SUPPLEMENTARY APPENDIX

Gene therapy cures the anemia and lethal bone marrow failure in a mouse model of RPS19-deficient Diamond-Blackfan anemia

Pekka Jaako,[1] Shubhranshu Debnath,[1] Karin Olsson,[1] Ute Modlich,[2,3] Michael Rothe,[2] Axel Schambach,[2] Johan Flygare,[1] and Stefan Karlsson[1]

[1]Molecular Medicine and Gene Therapy, Lund Strategic Center for Stem Cell Biology, Lund University, Sweden; [2]Institute of Experimental Hematology, Hannover Medical School, Germany; and [3]LOEWE Research Group for Gene Modification in Stem Cells, Paul-Ehrlich-Institute, 63225 Langen, Germany

 doi:10.3324/haematol.2014.111195
Manuscript received on May 21, 2014. Manuscript accepted September 8, 2014.
Correspondence: Stefan.Karlsson@med.lu.se

Supplementary Methods

Blood and bone marrow analysis

Peripheral blood was collected from the tail vein into Microvette tubes (Sarstedt) and analyzed using Sysmex XE-5000. Erythrocytes were lysed using ammonium chloride for 10 minutes at room temperature. To evaluate the contribution towards various blood lineages following BM transplantation, samples were stained with the following antibodies for 30 minutes on ice in the dark: CD45.1 (Biolegend,110730), CD45.2 (eBioscience, 47-0454-82), B220 (Biolegend, 103208), B220 (Biolegend,103212), CD3 (Biolegend,100312), CD11b (Biolegend,101208), Gr1 (Biolegend,108408). Experiments were performed using a FACSCanto™ II cytometer (Becton Dickinson) and analyzed by FlowJo software (Tree Star, v9.3.3). FACS analysis of the myeloerythroid compartment in BM was performed as previously described (Jaako, 2011; Jaako, 2012). Briefly, BM cells were isolated by crushing femur and tibia in PBS containing 2% FCS (GIBCO). Fresh cells were stained with following antibodies: CD41 (eBioscience, 12-0411-83), GR1 (Biolegend, 108410), CD11b (Biolegend, 101210), B220 (Biolegend, 103210), CD3 (Biolegend, 100310), Ter119 (eBioscience, 25-5921-82), CD150 (Biolegend, 115910), c-Kit (eBioscience, 47-1171-82), Endoglin (Biolegend, 120404) and Sca-1 (Biolegend, 122520). Streptavidin was purchased from Life Technologies (Q10101MP). Propidium iodide (Life Technologies) was used to exclude dead cells. Experiments were performed using a FACS Aria cell sorter (Becton Dickinson) and analyzed by FlowJo software (Tree Star, v9.3.3).

Expression analyses

Total RNA was isolated from cultured cells after two days using the RNAeasy mini kit (Qiagen) and cDNA transcribed with SuperScript III reverse trancriptase (Life Technologies). *Rps19* mRNA was quantified by real-time PCR using the SYBR GreenER™ system (5'-GCAGAGGCTCTAAGAGTGTGG-3'; 5'-CCAGGTCTCTCTGTCCCTGA-3') according to manufacturer's instructions (Life Technologies, 11761-500).

LAM-PCR and insertion site analysis

Linear amplification-mediated PCR (LAM-PCR) (Schmidt M et al., Nat Methods. 2007) was performed on 300 ng of whole BM genomic DNA using 5 units of the enzyme Tsp509I, TaqIα and HaeIII. Samples were subject to linear amplification, before they were split in three separate digestion reactions. After restriction enzyme specific linker ligation, the samples were unified again for the nested PCR reactions. Barcoding, sequencing and bioinformatic analysis using custom Perl scripts, HISAP and MAVRIC (http://mavric.erasmusmc.nl/) were performed as described previously (Rittelmeyer and Rothe et al., Hepatology 2013).

Statistical analyses

One-way ANOVA with Tukey's multiple-comparison test was used to determine statistical significance using Prism (version 6, GraphPad Software).

Supplementary Table 1. Comparison of estimated and observed number of integration sites.

