MYB bi-allelic targeting abrogates primitive clonogenic progenitors while the emergence of primitive blood cells is not affected

Zahir Shah,[1,2,3] Elena S. Philonenko,[1,2,4] Vasily Ramensky,[5,6] Chenyu Fan,[1,2,3] Cuihua Wang,[1,2] Hanif Ullah,[1,2,3] Baoyun Zhang,[1,2] Pavel Volchkov[5] and Igor M. Samokhvalov[1,2]

[1]Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Science, Guangzhou, China; [2]Key Laboratory of Regenerative Biology, Chinese Academy of Science, Guangdong Provincial Key Laboratory of Stem Cells and Regenerative Medicine, Guangzhou, China; [3]University of Chinese Academy of Science, Beijing, China; [4]Vavilov Institute of General Genetics, Russian Academy of Science, Moscow, Russia; [5]Moscow Institute of Physics and Technology, Dolgoprudny, Moscow, Russia and [6]National Medical Research Center for Preventive Medicine, the Ministry of Healthcare of the Russian Federation, Moscow, Russia

 doi:10.3324/haematol.2020.249193

Received: February 4, 2020.

Accepted: July 28, 2020.

Pre-published: July 30, 2020.

Correspondence: *IGOR M. SAMOKHVALOV* - igor@gibh.ac.cn

Supplementary Appendix

MYB bi-allelic targeting abrogates primitive clonogenic progenitors while the emergence of primitive blood cells is not affected

Zahir Shah[1,2,3], Elena S. Philonenko[1,2,4], Vasily Ramensky[5,6], Chenyu Fan[1,2,3], Cuihua Wang[1,2], Hanif Ullah[1,2,3], Baoyun Zhang[1,2], Pavel Volchkov[5], and Igor M. Samokhvalov[1,2,*]

[1]Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China; [2]Key Laboratory of Regenerative Biology, Chinese Academy of Sciences; Guangdong Provincial Key Laboratory of Stem Cells and Regenerative Medicine, Guangzhou 510530, China; [3]University of Chinese Academy of Sciences, Beijing 100049, China; [4]Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow 119333, Russia; [5]Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region 141701, Russia; [6]National Medical Research Center for Preventive Medicine, the Ministry of Healthcare of the Russian Federation, Moscow 101990, Russia;

Supplemental materials and methods

Cell lines and hematopoietic differentiation of hESC lines

The human ESC line used in this study was H1 (NIH code WA01). The cell line was grown and passaged in the absence of antibiotics, and regularly tested for the presence of mycoplasma contamination. The cells were kept undifferentiated on Matrigel-coated plates (Corning Matrigel, Growth Factor Reduced, #354230) in mTeSR1 medium (STEMCELL Technologies, Vancouver, Canada), and the culture medium was supplemented with 1 μM Thiazovivin (GIBH CAS, China) for improved cell recovery during the first 24 hours post-passage. Before transfection and hematopoietic differentiation, hESCs were subjected to at least three short passages (2-3 days) at a seeding density of 4-6 × 10^5 cells per one well of a standard 6-well plate. The cells were typically passaged at 70-80% confluence using TrypLE solution (Life Technologies, Carlsbad, CA) at 37^0C and minimal digestion time.

To initiate hematopoietic development, hESCs were subjected to a modified planar differentiation protocol in which embryoid bodies (EBs) briefly formed in AggreWell^{TM}400

plates (STEMCELL Technologies) were allowed to attach under normal gravity to surfaces that were coated with extracellular matrix proteins. The attachment was accomplished during 48 hours in mTeSR1 medium supplemented with 10 µM Thiazovivin (GIBH CAS), 2 ng/mL BMP4, and 50 ng/mL hVEGF165 (both PeproTech, Rocky Hill, NJ). After the EB attachment was accomplished, the differentiation was continued in a Stemline®II SFM (Sigma-Aldrich, St. Louis, MO) – based medium supplemented with hVEGF165 (PeproTech). Attached EBs formed a confluent layer of differentiating cells during the next several days. All cells were incubated at 37^0C in a mixture of 5% carbon dioxide and 95% air in a humidified atmosphere. To stimulate the primitive hematopoietic progenitors, Activin A (PeproTech) was added to a final concentration of 1 ng/mL on Day 0 - Day 2 of differentiation. For the Activin A experiments, we increased the EB plating density two-fold to three-fold to strengthen the effect. To suppress primitive hematopoiesis, SB-431542 (TOCRIS, Bio-Techne, Minneapolis, MN) was added to Day 2 - Day 4 hESC differentiating cultures at a final concentration of 2 to 6 µM.

TALEN design and assembly

Six candidate TALEN pair sequences were predicted using TAL Effector Nucleotide Targeter 2.0 tool, https://tale-nt.cac.cornell.edu/node/add/talen. The TALENs were assembled using the Golden Gate TALEN[1] and TAL Effector Kit 2.0 (Kit #1000000024) purchased from the Addgene repository, https://www.addgene.org/taleffector/goldengatev2/. To check the efficiency of designed TALENs, the *MYB* recognition sequences for candidate TALENs were cloned into the eGFP rescue reporter vector[2]. The resulting reporter construct was co-transfected with two candidate TALEN plasmids into HEK293 cells, and the restoration of the eGFP gene due to TALEN-induced internal homologous recombination was measured by flow cytometry and fluorescence microscopy. The best pair of TALENs were selected based on the highest level of induced eGFP fluorescence.

Generation of targeted hESC lines

Human ESCs were transfected by electroporation with the linearized targeting construct and a pair of selected TALEN plasmids. In a typical transfection experiment, 1×10^6 cells were electroporated with 10 µg of the linearized targeting construct and 10 µg of each TALEN plasmid in mTeSR1 medium containing 1 mM Thiazovivin in two consecutive pulses at 240 V

and 500 µF. After electroporation, the cells were seeded on Matrigel-coated 6-well plate in mTeSR1 + 1 µM Thiazovivin at a density of 2.5×10^5 cells/well. The medium was changed daily and drug selection was performed with 500 µg/mL of puromycin after 48-72 hours of post-electroporation recovery. To control the acquisition of the drug resistance, the same procedure was performed in the same conditions without the targeting construct. Typically, the drug selection continued for 3-4 days until all cells perished in a control dish and individual puromycin-resistant colonies appeared in the targeted sample. The colonies were grown until they reached the size of about 1000 cells/colony. Individual colonies were picked up under a microscope in an aseptic environment, dissociated briefly by TrypLE, and further cultured in individually-labeled Matrigel-coated wells of 24-well plates in mTeSR1 + 1 µM Thiazovivin. Genomic DNA of the targeted hESC clones was isolated by phenol-chloroform extraction and used for PCR pre-screening with optimized primer pairs. The PCR-positive clones were further analyzed by Southern blot hybridization.

Southern blot hybridization

Southern blotting was carried out according to a protocol supplied by the manufacturer of the Nylon$^+$ membrane (BrightStar^{TM}-Plus, Invitrogen by Thermo Fisher Scientific, Waltham, MA) with slight modifications. The DNA probe labeling and chemiluminescent signal development were carried out with the use of North2South Biotin Random Prime DNA Labeling Kit (Thermo Fisher Scientific, Cat. No. 17075) according to the manufacturer's recommendations. Hybridization was performed in the Church-Gilbert buffer at 65^0C. The images were captured by ChemiScope 5300 (CLiNX, China).

Karyotyping

Karyotyping by G-banding was performed according to the standard protocol. Forty randomly selected metaphase spreads for each cell line were counted and stained using standard G-banding analysis at 400-band resolution.

