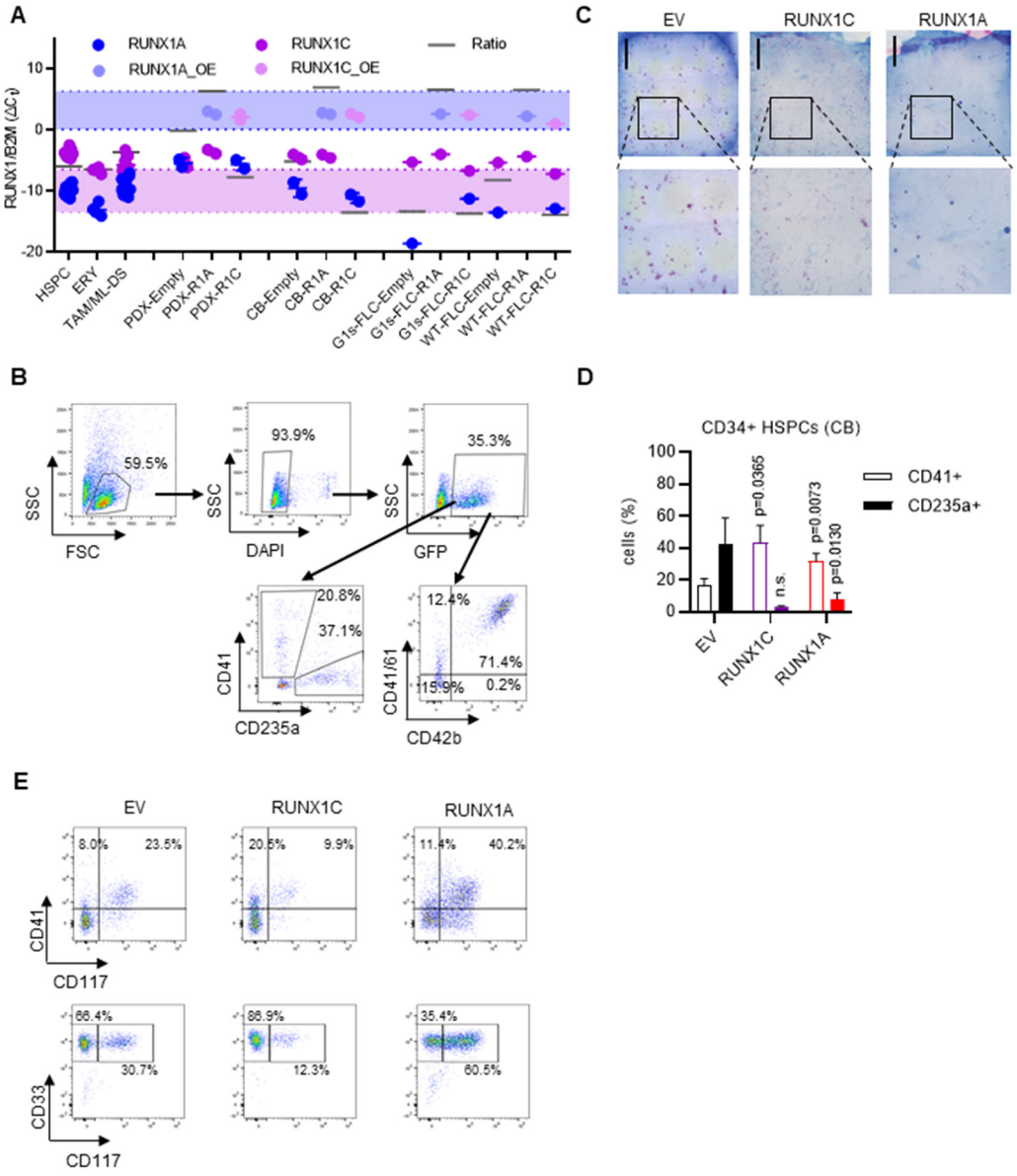
(B) Western blot of Cas9-knock-in Ter119$^-$ FLCs that were lentivirally transduced with *Gata1* (sgGata1s) or control (sgCtrl) sgRNAs, as well as with the indicated cDNA or empty vector control (EV).

(C) Western blot of *Gata1s* fetal liver cells (*Gata1s*-FLC) after doxycycline induced HA-GATA1/GATA1s expression or in EV control expressing cells using the indicated antibodies.

(D) IGV snapshots of the *RUNX1* gene locus, showing occupancy of HA-GATA1 and HA-GATA1s in *Gata1s*-FLCs after doxycycline induced HA-GATA1/GATA1s expression or in EV control expressing cells. Data were generated by CUT&RUN. The *RUNX1* promoters are indicated with P1 and P2. The tracks display coverage (RPKM) (left). Scale and chromosome location are shown (top).

(E) Expression of *RUNX1A* and *RUNX1C* normalized to the expression of β2-microglobulin (*B2M*) in cord blood (CB) HSPCs expressing Cas9 and sgRNAs against *GATA1* (sgGATA1) or control sgRNAs (sgCtrl) at days 0 and 5 after transduction (left). Data from at least 3 replicates are shown. Ratio of *RUNX1A* to *RUNX1C* (right; n≥3, one-way ANOVA compared to sgCtrl at day 0).

(F) *RUNX1* isoform expression in sorted K562 cells 4 days after transduction with *GATA1*, *GATA1s* or EV (n=3, two-tailed unpaired t-test, *$P_{t\text{-}test}$<0.05, **$P_{t\text{-}test}$<0.01).