Treatment group	Transplanted cells	Transduction efficiency	Transduced cells	Possibly tdx HSC (1 in 1000)	Mean VCN of 3	Mappable IS with 3E	After clustering (mean IS number)	Percent of theoretically mappable IS	Potential new IS by repeat of LAMs
Control RPS19	600000	50,0%	300000	300	900	675	136	20,2 %	555
D/+ RPS19	900000	50,0%	450000	450	1350	1013	209	20,7%	905

Most deep sequencing data for gene therapy experiments available in the literature are related to the follow-up of treated patients and analysis of the starting pool of integrations prior to transplantation. In these cases, thousands of different integration sites are mapped, something that is specific to the human setting, where we deal with much more input cells used for transplantation. In mice, however, the amount of long-term (LT) repopulating hematopoietic stem cells (HSC) is around 1 in 10000 nucleated bone marrow cells (Challen et al., Cytometry A 2009). In our experiments, the c-kit enrichment increased the number of LT-HSC to around 1 in 1000 isolated cells. Supplementary Table 1 shows how we would calculate the expected number of detectable insertions in our mice according to the expected number of stem cells in the grafts. We explain this by looking at the Control group:

In the Control group, we had one mouse from the experiment 1 (1x10^6 cells transplanted) and four mice from experiments 2 and 3 (both 5x10^5 cells transplanted). For calculation, we chose the mean of 6x10^5 cells transplanted cells in this group. With a transduction efficiency of around 50% (Supplementary Figure 2 for transduction efficiencies), we likely had around 3x10^5 cells transduced. Of those, approximately 300 HSC could have been hit. With a mean vector copy number (VCN) of 3 in the animals (Supplementary Figure 6), we could have possibly screened for 900 insertion sites (IS) per animal. From previous experiments using LAM-PCR (Gabriel et al. Nature Medicine 2009; Rittelmeyer and Rothe et al., Hepatology 2013) we know that is possible to map around 75% of all insertions with 3 Enzymes, in our case 675 integrations. We observed a mean of 136 insertions per animal in the Control group, which is 20.2 % of the theoretically possible sites. This was also true for the D/+ group for which we mapped 209 insertions per mouse (20.9 %). As mentioned in Supplementary Figure 6C, we see on average between 82-86% of low read insertions in the samples. In the standard LAM-PCR protocol only 1/5th of the DNA sample is used as input for the following nested PCR steps. Assuming that we would have detected a certain amount of insertions again due to increased abundance indicated by high read counts, we can also expect

to find even more new insertions with low read counts per mouse. From the experience of previous experiments we know that the ratio of background to high read counts would not be affected by this.

Putting these calculations into perspective, we can assume a polyclonal to oligoclonal situation in these mice because malignant clonal dominance would have resulted in a much lower number of different integration sites. From our estimations, it seems as if there was no major reduction in clonality. We cannot exclude the likely situation of certain clones contributing more actively to hematopoiesis at the time of measurement. This can lead to an overrepresentation of clones within a polyclonal background (Supplementary Figure 6D). With our experiments we aimed to detect clonal dominance and not for mapping of all possible integration sites in the mice.

Supplementary Table 2. Transcription start site (TSS) distance analysis using MAVRIC (http://mavric.erasmusmc.nl/).

Treatment group	% in gene	- 10 kb to TSS	+ 10 kb to TSS	+/- 10 kb to TSS in %
Control RPS19	65,1%	7,3%	14,2%	21,5%
D/+ RPS19	49,8%	7,0%	8,1%	15,1%
Lentivirally transduced human CD34+ cells	84,5%	6,5%	12,9%	19,4%
Gammaretrovirally transduced human CD34+ cells	62,9%	16,7%	21,8%	38,6%

Supplementary Table 3. Cancer gene list overlap using MAVRIC (http://mavric.erasmusmc.nl/).

