

Supplemental Materials

Supplemental Methods

Guide RNA and ribonucleoprotein complex production:

Guide RNAs were designed using Benchling(<https://benchling.com>) and *in vitro* transcribed with the TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific, Waltham, MA). These gRNAs were purified using the RNeasy Mini Kit (QIAGEN, Venlo, Netherlands). For generation of ribonucleotide protein (RNP) complexes, gRNAs were premixed with Cas9 protein purchased from PNA Bio (Newbury Park, CA) at a ratio of 1:2 at room temperature for 10 minutes before electroporation.

Transfection and transplantation of HSPC:

All animal studies were conducted according to protocols approved by the National Heart, Lung and Blood Institute Animal Care and Use Committee. Rhesus macaque CD34⁺ cells were collected via apheresis from the blood following mobilization with G-CSF and plerixafor, and immunoselected for CD34 expression as described¹² via column purification. CD34⁺ cells were resuspended and aliquoted at a concentration of 5x10⁶ cells per 750µl of Opti-MEM (Lonza, Basel, Switzerland). 15-30µg of gRNA complexed with 30-60µg of Cas9 protein were added into each aliquot and electroporation with a single pulse of 400V for 5msec was performed using the BTX ECM 830 Square Wave Electroporation System (Harvard Apparatus, Holliston, MA). The electroporated cells were pooled and incubated in X-VIVO™ 10 (Lonza) supplemented with 1% Human Serum Albumin (Baxter, Deerfield, IL), SCF 100ng/mL, FLT3L 100ng/mL and TPO 100ng/mL (all purchased from PeproTech, Rocky Hill, NJ) at 32°C overnight, then collected, washed and reinfused into the autologous macaque following total body irradiation (500 rad/day for 2 days).

Flow cytometry

PBMNCs (peripheral blood mononuclear cells) and granulocytes were separated from macaque PB through density gradient centrifugation using Lymphocyte Separation Medium (MP Biomedicals, Santa Ana, CA). The purified cells had red blood cells removed by incubation with ACK lysing buffer (Quality Biological, Gaithersburg, MD) and were then stained with the following antibodies: FLAER (Cedarlane Lab, Burlington, Canada), CD45-V450 (BD Biosciences, San Jose, CA),

CD24-APC-Cy7 (Abcam, Cambridge, UK), CD33-PE (Miltenyi Biotec, Bergisch Gladbach, Germany), CD14-pacific blue (Thermo Fisher Scientific), CD3-APC-Cy7, CD20-PE-Cy5, CD11b-BV605, and CD45-Alexa Fluor 700 (BD Biosciences). Flow cytometric analysis and/or sorting were performed on a FACSARIA-II flow sorter (BD), and data were analyzed using FlowJo software (Tree-star).

Targeted deep sequencing

Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) and libraries were constructed by amplifying the gRNA targeting region containing the flow cell linker adaptors and 10bp unique index barcodes. Libraries were sequenced on the Miseq instrument (Illumina, San Diego, CA) with 300bp paired-end reads. A consensus sequence was generated from pre-transplantation data from each macaque using a custom Python pipeline to evaluate particular single nucleotide polymorphisms, and sequencing depth was between 70,000 and 300,000 reads/gRNA target site. The resulting fastq files from multiple time points were aligned to the appropriate consensus sequence using CRISPResso (<http://crispresso.rocks>) to quantify the number of reads for each sequence. Through custom Python code, sequences that do not vary significantly in frequency ($<|0.2\%$) compared to the intact sample and non-specific sequences that appear only once across all time points were excluded from further analysis, and a 20bp window around binding of the gRNA was applied. The custom Python pipeline used in this study is available at <https://github.com/shint3/PNH-CRISPR-analysis>. Heatmaps were generated using a custom R application (Foundation for Statistical Computing, Vienna, Austria).

Reagents and resources

gRNAs

Name	Sequence	PAM
PIG-A #1	AGGGTGTAGCTGAGGCACGG	TGG
PIG-A #2	CTACCCAAATATGGGAGGCG	TGG
PIG-A #3	TGAGGCCATTAGTGAGGTAA	CGG
AAVS1 #1	GTGGCCCCACGGTGGGGTGG	AGG
AAVS1 #2	GGGGCCACTAGGGACAGGAC	TGG
AAVS1 #3	TGACAAAAAGCCCCATCCTT	AGG