Supplemental Figure 3. Increased RUNX1A:RUNX1C ratio induces transformation of fetal and neonatal CD34⁺ HSPCs

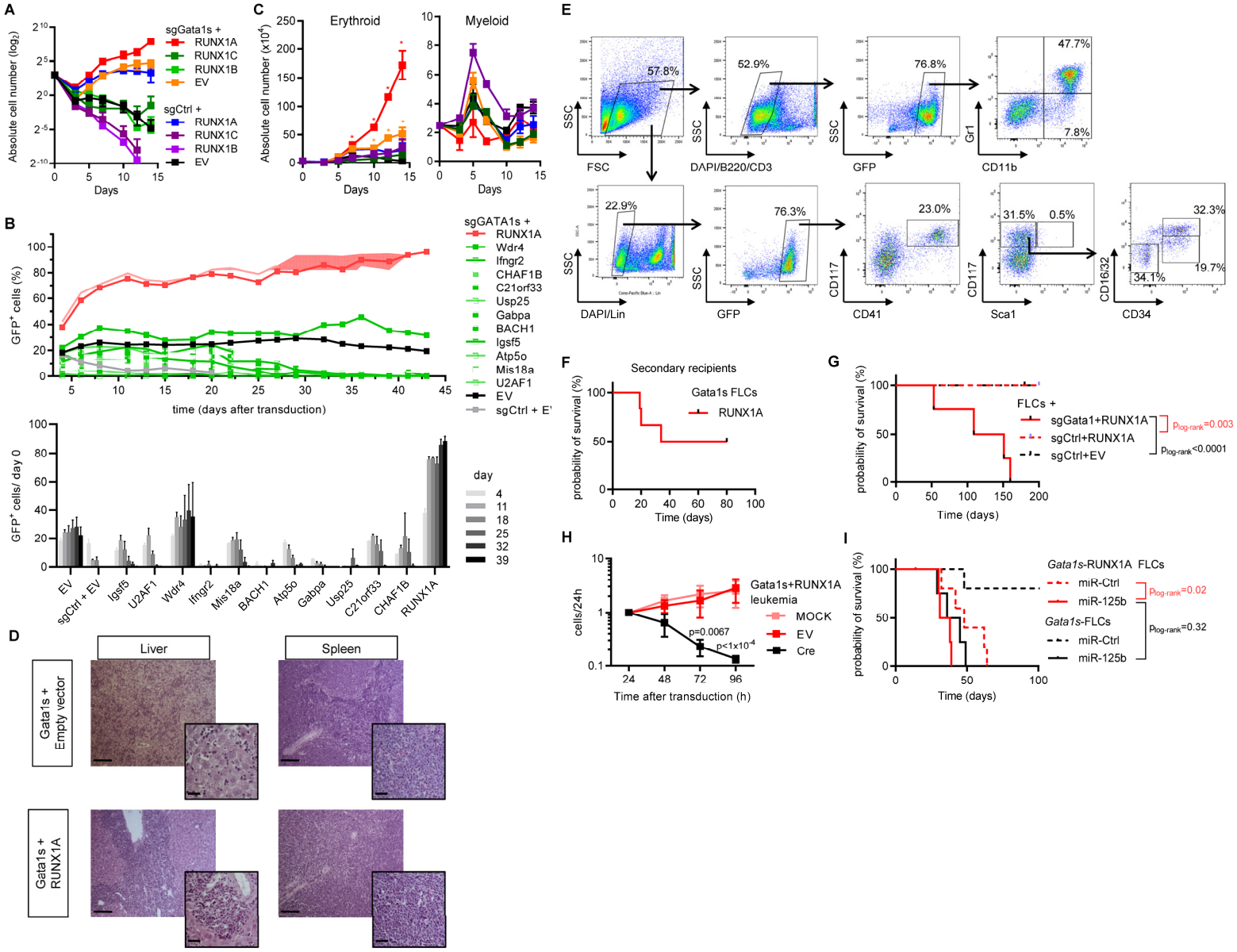
(A) Expression of *RUNX1A* and *RUNX1C* (endogenous or exogenous [OE]) normalized to the expression of β2-microglobulin (*B2M*) in CD34⁺ HSPCs, erythrocytes and TAM/ML-DS samples compared to ML-DS-PDX, neonatal (CB) CD34⁺ HSPCs as well as wild-type (WT-FLC) and *Gata1s* fetal liver cells (G1s-FLC) transduced with *RUNX1A*, *RUNX1C* or the empty vector control. The *RUNX1A*/*RUNX1C* ratios are indicated as horizontal lines. The colored areas indicate the range of *RUNX1A*/*RUNX1C* ratios between erythrocytes and G1S-FLCs (purple) and EV-and *RUNX1A*- transduced ML-DS PDX (blue).

(B-E) Fetal (FL) or neonatal (CB) CD34⁺ HSPCs were lentivirally transduced with *RUNX1A*, *RUNX1C* or control (empty) vectors. (B) Gating strategy applied in flow cytometry analysis. Data from the control (empty vector; EV) cells are shown.

(C) Photographs of CFU-Mk assays using CB CD34⁺ HSPCs (scale bar 5mm).

(D) Percentage of transduced CB HSPCs expressing CD41/CD235 cell surface markers on day 7 of megakaryocytic/erythroid differentiation. Data from three independent replicates are shown as mean±s.d. Two-way ANOVA. $^{**}P_{ANOVA}<0.01$.

