Supplemental Material and Methods

Patient samples

DNMT3A R882^{mut} and NPM1^{mut} human AML samples were selected from the AML-SG and SAL biorepositories (#148/10 and EK98032010, respectively).
Pseudonymized left-over samples from healthy bone marrow donors (rinse-back from bone marrow filters) were used with written informed consent in accordance with vote #329/10 (Ethics Committee of Goethe University Medical Center). Healthy bone marrow samples were processed as previously described [1]
Patient characteristics are summarized in Supplementary Table 1A and 1B.

Flow cytometry and FACS

For primary AML or nBM samples staining was performed in PBS+2%FCS using fluorescence conjugated antibodies targeting human CD3 (Biolegend), human CD19, human CD235a (eBiosciences), human CD20 (BD Biosciences), human CD34 (eBiosciences), human CD38 (Life Technologies), human GPR56 (Biolegend). Recombinant biotinylated NKG2D-Fc chimera was added to the mix to stain all NKG2D ligands[2]. DAPI (Sigma-Aldrich) or 7-AAD (BD Bioscience) were used for dead cell exclusion. Cells were FACS-sorted or analyzed using the gating strategy showed in Figure 1C. Briefly, after morphological gate and duplets exclusion, live cells were selected. We therefore further selected for SSC^{dim} Lineage^{low} cells (with lineage positivity defined as CD3+CD20+CD19+CD235a+). Finally, we defined 5 subpopulations based on CD34, GPR56 and NKG2DL expression: CD34-GPR56+NKG2DL-, CD34+GPR56+NKG2DL-, CD34-GPR56-NKG2DL-, CD34+GPR56+NKG2DL+, CD34-GPR56-NKG2DL-.
For PDX samples staining was performed in PBS+2%FCS using anti-human CD45 (BD), anti-human CD33 (Life Technology), anti-human CD19 (eBioscience) and anti-murine CD45 (Biolegend) antibodies. For PDX samples, after morphological gate and exclusion of duplets and dead cells, we gated on mCD45-hCD45+hCD33+ or mCD45-hCD45+hCD19+.
Samples were FACS sorted using BD FACS Aria Fusion (BD Biosciences) or analyzed using BD LSR Fortessa or BD LSR II (BD Biosciences).

Mice

NOD/SCID/IL2Rγ^{null} (NSG) mice (Jackson Laboratory, Bar Harbor, ME, USA) were maintained under pathogen-free conditions according to the German federal and state regulations (approved by the Regierungspräsidium Karlsruhe, Tierversuchsantrag number G43/18 and Z110/02).

Xenotransplantation assay

For xenotransplantation experiments using bulk samples, primary patient samples were thawed, washed in IMDM+10%FCS and treated with DNAse I (Roche) for 10' at 37°C. Cells were then washed and CD3+ cells depletion was performed according to manufacturer instructions using human CD3 conjugated microbeads from Miltheny (130-050-101). CD3- cells were then washed, counted and transplanted intrafemorally in 8-12 weeks old female NSG mice one day after sub-lethal irradiation (175 cGy).

For xenotransplantation experiments using FACS-sorted populations, primary patient samples (AML or healthy bone marrow donors) were thawed, washed in IMDM+10%FCS and treated with DNAse I (Roche) for 10' at 37°C. Cells were then washed and stained with fluorescence conjugated antibodies as described above. After FACS sorting, cells were resuspended in PBS and transplanted intrafemorally[3] (between $2.2x10^3$ and $1x10^5$ cells per mouse) into 8-12 weeks old female NSG mice one day after sub-lethal irradiation (175cGy).
For secondary xenotransplantation experiments, primary Patient Derived Xenografts from $CD34^-GPR56^+NKG2DL^-$ or $CD34^+GPR56^+NKG2DL^-$ transplanted primary samples were thawed and enriched for human cells by FACS sorting after staining for human CD45 and murine CD45, as described above. Cells (between $6x10^5$ and $4x10^6$ cells per mouse) were then resuspended in PBS and transplanted intrafemorally in 8-12 weeks old female NSG mice one day after sub-lethal irradiation (175cGy).
Human engraftment was monitored from 8 to 50 weeks by peripheral blood sampling or bone marrow aspiration (every 2 week or 6 weeks, respectively) or at signs of distress. Engraftment was defined as $\geq 0.1\%$ human cells in murine bone marrow measured by flow cytometry (BD LSR Fortessa or BD LSR II, BD Biosciences) using anti-human CD45 (BD), anti-human CD33 (Life Technology), anti-human CD19 (eBioscience) and anti-murine CD45 antibodies (Biolegend).

