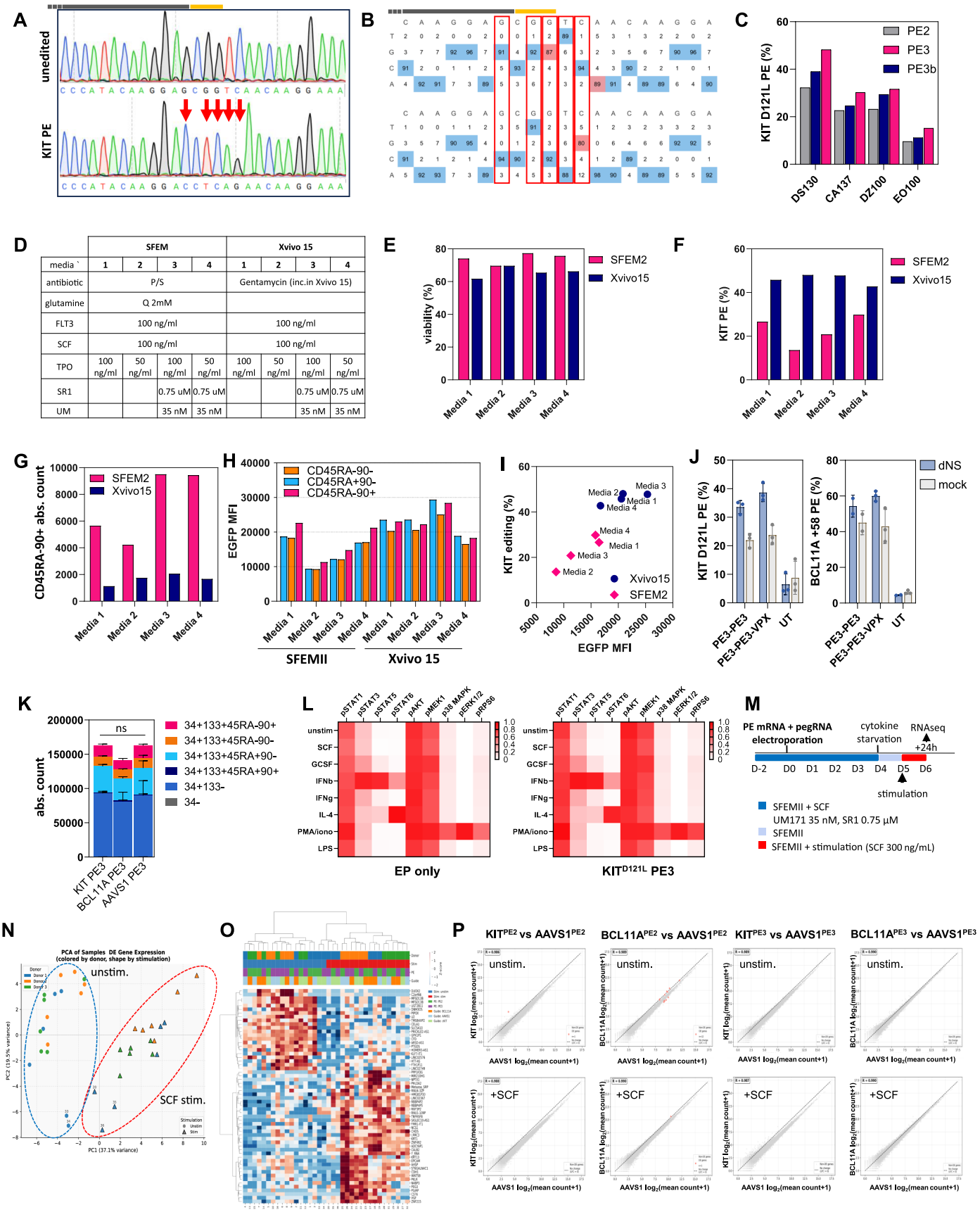
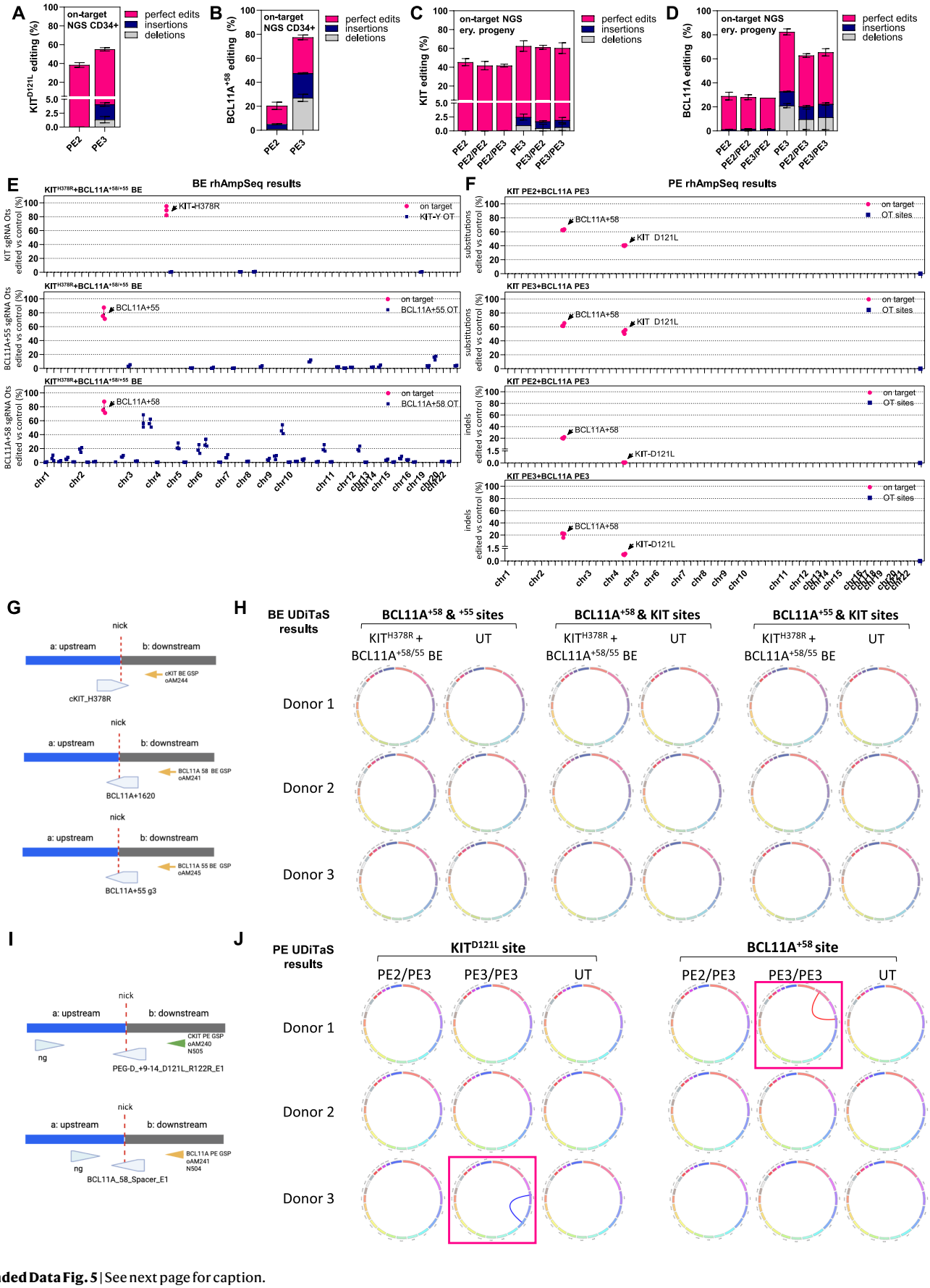


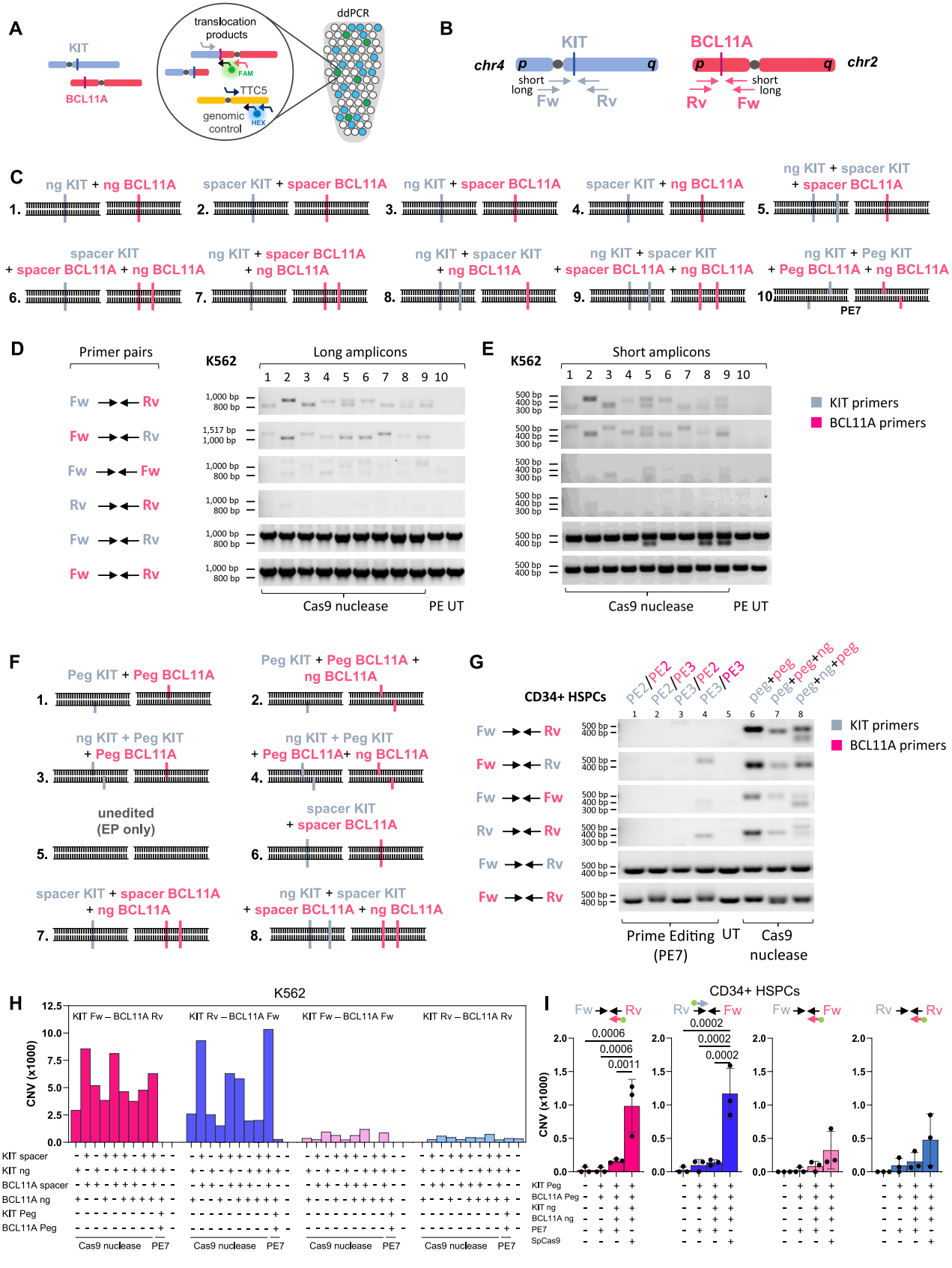
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Optimization of KIT PE efficiency and culture conditions for HSPC editing. (**A**) Sanger sequencing traces showing the introduction of the KIT D121L mutation by prime editing (PE3). Additional base mismatches are introduced to evade the cellular mismatch repair (MMR) and increase editing efficiency. Spacer and PAM are highlighted by the grey and yellow bars, respectively. (**B**) Quad-plot indicating the base changes at each position within the KIT^{D121L} pegRNA, with highlighted editing of the target bases (red squares). Spacer and PAM are highlighted by the grey and yellow bars, respectively. (**C**) Electroporation pulse optimization for efficient delivery of PE mRNA into CD34+ HSPCs using Lonza-4D Nucleofector. Bar plot reporting editing efficiencies for different prime editing strategies (PE2, PE3, PE3b) and four electroporation pulses (DS130, CA137, DZ100, EO100). (**D**) CD34+HSPC culture media formulations based on StemCell StemSpan SFEMII or Lonza X-VIVO15 tested for PE optimization. (**E**) Viability of CD34+ HSPCs 24 h after electroporation with PE mRNA and pegRNA, cultured in the 4 different media described in panel D. (**F**) KIT PE efficiencies across different media compositions in SFEMII and X-VIVO15. (**G**) Absolute