(E) Flow cytometric analysis of CD41/CD117 (upper panel) and CD33/CD117 surface marker expression on day 10 of erythroid-megakaryocytic differentiation. Representative dot plots are shown. The percentage of cells belonging to each immunophenotype is indicated in the corresponding gate.


A
Absolute cell number (log2)
Days
sgGata1s +
RUNX1A
RUNX1C
RUNX1B
EV
sgCtrl +
RUNX1A
RUNX1C
RUNX1B
EV
C
Erythroid
Myeloid
Absolute cell number (x10^4)
Days
B
GFP+ cells (%)
time (days after transduction)
sgGATA1s +
RUNX1A
Wdr4
Ifngr2
CHAF1B
C21orf33
Usp25
Gabpa
BACH1
Igsf5
Atp5o
Mis18a
U2AF1
EV
sgCtrl + EV
GFP+ cells/ day 0
day
4
11
18
25
32
39
EV
sgCtrl + EV
Igsf5
U2AF1
Wdr4
Ifngr2
Mis18a
BACH1
Atp5o
Gabpa
Usp25
C21orf33
CHAF1B
RUNX1A
D
Liver
Spleen
Gata1s + Empty vector
Gata1s + RUNX1A
E
57.8%
52.9%
76.8%
47.7%
7.8%
SSC
FSC
DAPI/B220/CD3
GFP
Gr1
CD11b
22.9%
76.3%
23.0%
31.5%
0.5%
32.3%
19.7%
34.1%
DAPI/Lin
CD41
Sca1
CD16/32
CD34
F
Secondary recipients
Gata1s FLCs
RUNX1A
probability of survival (%)
Time (days)
G
FLCs +
sgGata1+RUNX1A
sgCtrl+RUNX1A
sgCtrl+EV
plog-rank=0.003
plog-rank<0.0001
probability of survival (%)
Time (days)
H
cells/24h
Gata1s+RUNX1A leukemia
MOCK
EV
Cre
p=0.0067
p<1x10-4
Time after transduction (h)
I
Gata1s-RUNX1A FLCs
miR-Ctrl
miR-125b
plog-rank=0.02
Gata1s-FLCs
miR-Ctrl
miR-125b
plog-rank=0.32
probability of survival (%)
Time (days)

Supplemental Figure 4. RUNX1A synergizes with Gata1s in leukemic transformation of murine fetal liver cells

Cas9-knock-in Ter119$^-$ FLCs were lentivirally transduced with *Gata1* (sgGata1s) or control (sgCtrl) sgRNAs, as well as with the indicated cDNA or empty vector control (EV).

(A) Number of fluorescent cells after transduction with *RUNX1A*, *RUNX1B, RUNX1C* or control (EV) vectors and culture in media promoting megakarycytic differentiation. Data from 2 replicates are shown as mean±s.d. Two-way ANOVA. *P_{ANOVA}<0.05, **P_{ANOVA}<0.01.

(B) Growth curves (top) and bar graph (bottom) after transduction with cDNAs of Hsa21 candidate genes from the sgRNA drop out screen (selected based on effect size and described function). Percentage of fluorescent cells in cDNA transduced samples. Data from 4 replicates are shown as mean±s.d.

(C) Number of fluorescent cells after transduction with *RUNX1A*, *RUNX1C* or control (EV) vectors and culture in media promoting erythroid (left) or myeloid (right) differentiation. Data from 2 replicates are shown as mean±s.d. Two-way ANOVA. *P_{ANOVA}<0.05, **P_{ANOVA}<0.01.

(D) Representative microscopic images of liver and spleen sections from the indicated experimental groups (2 μm sections, H&E stained, 100-fold magnified, scale bar 500μm, highlighted segment 400-fold magnified, scale bar 80μm).