Flow cytometry, cell sorting, and progenitor analysis

For the flow cytometry and cell sorting, we used anti-human monoclonal antibodies from Becton Dickinson (BD Life Sciences, Franklin Lakes, NJ). All cell sorting procedures were performed

either on BD FACSAria II or MoFlo Astrios (Beckman Coulter, Brea, CA) machines. The flow cytometry analyses were done on BD Accuri C6 Plus and BD LSRFortessa. The data were analyzed with FlowJo V10 (FlowJo LLC, BD) and BD Accuri C6 Plus software. For the flow cytometry analysis, cells were harvested using the standard TrypLE procedure, resuspended in the Basal Hematopoietic Medium (StemLine®II, 1 × Glutamax™, 1 × 2-ME, 1 × NEAA), and resulting cell suspension was incubated in CO_2-incubator for at least 1 hour to restore cell surface antigens. Next, the cells were spun down and resuspended in the cold FACS Buffer (1 × D-PBS – Ca – Mg, 5% FCS, 20 mM HEPES pH 7.2-7.5) containing 5% normal human serum at a density of 1×10^6 cells per 100 µL and pre-incubated on ice for at least 10 min. Typically, cells were incubated with antibodies on ice for 20–30 min in the dark. The unbound antibodies were washed twice with 1mL of the cold FACS Buffer. For gating out dead and apoptotic cells the washed cell suspensions were incubated with 7-AAD (BioLegend, San Diego California) for analysis by BD Accuri C6 Plus or DAPI for the BD LSRFortessa cytometry and sorting. For the cell sorting, washed cells were resuspended in the FACS Sorting Buffer (1 × HBSS, 2% BSA, 25 mM HEPES pH7.2-7.5, 1 mM EDTA), filtered through 40 µm Cell Strainer, sorted into the Basal Hematopoietic Medium and washed twice by DMEM/F12 medium before downstream applications.

Hematopoietic progenitor assay was performed in the serum-free methylcellulose medium SF H4436 (STEMCELL Technologies) according to the manufacturer's recommendations in duplicates for at least two different cell densities of each input cell population. The colonies were grown for 16-18 days at 37^0C in the 5% CO_2 atmosphere. The hematopoietic progenitors were retrospectively classified into different types of Colony Forming Unit - Cells (CFU-Cs): BFU-E^P (Burst Forming Unit – Erythroid, Primitive), CFU-E^P (Colony Forming Unit – Erythroid, Primitive), CFU-Myeloid (Colony Forming Unit – Macrophage, Granulocyte, and Granulocyte-Macrophage), CFU-Mix^P (Colony Forming Unit – Mixed lineage, Primitive). May-Grünwald histochemical staining of individual colonies was performed to confirm the colony identity.

MACS and OP9-DL4 co-culture for T lineage differentiation

On Day 13 of the primary hematopoietic differentiation, SB (2 µM) - treated cells were harvested with TrypLE and $CD45^+CD34^{low/-}$ or $CD34^+CD45^{low/-}$ cells were isolated using

magnetic-activated cell sorting (MACS) with CD45 or CD34 MicroBeads, human, (Miltenyi Biotec, Auburn, CA) according to the manufacturer's protocol.

To generate stromal cells expressing DL4, low-passage OP9 cells were transfected with a PiggyBac transposon vector bearing human *DLL4* cDNA under the EF-1α promoter and a *Puro*^{R} gene cassette. Puromycin-resistant OP9-DL4 cells were selected and maintained in α-MEM supplemented with 20% FCS (Gibco^{TM}, Life Technologies). A total of 2–5 × 10^5 cells of MACS-enriched WT hESC- or DKO hESC-derived fractions ($CD45^+CD34^{low/-}$ and $CD34^+CD45^{low/-}$) were added to the individual well of a 6-well plate containing OP9-DL4 cells and cultured in the OP9 differentiation medium (α-MEM, 20% FBS, and 50 μM 2-mercaptoethanol) supplemented with rhSCF (5 ng/mL), rhIL-7 (5 ng/mL), and rhFLT3-L (5 ng/mL) (PeproTech). Every 7 days, differentiating human cells were placed on the fresh OP-DL4 stroma.

Immunocytofluorescence

For immunocytochemical studies, cells were fixed in 4% paraformaldehyde for 20 min at room temperature. Fixed cells were incubated for 1 hour in the Blocking Buffer (1 × D-PBS, 0.2% Triton X 100, 0.1% Tween-20, 2% goat or donkey serum, and 2% FBS). Primary antibodies were diluted in 1 × D-PBS, 0.1% Tween-20, and applied for 1–3 hours at room temperature or overnight at 4^0C. Slides were washed 3 × 3 min with 1 × D-PBS – 0.1% Tween-20 solution, and incubated with Alexa Fluor 546- or Alexa Fluor 488-conjugated either goat anti-mouse or goat anti-rabbit or donkey anti-goat (all from Invitrogen) secondary antibodies used at a dilution of 1:1000. Nuclei were stained with DAPI (Sigma-Aldrich). Slides were mounted in the Vectashield® mounting medium (Vector Laboratories, Burlingame, CA) and examined under the Confocal Microscope LSM800 (Carl Zeiss AG, Oberkochen, Germany).

Western blotting

Total protein was extracted from the differentiated cells using RIPA buffer (Beyotime, China, P001B). The lysate was boiled in an SDS loading buffer, resolved by SDS-PAGE through 12% polyacrylamide gel, and transferred to polyvinylidene difluoride (PVDF) membrane. The membrane was incubated overnight at 4°C with monoclonal mouse anti-MYB primary antibodies (1:100 dilution**,** Santa Cruz Biotechnology, Santa Cruz, CA, sc-74512**;** or 1:100

dilution, Millipore, Burlington, MA, #05-175), and rabbit anti-actin (1:1000 dilution; Beyotime, AA1281) after blocking with the TBST buffer (1 × Tris-buffered Saline, 0.1% Tween 20) containing 5% nonfat milk at room temperature for 2-3 hours. The membrane was washed three times with the TBST buffer and incubated with goat anti-mouse (1:1000 dilution, Beyotime, A0216) or goat anti-rabbit secondary antibodies (1:1000 dilution, Beyotime, A0208). After incubation at room temperature for 50 min, the membrane was washed 3 × 10 min in the TBST buffer and the bands were developed with the ECL reagents (Beyotime, P0018).

May-Grünwald staining

Sorted cells ($0.5\text{-}1.0 \times 10^5$) were washed and resuspended in 100 μL 1 × D-PBS. Cells were spun onto polylysine-coated slides at 500 rpm for 5 min, air-dried, and stained with May-Grünwald stain (Sigma-Aldrich, MG500) according to the manufacturer's protocol. The slides were washed with dH_2O, air-dried, and mounted for examination by light microscopy.

Real-time RT-PCR

Total RNA was extracted using PureLink™ RNA Mini Kit (Invitrogen, Thermo Fisher Scientific), and reverse transcribed using SuperScript™ IV First-Strand Synthesis System (Thermo Fisher Scientific) according to the protocol supplied by the manufacturer. Quantitative RT–PCR was performed using SYBR® Premix Ex Taq™ on Agilent MX3000P qPCR System (Agilent Technologies, Santa Clara, CA) or Bio-Rad CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad, CA, USA). The cycling conditions were: 15 min at 95^0C, then 40 cycles at 95^0C for 10 sec, and 60^0C for 30 sec, followed by a finishing stage, 95^0C for 60 sec, 55^0C for 30 sec, and 95^0C for 30 sec. A dissociation/melting curve was generated for every PCR product. ROX referencing dye was added for the normalization of fluorescence intensity of qPCR reporter dye (SYBR Green). All experiments were performed in triplicates and the gene expression was normalized against *ACTB* for each experiment. The administration and analysis software, MxPro qPCR or CFX manager, were used to run the Real-time PCR machines and to draw analysis diagrams.

