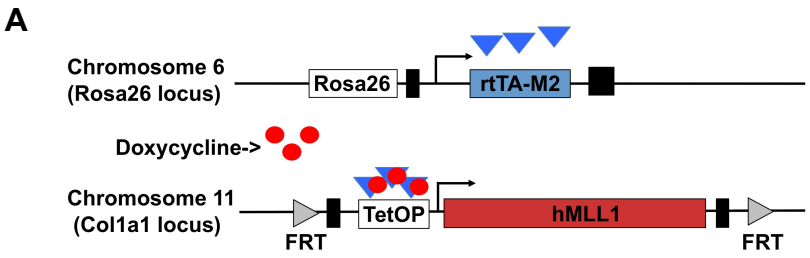
Stem Cell Reports, Volume *14*

Supplemental Information

Enhancing Hematopoiesis from Murine Embryonic Stem Cells through MLL1-Induced Activation of a Rac/Rho/Integrin Signaling Axis

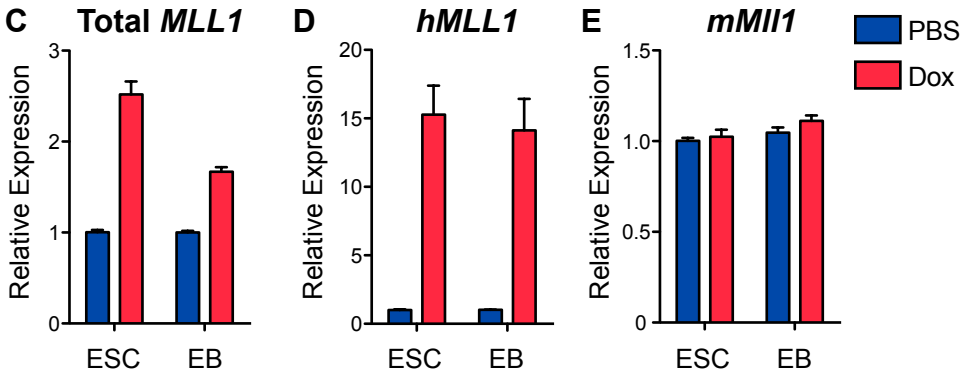
Weiwei Yang, G. Devon Trahan, Elizabeth D. Howell, Nancy A. Speck, Kenneth L. Jones, Austin E. Gillen, Kent Riemondy, Jay Hesselberth, David Bryder, and Patricia Ernst

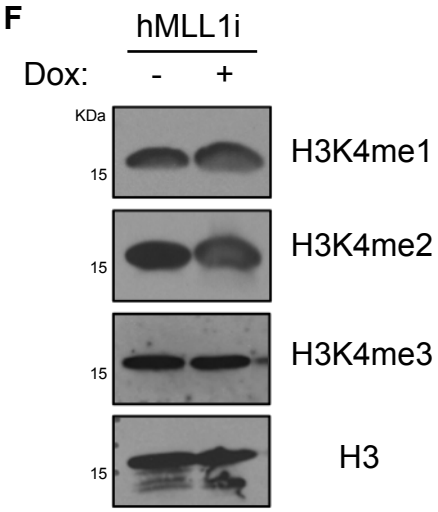
S1. Establishment and validation of hMLL1 inducible ES cell lines (related to Figure 1)

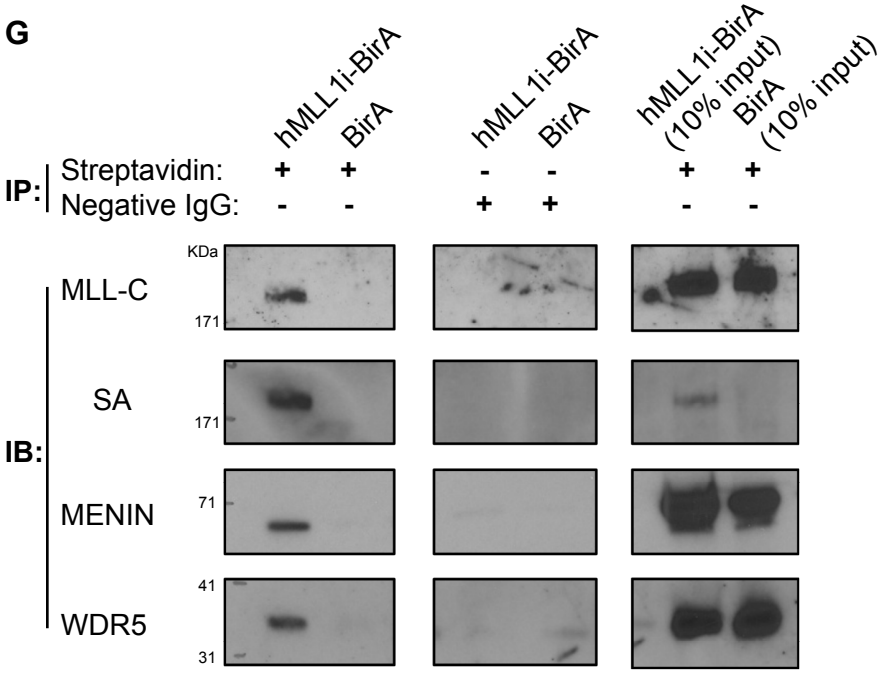


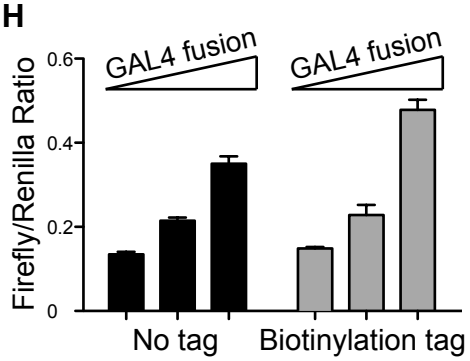
B

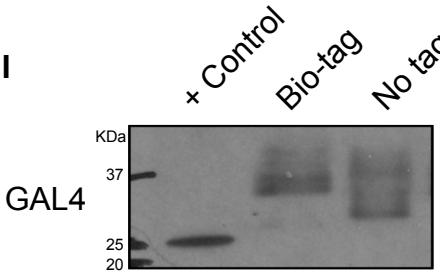
Rounds of Screening	Clone Number
Hygromycin Selection	69
Morphology Screen	52
PCR Screen	4
Total Efficiency	5.80%

