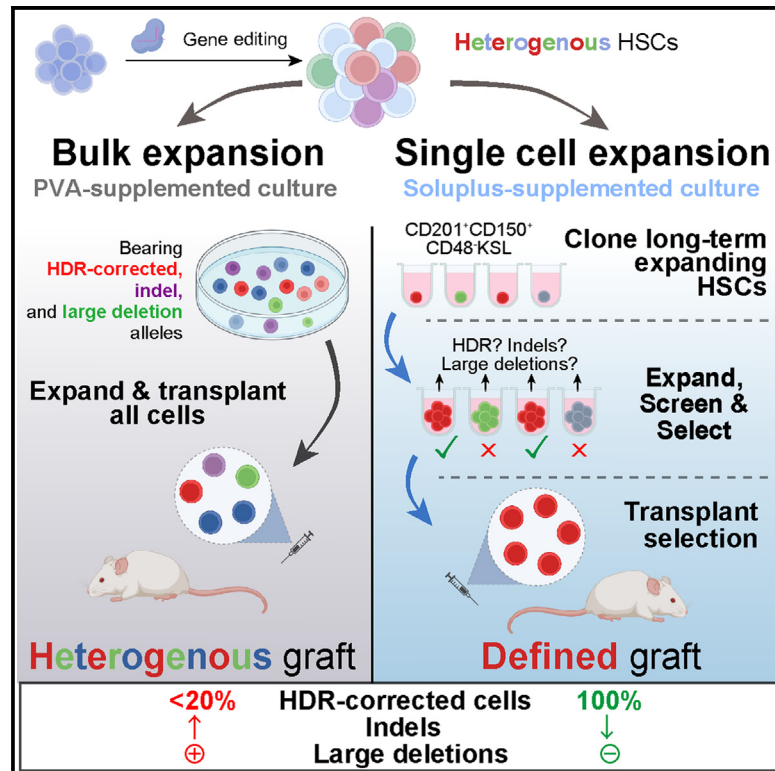# Controlling genetic heterogeneity in gene-edited hematopoietic stem cells by single-cell expansion

Graphical abstract

Authors

Hans Jiro Becker, Reiko Ishida,
Adam C. Wilkinson, ..., Arinobu Tojo,
David G. Kent, Satoshi Yamazaki

Correspondence

jiro.becker@md.tsukuba.ac.jp (H.J.B.),
y-sato4@md.tsukuba.ac.jp (S.Y.)

In brief

Becker and colleagues leverage a single-cell expansion system to assemble genetically defined grafts from heterogeneous gene-edited murine hematopoietic stem cell (HSC) cultures. The expansion platform presented here enables direct selection of *bona fide* HDR-corrected HSCs and the exclusion of clones bearing collateral modifications (indels, large deletions) from transplantation.

Highlights

- Gene editing frequently produces unwanted mutations in transplanted HSC grafts

- *Ex vivo* expanding HSCs are enriched in the CD201$^+$CD150$^+$CD48$^-$KSL population

- A polymer-supplemented culture system enables single-cell expansion of edited HSCs

- Expanded HSC clones can be screened to exclude specific clones from transplantation

CelPress

Cell Stem Cell

Resource

Controlling genetic heterogeneity in gene-edited hematopoietic stem cells by single-cell expansion

Hans Jiro Becker,[1,2,*] Reiko Ishida,[2] Adam C. Wilkinson,[3] Takaharu Kimura,[1] Michelle Sue Jann Lee,[4] Cevayir Coban,[4] Yasunori Ota,[5] Yosuke Tanaka,[6] Meike Roskamp,[7] Tsubasa Sano,[8] Arinobu Tojo,[9] David G. Kent,[10] and Satoshi Yamazaki[1,2,11,*]

[1]Laboratory for Stem Cell Therapy, Faculty of Medicine, Tsukuba University, Tsukuba 305-8577, Japan
[2]Division of Stem Cell Biology, Center for Stem Cell Therapy, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan
[3]MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DS, UK
[4]Division of Malaria Immunology and International Vaccine Design Center, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan
[5]Department of Pathology, Research Hospital, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan
[6]International Research Center for Medical Sciences, Kumamoto University, Kumamoto City 860-0811, Japan
[7]Pharma Solutions, Nutrition & Health, BASF SE, Carl-Bosch-Strasse 38, 67056 Ludwigshafen am Rhein, Germany
[8]Pharma Solutions, Nutrition & Health, BASF Japan Ltd, Tokyo 103-0022, Japan
[9]Tokyo Medical and Dental University, Tokyo 113-8510, Japan
[10]York Biomedical Research Institute, Department of Biology, University of York, Wentworth Way, York YO10 5DD, UK
[11]Lead contact
*Correspondence: jiro.becker@md.tsukuba.ac.jp (H.J.B.), y-sato4@md.tsukuba.ac.jp (S.Y.)
https://doi.org/10.1016/j.stem.2023.06.002

