SUPPLEMENTARY APPENDIX

The ASXL1-G643W variant accelerates the development of CEBPA mutant acute myeloid leukemia

Teresa D'Altri,[1,2,3] Anna S. Wilhelmson,[1,2,3] Mikkel B. Schuster,[1,2,3] Anne Wenzel,[1,2,3] Adrija Kalvisa,[1,2,3] Sachin Pundhir,[1,2,3] Anne Meldgaard Hansen[1,2,3] and Bo T. Porse[1,2,3]

[1]The Finsen Laboratory, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen; [2]Biotech Research and Innovation Centre (BRIC), University of Copenhagen and [3]Danish Stem Cell Center (DanStem) Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

 doi:10.3324/haematol.2019.235150

Received: August 16, 2019.

Accepted: March 19, 2020.

Pre-published: May 7, 2020.

Correspondence: *BO PORSE* - bo.porse@finsenlab.dk

Supplementary information:

The ASXL1-G643W variant accelerates the development of CEBPA mutant Acute Myeloid Leukemia

Teresa D'Altri, Anna S. Wilhelmson, Mikkel B. Schuster, Anne Wenzel, Adrija Kalvisa, Sachin Pundhir, Anne Meldgaard Hansen and Bo T. Porse

Content:

Supplemental Methods

Additional information regarding the generation of the Asxl1^{G643W} knock-in mice

The two *Asxl1* target sequences (5'-GGCCACCACT GCCATCGGAG-3' and 5'-GTGGTAACCT CTCGCCCCTC-3') were designed using the online tool (https://gt-scan.csiro.au). In order to generate the two guide RNA-expressing constructs, two pairs of complementary oligos with overhangs (5'-ACACCGGCCA CCACTGCCAT CGGAGG-3' and 5'-AAAACCTCCG ATGGCAGTGG TGGCCG-3', 5'- ACACCGTGGT AACCTCTCGC CCCTCG-3' and 5'-AAAACGAGGG GCGAGAGGTT ACCACG-3') were annealed and cloned into the pSpCas9n-2A-Puro vector (Addgene #48141) using the BsbI restriction site as previously described [1]. The sequence of the 141-mer ssDNA template is the following: (5'-AGGACCCTCG CAGACATTAA AGCCCGTGCT TTGCAGGCTC GGGGAGCGAG AGGTTACCAC TGCAATCGAG AGACGGCCAC CACTGCCATT GGGGGAGGGG GGTGGCCCGG GTGGATCCGG CAGTGGGGCC ATCGATGAGG G-3').

The following primers were used for genotyping: Asxl1wtF (5'-GGCCCGGGTG GATCC-3'), Asxl1mutF (5'-GGCCCGGGTG GAGGT-3') and Asxl1CommonR (5'-ACTGGAGTTT GGGAGGACAG-3').

Additional information regarding in vivo AML development

To assess the impact of the *Asxl1^{G643W}* allele on the development of CEBPA mutant AML, we used our *Asxl1^{G643W}* knock-in, the *Cebpa$^{p30/+}$*, the *Cebpa$^{fl/+}$* and Mx1-Cre lines to generate *Asxl1^{G643W}* / *Cebpa* mutant compound lines [2]. The following genotypes were used for experiments: *Asxl1$^{+/+}$; Cebpa$^{fl/+}$*, *Asxl1$^{G643W/+}$; Cebpa$^{fl/+}$*, *Asxl1$^{G643W/G643W}$; Cebpa$^{fl/+}$*, *Asxl1$^{+/+}$; Cebpa$^{fl/p30}$*; *Mx1Cre, Asxl1$^{G643W/+}$; Cebpa$^{fl/p30}$*; *Mx1Cre* and *Asxl1$^{G643W/G643W}$; Cebpa$^{fl/p30}$*; *Mx1Cre*.

Competitive transplant

BM cells were collected, filtered and counted. 500,000 live and nucleated donor cells (*Asxl1$^{+/+}$*, *Asxl1$^{G643W/+}$* and *Asxl1$^{G643W/G643W}$*; CD45.2) and 500,000 competitor cells (*Asxl1$^{+/+}$*, CD45.1) were mixed and transplanted into cohorts of eight lethally irradiated recipients. Recipient mice were analysed four months later for blood

chimerism. For the analysis, the output CD45.2/CD45.1 ratios were divided by the input ratios of the injection mix to obtain output-to-input ratios.