counts of CD45RA-CD90+ cells after PE and culture with the different media compositions. (**H**) EGFP Median Fluorescence Intensity (MFI) as a surrogate marker of transfection efficiency and mRNA translation in different progenitor subpopulations gated on CD45RA and CD90 subsets, according to media composition. (**I**) Correlation between EGFP MFI (see panel H) and KIT^{D121L} PE efficiency across tested conditions. (**J**) Quantification of KIT^{D121L} (left) and $BCL11A^{+58}$ (right) PE efficiency in human $CD34^{+}$ HSPCs using the PE3 strategy with or without co-electroporation of VPX mRNA and nucleoside supplementation (dNS). Mean±SD. N = 3 donors. (**K**) Absolute counts of KIT prime edited HSPCs vs AAVS1 control during in vitro expansion culture in the presence of SCF. Stacked bar plots report the different progenitor subpopulations by flow cytometry. Mean ± SD. N = 3 donors. Two-way ANOVA with multiple comparisons. (**L**) Phospho-flow analysis of CD34+ HSPCs-derived myeloid culture comparing non-edited, electroporation only (EP only) and KIT^{D121L} PE3-edited conditions. CD34+ cells were thawed, edited and cultured in myeloid-differentiation media, before overnight starvation and challenge with different stimuli. The Median Fluorescence Intensity (MFI) of the Fluorescence Minus One (FMO) control was subtracted from each replicate value for each marker to account for background staining. MFI