SUMMARY

Gene editing using engineered nucleases frequently produces unintended genetic lesions in hematopoietic stem cells (HSCs). Gene-edited HSC cultures thus contain heterogeneous populations, the majority of which either do not carry the desired edit or harbor unwanted mutations. In consequence, transplanting edited HSCs carries the risks of suboptimal efficiency and of unwanted mutations in the graft. Here, we present an approach for expanding gene-edited HSCs at clonal density, allowing for genetic profiling of individual clones before transplantation. We achieved this by developing a defined, polymer-based expansion system and identifying long-term expanding clones within the CD201$^+$CD150$^+$CD48$^-$c-Kit$^+$Sca-1$^+$Lin$^-$ population of precultured HSCs. Using the *Prkdc*^{scid} immunodeficiency model, we demonstrate that we can expand and profile edited HSC clones to check for desired and unintended modifications, including large deletions. Transplantation of *Prkdc*-corrected HSCs rescued the immunodeficient phenotype. Our *ex vivo* manipulation platform establishes a paradigm to control genetic heterogeneity in HSC gene editing and therapy.

INTRODUCTION

The rapid adoption of engineered nucleases has put hematopoietic stem cells (HSCs) at the center of gene editing applications. The ability to functionally interrogate genes by introducing or correcting mutations at precise loci has greatly advanced our understanding of HSC biology and has enabled curative approaches for genetic diseases. CRISPR-Cas9 currently represents the most widespread system for gene editing of the hematopoietic system.[1] A target-specific guide RNA (gRNA) directs the Cas9 endonuclease to a genomic site of interest, where it induces a DNA double-strand break (DSB). The subsequent engagement of the cell-intrinsic DNA damage repair (DDR) machinery can be exploited to create targeted modifications in HSCs.[2–4] Since mutagenic repair (e.g., non-homologous end joining [NHEJ]) takes precedence in primitive HSCs,[5] a phenomenon closely tied to their dormant phenotype, random small in-

sertions and deletions (indels) represent the most common editing outcome.[2,6] Furthermore, a string of recent reports have uncovered previously underappreciated lesions, such as kilo- and megabase-scale deletions as well as chromothripsis, illustrating the potential risk of Cas9-based gene editing.[7–9] In contrast, correction via templated repair (i.e., homology-directed repair [HDR]) among long-term (LT)-HSCs remains inefficient.[5,10] Off-target mutations may also raise concerns about genotoxicity in edited cells.[11] Together, these unwanted mutations may confound the effects of the targeted gene edit and represent incalculable risks in basic and translational research settings. Apart from gene editing, maintaining HSC self-renewal in LT cultures required by gene editing protocols remains challenging.[12] Consequently, edited and bulk-expanded HSC cultures contain genetically and functionally heterogeneous populations and only include a low fraction of functional HSCs with the desired genetic modifications.

Expansion of single, *bona fide* HSCs would overcome this limitation by enabling direct profiling of on- and off-target editing outcomes, allowing for selective transplantation only of clones with a defined mutational pattern. However, current protocols do not allow for expansion of HSCs at clonal density to the extent necessary for transplantation. Clonal expansion technologies for embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) have been major drivers for advances in the biology and translational research of PSCs, yet the generation of functional HSCs from PSCs remains a major hurdle.

We recently reported on a serum-free, polyvinyl alcohol (PVA)-based HSC expansion protocol that permits up to 899-fold expansion of HSCs over a period of 4 weeks.[13] Here, we use this protocol to show that bulk expansion produces a genetically heterogeneous graft with on- and off-target indels as well as large deletions (LDs). Addressing this issue, we present a system that supports single-cell expansion of edited HSCs and define a phenotype that assists in selecting precultured clones with LT expansion potential. We apply this system to a gene correction model of severe combined immunodeficiency (SCID), demonstrating the feasibility of single-cell expansion for sequence-based selection of edited HSC clones. Lastly, we use a human hemoglobin beta (HBB) editing model to underline the conceptual advantage of clonal expansion in a translational context.

RESULTS

Gene-edited HSCs correct Prkdc[scid] immunodeficiency but bear on- and off-target indels

The immunodeficient phenotype in CB17/SCID mice is caused by a T to A mutation in the *Prkdc* gene (*Prkdc*[scid]), p.Y4046X), leading to functional loss of its product, DNA-dependent protein kinase catalytic subunit (DNA-PKcs, Figure 1A).[14] DNA-PKcs is indispensable for the resolution of DNA DSBs during V(D)J recombination, which is reflected in the absence of functional B and T cells in CB17/SCID mice.