Additional information regarding flow cytometry analysis and cell sorting

For blood analysis, the following antibody cocktail was used: CD3-FITC, B220-PerCP, Mac1-APC and Gr1-PE (BD Bioscience).

For bone marrow analysis, the following antibodies were used: CD45.1-FITC and CD45.2-PE (BD Bioscience), Flt3-PE-CF594 (e-Bioscience), CD150-BV650 (e-Bioscience), Mac1-PE-Cy5 (e-Bioscience), Gr1-PE-Cy5 (BD Bioscience), B220-PE-Cy5(BD Bioscience), CD3- PE-Cy5 (BD Bioscience), Ter119-PE-Cy5(BD Bioscience), Sca1-APC (BD Bioscience), CD48-PE-CF7 (e-Bioscience), c-Kit-Alexa780 (e-Bioscience), Mac1-APC (BD Bioscience), Gr1-PE (BD Bioscience), Ter119-FITC (BD Bioscience), FcgRII/III-Alexa700 (Invitrogen). To assess BM chimerism in the transplantation experiments, blood or BM cells were stained with CD45.1-FITC and CD45.2-PE (both BD Bioscience).

The following marker combinations were used to define the analysed populations: LT-HSCs (Ter119-, B220-, CD3-, Mac1-, Gr1-, Sca1+, c-Kit+, Flt3-, CD48-, CD150+); ST-HSCs (Ter119-, B220-, CD3-, Mac1-, Gr1-, Sca1+, c-Kit+, Flt3- CD48-, CD150-); MPP2 (Ter119-, B220-, CD3-, Mac1-, Gr1-, Sca1+, c-Kit+, Flt3-, CD48+, CD150+); MPP3 (Ter119-, B220-, CD3-, Mac1-, Gr1-, Sca1+, c-Kit+, Flt3-, CD48+, CD150-); MPP4 (Ter119-, B220-, CD3-, Mac1-, Gr1-, Sca1+, c-Kit+, Flt3+); CFUs (Ter119-, B220-, CD3-, Mac1-, Gr1-, c-Kit+, Sca1-, CD105+, CD150-); GMPs (Ter119-, B220-, CD3-, Mac1-, Gr1-, c-Kit+, Sca1-, FCgRII/III+, CD150-); MkPs (Ter119-, B220-, CD3-, Mac1-, Gr1-, c-Kit+, Sca1-, CD41+, CD150+).

For assessment of myeloid progenitor populations in leukemic animals, the following strategy was applied: BM cells isolated as described above were incubated with anti-CD45.2-biotin (BD Pharmingen), washed once and stained with the following antibody cocktail: Mac1-FITC (BD Pharmingen), CD41-PE (eBioscience), Gr1-PECy5 (Invitrogen), B220-PE-Cy5 (eBioscience), CD3e-PE-Cy5 (eBioscience), Sca1-PerCP-Cy5.5 (eBioscience), CD105-PE-Cy7 (BioLegend), CD150-APC (Biolegend), FcgRII/III-Alexa700 (Invitrogen), c-Kit-APC-eFlour780 (Invitrogen), CD45.1-eFlour450 (eBioscience), Streptavidin-Q-dots655 (Life Technologies). Donor-derived cells were gated as $CD45.1^{-}$ $CD45.2^{+}$.

Cell sorting was performed using a FACSAriaI whereas analytical stains were analysed using LSRII or FACSCanto instruments (all BD Biosciences). All analyses were carried out using the FlowJo software (BD Bioscience).

Quantitative reverse transcription PCR

RNA was extracted using the RNAeasy Mini kit (Qiagen) following the manufacturer's instructions. cDNA was synthesized using the ProtoScript cDNA synthesis kit (New England Biolabs). Real-time PCR was performed in a LightCycler 480 (Roche) using the SYBR Green I PCR Master Mix (Roche). The following primers were used: mAsxl1F (5'-CAGCCCACTA AAGAGGAGCC-3'), (5'-TCCGGGGGCA TATCTGGTAA -3'), mTraipF (5'- CAAAGTGCTGACCAGGAGATCA-3), mTraipR (5'-CTCATTGGTCGCCGGAGG-3'), mFhl3F (5'-CTCGTCCAAGTGCCAGGAAT-3'), mFhl3R (5'-CCAGCTGCTGCCCTTGTATT-3'), mClcnkaF (5'-CTGGTCCTGGCTTCAGTCAC-3'), mClcnkaR (5'-AGAGAGTGCAAGTGTTCCGC-3'), mAnglt7F (5'-CGTCTTCAGCACCAAGGACA-3'), mAnglt7R (5'-GCAGCAGTTGTACCAGTAGCC-3'), mActinF (5'-CTCTTCCAGC CTTCCTTCCT-3') and mActinR (5'- TGCTAGGGCT GTGATCTCCT-3').