was normalized between 0 and 1 to enable plotting of different parameters on the same scale. N = 3 biological replicates (independent donors). Representative flow plots for all markers and stimuli are reported as Supplementary Figs. 6 and 7. (**M**) Experimental layout for RNAseq of KIT^{D121L} epitope edited CD34+ HSPCs. Cells were thawed and edited as previously described, cultured for 4 days to allow substitution of the receptor protein on the cell surface, starved for cytokines (overnight) and stimulated with SCF. Cells were harvested at 24 h after the start of stimulation for RNAseq (Plasmidsaurus), respectively. The experiments were performed in biological replicates (N = 3 independent HSPCs donors). (**N**) PCA plot of the first two principal components for the RNAseq dataset from the experiment described in (M). Samples are labeled by stimulation and donor origin. (**O**) Heatmap reporting the significant DEGs resulting from the SCF stimulation vs unstimulated comparison. Each column corresponds to a sample (see mapping reported above the heatmap and legend on the right), while rows correspond to transcripts. As depicted by the dendrogram, unbiased clustering shows predominant effects of the stimulation, while editing conditions are distributed within each cluster. Full comparisons are available in Supplementary Table 4. (**P**) RNA-seq analysis of KIT^{D121L} epitope- or BCL11A- edited CD34+ HSPCs, with or without stimulation with SCF, compared to AAVS1 control edited cells. log-scale scatter plot of the mean gene expression values of the AAVS1 control and KIT- or BCL11A- edited cells, either at the baseline (top) or after stimulation with SCF (bottom). N = 3 biological replicates. Unstim., unstimulated.



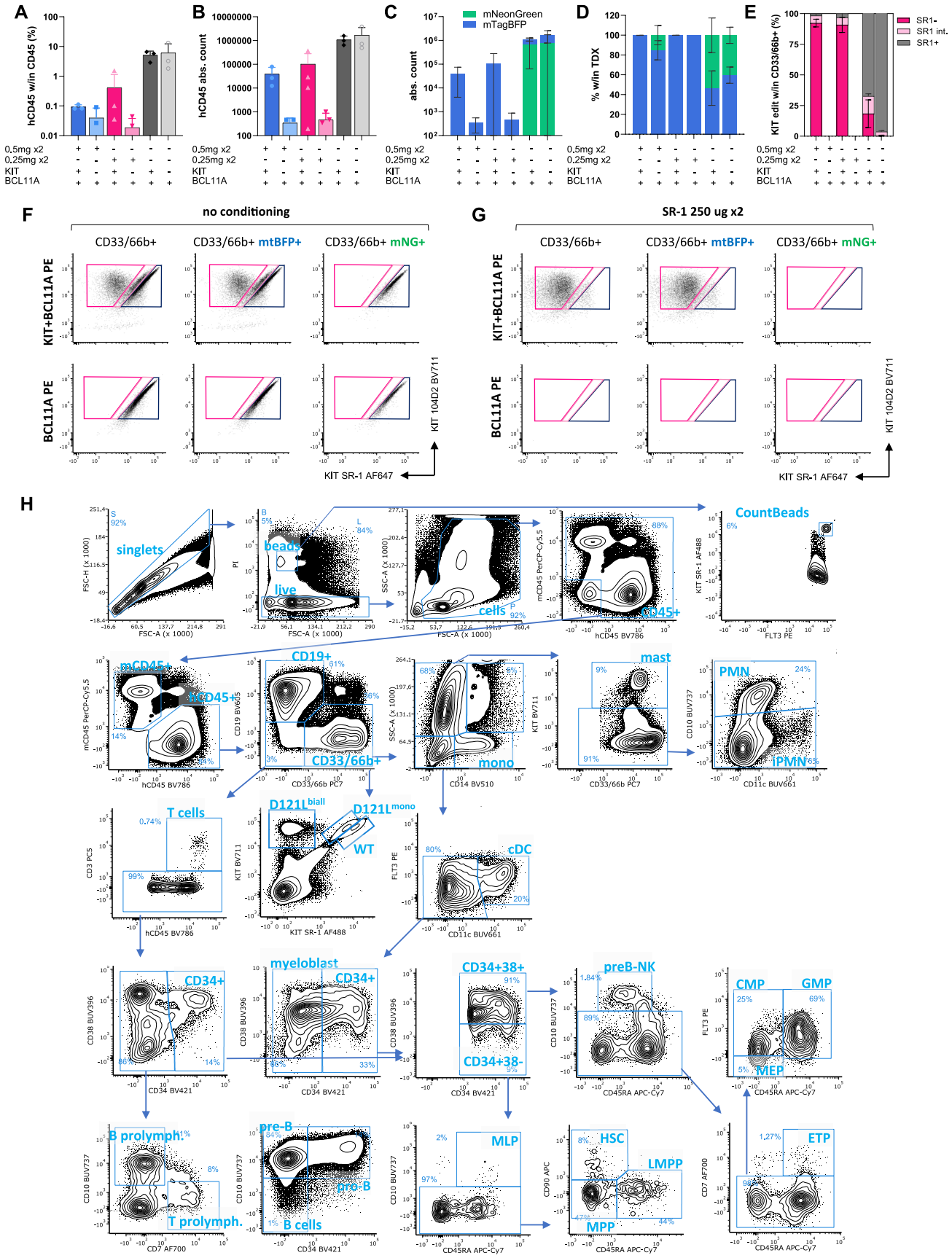
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | On-target and off-target genomic safety analyses of PE and BE HSPCs. (**A**, **B**) NGS amplicon sequencing of on-target prime editing sites on CD34+ HSPCs edited for either KIT or BCL11A, during in vitro expansion culture, according to PE approach. Bar plots report results for KIT (A), and BCL11A (B). Perfect edits entail precise introduction of pegRNA specified genomic changes. Mean ± SD. N = 3 biological replicates. (**C**, **D**) NGS amplicon sequencing of on-target prime editing sites on in vitro differentiated erythroid cells derived from CD34+ HSPCs edited for either KIT, BCL11A or both, according to PE approach. Bar plots report results for KIT KIT (C), and BCL11A (D). Perfect edits entail precise introduction of pegRNA specified genomic changes. Mean ± SD. N = 3 biological replicates. (**E**) rhAmpseq targeted NGS re-sequencing of genomic sites showing statistically significant editing in KIT^{H378R}, $BCL11A^{+55/+58}$ triplex BE HSPCs as compared to unedited controls. The editing