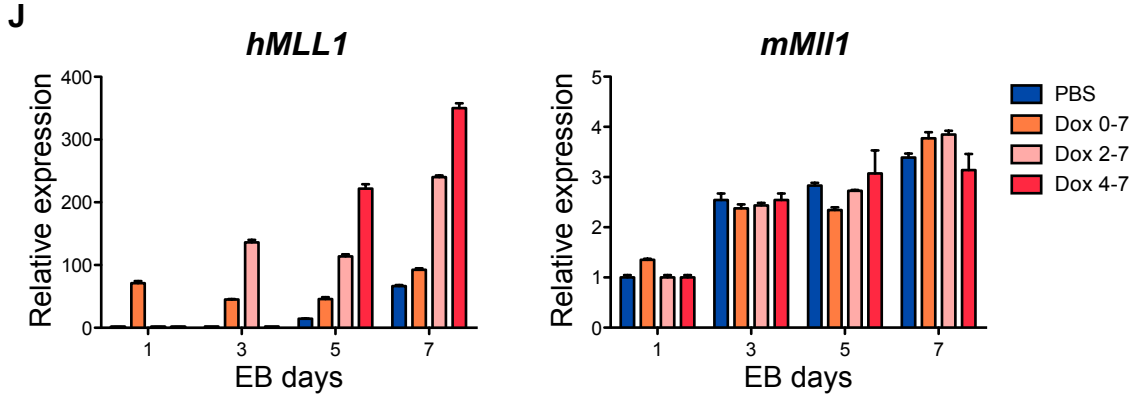




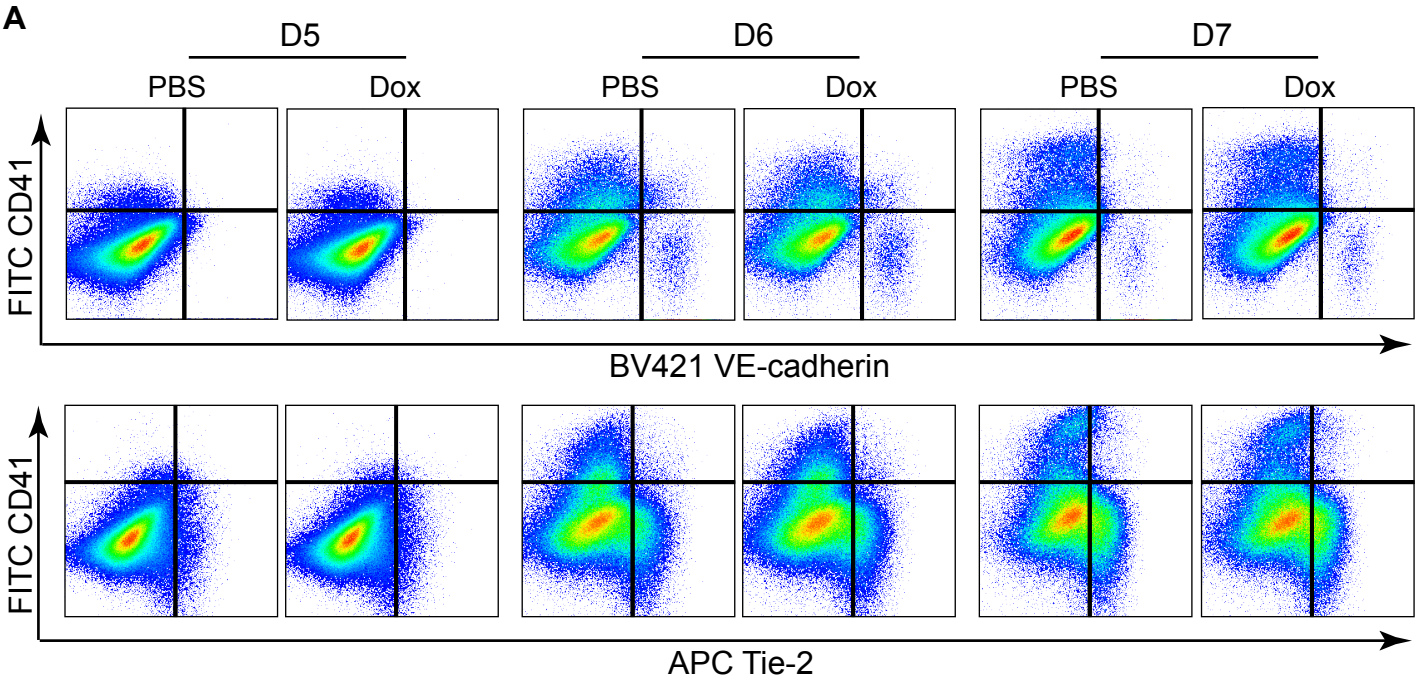


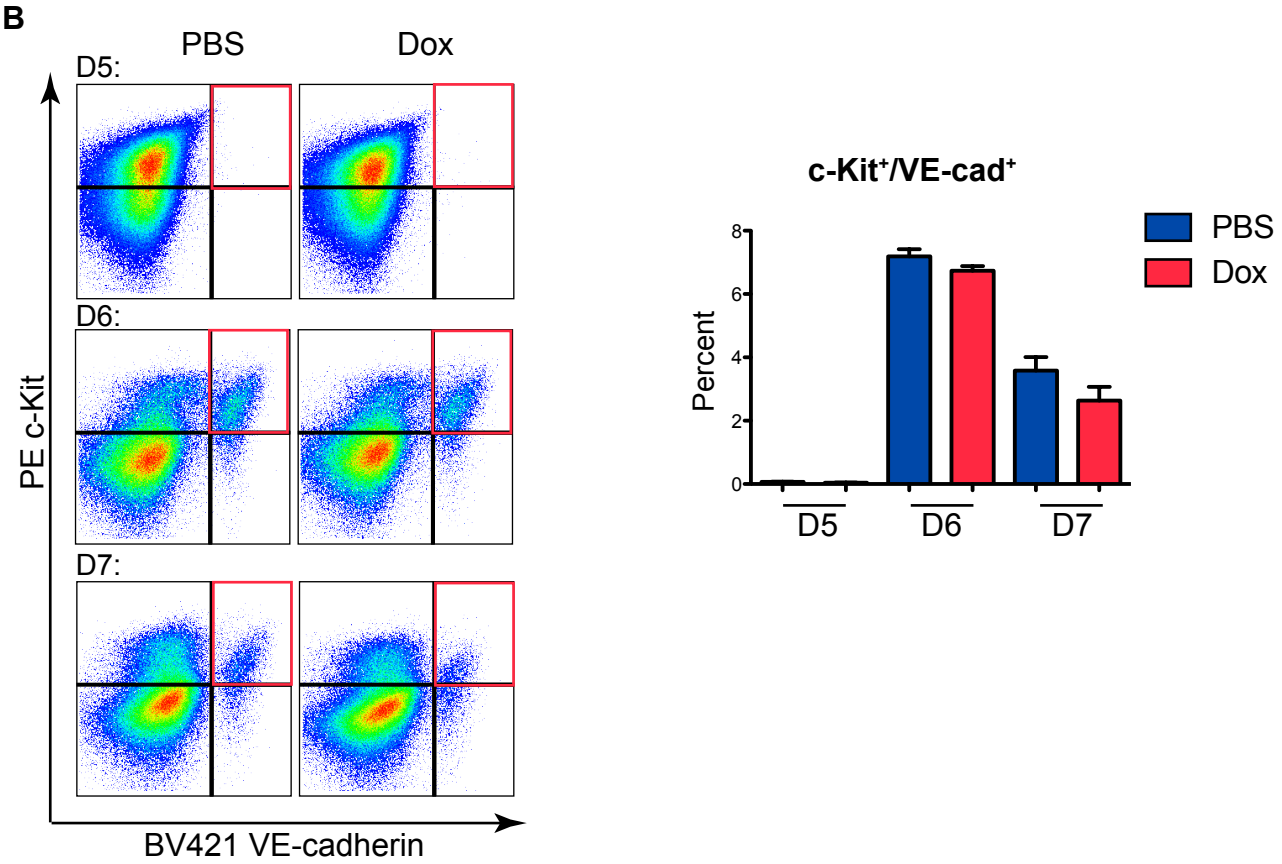




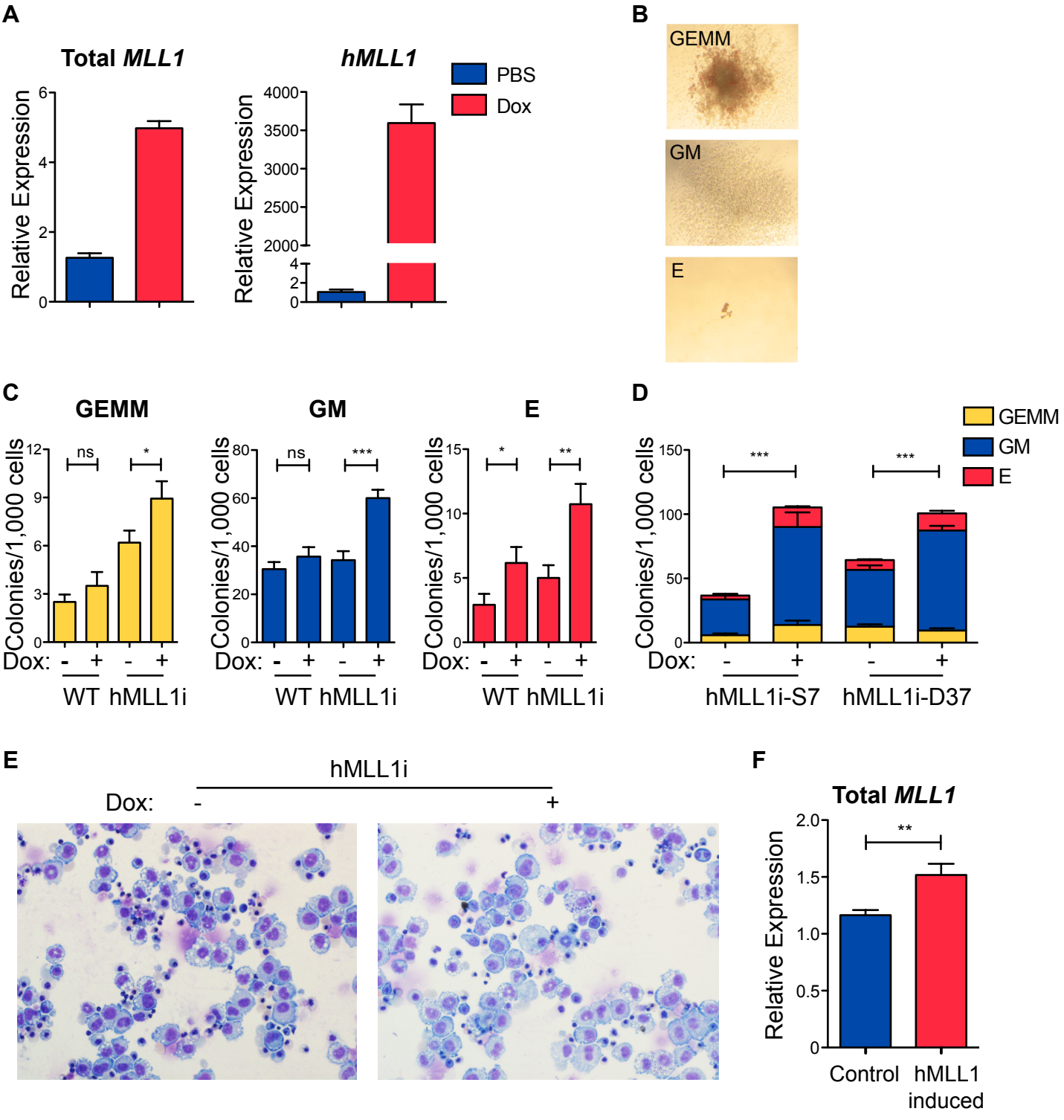


S2. Endothelial and hematopoietic specification is not affected by hMLL1 induction (related to Figure 2)