Primers for qRT-PCR were:

cmybF 5’-GCTTGGACAGAAAGAAAAGAAACT-3’

cmybR 5’-ACAAAAAAACCACAATAAAAAATG-3’

hActin-RT-F 5’-CCCAGAGCAAGAGAGG-3’

hActin-RT-R 5’-GTCCAGACGCAGGATG-3’

HBB-F 5’-TCTGTCCACTCCTGATGCTGTTATG-3’

HBB-R 5’-CCGAGCACTTTCTTGCCATGA-3’

HBG1-F 5’-TCACAGAGGAGGACAAGGCTACTAT-3’

HBG1-R 5’-CCTATCCTTGAAAGCTCTGAATCAT-3’

HBE1-F 5’-ATGGTGCATTTTACTGCTGAGG-3’

HBE1-R 5’-GGGAGACGACAGGTTTCCAAA-3’

RNA-sequencing library preparation and data analysis

Around 1-2 × 10^6 cells were sorted in the FACS Sorting Buffer on Day 6 or Day 12 of regular differentiation using FACSAria II cell sorter (BD Biosciences). Sorted cells were lysed with the Trizol reagent and stored at -80ºC. A total of 31 samples (biological repeats) were sequenced with a HiSeq2500 platform (Illumina, San Diego, CA) for single-end reads of length 50 bp, and HiSeq3000 for paired-end reads of length 150 bp at the RiboBio genomic sequencing laboratory (www.ribobio.com), Guangzhou, China. The sequenced samples were checked with the FastQC v0.11.5 software, trimmed, and filtered with the BBduk tool from the BBmap package. Trimmed reads were mapped to the GRCh38 human genome with Gencode v.24 gene annotation[3] using the --quantMode GeneCounts option of the STAR[4] v.2.5.3ab software. Approximately 15,000 pseudogenes and 552 rRNA genes were excluded from further analysis. The sample clustering was checked with the PlotMDS function of the edgeR software package[5] as described in the edgeR manual. EdgeR was used to convert raw read per gene counts to CPM (counts per million reads), RPKM (reads per kilobase per million reads), and Log_2RPKM matrices. The gene ontology analysis was performed using Metascape[6] (www.metascape.org) and default accumulative hypergeometric statistical test. All bar plots representing the mRNA expression

were generated with GraphPad Prism 5.0 plus and represented as (mean ± SEM). P-values were calculated using one-way ANOVA with post-Tukey's test for multiple comparisons.

Statistical analysis

All data presented in this paper is from at least 3 independent biological repeats except for the day 6 RNA-seq data of 2 biological repeats. All graphs were plotted using the GraphPad Prism and R packages. Student's t-test was used in comparing data between two groups of variables to test for significance. Data are the mean ± SD ($n \geq 3$ biological replicates with 3 technical replicates each). One-way ANOVA analysis with Tukey's test for multiple comparisons was applied to more than 2 groups of variables, where ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$. The ANOVA data are presented as the mean ± SEM ($n = 3$ biological replicates with 3 technical replicates each).

Data availability

The RNA-seq data described in this study were deposited to NCBI Gene Expression Omnibus with the gene accession number GEO: GSE125761.

Supplementary References

1. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39(12):e82.

2. Ma N, Liao B, Zhang H, Wang L, Shan Y, Xue Y, et al. Transcription activator effector nuclease (TALEN)-mediated gene correction in integration-free β-thalassemia induced pluripotent stem cells. J Biol Chem. 2013;288(48):34671-79.

3. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for the ENCODE project. Genome Res. 2012;22(9):1760-74.

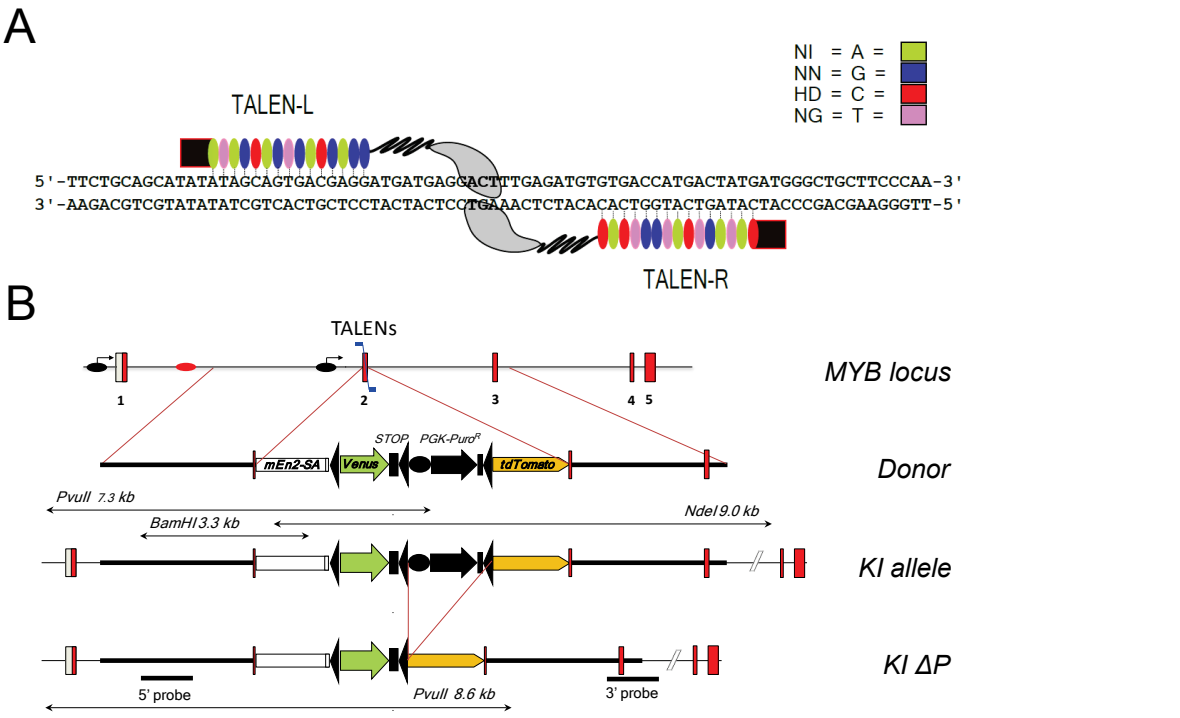
4. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15-21.

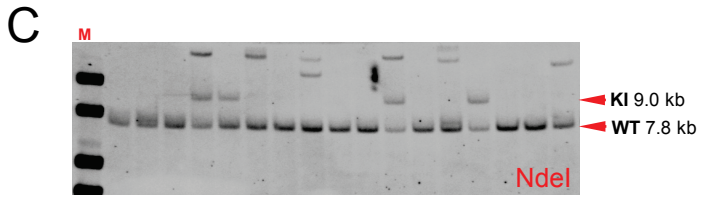
5. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26(1):139-140.