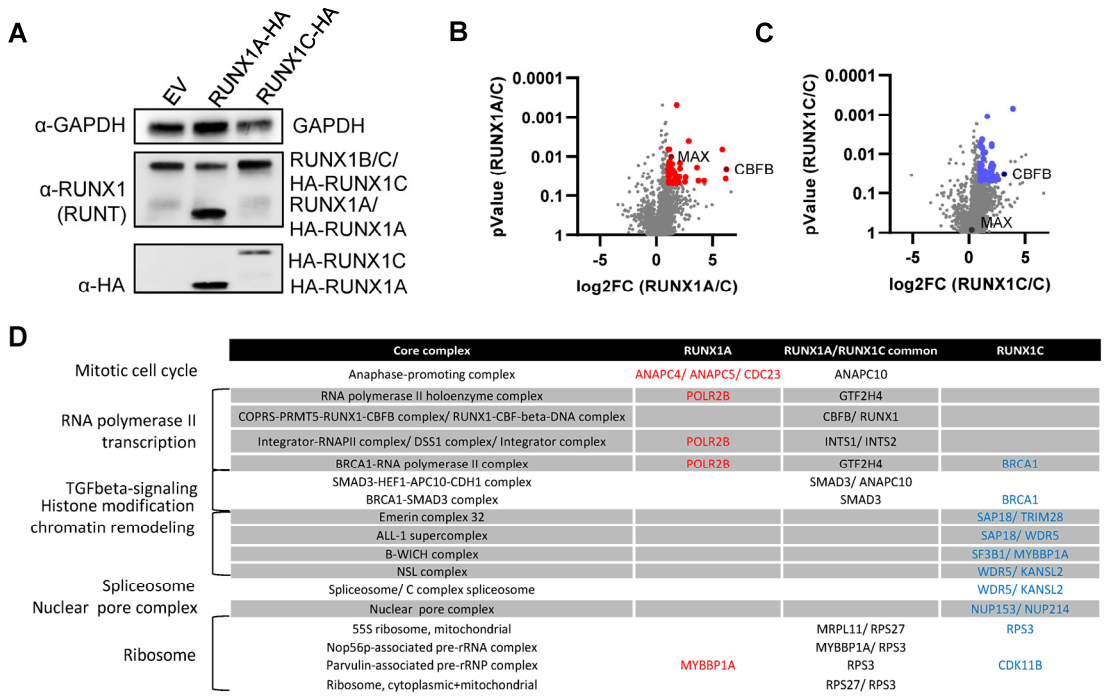
(E) Gating strategy applied in flow cytometric analysis of transduced FLCs isolated from the bone marrow of diseased mice. Data from the control (empty vector; EV) cells of healthy recipients are shown. Cells were first gated on the forward scatter (FSC)/side scatter (SSC) plot (upper far left). Live cells negative for lymphoid lineage markers (B220/CD3) (upper middle left) were gated to detect GFP$^+$ cells (upper right). These cells were further gated to detect myeloid-specific markers (CD11b/Gr1, upper right). Live cells negative for mature myeloid lineage markers (Lin, lower far left) were gated to detect GFP$^+$ cells (lower left). Cells were further gated for CD41/CD117 (lower middle), Sca1/CD117 (lower right), and CD16/32/CD34 (bottom right).

(F) Kaplan-Meier survival curve of secondary recipients transplanted with BM-derived leukemic cells from 3 primary recipients.

(G) Kaplan-Meier survival curve of mice transplanted with either FLCs transduced with sgCtrl+empty vector (EV; n=11), sgCtrl+RUNX1A (n=5), sgGata1+EV (n=12) or sgGata1+RUNX1A (n=4; log-rank test).

(H) Number of BM leukemic cells derived from diseased mice (normalized to 24h) after transduction with a Cre recombinase-expressing vector. Data from three replicates are shown as mean±s.d. Two-way ANOVA.

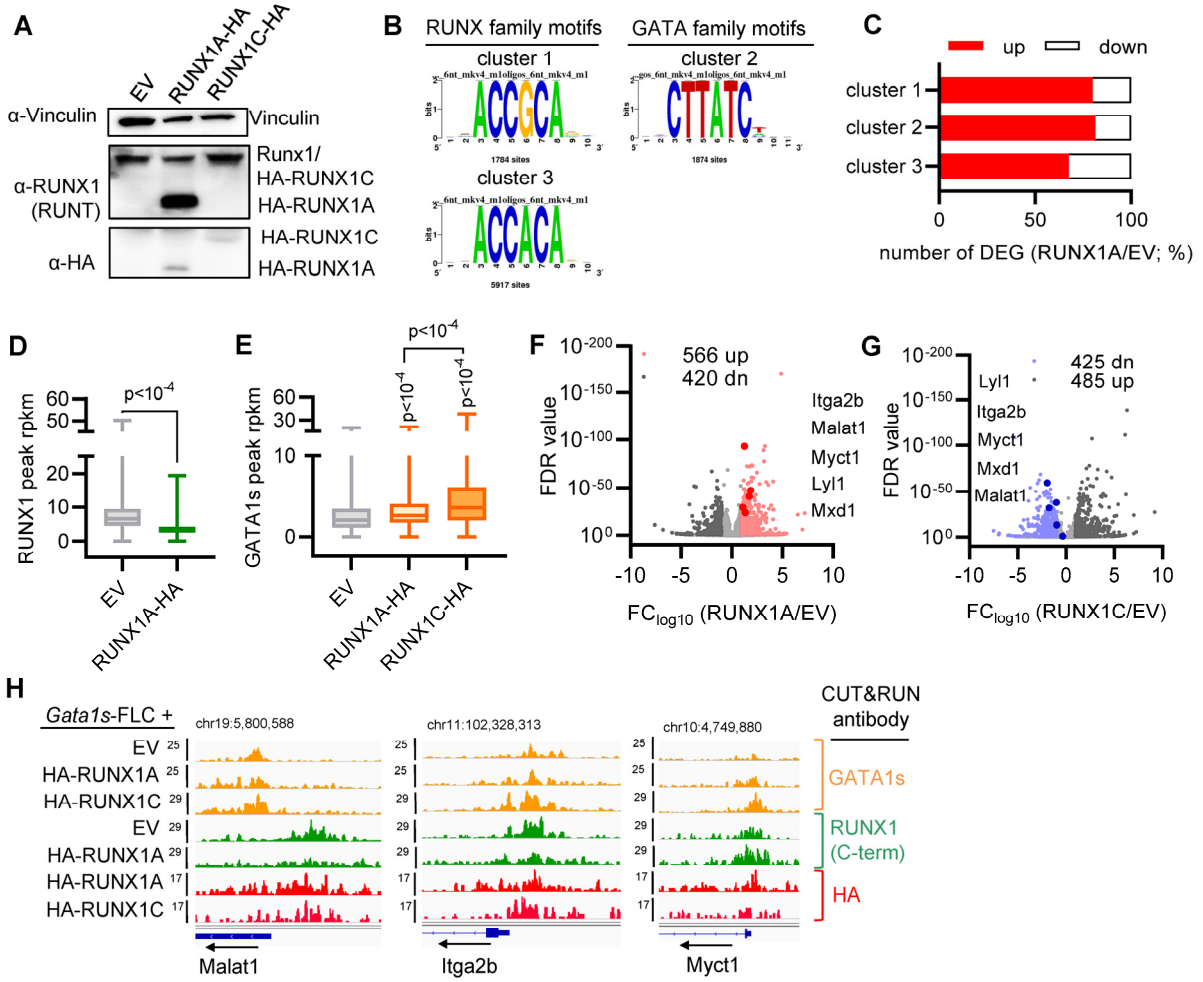
(I) Kaplan-Meier survival curve of mice transplanted with either *Gata1s*-FLCs or *RUNX1A-Gata1s*-FLCs expressing miR-125b or control (miR-Ctrl) (n=4-5; log-rank test).