To determine whether the *Prkdc*[scid] phenotype can be corrected with gene-edited and bulk-expanded HSCs and to assess the levels of indels generated in the process, we designed a gene-editing protocol based on our previously established *ex vivo* HSC expansion platform (Figure 1B).[13] CD201[+]CD150[+]c-Kit[+]Lin[−] cells from CB17/SCID mice were cultured in PVA-based medium (PVA-HSC) for 3 days (Figure S1A). We included CD201 (endothelial protein C receptor, EPCR) in our isolation panel since the commonly employed marker stem cell antigen 1 (Sca-1) is known to be poorly expressed on hematopoietic cells of non-C57BL/6 mouse strains and because CD201 has shown to be a reliable marker in BALB/c mice, from which the CB17/SCID strain is derived.[15] Cas9 ribonucleoprotein (RNP) complexes and a corrective HDR template were delivered into HSCs 3 days after isolation. 1 week after editing, the majority of alleles contained indels (*Prkdc*[indel], 48%), while 26% had incorporated the HDR donor sequence (*Prkdc*[HDR], Figure 1C). The HDR-corrected fraction increased over the course of the culture, likely reflecting a selective advantage of *Prkdc*-proficient cells over indel-bearing cells. HDR frequencies were lowest (11% ± 2%) in the most stringently defined HSC population (CD201[+]CD150[+]KL) and increased in fractions with lower HSC

enrichment, in line with previous reports (Figure S1B).[2] 1 week post-editing (day 10 of culture), most cells in the expansion cultures remained c-Kit[+] and Lin[−], with a majority also expressing CD150 (Figure 1D). Although the initial starting population of CD201[+]CD150[+]KL cells represented only 14.7% of expanded HSCs, absolute quantification revealed an 8.9-fold expansion (Figure 1E).

To validate functional recovery of edited SCID HSCs, we transplanted 0.5×10^6 expanded bulk HSC cultures into irradiated CB17/SCID recipients 7 days post-editing. B220[+] B cells as well as CD4[+] and CD8[+] T cells could be detected in peripheral blood (PB) samples from 4 weeks post-stem cell transplantation (SCT) (Figure 1F). Spleens of transplanted mice contained high fractions of B and T cells (B220[+]: 36%, CD4[+]: 16%, CD8[+]: 6% of splenocytes, Figure S1C). We further found that thymocytes of transplanted mice were abundant with CD4[+]CD8[+] double positive (DP), CD4[+], and CD8[+] single positive (SP) cells (Figure S1D) and thymus histology showed cortical and medullary regions (Figure S1E). The distributions of lymphocyte populations in the spleen and thymus were similar to those in age matched CB17/WT mice, suggesting orthotopic development of B and T lymphocytes. Secondary transplantations confirmed that LT-HSCs had been successfully edited in our gene correction model (Figure 1F).

As mentioned above, the high frequency of indel and SCID alleles in the transplanted HSPC population is a key limitation of this straightforward bulk expansion approach (Figure 1C). To check how this distribution was reflected in mature cell lineages, we sequenced PB cells 16–20 weeks post-SCT. As expected, *Prkdc*[HDR] frequencies were high in lymphocytes (B220[+]: 69%, CD4[+]: 70%, CD8[+]: 63%), suggesting at least monoallelic correction in these populations (Figure 1G). Since noncorrected cells fail to complete lymphocyte development in this model, their high prevalence in the transplanted graft did not obstruct the rescue of these mature compartments. By contrast, myeloid cells, which are not subject to the same selective pressure, showed a high rate of on-target indels (*Prkdc*[indel], 52%) and a low frequency of *Prkdc*[HDR] alleles (14%, Figure 1G). Off-target analysis of bulk-expanded HSCs showed a low but substantial prevalence of indels (Figure 1H; Table S1). Recent studies suggest that Cas9 gene-editing produces large on-target deletions that are not captured by conventional amplicon sequencing.[7] To check for these LDs, we adopted a droplet digital PCR (ddPCR)-based approach to assess the copy numbers of five regions spanning 1 kb around the cut site (Figure S1F).[16] Indeed, we detected a significant drop of copy numbers not only directly around the cut site, but also several hundred bp away, indicative of LD events in the bulk-expanded population (Figure 1I).