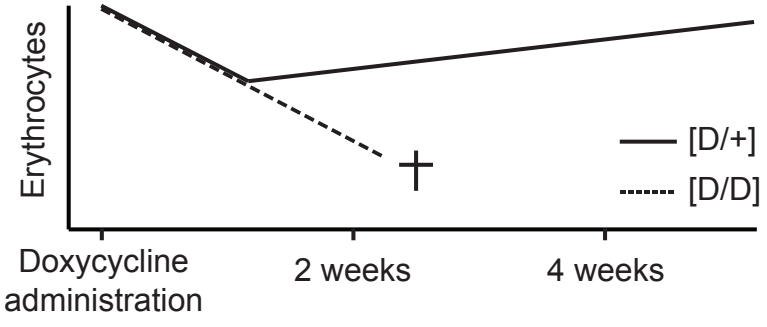
Treatment group	Unique gene symbols	Element of Cancer Lists	% Cancer list overlap
Control RPS19	274	51	18,6%
D/+ RPS19	444	72	16,2%
Lenti - control	2137	369	17,3%
Gamma - control	1245	266	21,4%

Cancer Gene List Overlap:
Network of Cancer Genes http://bio.ieo.eu/ncg3/index.html
Collection of Cancer Gene Lists http://www.bushmanlab.org/links/genelists

Supplementary Table 4. List over the 20 most common insertion sites. Insertion site analysis was performed on DNA isolated from unfractioned bone marrow cells of five recipient mice transplanted with transduced control or D/+ hematopoietic cells. Integration sites were amplified by LAM-PCR and sequences were retrieved by high throughput 454-sequencing. CIS=common insertion site, RF=recurrently found in other studies with this vector type; Tx=transplantation; Co=control.

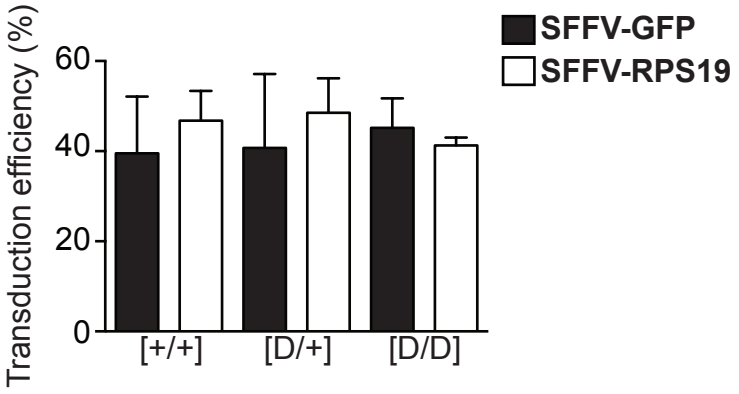
CIS #	Gene Symbol	Hits	Range (kb)	Animals	Different TX	Top Ranked insertion
1	Plcb4	2	3.9	Co2, D/+1	yes	Top 2 in animal D/+1
2	Ksr2	3	0.4	D/+2, D/+3, D/+4	no	Top 6 in animal D/+4; RF
3	Fzd9	2	21.9	Co2, D/+2	yes	
4	Sox5	2	0.2	Co2, D/+2	yes	
5	Olfr1344	4	0.2	Co1, D/+1, D/+2, D/+4	yes	
6	Lim2	3	0.7	Co1, D/+1, D/+2	yes	
7	Gm6816	4	0.2	Co1, Co3, D/+2, D/+3	yes	
8	Usp3	2	0.3	Co3, Co5	no	Top 3 in animal Co5
9	Sfi1	2	7	Co3, D/+2	yes	RF
10	Klhl29	3	0.2	D/+1, D/+3, D/+5	yes	
11	Prl	7	1.2	Co3, Co5, D/+1, D/+2, D/+3, D/+4	yes	RF
12	Fars2	3	0.2	Co3, D/+2, D/+3	yes	RF
13	4930452B06Rik	3	0.7	Co4, D/+2, D/+3	yes	RF
14	Tmem74	3	0.1	Co3, D/+2, D/+3	yes	Top 3 in animal Co3
15	Hdac7	3	0.2	Co1, D/+1, D/+2	yes	
16	Mir1941	3	31.5	Co2, D/+1, D/+2	yes	
17	Ndufv2	2	0.2	Co5, D/+1	yes	Top 3 in animal D/+1
18	Snx32	6	0.4	Co2, Co3, D/+1, D/+2, D/+3, D/+4	yes	RF
19	5830428H23Rik	3	0.2	Co3, D/+4, D/+5	yes	
20	Pir	2	0.7	Co2, Co4	yes	Top 1 in animal Co4