RNA sequencing analysis

RNA-seq reads were processed with the bcbio RNA-seq pipeline (https://github.com/bcbio/bcbio-nextgen) and the bcbioRNASeq R package (https://github.com/hbc/bcbioRNASeq). Transcript abundance estimates were obtained using Salmon [3], summarized to gene level and imported into R using tximport [4]. Differential gene expression analysis between *Asxl1*$^{+/+}$*; Cebpa*$^{\Delta/p30}$ and *Asxl1*$^{G643W/G643W}$*; Cebpa*$^{\Delta/p30}$ AML cells was performed using DESeq2 with standard parameters [5]. Shrunken log fold changes were used for visualization in the volcano plot and as ranking metric in gene set enrichment analysis (GSEA). GSEA was carried out using clusterProfiler [6], testing all gene ontology (GO) biological process gene sets containing 15-500 genes. Data has been deposited in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/gds) under the accession number GSE133314.

ChIP-seq analysis

ChIP-seq data from CEBPA mutant AML was derived from [7]. The mean ChIP enrichment between replicates was calculated using deepTools (v3.2.1) with options (bamCompare --operation mean) [8]. The mean bigWig plots were lifted over from mm9 to mm10 using CrossMap (v0.3.4) [9]. The heatmaps and line plots were created by deepTools computeMatrix (reference-point --upstream 5000 --downstream 5000 --missingDataAsZero) and plotHeatmap (--boxAroundHeatmaps yes --colorMap Blues --sortUsing mean).

For the promoter analysis, the expressed transcripts, overlapping between RefSeq and Ensembl were kept, and if a gene contained several isoforms, the longest coding sequence per gene was kept. The low-expressed genes were filtered at average of the normalized count values, taken over all samples (baseMean) > 10. Further, the groups of differentially expressed genes were defined at FDR-adjusted p-value cutoff 0.05 and log2 fold change (log2FC) 0. In contrast to differentially expressed genes, the neutral genes were defined by absolute value of log2FC < 0.01. The plots were visualized using R packages ggplot2[10] and ggpubr (https://CRAN.R-project.org/package=ggpubr).