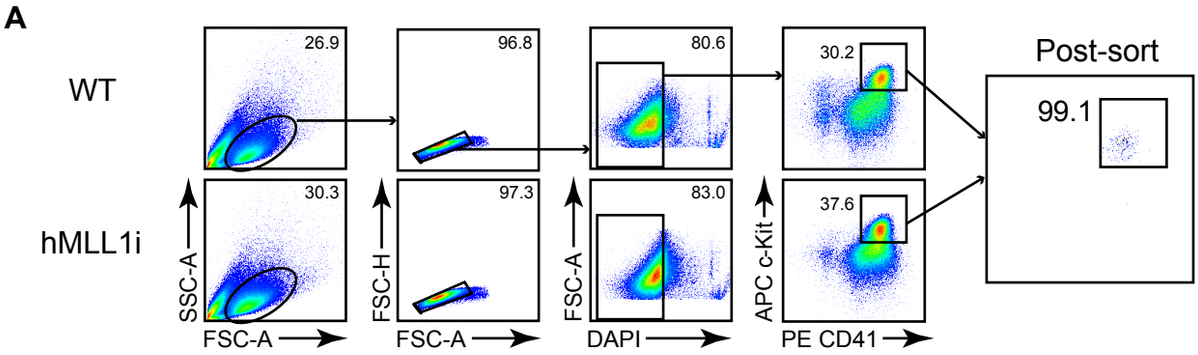


S3. MLL1 increase enhances hematopoietic progenitor activity (related to Figure 3)

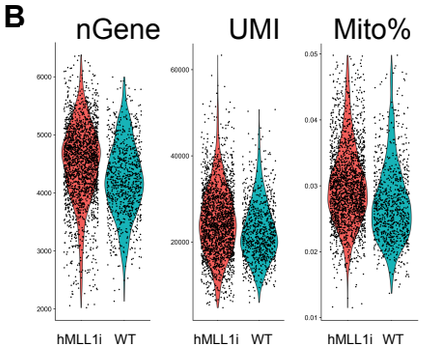


S4. Identification of three clusters from single cell-seq analysis (related to Figure 4)