Supplementary Figure 1



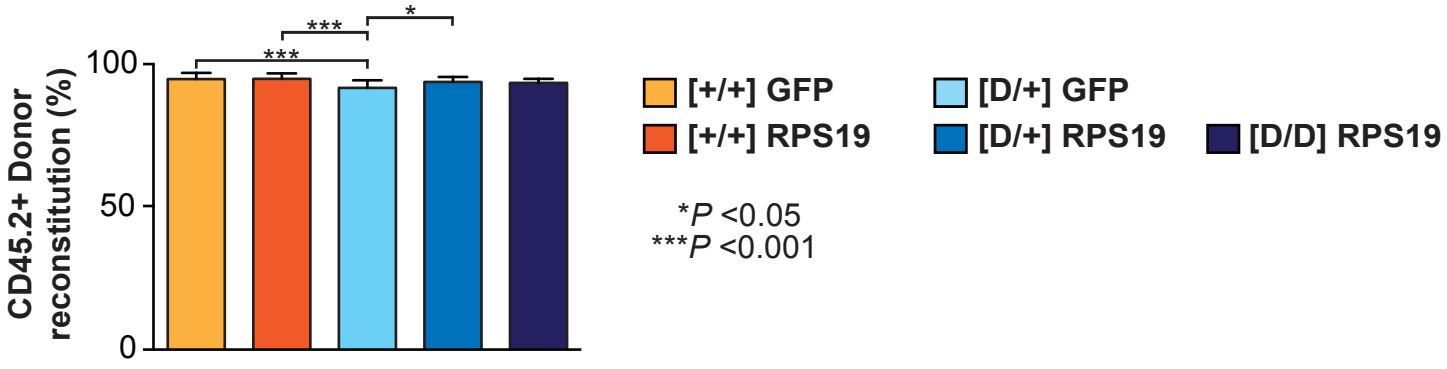
Supplementary Figure 1. Schematic representation of the erythroid phenotype in recipients transplanted with D/+ or D/D bone marrow. The graph illustrates how the number of erythrocytes decreases rapidly in the recipients with D/D bone marrow whereas the reduction of erythrocytes in recipients with D/+ bone marrow is milder and gradually compensated (Jaako et al., Blood 2011).

Supplementary Figure 2



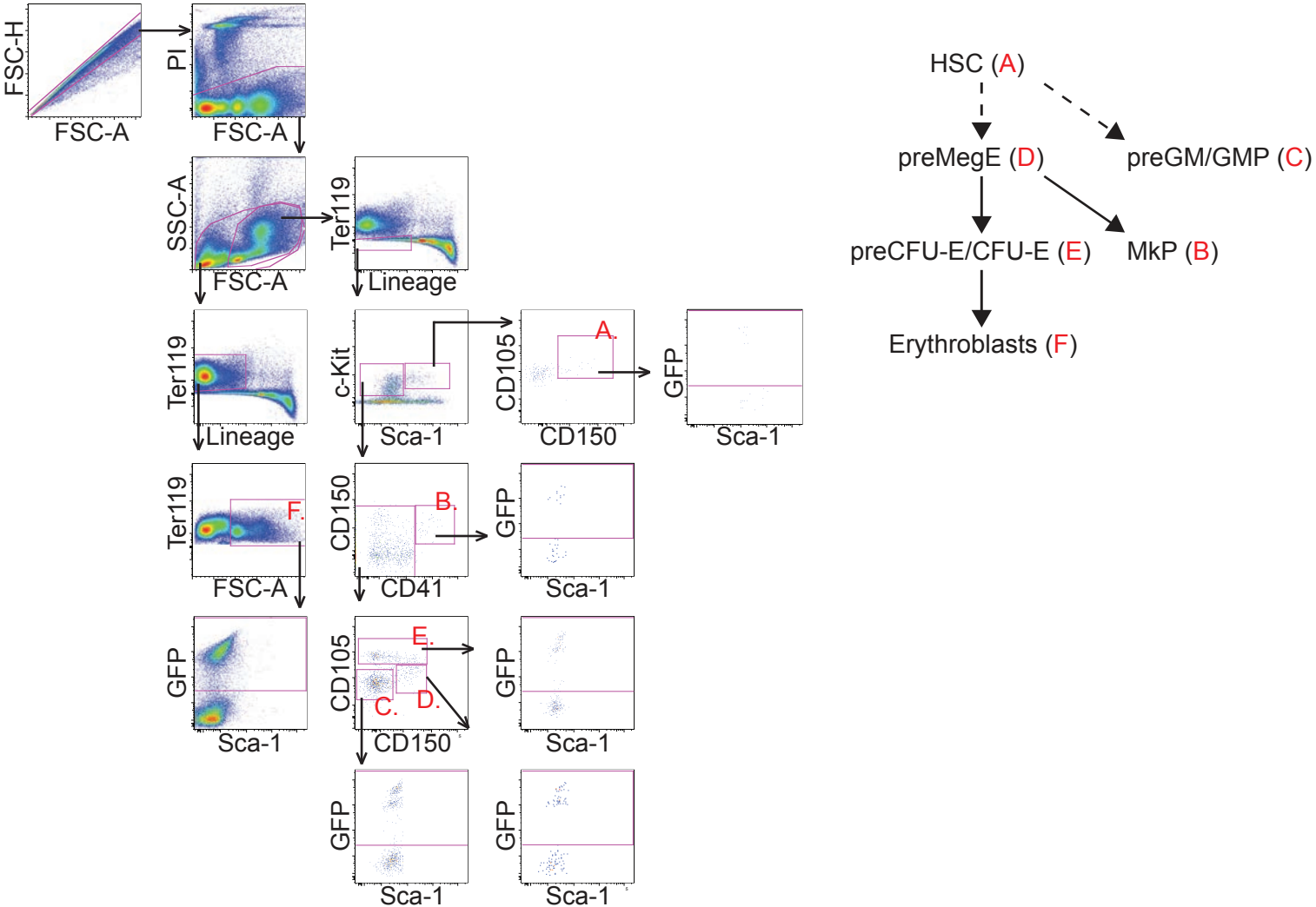
Supplementary Figure 2. Transduction efficiencies. Transduction efficiencies were determined based on the frequency of GFP+ cells one day after the transduction. Error bars represent standard deviation.