D

	Core complex	RUNX1A	RUNX1A/RUNX1C common	RUNX1C
Mitotic cell cycle	Anaphase-promoting complex	ANAPC4/ ANAPC5/ CDC23	ANAPC10	
RNA polymerase II transcription	RNA polymerase II holoenzyme complex	POLR2B	GTF2H4	
	COPRS-PRMT5-RUNX1-CBFB complex/ RUNX1-CBF-beta-DNA complex		CBFB/ RUNX1	
	Integrator-RNAPII complex/ DSS1 complex/ Integrator complex	POLR2B	INTS1/ INTS2	
	BRCA1-RNA polymerase II complex	POLR2B	GTF2H4	BRCA1
TGFbeta-signaling	SMAD3-HEF1-APC10-CDH1 complex		SMAD3/ ANAPC10	
Histone modification	BRCA1-SMAD3 complex		SMAD3	BRCA1
chromatin remodeling	Emerin complex 32			SAP18/ TRIM28
	ALL-1 supercomplex			SAP18/ WDR5
	B-WICH complex			SF3B1/ MYBBP1A
	NSL complex			WDR5/ KANSL2
Spliceosome	Spliceosome/ C complex spliceosome			WDR5/ KANSL2
Nuclear pore complex	Nuclear pore complex			NUP153/ NUP214
Ribosome	55S ribosome, mitochondrial		MRPL11/ RPS27	RPS3
	Nop56p-associated pre-rRNA complex		MYBBP1A/ RPS3	
	Parvulin-associated pre-rRNP complex	MYBBP1A	RPS3	CDK11B
	Ribosome, cytoplasmic+mitochondrial		RPS27/ RPS3	

Supplemental Figure 5. Distinct RUNX1A and RUNX1C protein interaction networks.
(A-C) Isolating HA-RUNX1A- or HA-RUNX1C-containing protein complexes from CMK cells.
(A) Western blot after doxycycline induced HA-RUNX1A/RUNX1C expression or in EV control expressing cells using the indicated antibodies.
(B-C) Volcano plots showing enriched proteins after co-immunoprecipitation (Co-IP) of either HA-tagged RUNX1A or RUNX1C. CMK cells were induced with doxycycline for 36h before Co-IP. Proteins with a log fold change of 1 and a p-value of 0.05 in RUNX1A (B; red) or RUNX1C (C; blue) compared to the empty control were deemed significantly bound. The known RUNX1 cofactors CBFbeta and Myc-association factor X (MAX) are indicated by name.
(D) Protein core complexes of RUNX1A (red) or RUNX1C (blue) co-bound proteins. The table shows complexes for selected co-bound proteins. Data were generated using the Corum database [34].



Supplemental Figure 6. RUNX1A affects gene regulation by displacing endogenous RUNX1C.

(A-H) CUT&RUN analyses of Cas9-knock-in Ter119$^-$ FLCs lentivirally transduced with *Gata1*-targeting sgRNAs, as well as with *HA-RUNX1A, HA-RUNX1C* or control (EV) vector.

(A) Western blot using the indicated antibodies.

(B) De novo motif discovery using sequences below RUNX1 peaks in CUT&RUN clusters 1-3. The most abundant motif in each cluster is shown. The number of the respective motif sites within each cluster are shown below.