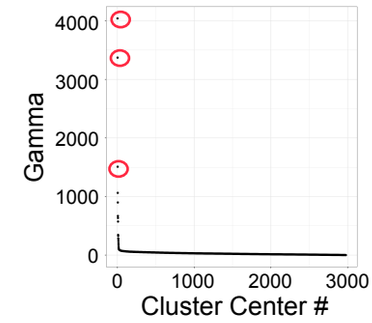
A



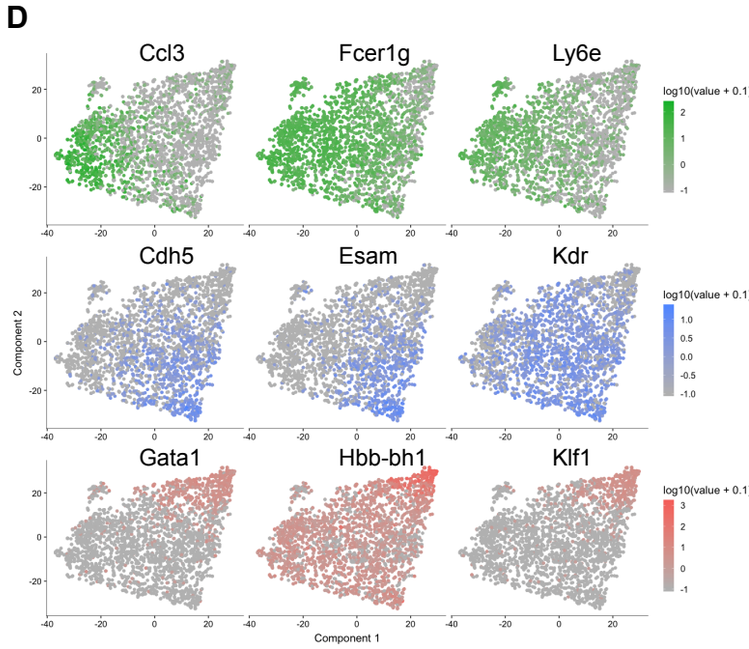
B



C



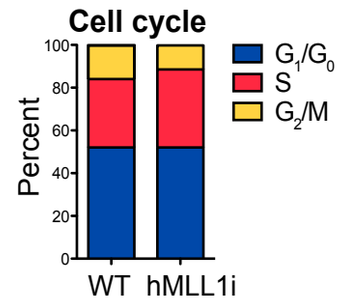
D



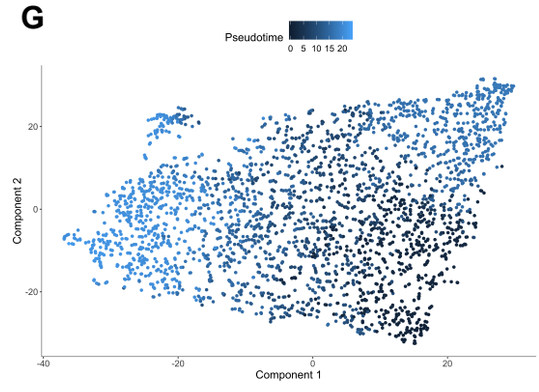
E

GO Terms	p-value
Cluster 1	
Immune system process	4.90E-13
Defense response	5.02E-12
Leukocyte activation involved in immune response	2.02E-10
Myeloid leukocyte differentiation	1.05E-08
Regulation of myeloid cell differentiation	2.61E-08
Cluster 2	
Erythrocyte differentiation	5.83E-05
Regulation of definitive erythrocyte differentiation	7.32E-05
Erythrocyte homeostasis	8.98E-05
Cluster 3	
Regulation of angiogenesis	1.61E-09
Vasculature development	3.70E-09
Regulation of vasculature development	4.40E-09
Positive regulation of angiogenesis	8.53E-09