Supplementary Figure 3



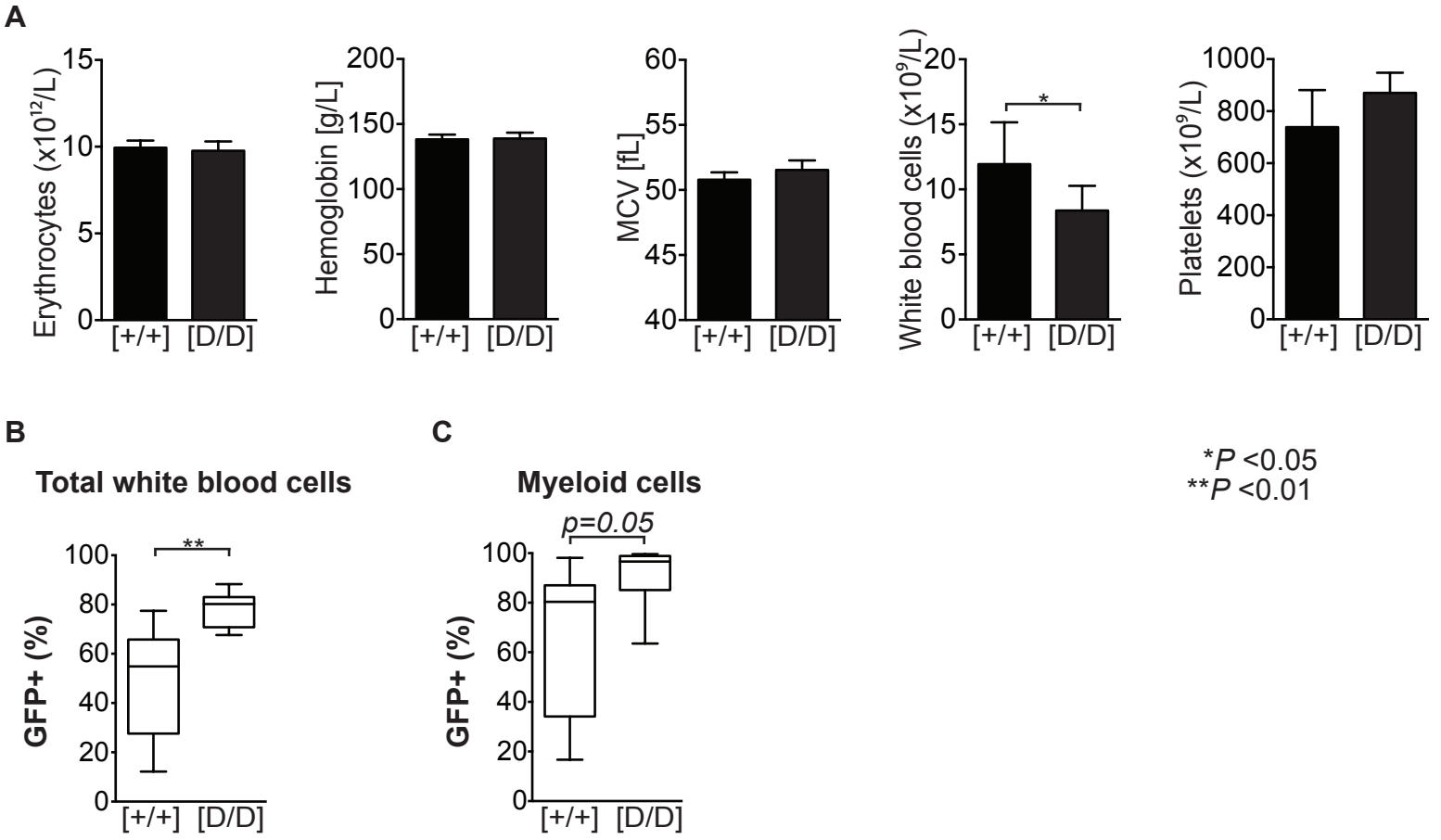
Supplementary Figure 3. Donor reconstitution based on the frequency of CD45.2+ cells in the blood after 4 months of doxycycline administration. Error bars represent standard deviation.