Colony formation assay
Primary patient samples (AML or healthy bone marrow donors) were thawed, washed in IMDM+10%FCS and treated with DNAse I (Roche) for 10' at 37°C. Cells were then washed and stained with fluorescence conjugated antibodies: anti-human CD3 (Biolegend), CD19 (eBioscience), CD235a (eBiosciences), CD20 (BD Biosciences), CD45 (Biolegend) CD34 (eBiosciences), GPR56 (Biolegend) together with simultaneous staining of all NKG2D ligands using a recombinant biotinylated NKG2D-Fc chimera as described[2]. 7AAD was used to exclude dead. After morphological gate and exclusion of duplets and dead cells, $Lineage^{Low}SSC^{dim}CD45^{dim}CD34\text{-}GPR56+NKG2DL\text{-}$, $Lineage^{Low}SSC^{dim}CD45^{dim}CD34+GPR56+NKG2DL\text{-}$, $Lineage^{Low}SSC^{dim}CD45^{dim}CD34\text{-}GPR56\text{-}NKG2DL\text{-}$, $Lineage^{Low}SSC^{dim}CD45^{dim}CD34+GPR56+NKG2DL+$ and $Lineage^{Low}SSC^{dim}CD45^{dim}CD34\text{-}GPR56\text{-}NKG2DL+$ cells were FACS-sorted, plated in triplicate in 1 ml of MethoCult™ H4434 Classic (Stem Cell Technologies) and incubated at 37 °C in a humidified incubator with 5% CO_2 (Thermo Scientific). Colonies were counted after 14 days.

DNA extraction
Primary PDX samples derived from $CD34^-GPR56^+NKG2DL^-$ or $CD34^+GPR56^+NKG2DL^-$ transplanted primary samples were thawed, washed in IMDM+10%FCS and treated with DNAse I (Roche) for 10' at 37°C. Cells were then washed and stained in PBS+2%FCS with fluorescence conjugated antibodies: anti-human CD45 (BD), anti-human CD33 (Life Technology), anti-human CD19 (eBioscience) and anti-murine CD45 antibodies (Biolegend). 7-AAD was used for dead cell discrimination. After morphological gate and exclusion of duplets and dead cells, mCD45-hCD45+hCD33+ and mCD45-hCD45+hCD19+ were FACS-sorted in PBS+2%FCS. Cells were then spun

down and DNA extraction was performed according to manufacturer instructions using QIAamp DNA Micro Kit (QIAGEN). DNA was quantified using Qubit dsDNA HS Assay Kit (Life Technologies).

Digital droplet PCR

Detection of DNMT3A p.R882C, DNMT3A DNMT3A p.R882H, NPM1 c.863_864insTCTG mutations was performed using ddPCR™ Mutation Assay dHsaMDS475153762 (Bio-Rad), ddPCR™ Mutation Assay dHsaMDV2010089 (Bio-Rad) and TaqMan® dPCR Liquid Biopsy Assays Hs000000064_rm (Thermo Fisher), respectively. QX200 Digital Droplet PCR system (Bio-Rad) was used.