Antibodies

Name	Vendor (Cat#)	Clone
FLAER: Alexa Fluor488 proaerolysin	Cedarlane Labs (# FL2S)	N/A
anti-CD24 APC-Cy7	Abcam (# ab197137)	SN3
anti-CD33 PE-Cy7	Miltenyi Biotec (# 130-091-732)	AC104.3E3
anti-CD45 V450	BD Biosciences (# 561291)	D058-1283
anti-CD14 Pacific Blue	Thermo Fisher (# MHCD1428)	Tuk4
anti-CD20 PE-Cy5	BD Biosciences (# 555624)	2H7
anti-CD3 APC-Cy7	BD Biosciences (# 557757)	SP34-2
anti-CD11b Brilliant Violet 605	BD Biosciences (# 562721)	ICRF44
anti-CD45 Alexa Fluor 700	BD Biosciences (# 561288)	D058-1283

Supplemental Figure 1

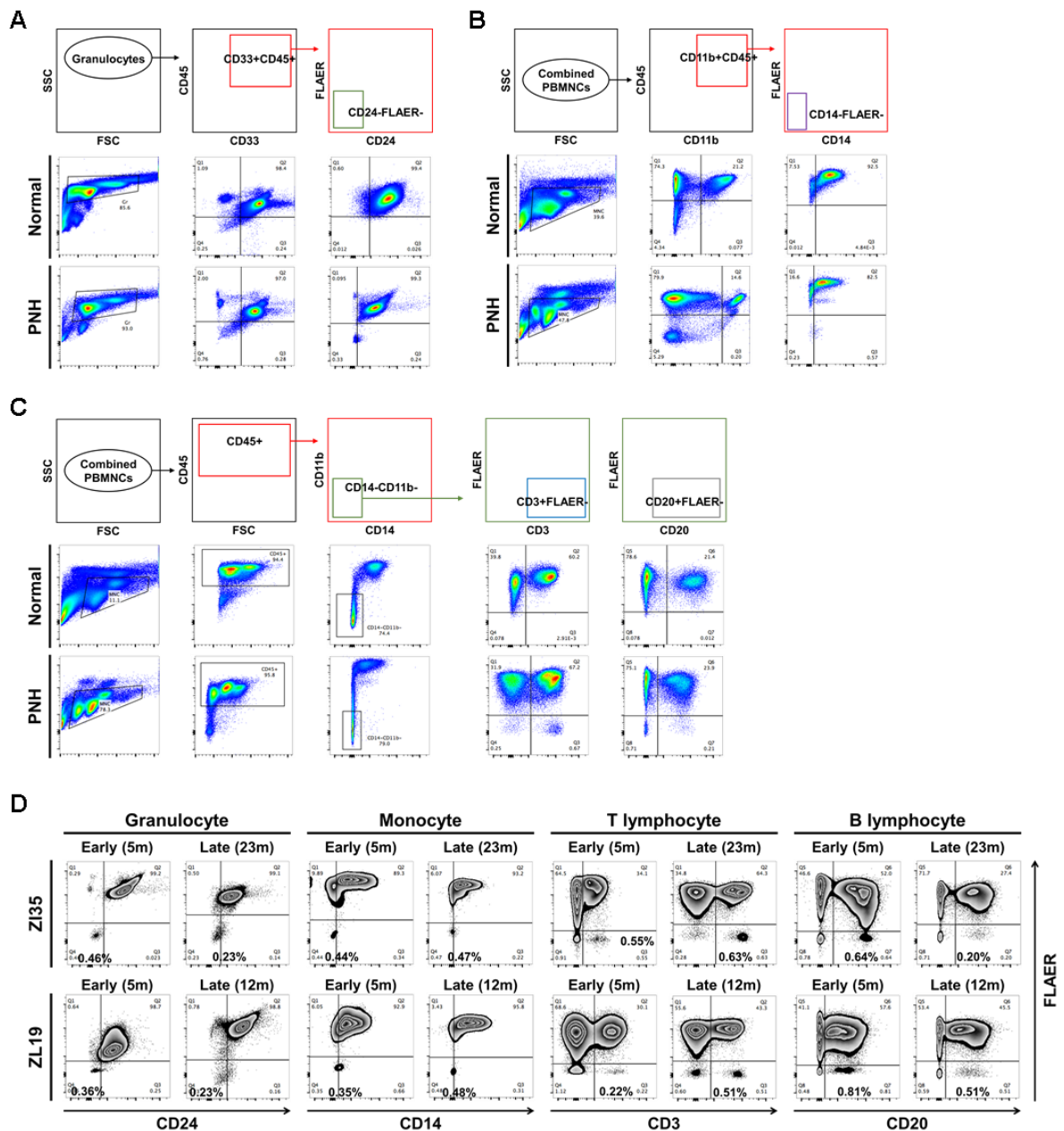


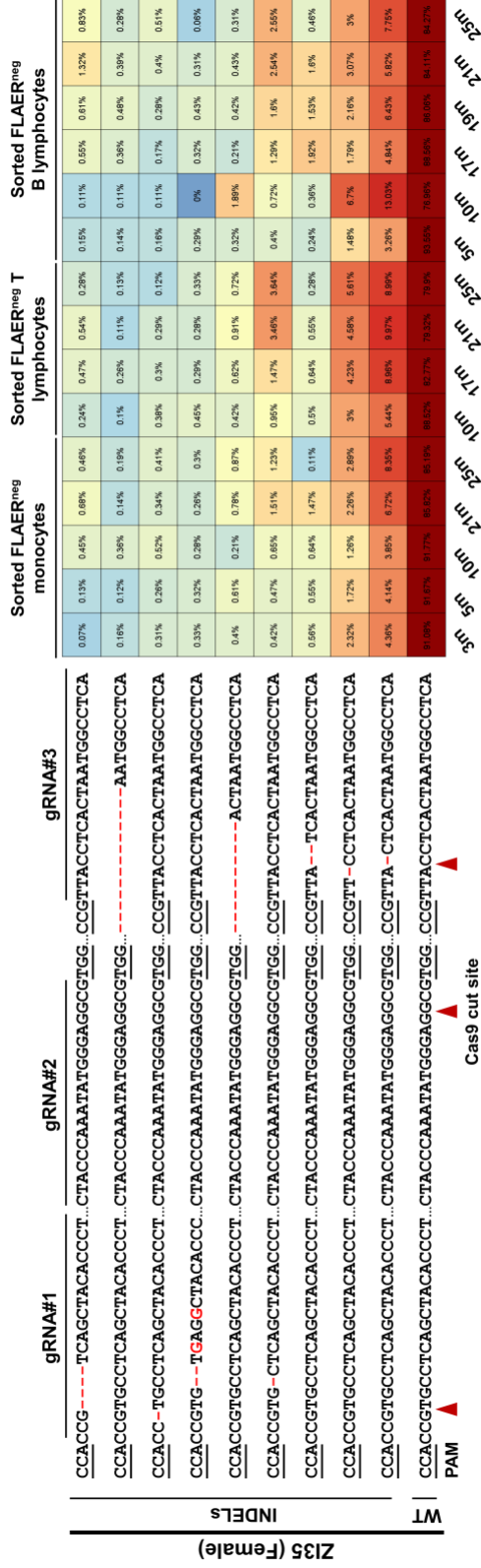
Figure S1. Strategy for detection and sorting of circulating macaque blood cells with the PNH phenotype.

(A-C) Schematic gating strategy (upper) and representative plot (lower) used for sorting of PNH granulocytes (A), monocytes (B), T and B lymphocytes (C) via FLAER and antibody staining.

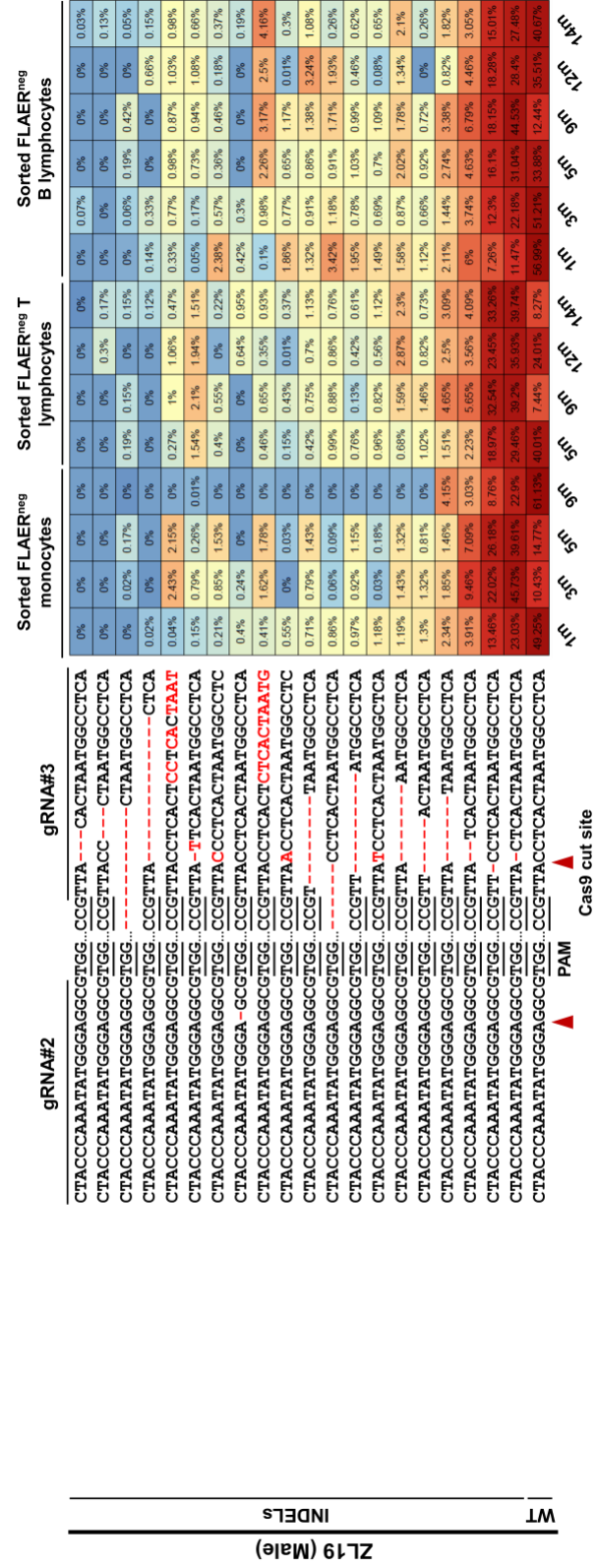
(D) Representative flow plots of PNH lineage cells from early (5 months) and late time points (12 or 23 months) post-transplantation in ZI35 and ZL19.