While gene-edited HSCs effectively reversed the *Prkdc*[scid] phenotype, these results indicate that most transplanted HSCs and their progeny contained unintended perturbations. The low allelic chimerism of *Prkdc*[HDR] and high abundance of *Prkdc*[indel] among myeloid cells demonstrate the challenge of ensuring that all hematopoietic cells are supplied by a genetically defined population of edited HSPCs. The potentially negative consequences of on-target indels has also recently been highlighted in other gene correction models.[17] This drawback inspired us to establish a single-cell HSC expansion system that would allow selection of edited, *bona fide* HSCs at the clonal level.
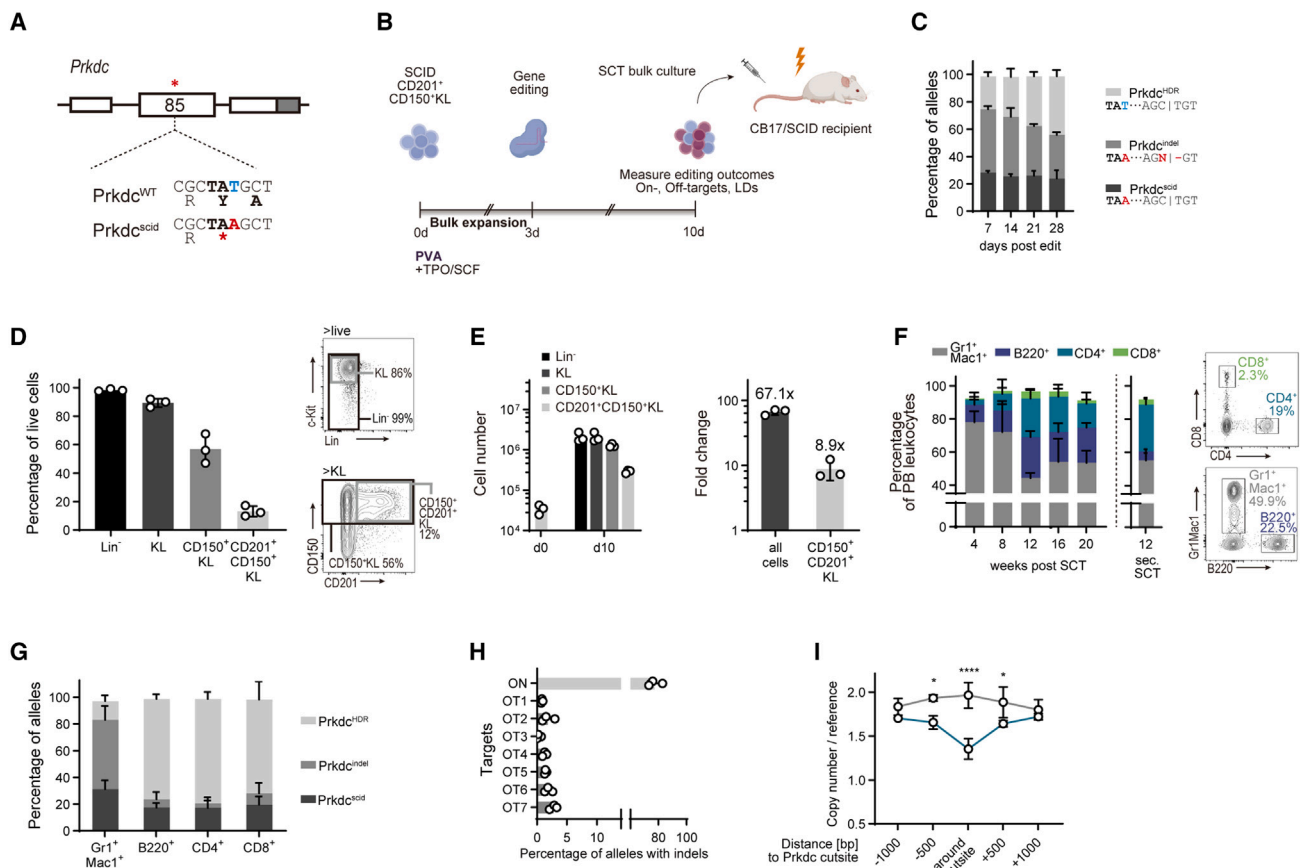
Figure 1. Autologous HSCT gene correction rescues the *Prkdc*^scid phenotype but introduces on-, off-target indels and large deletions

(A) Genomic context of the *Prkdc*^scid mutation in exon 85. White boxes: exons, gray box, 3′ UTR. * denotes location of *Prkdc*^scid mutation.

(B) Experimental scheme of the gene editing and HSC expansion model.

(C) Post-editing allele distribution at the *Prkdc* locus, assessed by inference of CRISPR edit (ICE) (n = 3 cultures).

(D) Fractions of immunophenotypically defined HSPC populations within cultures on day 10 of culture, 7 days post-editing. Percentage of all live cells (n = 3 cultures).

(E) Absolute cell numbers (left panel) and fold-change expansion (right panel) of cultured HSPCs, day 10 of culture.

(F) Left: frequencies of peripheral blood (PB) leukocytes as percentage of all live leukocytes (n = 3 groups, 3–4 mice per group). Plot next to dashed line shows frequencies 12 weeks post-secondary SCT (n = 5 mice). Right: representative fluorescence-activated cell sorting (FACS) plots 20 weeks post-transplant.

(G) Frequencies of *Prkdc* alleles in sorted PB cells 20 weeks post-SCT (n = 3 experiments, 3–4 mice per group).

(H) On- and off-target (OT) activity of the *Prkdc*-specific gRNA, assessed with tracking of indels by decomposition (TIDE). The seven highest scoring off-target sites, as predicted by COSMID, were interrogated. See Table S1 for detailed information about the off-target sites.