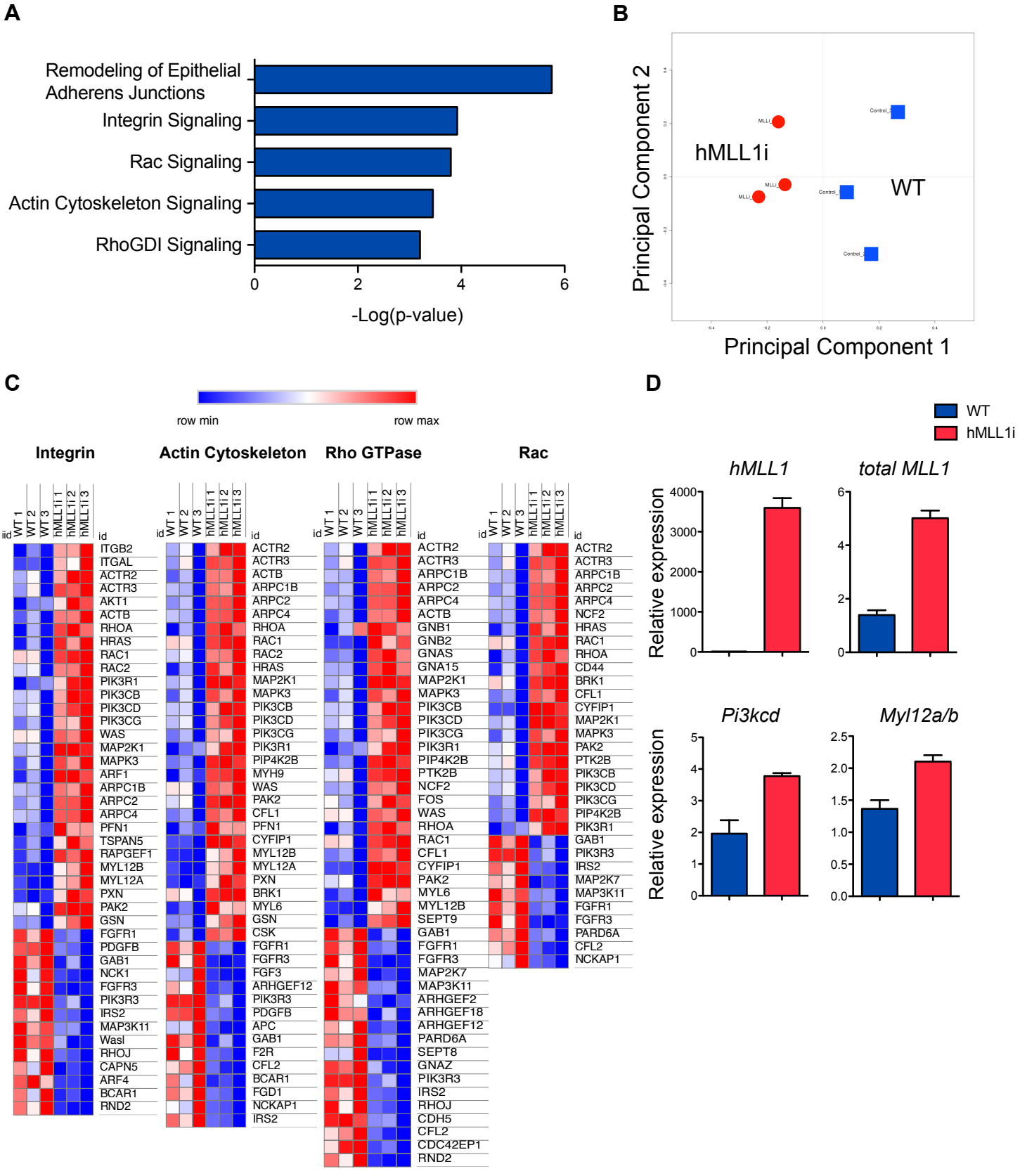
F



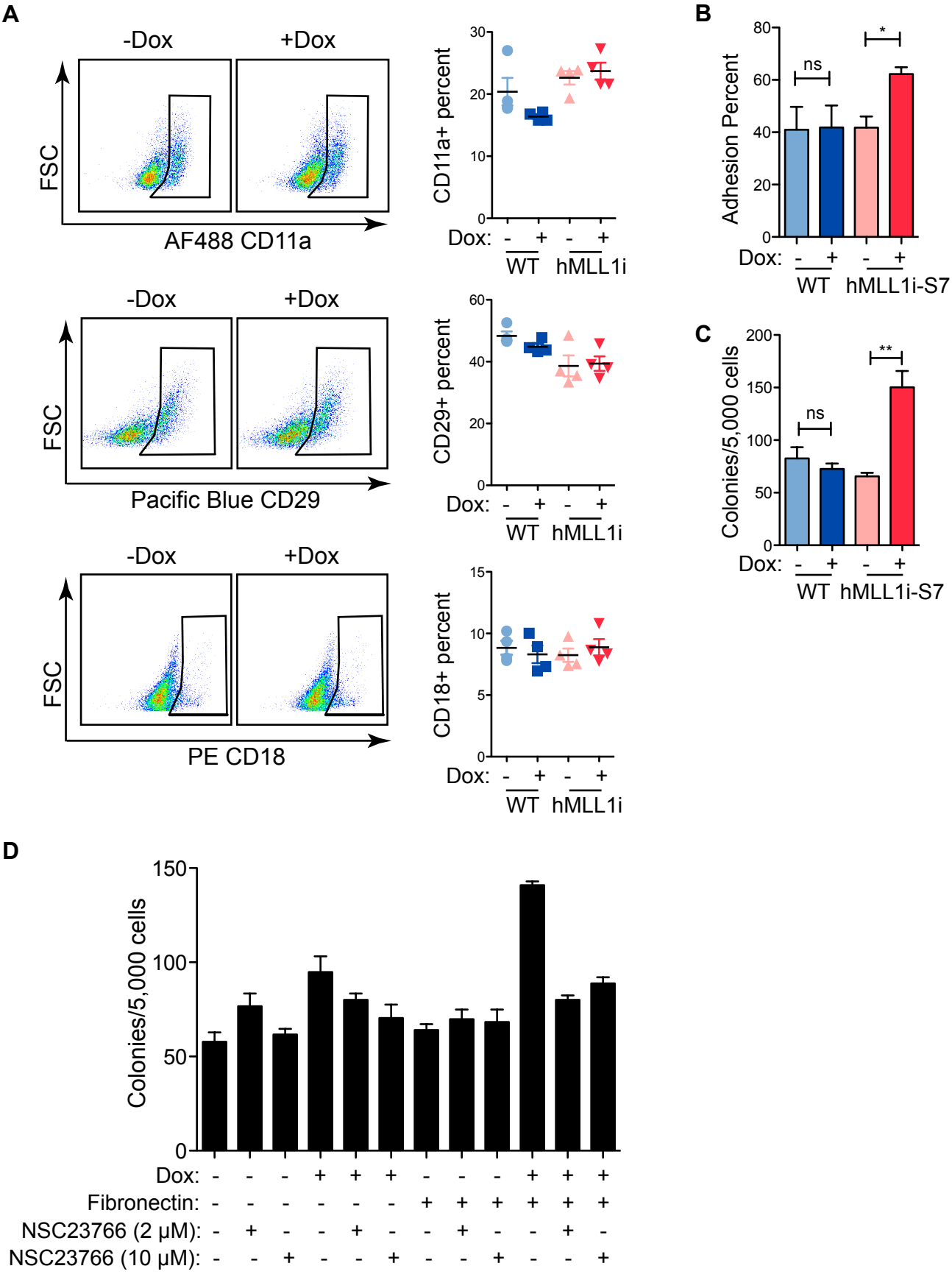
G



S5. hMLL1 induction activates a Rac/Rho/integrin signaling axis (related to Figure 5)



S6. Selective effect of hMLL1 induction on integrins and their activities (related to Figure 6)



Supplemental Figures:

Figure S1. Establishment and validation of hMLL1 inducible ES cell lines. Related to Figure 1. A) Schematic representation of the MLL1-inducible model. The human MLL1 (hMLL1) cDNA was integrated into the 3' UTR of the *Col1a1* locus under the control of a tetracycline operator (TetOP). Expression of the M2 reverse tetracycline transactivator (rtTA-M2) is driven by the *Rosa26* endogenous regulatory element. **B)** Screening strategy and numbers of clones obtained at each step. Hygromycin selection was performed at 200 µg/mL. The second round of screening selected visually against clones that exhibited differentiated morphology and the final round of screening was determined by PCR with the primers listed in Supplemental Experimental Procedures. Total efficiency refers to correctly targeted clones within the hygromycin-selected pool. **C-E)** Relative induction levels of *hMLL1* transcripts only, total *MLL1* transcripts (including both induced *hMLL1* and endogenous *mMLL1*) and mouse endogenous *Mll1* transcripts in ES and EB (day 6) samples. Data represent average expression (relative to *Gapdh*) +/- SEM, n=3 independent samples each. **F)** H3K4 methylation levels in ES cells. Cells were treated with doxycycline for 48 hrs before histone extraction. Total H3 was used as loading control. **G)** Biotin-tagged hMLL1 interacts with complex components including Menin and Wdr5. Nuclear extracts from BirA ligase stably transfected hMLL1i or control 293 lines were subjected to co-immunoprecipitation with streptavidin or IgG, followed by western blotting with antibodies against the MLL1 c-terminus, Menin and Wdr5. **H)** The presence of an internal biotinylation epitope does not disrupt transactivation by MLL1. Due to the introduction of an epitope tag near the transcriptional activation domain, we tested whether this tagged protein was still capable of transcriptional activation. Gal4-hMLL1 fusion fragments with or without the biotinylation tag were co-transfected with a luciferase reporter and a Renilla internal control plasmid into 293 cells. Luciferase measurements were performed 48 hrs after transfection. One representatve of two experiments is shown. Average normalized relative light units is shown +/- SEM . **I)** Western blotting showing Gal4-hMLL1 proteins. Nuclear extracts were prepared from transiently transfected cells. Bio-tag= Biotin tag. **J**) Relative induction levels of *hMLL1* and mouse *Mll1* in differentiating hMLL1-inducible EBs. Data represent average relative expression (relative to *Gapdh*) +/- SEM, n=3 biological replicates. Legend indicates duration of exposure to doxycycline or PBS control.

Figure S2. Endothelial and hematopoietic specification is not affected by hMLL1 induction. Related to Figure 2. A) Flow cytometry analysis of endothelial and hematopoietic markers c-Kit, VE-cadherin, Tie-2 and Cd41 on EBs at different days of EB differentiation. Experiments were performed using hMLL1 inducible cells. Doxycycline was supplied from day 4 to day 7. **B)** Development of hematopoietic cells from c-Kit+/VE-cadherin+ hemogenic endothelium. Experiments were performed using hMLL1 inducible cells. Dox was supplied from day 4 to day 7. Data is representative of three independent experiments and show averages +/- SEM.

Figure S3. MLL1 increase enhances hematopoietic progenitor activity. Related to Figure 3. A) qRT-PCR shows transcript levels of *hMLL1* and total *MLL1* in the sorted c-Kit+/Cd41+ population. Note that *hMLL1* and total *MLL1* transcript levels are higher in c-Kit+/Cd41+ hematopoietic progenitors than in crude ES/EBs. Data represents average expression (relative to *Gapdh*) +/- SEM, n=3 independent experiments. **B)** Representative colony morphologies from day 7 of the CFU assay. **C**) Colony numbers separated by hematopoietic colony type. GEMM = granulocyte-erythrocyte-monocyte-megakaryocyte; GM = granulocyte-macrophage; E = erythroid. **D)** The increase in hematopoietic potential is a consistent feature of hMLL1 induction. CFU assays were performed on sorted c-Kit+/Cd41+ cells from two additional hMLL1 inducible clones: S7 and D37. One thousand c-Kit+/Cd41+ EB cells were seeded per dish and quantified as in Figure 3. Experiments were performed twice with triplicate dishes each time using clone S7 and once with clone D37; bars show average colonies +/- SEM. **E)** Wright-Giemsa stain of cytospun cells from hMLL1-induced and control CFU plates after 7 days. **F**) Total *(hMLL1 + Mll1)* transcript induction in E9.5 yolk sac cells. Control embryo genotype = rtTA/+ and *hMLL1*/+ animals, n=6; *hMLL1* induced embryo genotype = hMLL1/+; rtTA/+, n=5.