difference (%) of BE samples vs controls for sites that showed significant editing (vs controls) is displayed. Off-target sites are separated by sgRNA (KIT, top; $BCL11A^{+55}$, middle; $BCL11A^{+58}$, bottom). N = 3 biological replicates (independent donors). (**F**) rhAmpseq targeted NGS re-sequencing of genomic sites showing statistically significant editing in KIT^{D121L} and $BCL11A^{+58}$ duplex PE HSPCs as compared to unedited controls. The editing difference (%) of PE samples vs controls for sites that showed significant editing (vs controls) is displayed. Off-target analyses are separated by PE approach (KIT PE2/BCL11A PE3, KIT PE3/BCL11A PE3) and whether substitutions (top 2 plots) or indels (bottom 2 plots) were detected. N = 3 biological replicates. (**G**) Design of UDiTaS (Uni-Directional Targeted Sequencing) strategy to detect translocations in multiplex BE CD34+ HSPCs. (**H**) Circos plot displaying UDiTaS results for KIT^{H378R}, $BCL11A^{+55/+58}$ triplex BE HSPCs. N = 3 biological replicates (independent donors). (**I**) Design of UDiTaS (Uni-Directional Targeted Sequencing) strategy to detect translocations in multiplex PE CD34+ HSPCs. (**J**) Circos plot displaying UDiTaS results for KIT^{H378R}, $BCL11A^{+55/+58}$ triplex BE HSPCs. Samples in which a translocation product was detected (1% threshold) are highlighted by a red outline. N = 3 biological replicates (independent donors).