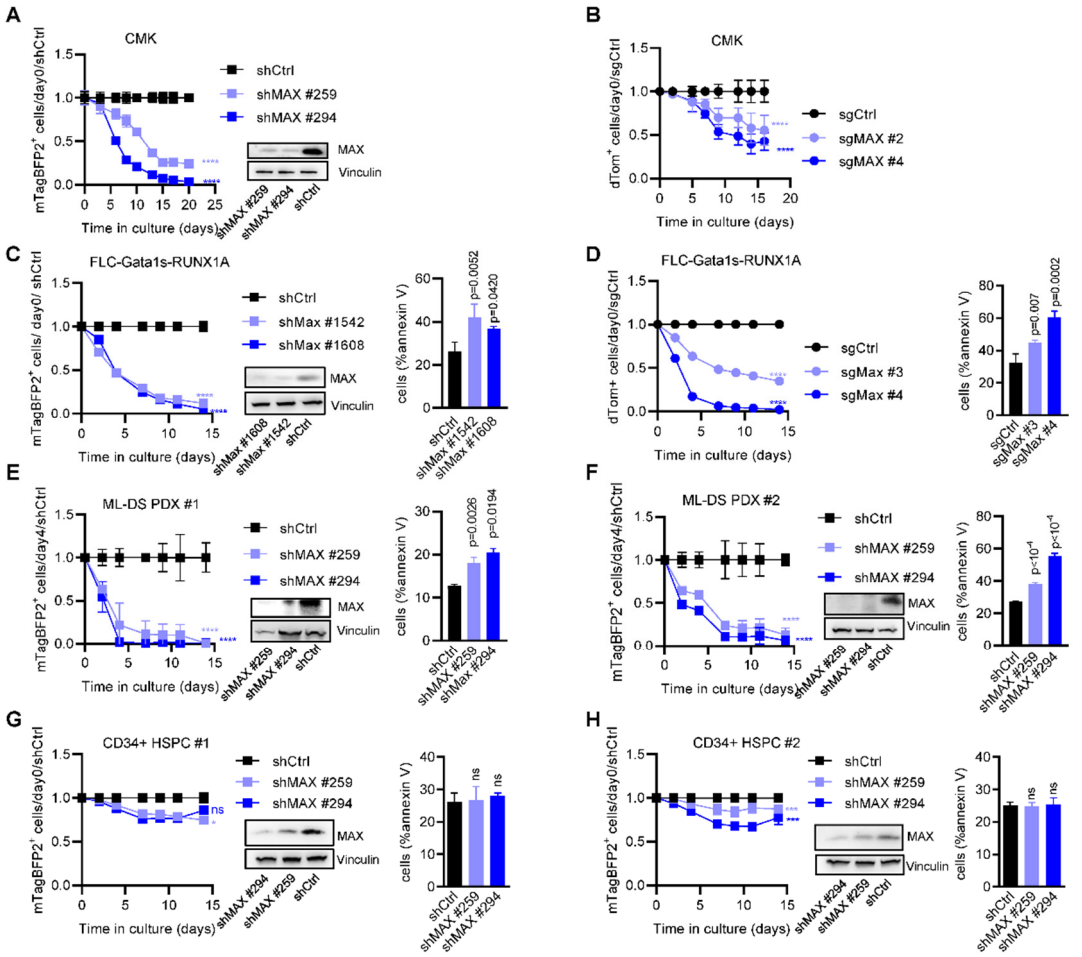
(C) Number of up- and downregulated genes within each CUT&RUN cluster upon doxycycline-induced *HA-RUNX1A* expression, normalized to the empty vector (EV) control (logFC >1; FDR < 0.1). Graphs depict the percentage of up- and downregulated genes normalized to the total number of deregulated genes per cluster.

(D) RUNX1 binding intensities (peak RPKM) at regulatory regions (enhancer and promoter) in RUNX1A-HA or empty vector control (EV) cells. (Two-way ANOVA).

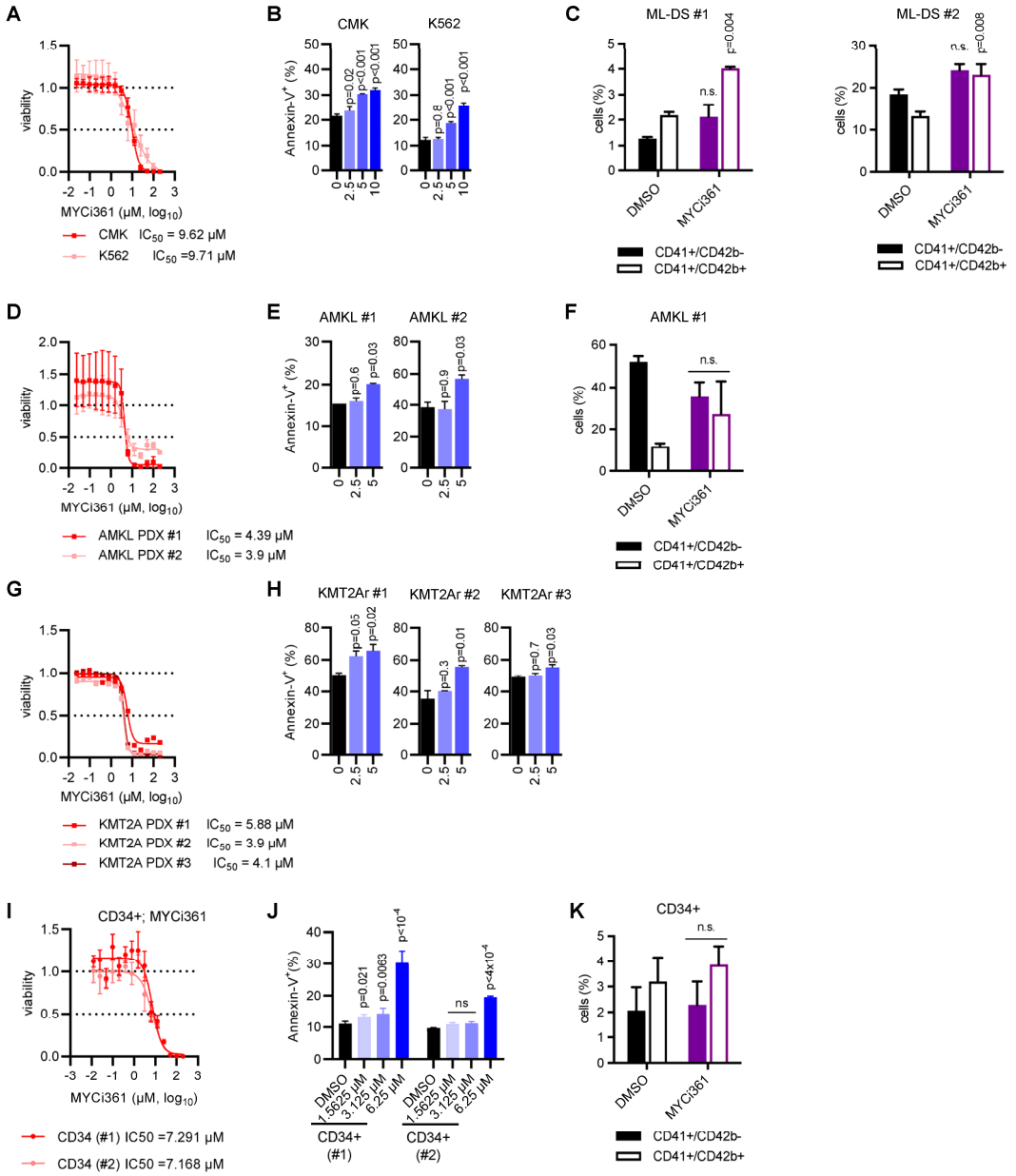
(E) GATA1s binding intensities (peak RPKM) at regulatory regions (enhancer and promoter) in RUNX1A-HA, RUNX1C-HA or empty vector control (EV) cells. (Two-way ANOVA).