Figure S4. Identification of three clusters from single cell-seq data. Related to Figure 4. A) Cd41+ enriched populations, gating strategy and post-sort analysis is shown for single cell sequencing experiments. **B)** Violin plots showing numbers of gene detected, unique molecular identifiers (UMI) and mitochondrial

gene read percentages post-filtering. **C)** Gamma plot of potential cluster centers from the corresponding t-SNE plot. Data suggests three major clusters circled in red using a gamma threshold of 1500 correpsonding to the three clusters shown in Figure 4. **D)** Marker gene expression for myeloid/innate immune (green), erythroid (red), and HE-like (blue) cells projected onto t-SNE reduced space. Color bars, log10 values of scaled UMI counts. **E)** Gene ontology terms obtained from PANTHER software and p-value of enrichment in the indicated cluster. **F)** Cell cycle analysis derived from analysis of cell cycle related genes (described in Supplemental Experimental Procedures) expressed in single cells. **G)** Developmental progression of cells, as defined by pseudotime, projected onto t-SNE reduced space. Cells are colored corresponding to the three clusters shown in Figure 4.

Figure S5. hMLL1 induction activates a Rac/Rho/integrin signaling axis. Related to Figure 5. A) IPA enrichment of top canonical pathways in cluster 3 "HE-like" population upon hMLL1 induction ordered by *p*-value. **B)** Principal component analysis shows the distribution of individual samples. **C)** Heatmaps of individual gene expression data driving IPA canonical pathways enriched in hMLL1-induced cells. **D)** Validation of gene expression by qRT-PCR using independently sorted samples. Data represents average expression (relative to *Gapdh*) +/- SEM from three independent experiments distinct from the samples used for RNA sequencing.

Figure S6. Selective effect of hMLL1 induction on integrins and their activities. Related to Figure 6. A) Determination of integrin cell surface expression by flow cytometry. Day 6 EB cells were dissociated and stained with antibodies detecting Cd11a (encoded by *Itgal*), Cd18 (encoded by *Itgb2*) and Cd29 (encoded by *Itgb1*). Quantification is shown below as average percentages +/- SEM, n= 4 biological replicates. Data shows one representative experiment of two. **B)** Increased adhesion upon hMLL1 induction reproduced with an additional hMLL1 inducible ES clone, S7. One representative experiment from two is shown as average percentage adherent cells after 24 hrs +/- SEM, n=3 biological replicates. **C)** Increased CFU frequency after fibronectin adhesion reproduced with an independent hMLL1i ES clone, S7. Methods were as described in Figure 6. Bar graph shows one representative experiment of two. Data represents average CFU +/- SEM, n = triplicate cultures. **D)** Primary CFU data and additional doses of Rac1 inhibitor (NSC23766). Data show colonies per 5,000 cells from one representative experiment of two. Shown are average colony numbers +/- SEM, n = triplicate cultures.

Table S1 Differentially expressed genes defining the 3 clusters identified from single cell sequencing data. Related to Figure 4 and Figure S4. Each of three tabs of the excel sheet show the complete list of differentially expressed genes varying by cluster for combined WT and hMLL1i samples. Differentially expressed genes for each cluster were determined using the Seurat FindConservedMarkers function (outlined in Supplemental Experimental Procedures). Normalized gene expression within cells from a particular cluster was compared to expression within cells from all other clusters. Genes with an absolute log fold change less than 0.25 or those expressed in less than 10 percent of cells in either population being compared were excluded from analysis. The pct.1 column represents the fraction of cells expressing at least one transcript of a gene in the cluster in question and pct.2 column represents the fraction of cells expressing that gene in all other clusters.

Supplemental Experimental Procedures:

Generation of hMLL1 inducible ES cell line. A human MLL1 cDNA (hMLL1) from the flag-tagged plasmid spFM11 (Jude et al., 2007) was first modified by introducing a 23 amino acid biotinylation tag (Beckett et al., 1999) within a Xho-SpeI fragment. This fragment was inserted into the full-length human cDNA which was then introduced into the tetracycline-inducible FRT site flanked vector pBS31 (Ugale et al., 2014). The resulting plasmid (pBS-hMLL1-s-bio) was electroporated into KH2 ES cells (Beard et al., 2006) together with the Flp-recombinase plasmid pOG44 to integrate the hMLL1 cDNA into the modified *Col1a1* locus (Figure S1A). Cells were subjected to hygromycin selection for 2 weeks. Correctly integrated clones were identified by screening genomic DNA (Figure S1B) with primers indicated below. The KH2 ES cells were a kind gift of Dr. David Bryder.

Embryos. To generate hMLL1 inducible adults and embryos, ES cells were injected into C57BL/6 blastocysts and chimeric mice identified (Mouse Genetics Core Facility, NJH, Denver), crossed to C57BL/6 mice (The Jackson Laboratory) and speed-backcrossed to generate congenic mice that are at least 95% C57BL/6 by genomic SNP assays before use. Mice were maintained in the animal facilities at University of Colorado, Anschutz Medical Campus. All animal studies were conducted in accordance with IACUC-approved animal protocols at University of Colorado, Anschutz Medical Campus and performed in accord with ethics regulations. Embryos were generated by timed matings between hMLL1;rtTA or C57BL/6 males and females. The day of vaginal plugging (checked at 8 am) was considered embryonic day (E) 0.5 and pregnant females were switched to doxycycline containing (625 mg/kg, Teklad) diet at the day of mating to induce hMLL1 in embryos.

Colony forming unit (CFU) assay. Sorted or Cd41+ enriched cells were seeded respectively at 1,000 or 5,000 cells per mL using M3434 semi-solid medium (StemCell Technologies) containing 1-2 µg/mL doxycycline (Dox). Colonies were scored and/or replated after 7 days culture at 37°C in a 5% CO_2 humidified incubator.

BrdU cell cycle analysis. Proliferation and cell cycle distribution analysis was performed using a bromodeoxyuridine (BrdU) Flow Kit. Sorted day 6 EB cells were culture in liquid medium for 2 days before a 30-minute incubation with BrdU was performed. Cells were then fixed, permeabilized and stained with APC anti-BrdU antibodies and 7-AAD according to the manufacturer's protocol (BD Biosciences).

Western blot. ES cells +/- Dox were harvested and subjected to nuclear/cytoplasmic fractionation using the Dignam procedure. Fifty µg of nuclear extract from each sample was resolved on 3-8% Tris-Acetate Nupage gels (Invitrogen), transferred to nitrocellulose and probed with C-terminal MLL1 rabbit polyclonal antibodies (Hsieh et al., 2003), anti-rabbit HRP conjugate, incubated with ECL reagent (Pierce) and exposed to x-ray film. For histone westerns, nuclear pellets were extracted with 0.25M HCl overnight at 4°C. Ten µg protein was resolved on 15% acrylamide gels, transferred to nitrocellulose and probed with anti-H3K4 mono, di and tri-methylation-specific antibodies (Cell Signaling and Abcam).