Supplemental Figure 2

A



B

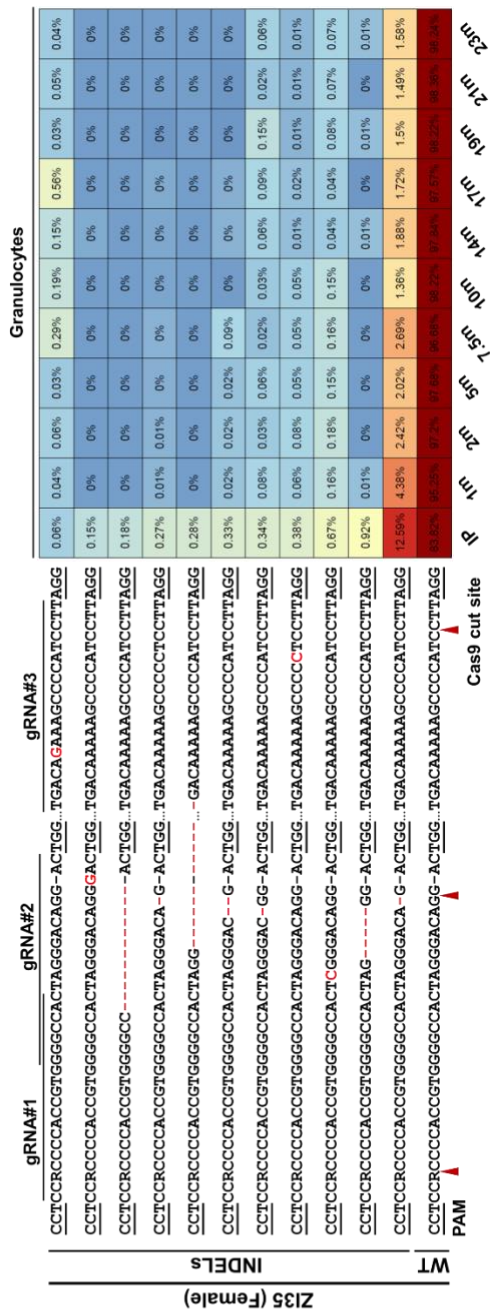
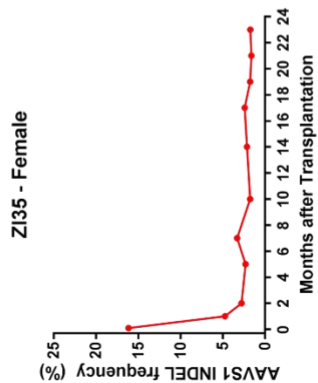


5

Figure S2. Quantitative tracking of *PIG-A* mutated clones in sorted FLAER^{neg} mature hematopoietic lineages overtime.