(F-G) Volcano plots showing up- and downregulated genes upon doxycycline-induced (F) *HA-RUNX1A* or (G) *HA-RUNX1C* expression in *Gata1s*-FLCs. Upregulated genes are shown as light red or blue dots, respectively. Representative genes with MYC:MAX or MAX binding motifs in their promoter regions are highlighted. Gene names are listed to the right according to their FDR values.

(H) IGV snapshots of the *Malat1*, *Itga2b* and *Myct1* gene promoters, showing occupancy of endogenous GATA1s, RUNX1 and HA-RUNX1A/RUNX1C in *Gata1s*-FLCs after doxycycline induced HA-RUNX1A/RUNX1C expression or in EV control expressing cells. The tracks display coverage (RPKM) (left). Scale and chromosome location are shown (top).



Supplemental Figure 7. RUNX1A exerts oncogenic effects via MYC:MAX.
(A) Percentage of shMAX (BFP2⁺) transduced CMK cells normalized to shCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA) (left). Representative Western blot confirming shRNA-mediated knockdown of MAX using anti-MAX and anti-Vinculin (loading control) antibodies.
(B) Percentage of sgMAX (dTomato⁺) transduced CMK cells normalized to sgCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA).
(C) Percentage of shMax (BFP2⁺) transduced *RUNX1A-Gata1s*-FLCs cells normalized to shCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA). Representative Western blot confirming shRNA-mediated knockdown of Max using anti-Max and anti-Vinculin (loading control) antibodies. Bar graph showing percentage of AnnexinV⁺ FLC-*Gata1s*-RUNX1A cells after shRNA-mediated knockdown of Max at day 2 (mean ± SD; n=3; one-way ANOVA).
(D) Percentage of sgMax (dTomato⁺) transduced *RUNX1A-Gata1s*-FLCs cells normalized to sgCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA). Bar graph showing percentage of AnnexinV⁺ FLC-*Gata1s*-RUNX1A cells after sgRNA-mediated knockout of Max at day 2 (mean ± SD; n=3; one-way ANOVA).
(E-F) Percentage of shMAX (BFP2⁺) transduced ML-DS PDX cells (PDX #1; PDX #2) normalized to shCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA). Representative Western blot confirming shRNA-mediated knockdown of Max using anti-MAX and anti-Vinculin (loading control) antibodies. Bar graph showing percentage AnnexinV⁺ ML-DS PDX cells (PDX #1; PDX #2) after shRNA-mediated knockdown of MAX at day 2 (mean ± SD; n=3; one-way ANOVA).
(G-H) Percentage of shMAX (BFP2⁺) transduced CD34⁺ HSPCs cells (#1 and #2) normalized to shCtrl and to day 0 (mean±s.d., n=3, one-way ANOVA). Representative Western blot confirming shRNA-mediated knockdown of MAX using anti-MAX and anti-Vinculin (loading control) antibodies. Bar graph showing percentage of AnnexinV⁺ CD34⁺ HSPCs cells after shRNA-mediated knockdown of MAX at day 2 (mean±s.d.; n=3; one-way ANOVA).



Supplemental Figure 8. Targeting MYC:MAX dimerization as therapeutic approach.
(A) Dose response curves for the MYC:MAX dimerization inhibitor MYCi361 in CMK and K562 cells 24h after treatment *in vitro*. The IC50-values are depicted. (B) Bar graph showing the corresponding percentage of AnnexinV$^+$ cells after treatment with the indicated doses of MYCi361 in comparison to the DMSO control (right; mean±s.d.; n=3; one-way ANOVA). (C) Percentage of ML-DS blasts expressing CD41/CD42b cell surface markers 5 days after MYCi361 treatment. Data from two independent replicates are shown as mean±s.d. Two-way ANOVA.
(D) Dose response curves for the MYC:MAX dimerization inhibitor MYCi361 in AMKL blasts derived from two patients 24h after treatment *in vitro*. The IC50-values are depicted. (E) Bar graph showing the corresponding percentage of AnnexinV$^+$ cells after treatment with the indicated doses of MYCi361 in comparison to the DMSO control (mean±s.d.; n=3; one-way ANOVA). (F) Percentage of AMKL blasts expressing CD41/CD42b cell surface markers 5 days after MYCi361 treatment. Data from two independent replicates are shown as mean±s.d. Two-way ANOVA.
(G) Dose response curves for the MYC:MAX dimerization inhibitor MYCi361 in *KMT2Ar* blasts derived from three patients 24h after treatment *in vitro*. The IC50-values are depicted below the graph. (H) Bar graph showing the corresponding percentage of AnnexinV$^+$ cells after treatment with the indicated doses of MYCi361 in comparison to the DMSO control (mean±s.d.;