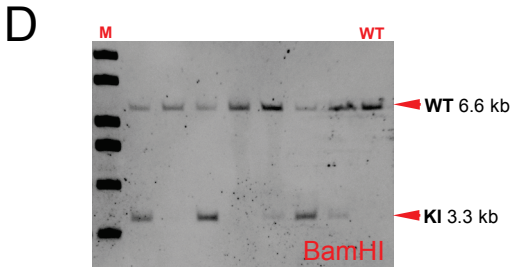
6. McCarthy DJ, Chen Y, Smyth GK Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40(10):4288-97.

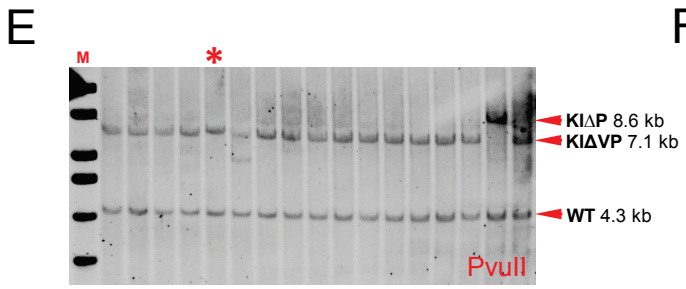
7. Tripathi AK, Singh K, Pareek A, Singla-Pareek SL. Histone chaperones in Arabidopsis and rice: genome-wide identification, phylogeny, architecture, and transcriptional regulation. BMC Plant Biol. 2015;15:42.

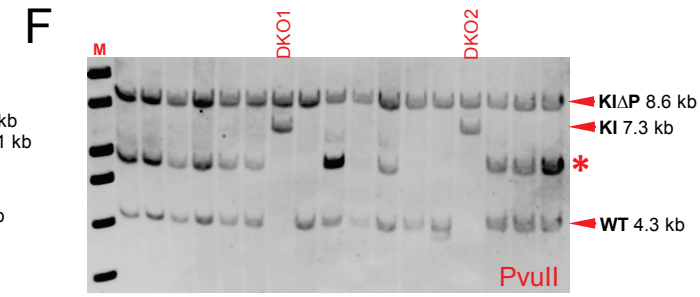
Figure S1


A
NI = A
NN = G
HD = C
NG = T
TALEN-L
5'-TTCTGCAGCATATATAGCAGTGACGAGGATGATGAGGACTTTGAGATGTGTGACCATGACTATGATGGGCTGCTTCCCAA-3'
3'-AAGACGTCGTATATATCGTCACTGCTCCTACTACTCCTGAAACTCTACACACTGGTACTGATACTACCCGACGAAGGGTT-5'
TALEN-R
B
TALENs
MYB locus
1
2
3
4 5
STOP
PGK-Puro^R
mEn2-SA
Venus
tdTomato
Donor
PvuII 7.3 kb
NdeI 9.0 kb
BamHI 3.3 kb
KI allele
KI ΔP
5' probe
PvuII 8.6 kb
3' probe


C
M
KI 9.0 kb
WT 7.8 kb
NdeI


D
M
WT
WT 6.6 kb
KI 3.3 kb
BamHI


E
M
*
KIΔP 8.6 kb
KIΔVP 7.1 kb
WT 4.3 kb
PvuII


F
M
DKO1
DKO2
KIΔP 8.6 kb
KI 7.3 kb
*
WT 4.3 kb
PvuII

Supplementary Figure 1. TALEN design, MYB targeting strategy, and Southern blot analysis of the targeted hESC clones.

(A) Scheme of MYB TALENs and their recognition sequences in exon 2 of the MYB gene. TALE repeat domains are colored to indicate the identity of the RVD; the RVD code is shown.

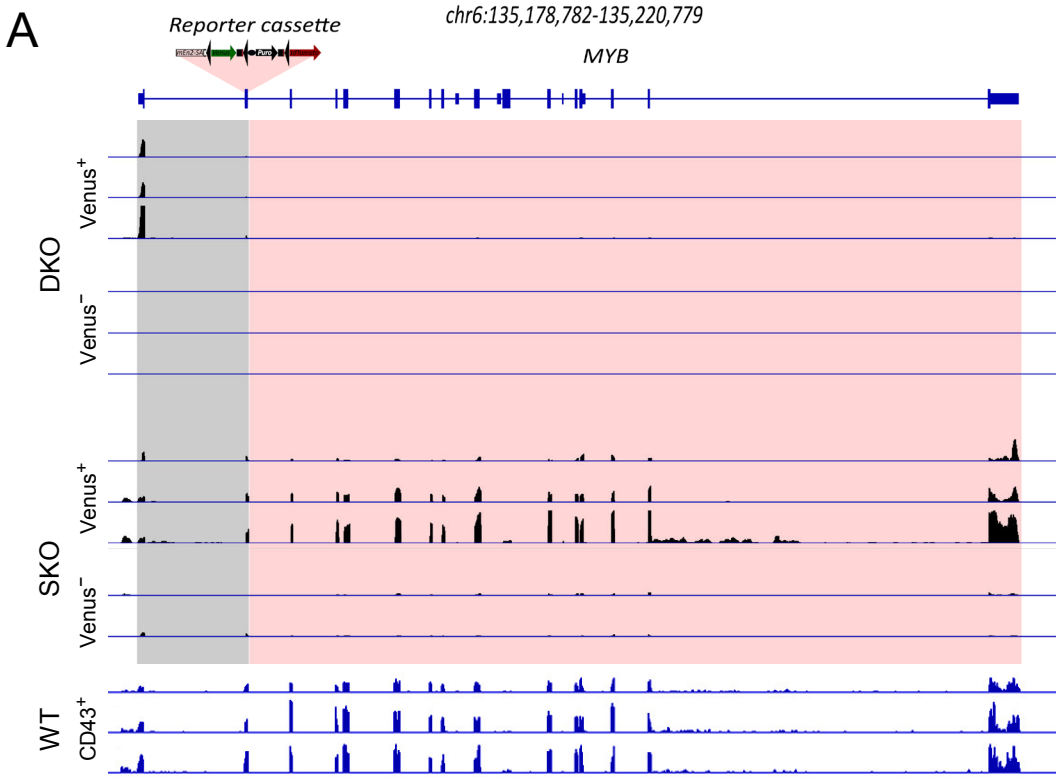
(B) Schematics of the *MYB* targeting. MYB exons are shown as red rectangles, *MYB* promoters – as black ovals with arrows. The red oval represents the *MYB* intronic transcription attenuation site. Positions of two DNA probes for Southern blot hybridization and the size of the indicative DNA fragments are shown along the targeted alleles. *LoxP* sites are shown as black triangles. KIΔP denotes the targeted allele which sustained the Cre-deletion of the *PGK-Puro*^{R} gene cassette.

(C and D) Selection of the properly targeted H1 hESC clones by Southern blot hybridization with two DNA probes and two restriction digests that are shown in (B). Here and in other hybridization panels, the restriction endonucleases used for digestion of genomic DNA are shown in the lower right corner of each panel; the sizes (in kilobases) of the indicative DNA fragments that hybridize with either 5' or 3' probe are shown next to the red arrowheads. Here and in other Southern panels, M = molecular-weight size marker; WT = genomic DNA extracted from unmodified H1 hESCs.

(E) Selection of the *MYB*-targeted hESC clones that sustained the Cre-deletion of the *PGK-Puro*^{R} gene cassette (the KIΔP allele). KIΔVP designates a targeted *MYB* allele in which both the *Venus* gene and the *PGK-Puro*^{R} gene cassette were deleted. The red asterisk indicates a clone that sustained the deletion of the *Venus* gene only.