RNA extraction

Primary patient samples (AML or healthy bone marrow donors) were thawed, washed in IMDM+10%FCS and treated with DNAse I (Roche) for 10' at 37°C. Cells were then washed and stained with fluorescence conjugated antibodies: anti-human CD3 (Biolegend), CD19 (eBioscience), CD235a (eBiosciences), CD20 (BD Biosciences), CD45 (Biolegend) CD34 (eBiosciences), GPR56 (Biolegend) together with simultaneous staining of all NKG2D ligands using a recombinant biotinylated NKG2D-Fc chimera as described[2] were used for the gating $Lineage^{Low}SSC^{dim}CD45^{dim}CD34$-GPR56+NKG2DL- and $Lineage^{Low}SSC^{dim}CD45^{dim}CD34$+GPR56+NKG2DL-. Staining for NK cells markers using CD94 (BD Bioscience) and CD56 (Biolegend) was performed to check that sorted subpopulations were excluded of NK cells. $1x10^3$ to $50x10^3$ cells from each subpopulation were directly sorted into RNA lysis buffer and RNA was isolated according to the instructions (Arcturus PicoPure RNA Isolation Kit, Life Technologies, Invitrogen). RNA integrity was checked with Bioanalyzer using Agilent RNA 6000 Pico Reagents (Agilent).

RNAseq library preparation

cDNA was generated using the SMART Seq v4 ultra-low RNA kit (Takara). Libraries were produced using NEBNext® ChIP-Seq Library Prep Master Mix Set for Illumina® (New England Biolalbs). Libraries were sequenced paired-end with read a length of 125 bp using a HiSeq2000 V4 (Illumina).

Statistical analysis

Data represent mean ±SD. *P*-values are derived via application of unpaired t-test. Multiple testing correction was performed according to Benjamini and Hochberg. Significance is depicted as: no significance = ns, $P<0.05$ *, $P<0.01$ **, $P<0.001$ ***, $P<0.0001$ **** according to statistical tests in each figure legend.

RNAseq analysis

Bcl2fastq2 2.20 was used for conversion. Reads were trimmed for adapter sequences and aligned to the 1000 Genomes Phase 2 assembly of the Genome Reference Consortium human genome (build 37, version hs37d5) with STAR (v2.5.3a)[4]. For alignment, the following parameters were used: alignIntronMax: 500.000, alignMatesGapMax: 500.000, outSAMunmapped: Within, outFilterMultimapNmax: 1, outFilterMismatchNmax: 3, outFilterMismatchNoverLmax: 0.3, sjdbOverhang: 50, chimSegmentMin: 15,

chimScoreMin: 1, chimScoreJunctionNonGTAG: 0, chimJunctionOverhangMin: 15. GENCODE gene annotation (GENCODE Release 19) was used for building the index. BAM files were sorted using SAMtools (v1.6)[5] and duplicates were marked with Sambamba (v0.6.5)[6]. Raw counts were generated using featureCounts (Subread version 1.5.3)[7].
For calculation of normalized counts, mtRNA, tRNA, rRNA as well as all transcripts from the Y- and X-chromosome were removed and subsequently normalization was performed in analogy to TPM (transcripts per million).
For downstream analysis, counts were processed using a variance-stabilizing transformation obtained from DESeq2[8] and corrected for batch effects using batchelor[9].
Analysis of differential expression was performed using DESeq2. The lfcshrink function was used to define differentially expressed genes (abs(log2FC) ≥1, p.adj ≤ 0.05). Log2 fold changes (non-shrinked) were used for GSEA analysis with clusterProfiler[10] and the Molecular Signatures Database v7.4[11] as reference gene sets. Cell types were inferred by deconvolution via non-negative least square (implemented in the R package nnls v1.4) with previous integration and feature selection using the package Seurat. Data handling was performed in R (v3.6.0) using RStudio (v1.4). The LSC17 signature score was calculated on log2-transformed TPM (transcript per million). Student's t-test was used to calculate significance ($P < 0.001$: ***, $P < 0.01$: **, $P < 0.05$: *, no significance: NS).

Data availability

Data are deposited in the European Genome-Phenome Archive (EGA) with the following accession numbers: Study: EGAS00001006527
Data set: EGAD00001009293

Supplementary Items legends

Supplementary Table 1A
Summary table of diagnostic primary human AML samples and bone marrow from healthy donors.

Supplementary Table 1B
Summary table including AML patients clinical information.

Supplementary Table 2
Summary table showing number of cells transplanted, time elapsed from transplantation to endpoint, fraction of hCD45+, hCD33+ and hCD19+ cells in primary PDX transplanted with bulk primary patient AML samples.