Generation of MLL1 293 Flp-in cell line. The full-length *hMLL1* fragment from spFM11 was cloned into pcDNA5/FRT vector (Invitrogen) and cotransfected with pOG44 (Invitrogen) plasmid into the 293 Flp-in cell line (Invitrogen). Stable transfectants were isolated using hygromycin B at 200 µg/mL. Single clones were isolated and expanded. Correctly integrated clones were identified by PCR and lack of beta-galactosidase activity.

Co-immunoprecipitation. Nuclear extract was prepared from BirA-transfected hMLL1i or control 293 cell lines using the Dignam procedure. Two-hundred µg of nuclear extracts were combined with 20 µL pre-washed M280 streptavidin Dynabeads (Invitrogen) and incubated at 4ºC overnight. For a pull-down specificity control, same amount of lysate was incubated with 3 µg mouse IgG and combined with anti-mouse IgG Dynalbeads (Invitrogen) overnight at 4ºC on a rotating stand. Lysate-bead conjugates were washed three times in lysis buffer, eluted in NuPAGE LDS sample buffer (Invitrogen) and resolved by SDS-PAGE using 3-8% Tris-Acetate Nupage gels (Invitrogen). Co-precipitated protein was determined by western blot probed with antibodies as indicated below.

Transactivation assay. Gal4(1-147)-hMLL1 fusion fragments with or without the biotinylation tag were co-transfected with a luciferase reporter (TATA+Inr, (Zenzie-Gregory et al., 1992)) and a Renilla internal control plasmid into 293 cells. Cells were harvested and lysate was used to perform Dual-Luciferase reporter assays (Promega) according to the manufacturer's instructions. Emission from luciferase molecules was measured using Glomax multi detection system (Promega).

Adhesion assay. Plates were coated with 25 µg/mL fibronectin fragment (RetroNectin, Takara) or 10 µg/mL Vcam1 (R&D Systems) overnight at 4°C. Coated plates were blocked with 2% bovine serum albumin for 1 hour and washed three times with phosphate-buffered saline buffer before applying cells for a 30 minute incubation at 37° C, after which non-adherent viable cells were counted using a hemacytometer. The adherent fraction of cells was removed with enzyme-free Cell Dissociation Buffer (Gibco) and counted as above to determine their percentage in the total population. Cells pooled from both adherent and

suspension fractions were seeded into M3434 medium in 35 mm dishes to determine CFU.

Microscopy for EBs and immunofluorescence. Embryoid bodies were photographed using a Cannon Powershot S3 camera mounted to an Olympus CX41 microscope at 40-100x magnifications with a 1951 USAF resolution target to produce accurate scale bars. Enriched Cd41+ cells were subjected to a 24-hour liquid culture on fibronectin coated slides (Neuvitro Corporation). Cells were then fixed with Cytofix/Cytoperm buffer (BD Biosciences) and stained with rhodamine-conjugated phalloidin at 5 U/mL and Hoescht33342 (both from Invitrogen) for DNA content. Images were captured on an Olympus IX83 automated fluorescence microscope. Four representative images containing at least 20 cells per field were used for quantification of cell area using ImageJ software.

Whole-mount immunostaining. E10.5 embryos were fixed in 2% paraformaldehyde/PBS for 20 minutes on ice and then dehydrated in graded concentrations of methanol/PBS (50%, 100%) for 10 minutes each. Embryo trimming and immunostaining for confocal imaging was performed as described (Yokomizo and Dzierzak, 2010). The following primary antibodies were used: rat anti-mouse Cd117 (eBioscience, AB_467434, 1:250), rat anti-mouse Cd31 (BD Pharmingen, AB_396660, 1:500) and rabbit anti-human/mouse Runx1 (Abcam, AB_2049267, 1:250). Secondary antibodies were goat anti-rat Alexa Fluor 647 (Abcam, AB_141778, 1:500), goat-anti rat Alexa Fluor 555 (Invitrogen, AB_141733, 1:1000) and goat anti-rabbit Alexa Fluor 488 (Life Technologies, AB_2576217, 1:1000). Images were acquired on a Zeiss LSM 710 AxioObserver inverted microscope with ZEN 2011 software. The Zeiss LSM 710 is equipped with 488, 543 and 633 nm wavelengths. Images were processed with Fiji software (Schindelin et al., 2012).

Single cell sequencing RNA (scRNAseq), RNA sequencing (RNAseq) and bioinformatics. Day 6 embryoid bodies were dissociated as described above, Cd41+ cells were magnetically enriched, then incubated with c-Kit and Cd41 antibodies (Biolegend). Single cell suspensions were re-suspended in sorting buffer (Hank's Balanced Salt Solution buffer [HBSS] containing 2% FBS) containing 1 µg/mL diamidino-2-phenylindole (DAPI). Singlet-gated, DAPI-negative, c-Kit+/Cd41+ cells were sorted using a FACSAria Fusion. Cell purity was determined by post-sort re-analysis and was typically >90% (Figure S4A). Approximately 4,000-sorted cells were used to generate libraries and sequenced by the University of Colorado Cancer Center Genomics and Microarray core facility. Fastq files for each sample were processed using Cell Ranger 2.0.2. (10x Genomics) with mm10 as the reference genome. The resulting data was then aggregated and normalized using the Cell Ranger aggr pipeline. A total of 4,257 cells (hMLLi, n=2549; parent KH2 control, n=1708) remained following aggregation, with an average of 185,462 reads per cell and a median of 4,091 genes per cell (see Figure S4 for additional quality control).

Expression data for hMLLi and KH2 cells (n = 4,257) was imported into an R environment; the data was filtered and analyzed using the R packages Seurat (Butler et al., 2018) and Monocle (Qiu et al., 2017). Filtering based on unique gene counts less than 2,000, unique molecular identifier (UMI) counts greater than 80,000, or mitochondrial percentages < 1% or > 5% resulted in 2,977 cells in the final analysis, with a average of 24,207 UMI counts per cell and a median of 4,465 genes per cell (Figure S4B). To determine the identities of the clusters, we found it useful to regress out proliferation markers using the algorithm used by the Seurat::Cell Cycle Scoring function (Tirosh et al., 2016). Principal Component Analysis (PCA) was then performed and the top 16 principal components were selected for subsequent use on the basis that each explained at least 1% of the variance observed. The dimensionality was reduced to two dimensions using t-stochastic neighbor embedding (t-SNE)(Maaten and Hinton, 2008), with the 16 principal components being used as input. Cell clusters were demarcated via fast search and find of density peaks (Rodriguez and Laio, 2014) using a gamma threshold of 1,500. Differentially expressed genes were identified for each cluster using the two-sided Student's *t*-test provided by the Seurat FindConservedMarkers function. Normalized gene expression within cells from a particular cluster was compared to expression within cells from all other clusters. Genes with an absolute log fold change less than 0.25 or those expressed in less than 10 percent of cells in either population being compared were excluded from analysis. Following cluster designation, a likelihood ratio test using a generalized linear model was performed to identify genes that vary by cluster. Genes with an adjusted *p*-value of 0.05, an average greater than the bottom quintile of averages, and a dispersion higher than what would be expected using the DESeq model(Anders and Huber, 2010) were then selected to construct a developmental trajectory using DDRTree(Qiu et al., 2017).