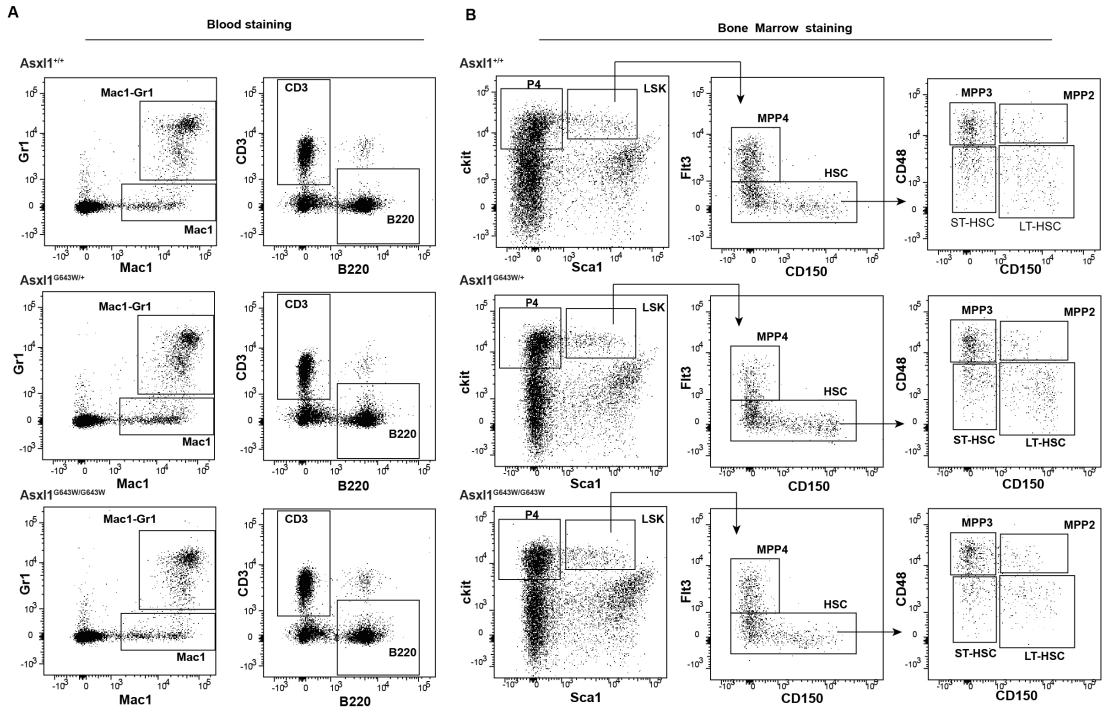
Supplemental References

1. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-2308.
2. Kirstetter P, Schuster MB, Bereshchenko O, et al. Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. Cancer Cell. 2008;13(4):299-310.
3. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-419.
4. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research. 2016;4:1521.
5. 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.
6. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics : a journal of integrative biology. 2012;16(5):284-287.
7. Jakobsen JS, Laursen LG, Schuster MB, et al. Mutant CEBPA directly drives the expression of the targetable tumor-promoting factor CD73 in AML. Science advances. 2019;5(7):eaaw4304.

8. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 2014;42(Web Server issue):W187-191.
9. Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics. 2014;30(7):1006-1007.
10. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York, 2016.