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Quantitative estimation of PE-induced translocations by digital-droplet PCR (ddPCR). (**A**) Schematic overview of the ddPCR strategy used to detect and quantify chromosomal translocation products involving the KIT and BCL11A loci. Translocations were detected through custom ddPCR assays using a FAM-fluorescent probe. TTC5 assay (HEX fluorescence) was used as internal genomic normalizer. (**B**) Genomic locations of on-target primers at the KIT and BCL11A sites, indicating the "short" and "long" amplicon configurations used for qualitative detection of CRISPR-induced translocations. (**C**) Editing conditions (gRNA combinations) evaluated in the initial screening experiment in K562 cells to assess detection of chromosomal translocations between KIT and BCL11A loci. All conditions were edited using CRISPR–Cas9 nuclease, except condition #10, which employed KIT+BCL11A PE3/PE3 editing using PE7. (**D**) Agarose gel electrophoresis of PCR products generated using long-amplicon primer combinations (indicated on the left) across the conditions shown in (C); the editor used (SpCas9 nuclease or PE7) is indicated below each condition. Uncropped gels are reported in Supplementary Fig. 5. (**E**) Agarose gel electrophoresis of PCR products generated using short-amplicon primer combinations (indicated on the left) across the conditions shown in (C); the editor used (SpCas9 nuclease or PE7) is indicated below each condition. Uncropped gels are reported in Supplementary Fig. 5. (**F**) Editing conditions evaluated in edited CD34+ HSPCs to assess detection of KIT–BCL11A translocations in primary cells. Conditions #1–4 were edited with PE7, whereas conditions #6–14 were edited using CRISPR–Cas9 nuclease. Positive-control translocation-generating conditions were performed in triplicate. (**G**) Agarose gel electrophoresis of PCR products generated using short-amplicon primer pairs (indicated on the left) across the conditions shown in (**F**); the editor used (SpCas9 nuclease or PE7) is indicated below each condition. PCR experiments reported in panels E-G were repeated twice (on K562 cells and CD34+ HSPCs on 3 donors). Uncropped gels are reported in Supplementary Fig. 5. (**H**) ddPCR quantification of KIT–BCL11A translocations in edited K562 cells used as positive controls. Primer combinations used for each assay are indicated above bar groups, with the green dot indicating the location of the fluorescent probe (see also Supp. Table 2), and corresponding gRNA/editor combinations are shown below the plot. (**I**) ddPCR quantification of KIT–BCL11A translocations in edited CD34+ HSPCs. Primer combinations and corresponding gRNA/editor conditions are indicated below the plots. Mean ± SD. N = 3 donors.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Secondary transplantation of epitope-edited HSPCs after non-genotoxic hematopoietic replacement. (**A**) Human chimerism assessed by flow cytometry (hCD45+ w/in total CD45+) in the BM of secondary recipients of the experiment in Fig. 5a. Mean ± SD. N = 3 mice per group. (**B**) Absolute counts of hCD45+ cells within the BM of secondary recipients transplanted with BM cells from the experiment in Fig. 5a. Mean ± SD. N = 3 mice per group. (**C**) Absolute counts of human cells within the BM of secondary recipients, according to fluorescent protein marking. Mean ± SD. N = 3 mice per group. (**D**) Relative composition of human CD45+ cells in the BM of secondary recipients, according to fluorescent protein marking. N = 3 mice per group. (**E**) Relative composition of KIT^{D121L} editing outcomes within myeloid cells (CD33/66+) as visualized by dual staining with SR-1 and 104D2 clones. Mean ± SD. N = 3 mice per group. (**F,G**) Representative flow cytometry plots of myeloid cells (pre-gated on hCD45+CD33/66b+) from the BM of secondary recipients stained with SR-1 and 104D2 anti-KIT mAb clones, according to treatment and editing conditions in the primary mice. Biallelic edited cells are highlighted by the purple gate, monoallelic by the pink gate, and unedited (wild-type) by the navy gate. (**H**) Representative flow cytometry gating strategy for bone marrow analysis of mice xeno-transplanted with edited HSPCs. Cells are pre-gated on singlets, live and physical parameter gates. Gated populations: total CD45+, human CD45+; total CD19+; total myeloid cells (CD33/66b+); monocytes (mono); mast cells (mast); total granulocytes; mature granulocytes (PMN, CD10+); immature granulocytes (iPMN, CD10-); classical dendritic cells (cDC); pro-B cells (proB); pre-B cells (preB); mature B cells; B-prolymphocytes (B-prolymph.); T prolymphocytes (T-prolymph.); myeloblasts; lin-CD34+ (CD33/66+ lin-CD34+ and CD33/66- lin-CD34+ are pooled for downstream gating); pre-B/NK; granulo-mono progenitors (GMP); common myeloid progenitors (CMP); megakaryocyte-erythroid progenitors (MEP); hematopoietic stem cells (HSC); multipotent progenitors (MPP); lymphoid-primed multipotent progenitors (LMPP); multipotent lymphoid progenitors (MLP). CountBeads are first gated on singlets as FSC-$A^{low}PI^{high}$ and then with two additional fluorescent parameters (plot in the top right corner). See also Supp.Table 3 for additional information on gating strategy and the list of flow cytometry antibodies used.