Bulk RNA sequencing was performed using sorted c-Kit+/Cd41+ pools of cells from KH2 or hMLL1 induced EB cells incubated with doxycycline from day 4 to day 6. Three separate differentiation experiments were performed with KH2 and hMLL1i differentiated in parallel. Illumina HiSeq libraries were prepared and sequenced by the Genomics and Microarray Core Facility at the University of Colorado Anschutz Medical Campus. Sequenced single-end reads were mapped to the mouse genome (mm10) by GSNAP, expression (FPKM) derived by Cufflinks2, and differential expression analyzed with ANOVA in R.

Gene ontology analysis of single cell RNA seq clusters. Differentially expressed genes (p-value < 0.05) in each cluster were ranked by fold change and top 50 genes were selected to perform gene ontology analysis on the PATHER website (http://www.pantherdb.org/) and confirmed with DAVID v6.8. Additional pathway analyses were performed using Ingenuity Pathway Analysis (Qiagen).

Antibodies used in this study

Western

Antigen	*Host species*	*Catalog*	*Company*
MLL1-C	rabbit polyclonal		in house
Nucleolin (C23)	rabbit polyclonal	sc-13057	Santa Cruz
Streptavidin-HRP	N/A	21130	Thermo Fisher
Wdr5	rabbit polyclonal	13105S	Cell Signaling
Menin	rabbit polyclonal	A300-105A	Bethyl
H3K4me1	rabbit polyclonal	ab8895	Abcam
H3K4me2	rabbit polyclonal	9725S	Cell Signaling
H3K4me3	rabbit polyclonal	ab8580	Abcam
pan-H3	rabbit polyclonal	ab1791	Abcam

Immunofluorescence

Antigen	*Conjugate*	*Catalog*	*Company*
Phalloidin	rhodamine	R415	Invitrogen

Flow Cytometry

Fluorochrome-antibody	*Clone*	*Company*
APC anti-mouse c-Kit	2B8	Biolegend
PE anti-mouse Cd41	MWReg30	Biolegend
FITC anti-mouse Cd41	MWReg30	Biolegend
PE anti-mouse Flk-1	Avas12	Biolegend
PE/Cy7 anti-mouse Pdgfra	APA5	Biolegend
PE anti-mouse c-Kit	2B8	Biolegend
APC anti-mouse Tie-2	TEK4	Biolegend
BV421 anti-mouse Cdh5	BV13	Biolegend
BV421 anti-mouse Cd45	30-F11	Biolegend
APC/Cy7 anti-mouse c-Kit	2B8	Biolegend

BV510 anti-mouse Cd41	MWReg30	Biolegend
FITC anti-mouse Cd34	RAM34	BD Pharmingen
PE/Cy7 anti-mouse Cd16/32	93	Biolegend
APC anti-mouse Mac-1	M1/70	Biolegend
PE anti-mouse Gr-1	R86-8C5	eBioscience
BV421 anti-mouse Sca-1	D7	Biolegend
PE anti-mouse Cd18	M18/2	Biolegend
APC anti-mouse Cd49d	R1-2	Biolegend
AF488 anti-mouse Cd11a	I21/7	Biolegend
Pacific Blue anti-mouse Cd29	HMb1-1	Biolegend

Primers used in this study

Genotyping

Col1-F	5'- TCCCTCACTTCTCATCCAGATATT
Col1-R	5'- AGTCTTGGATACTCCGTGACCATA
SApA-R	5'- GGACAGGATAAGTATGACATCATCAA

qRT-PCR

Gene	Forward primer	Reverse primer
hMLL1	CAGCCAGCCTCCAGTATCTC	TTCCCTTGCATAGGAGCAGT
total MLL1	GCATCTTCTGAGCCAGCAA	GAGGACCCCGGATTAAACAT
Brachyury	CCTCCCTTGTTGCCTTAGAGTAGTT	GCAGATTGTCTTTGGCTACTTTGTC
Flk1	CACCTGGCACTCTCCACCTTC	GATTTCATCCCACTACCGAAAG
Sox1	GCGATGCCAACTTTTGTATG	AGAGGGGATTGCGGTATAAA
Pax6	GTTCCCTGTCCTGTGGACTC	ACCGCCCTTGGTTAAAGTCT
Gata6	CTTGCGGGCTCTATATGAAACTCCAT	TAGAAGAAGAGGAAGTAGGAGTCATAGGGACA
Sox17	GGAGGGTCACCACTGCTTTA	AGATGTCTGGAGGTGCTGCT
Itga4	GAATCCAAACCAGACCTGCGA	TGACGTAGCAAATGCCAGTGG
Itgal	CCAGACTTTTGCTACTGGGAC	GCTTGTTCGGCAGTGATAGAG
Itgb2	CAGGAATGCACCAAGTACAAAGT	CCTGGTCCAGTGAAGTTCAGC
Rac1	ATGCAGGCCATCAAGTGTG	TAGGAGAGGGGACGCAATCT
Rac2	TGCAGGCCATCAAGTGTGTGGT	TAGAGCAGGCTGCAGGGGCGCTT
Rhoa	GCAGGTAGAGTTGGCTTTATGG	TTCTTGTTCCCAACCAGGATGA
Akt1	ATGAACGACGTAGCCATTGTG	TTGTAGCCAATAAAGGTGCCAT
Actb	ATGGATGACGATATCGCT	ATGAGGTAGTCTGTCAGG
Arp3	CAGGCTGAAGTTAAGCGAGGAG	CCTCCAAACCAGACTGCATACC
Pik3cd	ACCATCAGTGGCTCTGCGGTTT	GTGGTCTTCTGGGAACTCACCT
Myl12a/b	CACCATCCAGGAGGATTACC	CTTCAGGATGCGTGTGAACT

Supplemental References:

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology *11*, R106.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23-28.

Beckett, D., Kovaleva, E., and Schatz, P.J. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein science : a publication of the Protein Society *8*, 921-929.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology *36*, 411-420.

Hsieh, J.J., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Molecular and cellular biology *23*, 186-194.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell stem cell *1*, 324-337.

Maaten, L.v.d., and Hinton, G. (2008). Visualizing Data using t-SNE. Journal of Machine Learning Research *9*, 2579-2605.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature methods *14*, 979-982.

Rodriguez, A., and Laio, A. (2014). Machine learning. Clustering by fast search and find of density peaks. Science *344*, 1492-1496.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nature methods *9*, 676-682.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G.*, et al.* (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189-196.

Ugale, A., Norddahl, G.L., Wahlestedt, M., Sawen, P., Jaako, P., Pronk, C.J., Soneji, S., Cammenga, J., and Bryder, D. (2014). Hematopoietic stem cells are intrinsically protected against MLL-ENL-mediated transformation. Cell reports *9*, 1246-1255.

Yokomizo, T., and Dzierzak, E. (2010). Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. Development *137*, 3651-3661.

Zenzie-Gregory, B., O'Shea-Greenfield, A., and Smale, S.T. (1992). Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. The Journal of biological chemistry *267*, 2823-2830.