(I) Copy-number analysis of *Prkdc* probes against reference gene (n = 3). Two-way ANOVA with Sidak's multiple comparison test.

Error bars represent SD. *p < 0.05, ****p < 0.0001.

CD150⁺CD201⁺CD48⁻KSL cells contain clones with long-term expansion potential

Single-cell expansion of edited HSCs requires the identification of clones with prospective LT expansion potential within a population of precultured cells. Cell surface markers are particularly useful since they permit flow cytometric profiling and simultaneous cloning via fluorescence-activated cell sorting. However, HSC marker expression undergoes a dynamic shift over the course of *ex vivo* expansion.[18,19]

To address this issue, we leveraged index sorting analysis to identify HSC markers that predict LT expansion of HSC clones. Fresh CD34⁻CD150⁺c-Kit⁺Sca-1⁺Lin⁻(CD34⁻CD150⁺ KSL) HSCs were cultured for 10 days, after which KSL cells were subjected to index sorting. We used C57BL/6-derived HSCs for these experiments, since this was the background of mice used to optimize our HSC culture system and is widely used in the field.[13] Marker profiles of each sorted KSL cell were compared with HSC colony formation after 14 days. Expression of a total of six HSC markers within the KSL population, divided into two panels (CD34, CD48, and CD105; as well as CD135, CD150, and CD201), were evaluated (Figure 2A). Colony formation was observed in 17.1% of sorted KSL clones (set 1: 16.2%, set 2: 19.1%), mainly from clones within the CD48⁻, CD150⁺, and CD201⁺ KSL populations (Figure 2B). Quantification of expression levels confirmed significantly higher expression of CD150 and CD201 as well as lower expression of CD48 among colony-forming HSCs (Figures 2C and 2D). CD135 expression was lower in colony-forming HSCs (Figure 2C);
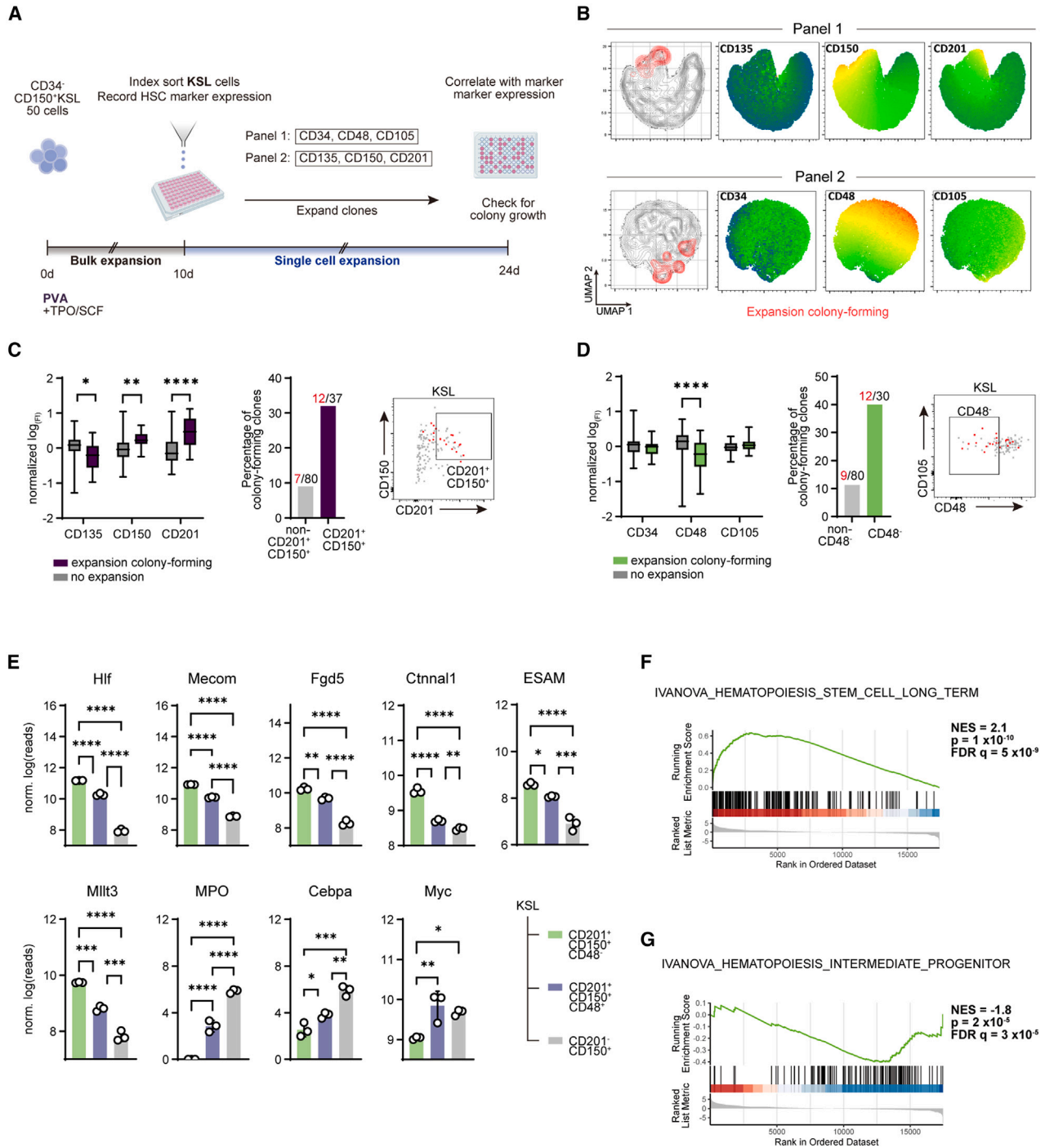
Figure 2. Identification of a surface marker combination for long-term (LT) expanding HSC clones