nature portfolio

Corresponding author(s): Genovese Pietro

Last updated by author(s): May 5, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Immunophenotypic analyses were performed on FACS Fortessa, Fortessa X-20 (BD Pharmingen) or using BDFACS Diva software. Cell sorting was performed on a BD FACS Melody or BD FACS Aria Fusion (BD Biosciences) using BD FACS Diva software v9.
Data analysis	EditR (Kluesner M, Nedveck D, Lahr W, Moriarity B. EditR: A method to quantify base editing via Sanger sequencing. The CRISPR Journal. 2018.); FCS express v6 (DeNovo Software); GraphPad Prism v9.4; Microsoft Excel 365. R software, R Core Team (2021), R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. For CFU analysis and counting, StemCell Stem Vision software was used (no version available). Code used to analyze CHANGE-seq data can be found at: https://github.com/tsailabSJ/changeseq. Code used to analyze CHANGE-seq-BE datasets can be found at: https://github.com/tsailabSJ/changeseq/tree/BE. The custom code used to conduct base-editing off-target quantification in rhAmpseq datasets is available at: https://github.com/tsailabSJ/MKSR_off_targets and https://github.com/YichaoOU/PE_off_target. The source code to analyze UDiTaS assays is available at https://github.com/editasmedicine/uditas.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All relevant data are included in the manuscript.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	CD34+ HSPCs from mobilized peripheral blood of 8 different, deidentified healthy donors (male and female) were obtained from the Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829). Plerixafor-mobilized peripheral blood from patients with sickle cell disease was procured after Boston Children's Hospital institutional review board (IRB) approval and patient-informed consent.
Population characteristics	De-identified healthy donors and patients affected by sickle cell disease
Recruitment	Cells obtained for research purpose only through commercial sources (Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829)) and Boston Children's Hospital.
Ethics oversight	Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829), Boston Children's Hospital institutional review board (IRB)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were determined by the availability of primary human cells and, for in vivo studies, by engraftment capacity and ethical constraints on animal use. Within these limits, cells were allocated evenly across conditions, and we aimed for ≥3–5 independent biological replicates per group, consistent with standard practice in HSPC editing and xenotransplantation studies. The selected sample sizes enabled estimation of variance and application of appropriate statistical tests. Observed effect sizes were large relative to within-group variability and were reproducible across independent experiments, supporting the adequacy of the sample sizes to detect biologically meaningful differences. No samples were excluded unless predefined technical quality criteria were not met (see below). In some in vivo experiments, such as secondary transplantation, the total number of available cells was more constrained and limited to what could be harvested from the primary recipients.
Data exclusions	No data or sample were excluded from analysis, unless the data point was missing due to technical reasons (sample availability or instrumentation acquisition errors). Outlier exclusion was applied to the in vivo barcode library fully analyzed data (retrived barcodes, Shannon H) and is fully described in the Figure legend/data description. N=1 outlier data point, likely due to sequencing issues. All these criteria were pre-established.
Replication	Number of biological replicates is specified for each experiment in figure legends. All attempts at replication were successful. Inferential techniques were applied in presence of adequate sample sizes (n ≥ 3-5), otherwise only descriptive statistics are reported.
Randomization	Mice were randomly distributed to each experimental group. For in vitro experiments, formal randomization was not applied, as experimental groups were defined by predetermined conditions (e.g., editing constructs, antibody treatments) rather than by assignment of interchangeable units. Cells from the same donor or culture were distributed evenly across all conditions and processed in parallel to minimize technical variability. Where applicable, treatments were applied simultaneously using identical reagents and conditions, and sample handling and readouts were performed in a consistent and unbiased manner. For experiments involving multiple donors or independent cultures, biological replicates were included and analyzed across conditions to control for inter-sample variability.