Supplementary Table 3
Summary table showing number of cells transplanted, time elapsed from transplantation to endpoint and frequency of hCD45+ cells in primary PDX transplanted with FACS-sorted CD34-GPR56+NKG2DL-, CD34+GPR56+NKG2DL-, CD34+GPR56+NKG2DL-, CD34-GPR56-NKG2DL- or CD34-GPR56-NKG2DL+ populations from either primary AML samples or healthy donors.

Supplementary Table 4
Summary table showing number of cells transplanted, time elapsed from transplantation to endpoint and frequency of hCD45+ cells in secondary PDX transplanted with hCD45+ cells obtained from CD34-GPR56+NKG2DL- and CD34+GPR56+NKG2DL- primary PDXs.

Supplementary Figure 1
a. Boxplot showing expression levels of CD34, GPR56 and NKG2DLigands (MICA, MICB, RAET1E, RAET1G, ULBP1, ULBP2, ULBP3) in CD34-GPR56+NKG2DL- (green) or CD34+GPR56+NKG2DL- (violet) FACS-sorted subpopulations from either healthy donors or primary AML patients. **b**. Boxplot showing the LSC104 score[12] for CD34+GPR56+NKG2DL- (violet) and CD34-GPR56+NKG2DL- (green) in primary AML patient samples and healthy donors. **c**. Venn Diagram showing overlap between Differentially Expressed Genes (DEGs) identified comparing CD34+GPR56+NKG2DL- from healthy donors versus CD34+GPR56+NKG2DL- from AML samples, and CD34+GPR56+NKG2DL- from healthy donors versus CD34-GPR56+NKG2DL- from AML samples. **d**. Boxplot showing expression levels of representative genes (CD96, CD97, IL1RAP, HAVCR2, IL2RA) upregulated in CD34+GPR56+NKG2DL- and CD34-GPR56+NKG2DL- from AML samples when compared to CD34+GPR56+NKG2DL- from healthy donors. **e**. Table summarizing the number of DEGs identified when comparing CD34+GPR56+NKG2DL- versus CD34-GPR56+NKG2DL- from healthy donors or CD34+GPR56+NKG2DL- versus CD34-GPR56+NKG2DL- from AML samples. **f**. Barplot of the top20 most enriched gene sets when GSEA is run on CD34+GPR56+NKG2DL- versus CD34-GPR56+NKG2DL- from healthy donors. **g**. Barplot of the top20 most enriched gene sets when GSEA is run on

CD34+GPR56+NKG2DL- versus CD34-GPR56+NKG2DL- from AML samples. **h.** Boxplot representing the inferred frequency of MEP and CLP obtained by deconvoluting bulk RNAseq from FACS-sorted primary AML and healthy donors with RNAseq data from Corces et al.[13] **i**. Boxplot representing the frequency of DNMT3A (left) and NPM1 (right) mutations quantified using RNAseq reads mapping on the wild type or mutated allele in CD34+GPR56+NKG2DL- (violet) and CD34-GPR56+NKG2DL- (green) in primary AML samples. **j**. Barplot showing the fraction of CD99- cells in CD34+GPR56+NKG2DL- (violet) and CD34-GPR56+NKG2DL- (green) in primary AML samples.

Additional references:

1. Kim-Wanner SZ, Luxembourg B, Schmidt AH, et al. Introduction of principles of blood management to healthy donor bone marrow harvesting. *Vox Sang*. 2020;115(8):802-812.
2. Paczulla AM, Rothfelder K, Raffel S, et al. Absence of NKG2D ligands defines leukaemia stem cells and mediates their immune evasion. *Nature*. 2019;572(7768):254-259.
3. Mazurier F, Doedens M, Gan OI, Dick JE. Rapid myeloerythroid repopulation after intrafemoral transplantation of NOD-SCID mice reveals a new class of human stem cells. *Nat Med*. 2003;9(7):959-963.
4. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
5. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
6. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. 2015;31(12):2032-2034.
7. 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.
8. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
9. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology*. 2018;36(5):421-+.
10. Yu GC, Wang LG, Han YY, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology*. 2012;16(5):284-287.
11. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740.
12. Ng SW, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 2016;540(7633):433-437.
13. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet*. 2016;48(10):1193-1203.

Supplementary Figure 1