(F) Southern hybridization screening of hESC clones with bi-allelic targeting of *MYB*. The red asterisk denotes a DNA fragment that corresponds to a major off-target or an internal cis-recombination event. Clones containing both KI and KIΔP alleles were selected for phenotyping.

Figure S2

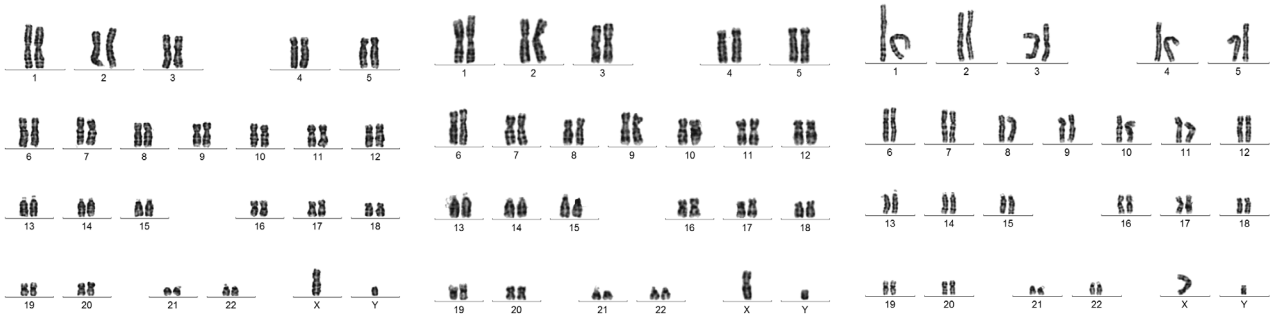


B

H1 MYB$^{+/-}$ clone SKO1

H1 MYB$^{-/-}$ clone DKO1

H1 MYB$^{-/-}$ clone DKO2

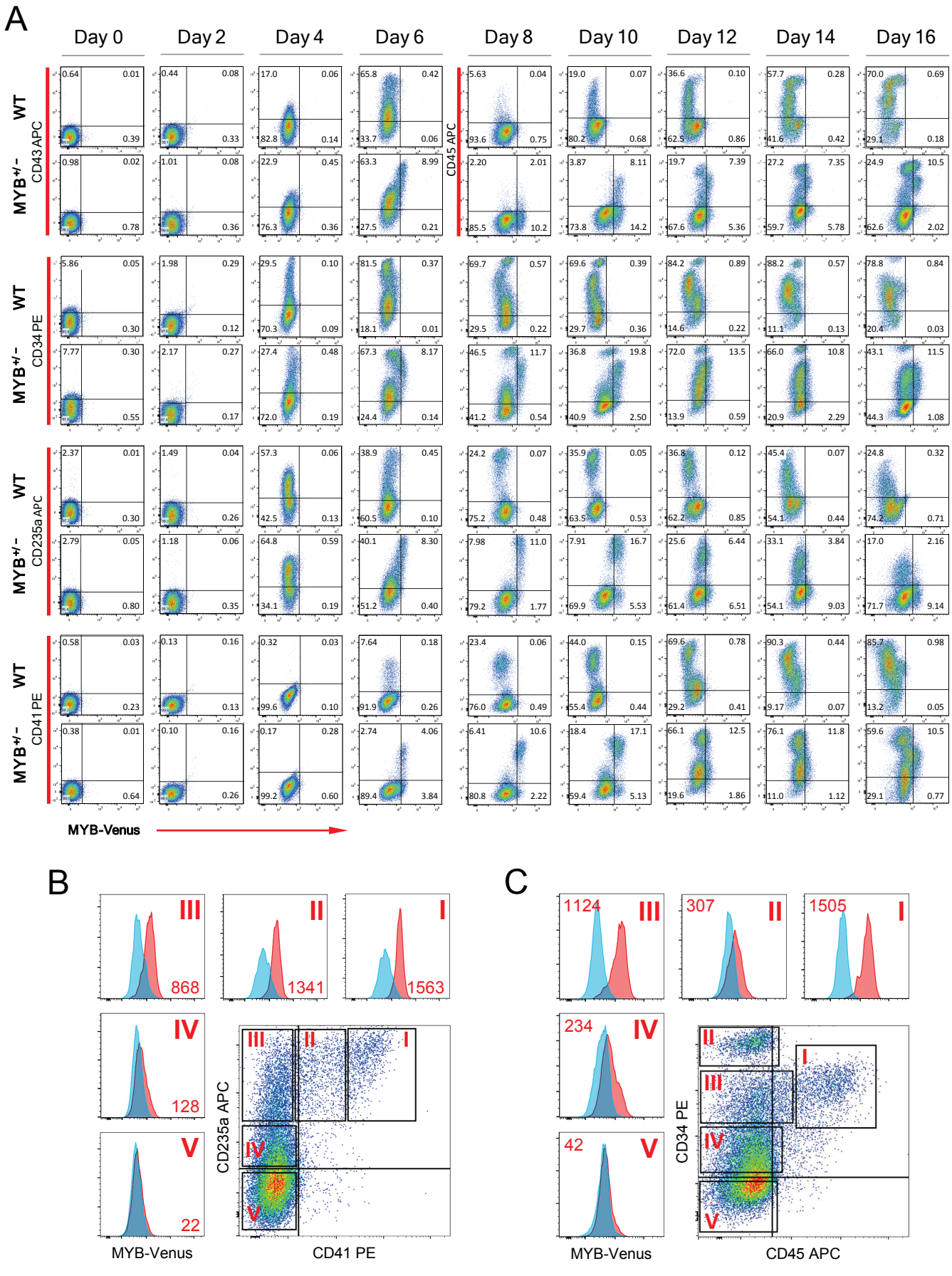


Supplementary Figure 2. Insertion of the reporter gene cassette into MYB locus arrests MYB transcription at exon 2 producing mutant cell lines are karyotypically normal.

(A) Exonal expression of MYB in H1-isogenic WT and mutant cell lines by RNA-seq transcriptome profiling. Shown are the data of several biological repeats for designated sorted cell populations.

(B) Karyotyping of mutant hESC lines with single knock-out (SKO1) and double knock-out (DKO1 and DKO2) of MYB.

Figure S3


A
Day 0
Day 2
Day 4
Day 6
Day 8
Day 10
Day 12
Day 14
Day 16
WT
MYB+/-
CD43 APC
CD45 APC
CD34 PE
CD235a APC
CD41 PE
MYB-Venus
B
III
II
I
868
1341
1563
IV
V
128
22
CD235a APC
CD41 PE
MYB-Venus
C
1124
307
1505
234
42
CD34 PE
CD45 APC
MYB-Venus

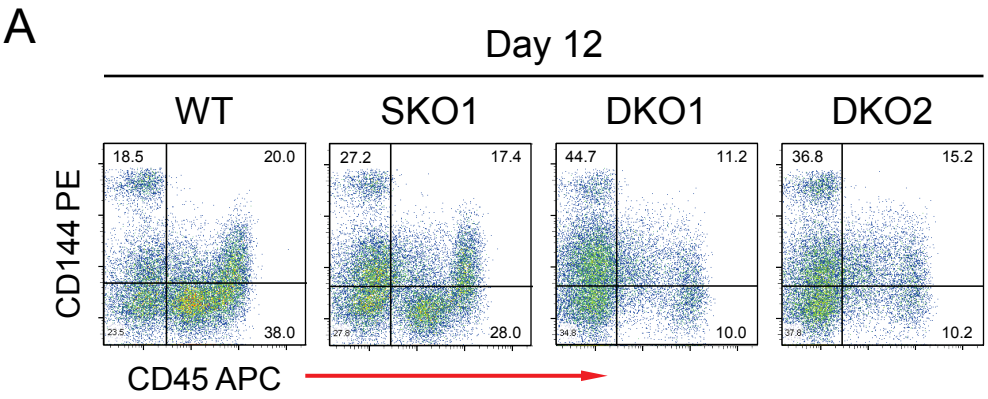
Supplementary Figure 3. Expression of the MYB-Venus reporter during hematopoietic differentiation of SKO hESCs.