Supplementary Figure 4.



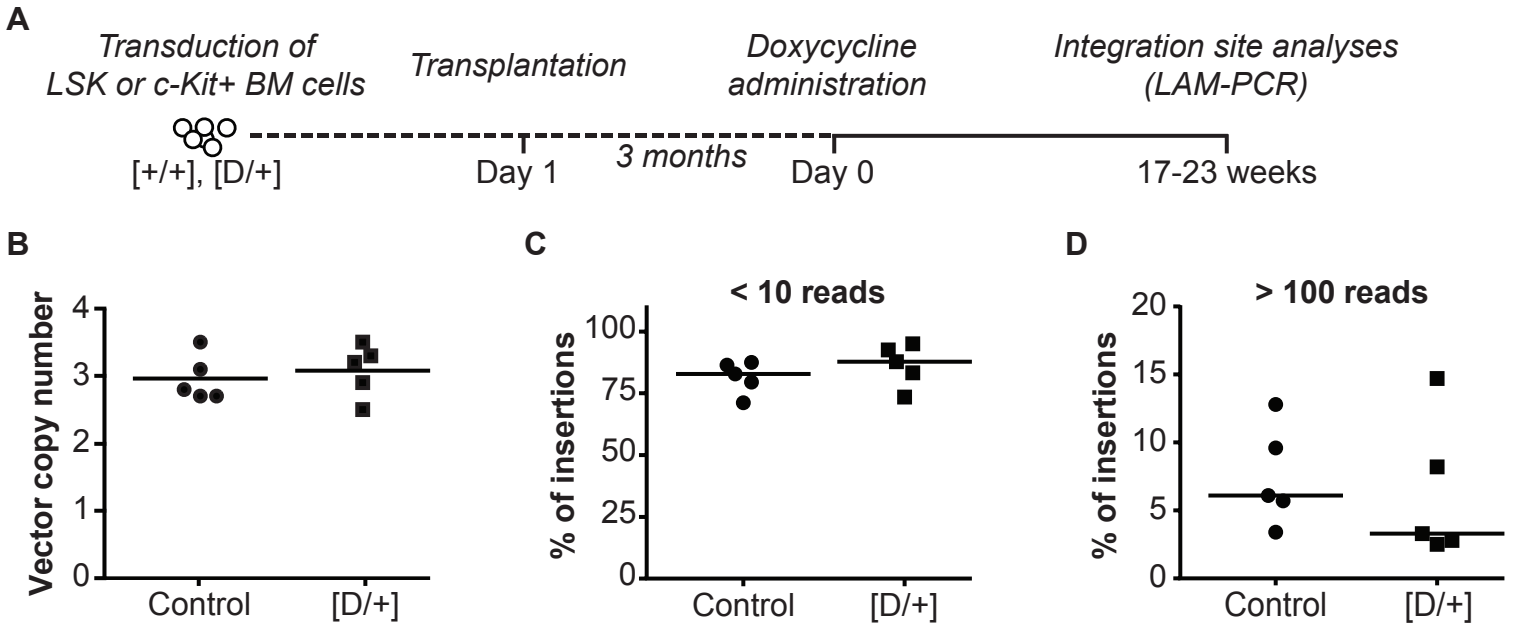
Supplementary Figure 4. The applied FACS strategy allowing the fractionation of myeloerythroid progenitors and erythroid precursors.