(A-B) Targeted deep sequencing of *PIG-A* target sites was conducted in sorted FLAER^{neg} monocytes, T cells and B cells. Color gradient in the heatmap shows the fractional read-count contribution of INDELS and the individual sequences of the most frequent INDELS are shown to the left. Each row indicates an individual INDEL type identified and each column represents a time point. PAM; protospacer adjacent motif, WT; wild-type, IP; infusion product

A



B

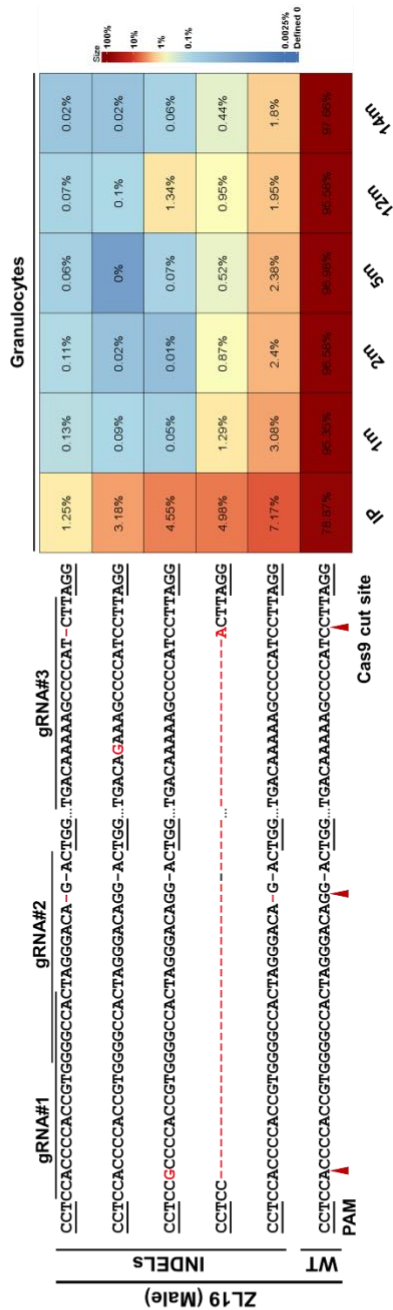
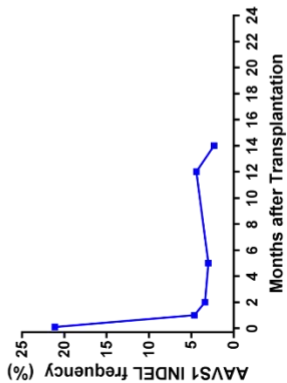


Figure S3. Tracking of AAVS1 mutation frequency and INDELS generated over time.

(A-B) Mutation frequency of AAVS1 “safe harbor” locus over time were followed in peripheral blood granulocytes from both ZI35-female (A) and ZL19-male (B) macaques by targeted deep sequencing. Heatmap and INDEL type retrieved are shown in the right panel, and the overall frequency of AAVS1 mutation post-transplantation is represented in the left as a line graph.

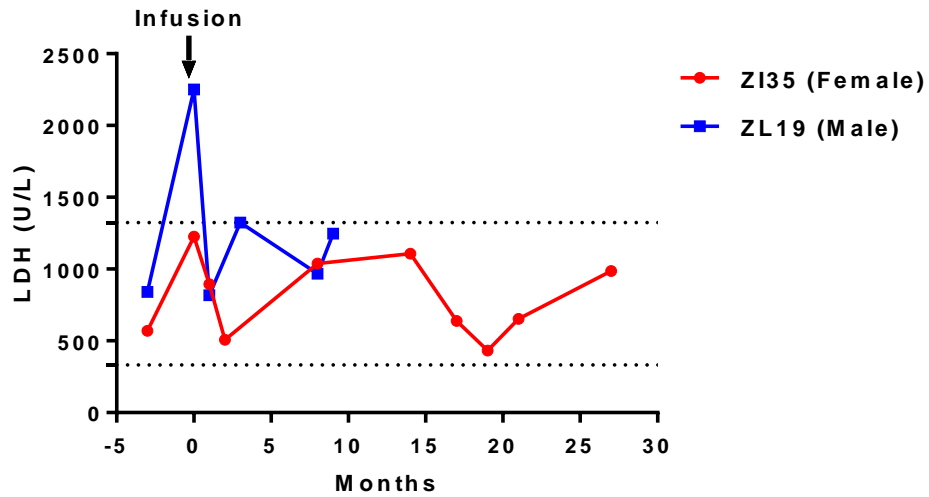


Figure S4. Lactic acid dehydrogenase (LDH) level following autologous transplantation.

Concentration of LDH was determined in the peripheral blood of ZI35 and ZL19 over time. Values are expressed as Unit per liter. Between the two dotted lines represents a normal range of the rhesus macaque (332-1324 U/L), and the time of transplantation was shown as an arrow.