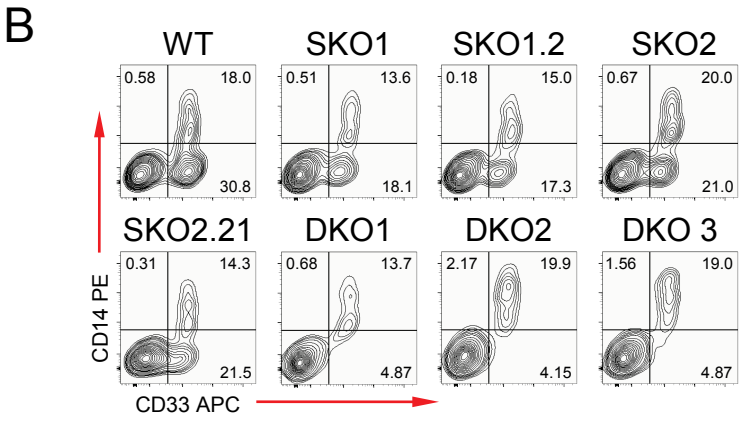
(A) Representative flow cytometry of the MYB-Venus expression during the course of hematopoietic differentiation of the SKO1 cell line. Total live cells were used for the analysis. Numbers in the dot plots here and elsewhere represent the percentages of cells within the corresponding quadrants.

(B) Representative analysis of MYB-Venus fluorescence across CD235a/CD41a – stained SKO1 total live cells on Day 6 of hematopoietic differentiation. Here and in (C), the correspondence between gated cell populations and MYB-Venus fluorescence intensity histograms is denoted by red Roman numerals. Bluish histograms show the fluorescence of WT H1 hESCs, the red ones – of the H1-isogenic SKO1 cell line. The ΔMFI values (in red) are shown in either lower right or upper left corners of the histograms.

(C) Representative analysis of MYB-Venus fluorescence across CD34/CD45 – stained SKO1 total live cells on Day 10 of hematopoietic differentiation. The MYB-Venus expression is drastically reduced in $CD34^{high}CD45^{-}$ endothelial and definitive hematopoietic precursor cells.

Figure S4


A
Day 12
WT
SKO1
DKO1
DKO2
18.5
20.0
38.0
27.2
17.4
28.0
44.7
11.2
10.0
36.8
15.2
10.2
CD144 PE
CD45 APC

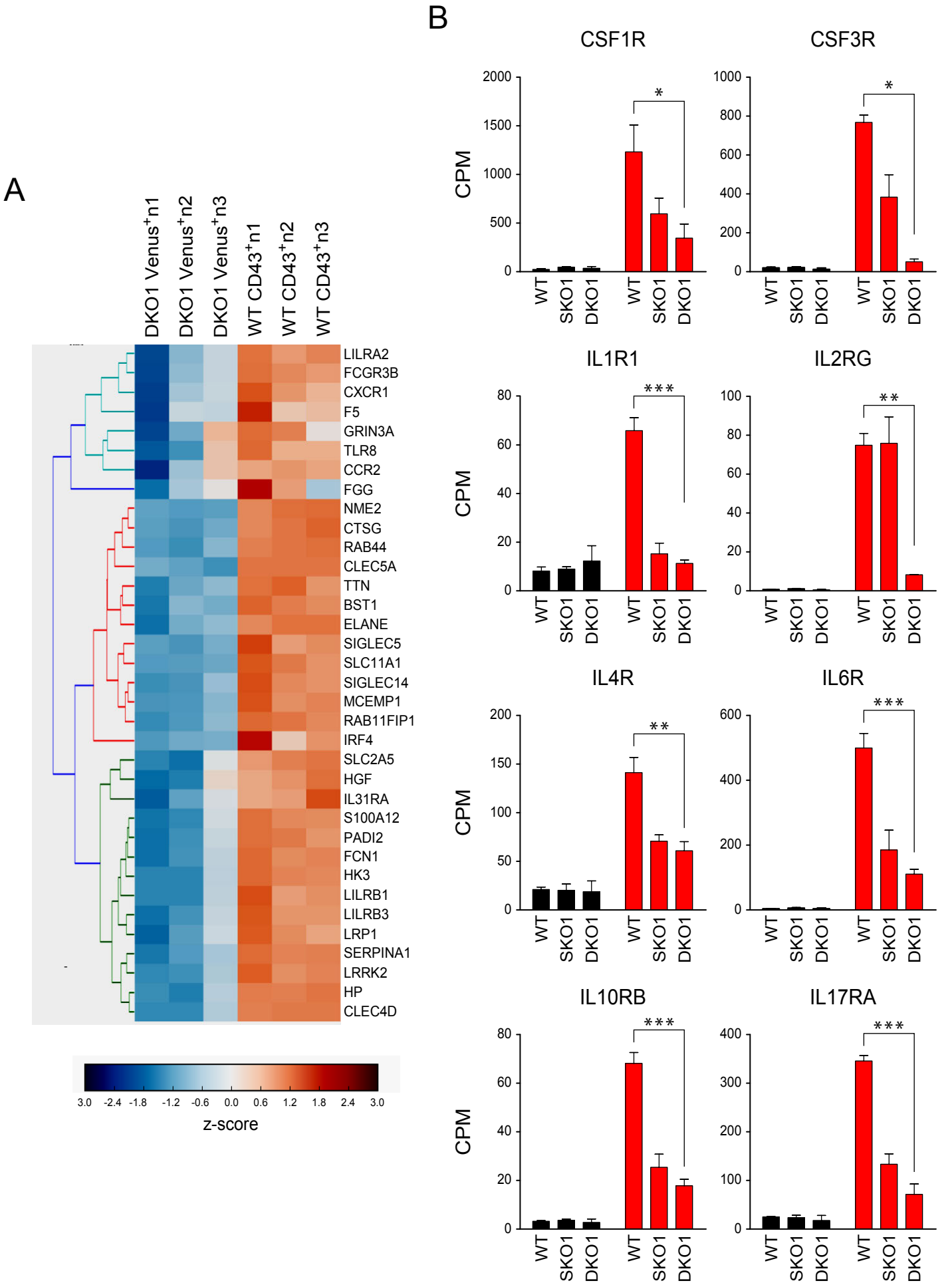

B
WT
SKO1
SKO1.2
SKO2
0.58
18.0
30.8
0.51
13.6
18.1
0.18
15.0
17.3
0.67
20.0
21.0
SKO2.21
DKO1
DKO2
DKO 3
0.31
14.3
21.5
0.68
13.7
4.87
2.17
19.9
4.15
1.56
19.0
4.87
CD14 PE
CD33 APC

Supplementary Figure 4. Defects in hESC-derived hematopoiesis caused by MYB gene targeting

(A) Representative flow cytometry of phenotypical changes in blood cell populations of the three MYB genotypes on Day 12 of differentiation.