(A) Experimental setup.

(B) Uniform manifold approximation and projection (UMAP) representation of sorted KSL clones with overlay of panel 1 (upper) and 2 (lower) surface markers. Expansion colony-forming clones are indicated in red.

(C and D) Quantification of markers associated with colony expansion. Left: fluorescence intensity (FI) measured at index sorting. Data presented as log-transformed and normalized to mean. Boxplots with whiskers showing minimum and maximum. Center: fraction of clones of the indicated phenotype showing LT expansion. Right: representative FACS sorting plots, LT-expanding clones indicated in red. (C) Panel 1 (n = 110 clones); (D) panel 2 (n = 117 clones). Multiple Mann-Whitney tests with FDR correction.

(E) RNA-seq expression profiles of select HSC- and progenitor-associated genes. Error bars represent SD. One-way ANOVA with Tukey's post-test.

(F and G) Gene set enrichment analysis (GSEA) of differentially expressed genes in CD201+CD150+CD48−KSL (F) and CD201−CD150+KSL (G) cells.

*p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

however, the small absolute difference in expression precludes the use of this marker for effective gating.

We next performed RNA sequencing to characterize the populations defined by these markers in 10-day bulk-expanded HSPCs (Figure S2A). Comparing global expression profiles, we found the greatest difference between CD201+CD150+CD48− and CD201−CD150+KSL cells, with CD201+CD150+CD48+KSL cells representing an intermediary phenotype (Figures S2B and S2C). This representation was mirrored in the expression profiles of canonical genes related to hematopoiesis: transcripts of HSC-associated genes, such as *Hlf*, *Mecom*, and *Fgd5*, were more abundant in CD201+CD150+CD48−KSL cells, whereas downstream progenitor-associated genes (*MPO*, *Cebpa*) were upregulated in CD201−CD150+KSL cells (Figure 2E). Enrichment analysis confirmed that the transcriptional phenotype of CD201+CD150+CD48−KSL cells was similar to that of LT-HSCs (Figure 2F), while CD201−CD150+KSL cells were similar to progenitor cells (Figure 2G). GO term enrichment pointed to a proliferating state of CD201−CD150+KSL cells, with several enriched mitosis- and translation-related pathways (Figure S2D), matching our previous observation that progenitor cells proliferate more rapidly than primitive HSCs in culture (Figure 1E).

These results indicate that CD150+CD201+CD48−KSL cells possess LT expansion potential and retain a transcriptional phenotype associated with *bona fide* HSCs after extended culture. We thus considered these cells suitable for single-cell cloning and expansion. However, we found that the CD201+CD150+CD48− expression profile was lost in single-clone-derived colonies generated from this population after 14 days, suggesting that repopulating activity had been compromised (Figure S2E). Since we have shown previously that HSC marker expression is preserved in bulk cultures even after 28 days of expansion,[13] we reasoned that the single-cell cloning step and expansion conditions, rather than the total length of *ex vivo* expansion, were not supported by our expansion system.

Soluplus is a superior alternative to PVA for single-cell HSC expansion

Having established that HSC activity among bulk cultured cells is enriched in the CD150+CD201+CD48−KSL population but that PVA-based culture conditions poorly supported their clonal expansion after re-sorting, we sought to improve clonal expansion culture conditions by screening alternative serum replacement compounds. We cultured 50 freshly isolated CD34−KSL cells in media supplemented with recombinant albumin, PVA and 7 different polymers and evaluated cell growth after 1 week. Of all compounds tested, only Soluplus led to comparable levels of proliferation as PVA and recombinant albumin (Figure S3A). Soluplus is an amphiphilic polyvinyl caprolactam-acetate polyethylene glycol (PCL-PVAc-PEG) graft copolymer approved for clinical use as a drug solubilizer.[20] We hypothesized that Soluplus, like PVA,[21] enhances the stability of cytokines in the culture medium. Indeed, we found that thrombopoietin (TPO) levels were higher in 3-day cultures if Soluplus is present compared with plain or PVA-supplemented medium (Figure S3B). To identify the most suitable concentration for HSC culture, we performed transplantations with HSCs grown in titrated concentrations of Soluplus. Although 16-week chimerism was highest in the 0.2% cohort, (Figure S3C), we selected

0.1% Soluplus for our expansion system since supplementation with 0.2% Soluplus led to mild precipitation during culture, obscuring the visibility of cells.