Supplemental Figures

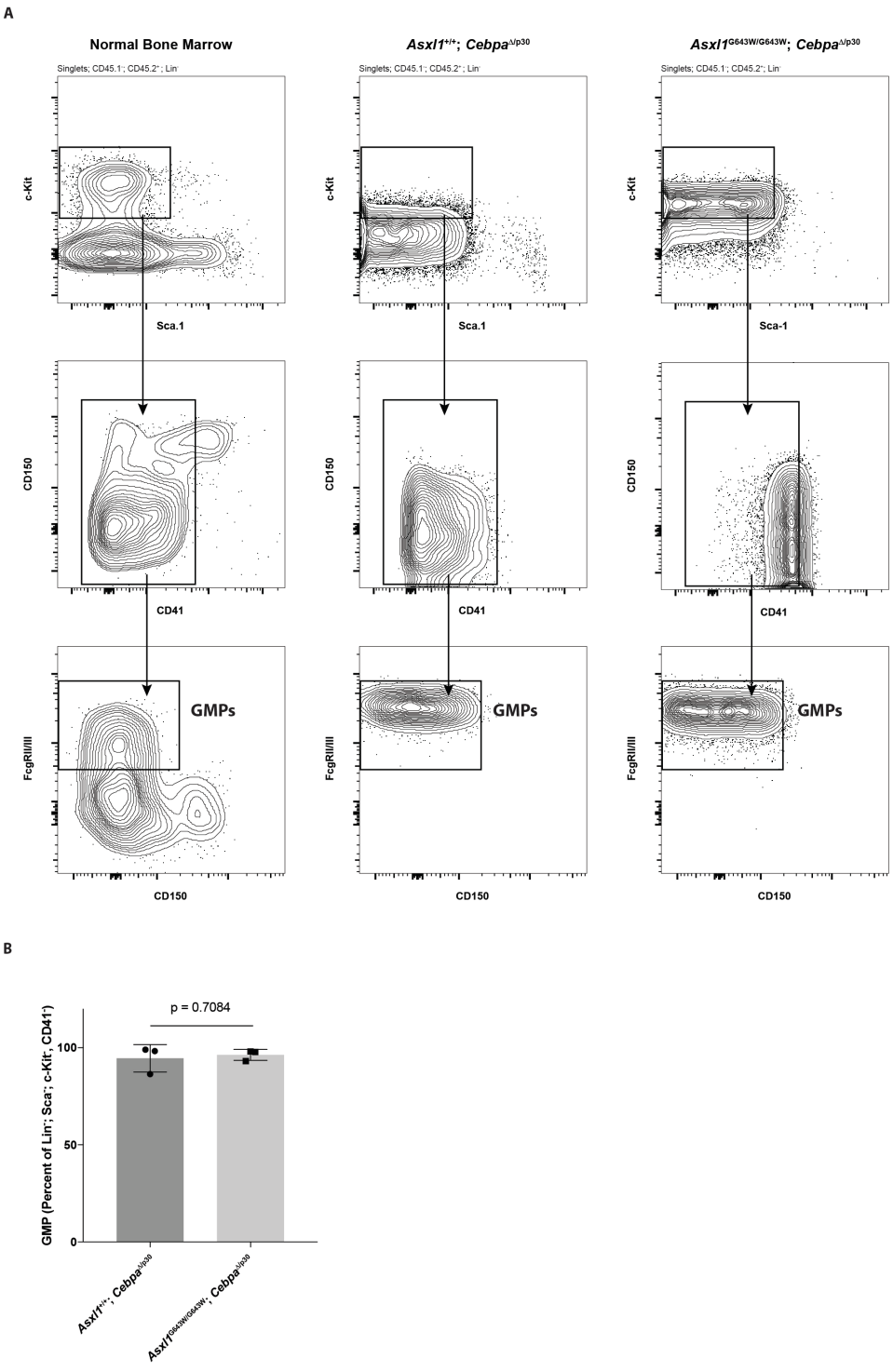
Supplementary Figure S1



Supplementary Figure S1 – related to Figure 2

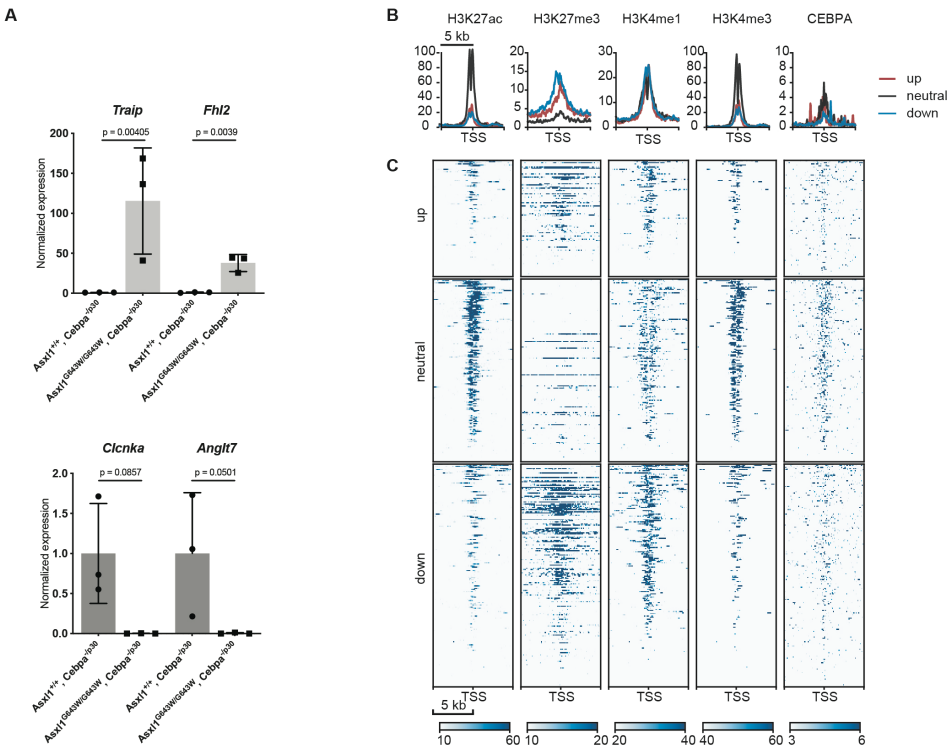
A. Representative FACS analysis of peripheral blood derived from 6-months old *Asxl1*$^{+/+}$, *Asxl1*$^{G643W/+}$ and *Asxl1*$^{G643W/G643W}$ mice.
B. Representative FACS analysis of bone marrow HSCs and MPPs subsets in 6-months old *Asxl1*$^{+/+}$, *Asxl1*$^{G643W/+}$ and *Asxl1*$^{G643W/G643W}$ mice.