(B) $CD14^{-}CD33^{+}$ immature granulocytes fail to emerge upon the MYB inactivation in hESCs. This cell population is also affected by MYB haploinsufficiency in the differentiated SKO cells. Total live cells were used for the analysis. Representative flow cytometry data of four independent experiments are shown.

Figure S5

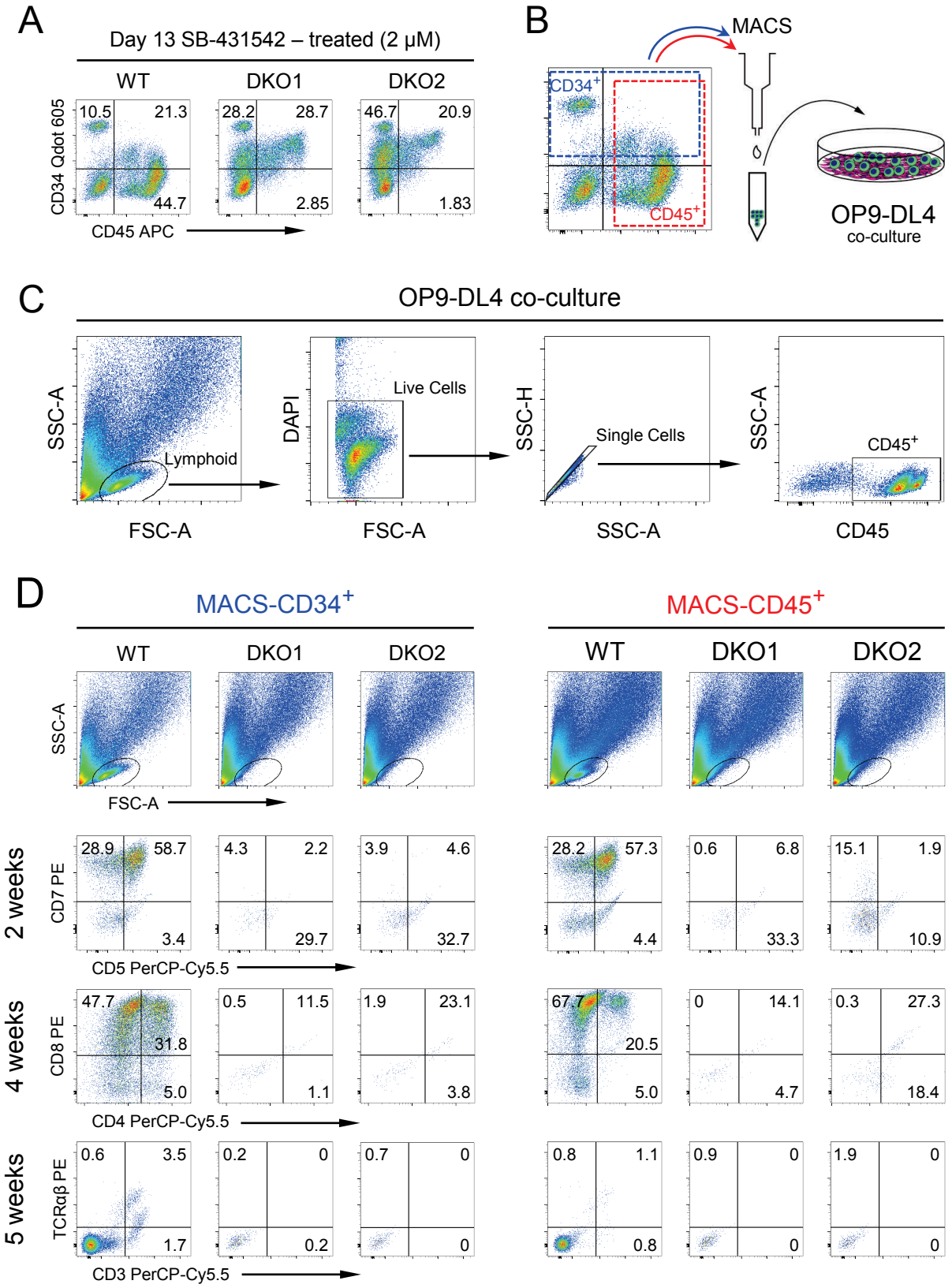

A
DKO1 Venus+n1
DKO1 Venus+n2
DKO1 Venus+n3
WT CD43+n1
WT CD43+n2
WT CD43+n3
LILRA2
FCGR3B
CXCR1
F5
GRIN3A
TLR8
CCR2
FGG
NME2
CTSG
RAB44
CLEC5A
TTN
BST1
ELANE
SIGLEC5
SLC11A1
SIGLEC14
MCEMP1
RAB11FIP1
IRF4
SLC2A5
HGF
IL31RA
S100A12
PADI2
FCN1
HK3
LILRB1
LILRB3
LRP1
SERPINA1
LRRK2
HP
CLEC4D
3.0 -2.4 -1.8 -1.2 -0.6 0.0 0.6 1.2 1.8 2.4 3.0
z-score
B
CSF1R
CSF3R
IL1R1
IL2RG
IL4R
IL6R
IL10RB
IL17RA
CPM
WT
SKO1
DKO1

Supplementary Figure 5. The expression of innate immunity genes is downregulated by the targeted inactivation of MYB.

(A) The heatmap shows the differential gene expression in DKO1 Venus^{+} versus WT CD43^{+} cells on Day 12 of hematopoietic differentiation. The 37 genes shown in the heatmap are from the GO term cluster "myeloid leukocytes activation" selected in the meta-enrichment analysis (Figure 4A). Gene expression levels are normalized by Z-score transformation across the RNA-seq experiments, with three independent biological repeats for each cell population. Of note, several prominent granulocyte/neutrophil genes (FCGR3B, CTSG, ELANE, SIGLEC5, S100A12, HK3, HP) are downregulated in MYB-null cells already on Day 12 of differentiation.

(B) MYB deficiency leads to selective downregulation of cytokine receptor gene expression. The CPM bar plots show the relative expression of several cytokine receptor genes in the differentiated cells of the three MYB genotypes. The mRNA levels of the cytokines receptors in MYB-Venus^{+} DKO1/SKO1 and CD43^{+} WT cell populations were measured on Day 6 (black bars) and Day 12 (red bars) by RNA sequencing. Data represent three independent biological repeats as (mean ± SEM), p-values were calculated using one-way ANOVA with Tukey's multiple comparison test, where ***p<0.001, **p<0.01 and *p<0.05.

Figure S6


A
Day 13 SB-431542 – treated (2 μM)
WT
DKO1
DKO2
CD34 Qdot 605
CD45 APC
10.5
21.3
44.7
28.2
28.7
2.85
46.7
20.9
1.83
B
MACS
CD34+
CD45+
OP9-DL4
co-culture
C
OP9-DL4 co-culture
SSC-A
Lymphoid
FSC-A
DAPI
Live Cells
FSC-A
SSC-H
Single Cells
SSC-A
SSC-A
CD45+
CD45
D
MACS-CD34+
MACS-CD45+
WT
DKO1
DKO2
SSC-A
FSC-A
2 weeks
CD7 PE
CD5 PerCP-Cy5.5
28.9
58.7
3.4
4.3
2.2
29.7
3.9
4.6
32.7
28.2
57.3
4.4
0.6
6.8
33.3
15.1
1.9
10.9
4 weeks
CD8 PE
CD4 PerCP-Cy5.5
47.7
31.8
5.0
0.5
11.5
1.1
1.9
23.1
3.8
67.7
20.5
5.0
0
14.1
4.7
0.3
27.3
18.4
5 weeks
TCRαβ PE
CD3 PerCP-Cy5.5
0.6
3.5
1.7
0.2
0
0.2
0.7
0
0
0.8
1.1
0.8
0.9
0
0
1.9
0
0

Supplementary Figure 6. MYB is required for the development of human T cells.