To validate the cell growth supporting properties of Soluplus, we directly compared single HSC expansion conditions using Soluplus and PVA. Freshly isolated, single CD34−CD150+KSL cells from C57BL/6 mice were cultured in individual wells on 96-well plates. After 19 days, expansion was evaluated by flow cytometric profiling including cell viability and HSC marker expression (Figure 3A). Cell viability was higher in clones cultured in Soluplus-supplemented medium, as measured by propidium iodide (PI) exclusion staining (Figure S3D). Accordingly, the percentage of cloned HSCs forming viable cell colonies (i.e., >20% live cells) was higher under Soluplus expansion conditions (Figure 3B). Furthermore, we found that Soluplus supplementation was associated with a higher retention of HSC marker expression. In particular, the fraction of CD201+CD150+KSL cells was higher in clones cultured in Soluplus-containing medium, suggesting that Soluplus was superior in expanding phenotypically primitive HSCs (Figures 3C and S3E).

We next asked if HSC clones cultured in Soluplus medium produce functional HSC grafts *in vivo*. Freshly isolated CD45.1+−CD34−CD150+KSL HSCs were cloned and cultured for 35 days. Three clones containing 35%, 17%, and 6% CD201+CD150+KSL cells were selected for split-clone transplantation into 10 to 15 CD45.2+ recipients against 5 × 10^5 WBM cells (Figure 3D). All recipients showed multilineage LT engraftment at ≥1% chimerism in PB samples despite the high number of recipients per clone (Figures 3E and 3F), though chimerism declined significantly over time in two of the three groups. This decline was pronounced in those mice receiving grafts with a smaller CD201+CD150+KSL fraction. Secondary transplantations from pooled bone marrow of highly chimeric mice showed stable engraftment of CD45.1+ cells in all recipients (Figures 3E and 3F).

To quantify the potential for single HSC expansion with Soluplus, we performed a limiting dilution assay (LDA) with a CD34−CD150+KSL clone expanded for 28 days (6.37 × 10^5 cells) and containing 84% of CD201+CD150+KSL cells (Figure S3F). We observed multilineage chimerism of ≥1% in all dose groups including from just 10 cells (in 2/5 recipients, Figures S3G and S3H). Based on these results, we estimated a mean HSC frequency of 1/18.9 (confidence interval [CI] 1/6.1–1/61.4) in the culture using extreme limited dilution analysis (ELDA, Figure 3G)[22] and determined that the initial HSC had expanded >33,000-fold (range 10,375- to 104,426-fold corresponding to frequency CIs) under our culture conditions. Thus, our results suggest that Soluplus is superior to PVA in supporting efficient expansion of single HSCs. Importantly, whole-exome sequencing (WES) of a single-clone-derived HSC colony expanded for 28 days did not reveal nonsynonymous mutations in critical genes (Table S2). Based on these encouraging results, we attempted to expand precultured and gene-edited HSC clones.