Supplementary Figure S2



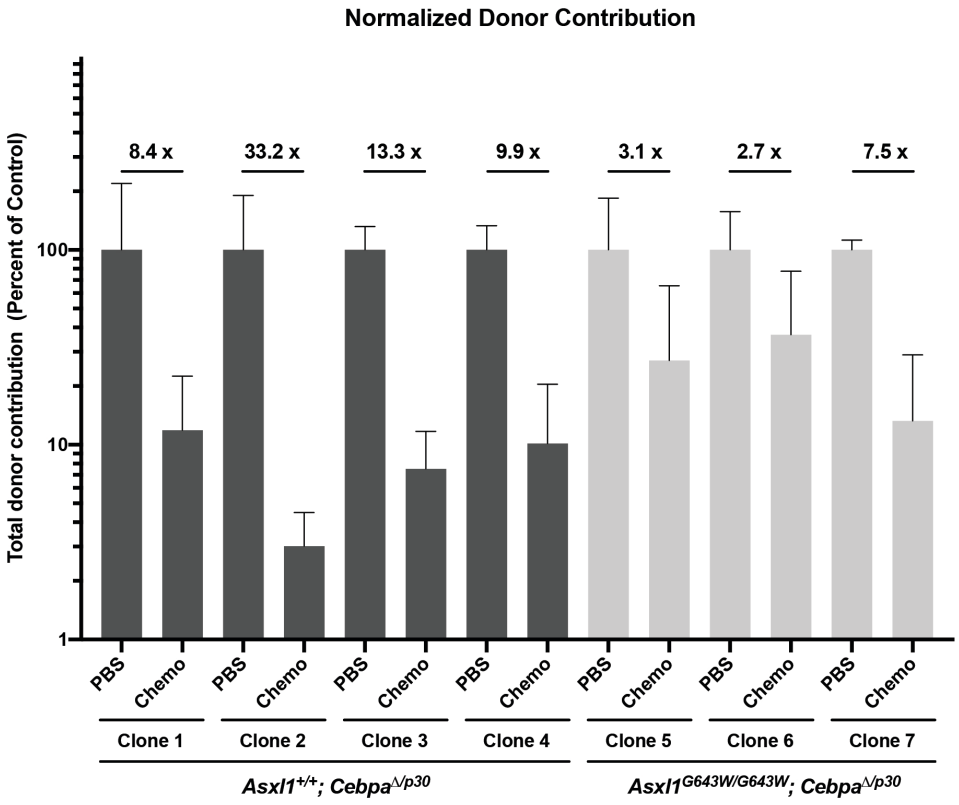
Supplementary Figure S2 – related to Figure 3

A. FACS analysis of *Asxl1*$^{+/+}$*; Cebpa*$^{\Delta/p30}$ or *Asxl1*$^{G643W/G643W}$*; Cebpa*$^{\Delta/p30}$ bone marrow isolated from transplanted mice. The plot shows that all c-Kit+ progenitors are GMPs, irrespectively of the leukemic genotype.
B. Quantification of the data from (A)



Supplementary Figure S3 – related to Figure 4

A. qRT-PCR of selected differentially expressed genes.
B. The ChIPseq signal surrounding differentially expressed genes between *Asxl1*$^{+/+}$*; Cebpa*$^{\Delta/p30}$ and *Asxl1*$^{G643W/G643W}$*; Cebpa*$^{\Delta/p30}$ leukemic samples represented by line plots. C. The data from (B) represented as a heatmap. Up: upregulated genes in *Asxl1*$^{G643W/G643W}$ vs. *Asxl1*$^{+/+}$ leukemic blasts (FDR 0.05, log2FC > 0, baseMean >10, n=105), down: downregulated genes in *Asxl1*$^{G643W/G643W}$ vs. *Asxl1*$^{+/+}$ (FDR 0.05, log2FC < 0, baseMean >10, n=201), neutral: neutral genes (-0.01 < log2FC < 0.01, baseMean >10, n=165) in *Asxl1*$^{G643W/G643W}$ vs. *Asxl1*$^{+/+}$.



Supplementary Figure 4 – related to Figure 5

Analysis of leukemic mice after chemotherapy/vehicle treatment indicating the relative numbers of leukemic cells in the blood. The data represent the behavior of seven different leukemic clones (n=12 recipients of each clone, half treated with vehicle and half with chemotherapy). To facilitate comparison of leukemic engraftment between clones, the vehicle samples (PBS) were set to 100 for each clone. The fold reductions in leukemic cell numbers following chemotherapy are shown for each clone. These values constitute the foundation for the data in Figure 5B. Error bars indicate standard errors for the 6 technical replicates for each condition (PBS/chemo) of each clone.

Legends to Supplemental Tables

Supplementary Table S1

Differentially expressed genes between (*Asxl1*$^{+/+}$*; Cebpa*$^{\Delta/p30}$ and *Asxl1*$^{G643W/G643W}$*; Cebpa*$^{\Delta/p30}$) AML (adjusted p-value<0.05).

Supplementary Table S2

GSEA representing differentially expressed pathways between *Asxl1*$^{+/+}$*; Cebpa*$^{\Delta/p30}$ and *Asxl1*$^{G643W/G643W}$*; Cebpa*$^{\Delta/p30}$ AML (q-value<0.05).