(A) Representative flow cytometry analysis of WT versus DKO cells at Day 13 of differentiation. To increase the number of definitive hematopoietic cells, the cell lines were treated by SB-431542 (at a final concentration of 2 μM) on Day 2 – Day4 of differentiation.

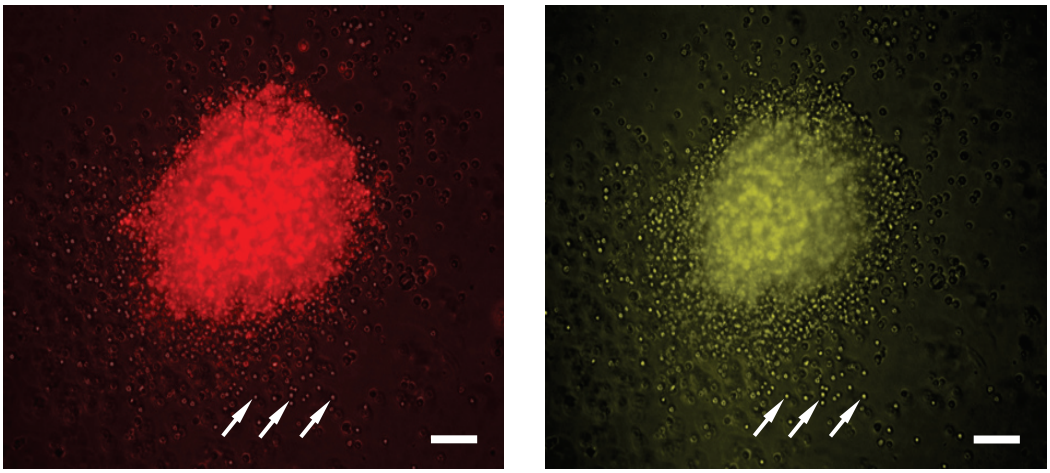
(B) Scheme of the T cell potential assay. MACS has been used to enrich $CD34^{+}CD45^{low/-}$ and $CD45^{+}CD34^{low/-}$ cell fractions that contain definitive hematopoietic precursor cells. The enriched cell populations were placed on preformed OP9-DL4 stroma and cultured for several weeks.

(C) Standard gating strategy for hESC-derived lymphoid cells arising during the co-culture with OP9-DL4 cells.

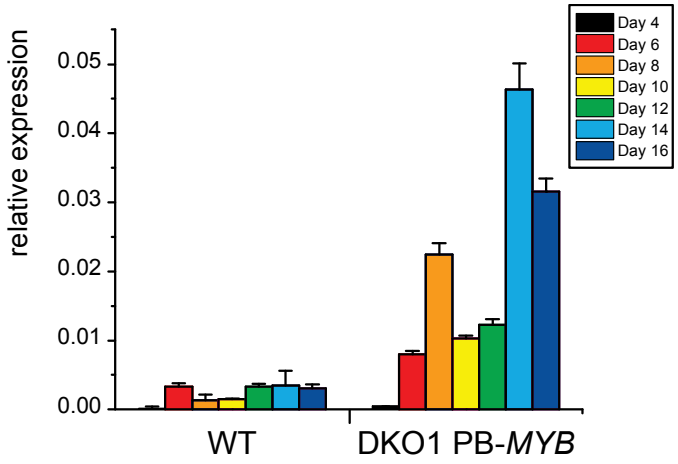
(D) Representative analysis of T cell potential of $CD34^{+}CD45^{low/-}$ and $CD45^{+}CD34^{low/-}$ cell fractions revealed the crucial role of MYB in T cell development. Cultures were harvested and the cells were analyzed on weeks 2, 4, and 5 of the co-culture.

Figure S7

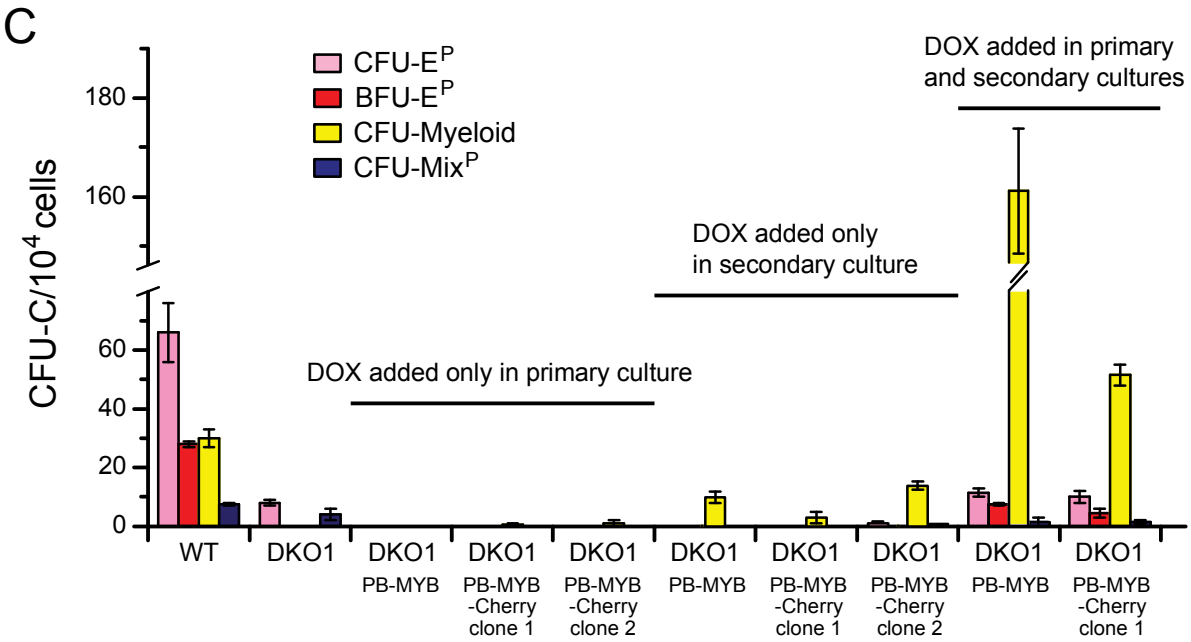
A



B



C



Supplementary Figure 7. The development of Day 12 clonogenic progenitors is rescued by forced expression of MYB in differentiating MYB-null hESCs.

(A) The MYB transgene expression is suppressed in differentiated cells at the fringes of a DKO1 PB-MYB-mCherry methylcellulose colony (white arrows, left panel), whereas the MYB-Venus expression is still expressed in these cells (white arrows, right panel). Day 6 CFU-mix colony is shown. DOX activation was performed at the primary differentiation of the hESCs and the clonogenic assay step.

(B) Comparative kinetics of the MYB mRNA expression in WT and DOX-activated DKO1 PB-*MYB* cells.

(C) The rescue of Day 12 hematopoietic progenitors. The most efficient recovery of the clonogenic progenitors is observed when MYB expression was induced at both the primary differentiation and during the clonogenic assay of differentiated DKO1 PB-*MYB* or DKO1 PB-*MYB-mCherry* cells. Data are presented as mean ± SD, n=3.