Blinding	Not relevant for objective measures.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	See attached Supplementary Table 3 for the complete list of antibodies
Validation	All primary antibodies were validated for their intended applications (flow cytometry) using a combination of manufacturer-provided validation and in-house testing. Commercial antibodies (BioLegend, BD Biosciences, R&D Systems) were selected based on manufacturer validation for flow cytometry applications, including specificity testing and recommended use conditions as reported on vendor datasheets and technical resources (see, e.g., BioLegend: https://www.biolegend.com; BD Biosciences: https://www.bdbiosciences.com; R&D Systems: https://www.rndsystems.com). These resources document validation of antibodies for flow cytometry and related applications. See In-house validation included titration of each antibody to determine optimal staining concentrations, and testing on human peripheral blood mononuclear cells and/or bone marrow samples from healthy donors to establish gating strategies and confirm expected staining patterns. Where applicable, specificity was further supported by the use of appropriate biological controls (e.g., target-positive and target-negative populations). Antibodies targeting KIT (Fab-79D, 104D2, SR-1, Ab55), which are central to this study, were additionally validated using human or murine cell lines engineered to overexpress the target gene via Sleeping Beauty transposase, confirming target-specific binding.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	K562, HEK-293T cells were a kind gift from the Biffi lab (Dana Farber Cancer Institute, Boston, US). NIH-3T3 were a kind gift from the Brendel lab (Boston Children's Hospital, Boston, US). BaF3 were a kind gift from the Armstrong lab (Dana Farber Cancer Institute, Boston, US). All these cell lines were originally obtained through ATCC
Authentication	None of the cell lines used were authenticated (in this work cell lines were used only to express and study transgenes).
Mycoplasma contamination	All cell lines were tested periodically for mycoplasma contamination and found negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	NOD.Cg-KitW-41J Tyr + Prkdcscid Il2rgtm1Wjl/ThomJ female mice (aka. NBSGW, Jax strain #026622). Age of the mice at experiment initiation was 6-8 weeks.
Wild animals	The study did not involve wild animals.
Reporting on sex	Only female mice were included for in vivo experiments due to their superior ability to support the engraftment of human CD34+ hematopoietic stem and progenitor cells (Notta F, Doulatov S, Dick JE. Engraftment of human hematopoietic stem cells is more efficient in female NOD/SCID/IL-2Rgc-null recipients. Blood. 2010;115(18):3704–3707)
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All animal experiments were performed in accordance to regulations set by the American Association for Laboratory Animal Science and an Institutional Animal Care and Use Committee (IACUC) approved protocol (DFCI#21-002).

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cell lines: harvested, washed, incubated with Fc-block reagent and stained at 4C for 15-30 min, then washed. Mouse peripheral blood: collected, lysed with ACK reagent for 10 minutes at room temperature, then washed (repeat 2x lysis and wash); incubated with human and murine Fc-block, stained at room temperature for 15 minutes, then washed. Bone marrow, spleen: collected, washed, lysed with ACK reagent for 5 minutes at room temperature and washed; incubated with human and murine Fc-block, stained on ice for 35 minutes, then washed. BD Brilliant Stain buffer was included when multiple Brilliant Violet or UV dyes were co-stained. Viability staining was performed using Live/Dead yellow, 7-AAD, Propidium Iodide and AnnexinV staining depending on the experiment. All wash steps were performed with PBS + 2% FBS or PBS + 2% FBS + 2mM EDTA.
Instrument	BD Fortessa flow cytometer, BD Fortessa X-20, BD FACS Melody sorter, BD Aria Fusion sorter.
Software	BD Diva (acquisition, sorting), BD Chorus (sorting), FCS Express v6 or v7 (analysis).
Cell population abundance	See Extended Data FIG.5 for gating strategy on mouse BM samples. We also directly report absolute counts of all hematopoietic populations from in vivo experiments in the Main Figures and Source Data.
Gating strategy	See Methods for details. For most analyses, gating strategy included FSC-A vs FSC-H to exclude doublets (diagonal gating), followed by viability through exclusion of PI/7AAD+ events, followed by physical parameters to detect cells with appropriate physical features according to the specific sample being analyzed (FSC-A vs SSC-A). Subsequent gating is experiment-specific and detailed in Supplementary Data Table 2, Extended Data Figure 1H, 7H and Supplementary Data FIG.6-7.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.