n=3; one-way ANOVA). Data from three independent replicates are shown as mean±s.d. Two-way ANOVA.
(I) Dose response curves for the MYC:MAX dimerization inhibitor MYCi361 in $CD34^+$ HSPCs from two donors 24h after treatment *in vitro*. The IC50-values are depicted. (J) Bar graph showing the percentage of $AnnexinV^+$ cells after treatment with the indicated doses of MYCi361 in comparison to the DMSO control (mean±s.d.; n=3; one-way ANOVA). (K) Percentage of $CD34^+$ HSPCs expressing CD41/CD42b cell surface markers 5 days after MYCi361 treatment. Data from two independent replicates are shown as mean±s.d. Two-way ANOVA

Supplemental Tables

Supplemental Table 1. List of reagents and resources

Supplemental Table 2. Patient characteristics

Supplemental Table 3. Hsa21 CRISPR/Cas9 sgRNA library and screen results

Supplemental Table 4. GSEA of sgGata1s- or sgCtrl-FLCs transduced with *RUNX1A*, *RUNX1C* or EV

Supplemental Table 5. GSEA of murine ML-DS-like leukemia samples compared to HSPCs

Supplemental Table 6. RUNX1A-HA lost and bound genomic regions

Supplemental Table 7. Transcription factor overrepresentation analysis of CUT&RUN peaks

Supplemental Information References

1. Reimer J, Knoss S, Labuhn M, et al. CRISPR-Cas9-induced t(11;19)/MLL-ENL translocations initiate leukemia in human hematopoietic progenitor cells in vivo. *Haematologica*. 2017;102(9):1558-1566.
2. Weber K, Bartsch U, Stocking C, Fehse B. A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. *Mol Ther*. 2008;16(4):698-706.
3. Adams FF, Heckl D, Hoffmann T, et al. An optimized lentiviral vector system for conditional RNAi and efficient cloning of microRNA embedded short hairpin RNA libraries. *Biomaterials*. 2017;139:102-115.
4. Bhayadia R, Krowiorz K, Haetscher N, et al. Endogenous Tumor Suppressor microRNA-193b: Therapeutic and Prognostic Value in Acute Myeloid Leukemia. *J Clin Oncol*. 2018;36(10):1007-1016.
5. Emmrich S, Rasche M, Schoning J, et al. miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFbeta and Wnt signaling. *Genes Dev*. 2014;28(8):858-874.
6. Gialesaki S, Mahnken AK, Schmid L, et al. GATA1s exerts developmental stage-specific effects in human hematopoiesis. *Haematologica*. 2018;103(8):e336-e340.
7. Li W, Xu H, Xiao T, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol*. 2014;15(12):554.
8. Labuhn M, Adams FF, Ng M, et al. Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. *Nucleic Acids Res*. 2018;46(3):1375-1385.
9. Rongvaux A, Willinger T, Martinek J, et al. Development and function of human innate immune cells in a humanized mouse model. *Nat Biotechnol*. 2014;32(4):364-372.
10. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094-3100.
11. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
12. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923-930.
13. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
14. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.
15. Frankish A, Diekhans M, Ferreira AM, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*. 2019;47(D1):D766-D773.
16. Nowicka M, Robinson MD. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Res*. 2016;5:1356.
17. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
18. Aken BL, Achuthan P, Akanni W, et al. Ensembl 2017. *Nucleic Acids Res*. 2017;45(D1):D635-D642.
19. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621-628.
20. 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 U S A*. 2005;102(43):15545-15550.
21. Schwarzer A, Emmrich S, Schmidt F, et al. The non-coding RNA landscape of human hematopoiesis and leukemia. *Nat Commun*. 2017;8(1):218.
22. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife*. 2017;6.
23. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools. *Elife*. 2019;8.

24. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589.
25. Gao T, Qian J. EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Res*. 2020;48(D1):D58-D64.
26. Ramirez F, Ryan DP, Gruning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*. 2016;44(W1):W160-165.
27. Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*. 1995;23(23):4878-4884.
28. Cartharius K, Frech K, Grote K, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*. 2005;21(13):2933-2942.
29. Thomas-Chollier M, Darbo E, Herrmann C, Defrance M, Thieffry D, van Helden J. A complete workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. *Nat Protoc*. 2012;7(8):1551-1568.
30. Fornes O, Castro-Mondragon JA, Khan A, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2020;48(D1):D87-D92.
31. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods*. 2009;6(5):359-362.
32. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res*. 2019;18(2):623-632.
33. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.
34. Giurgiu M, Reinhard J, Brauner B, et al. CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res*. 2019;47(D1):D559-D563.
35. Bolouri H, Farrar JE, Triche T, Jr., et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat Med*. 2018;24(1):103-112.