Supplementary Figure 5.



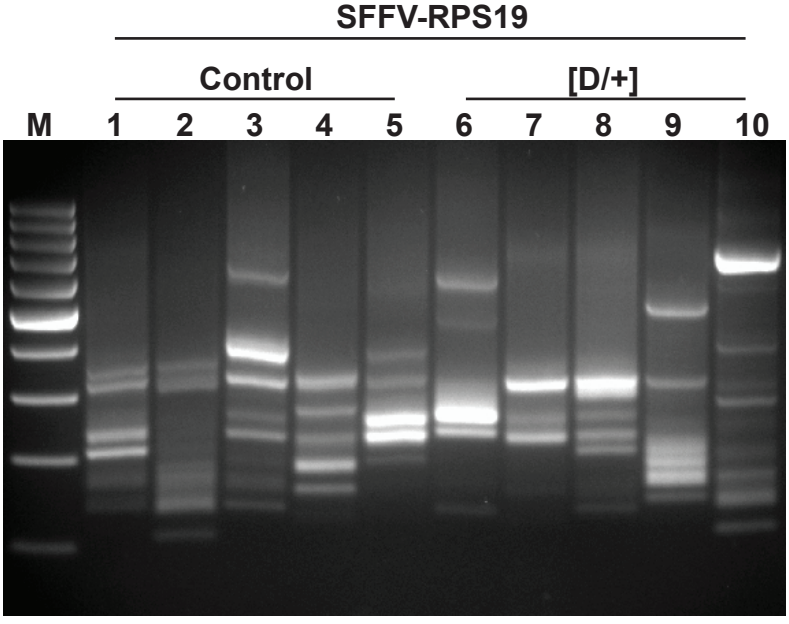
Supplementary Figure 5. (A) Erythrocyte number, hemoglobin concentration, mean corpuscular volume (MCV), white blood cell number and platelet number in secondary recipients reconstituted with SFFV-RPS19 control or D/D whole BM 10 weeks after transplantation . (B and C) The percentage of GFP+ total white blood cells and myeloid cells in the peripheral blood 10 weeks after transplantation. (n = 7 /group). Error bars in (A) represent standard deviation. Data in (B-C) are presented as a box whisker plot with minimum and maximum values

Supplementary Figure 6



Supplementary Figure 6. Gene-corrected Rps19-deficient cells sustain polyclonal hematopoiesis and show a typical lentiviral insertion profile. (A) Experimental strategy to validate the integration profile of the SFFV-RPS19 vector and clonal dynamics of the transduced cells. DNA was isolated from unfractioned BM cells of 5 recipient mice transplanted with transduced control or D/+ hematopoietic cells after 17-23 weeks of doxycycline administration. Integration sites were amplified by LAM-PCR and sequences were retrieved by high throughput 454-sequencing. (B) Vector copy number per cell in recipients with SFFV-RPS19 transduced control or D/+ BM cells. (C-D) Percentage of insertions with low read counts (<10 reads; C) or with high read counts (>100 reads; D) in recipients with SFFV-RPS19 transduced control or D/+ BM cells.

Supplementary Figure 7.



Supplementary Figure 7. LAM-PCR amplicons. The gel picture shows the amplicons generated by LAM-PCR.