Soluplus enables single-cell expansion of edited HSCs
To evaluate HSC gene editing and clonal expansion with single allele resolution, we developed a strategy that targets protein tyrosine phosphatase receptor type C (*Ptprc*), a cell surface
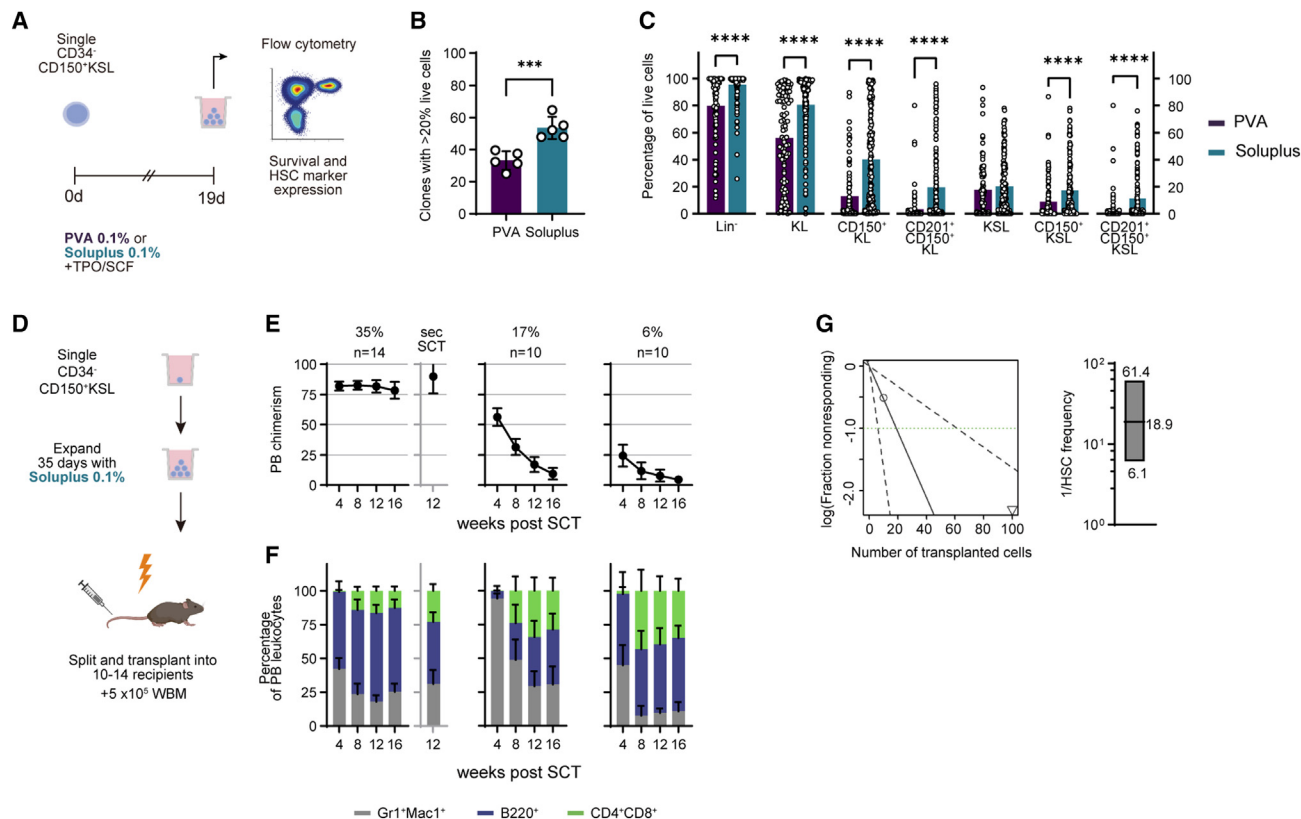
Figure 3. Optimization of polymer-based cultures for single-cell HSC expansion

(A) Scheme of experimental setup.

(B) Percentage of colonies with ≥20% live cells (n = 5 experiments). Unpaired, two-tailed t test.

(C) Percentage of phenotypic HSC populations in live colonies cultured in PVA (n = 94)- and Soluplus (n = 155)-based media. Multiple Mann-Whitney tests with FDR correction.

(D) Schematic of split-clone transplantation.

(E and F) Donor PB chimerism (E) and lineage distribution (F) in 3 recipient groups transplanted with split clones. Numbers over graphs in (E) represent percentage of CD201+CD150+KSL cells in the transplanted clone (%) and the number of recipients (n). Secondary SCT was performed with the group showing highest chimerism, data shown in graph with gray axis.

(G) Left: ELDA output of HSC frequency calculation. Right: boxplot represents calculated reciprocal mean, upper, and lower limits of HSC frequency.

See also Figure S3 and Table S2.

Error bars represent SD. ***p < 0.001, ****p < 0.0001.

protein. Two alleles of the *Ptprc* gene are common among major inbred mouse strains: *Ptprc*[a] and *Ptprc*[b], which code for CD45.1 (Ly5.1) and CD45.2 (Ly5.2), respectively. The CD45.1 allele is expressed in SJL/J and STS/A strains, while C57BL/6 and BALB/c strains share the CD45.2 allele.[23] Sequence diversion between these two alleles amounts to 12 base differences resulting in 5 amino acid substitutions.[24] The epitope of CD45.1- and CD45.2-binding antibody clones A20 and 104 is defined by a single-base difference at codon 302 (based on reference transcript GenBank:NM_001111316.2, Figure 4A).[25] We leveraged the low complexity of this single-nucleotide polymorphism (SNP) to simultaneously identify and clone gene-edited HSCs, followed by single-cell expansion and transplantation (Figure 4B). To this end, we knocked in the CD45.1-specific SNP variant (A → G, p.K302E) into the *Ptprc* gene of CD45.2+ HSCs (Figure S4A).

CD34−CD150+KSL HSCs from CD45.2+ C57BL/6 mice were cultured in Soluplus-HSC expansion medium for 3 days, after

which we targeted *Ptprc* for allele conversion. 4 days after editing, 20% (±5.8%) of cells had converted to the CD45.1+CD45.2− phenotype (Figure S4B). As with the SCID model, conversion rates were lower in the primitive CD201+CD150+KSL fraction (Figure S4C). We started 570 single-cell cultures from the CD45.1+CD201+CD150+KSL population. 14 days later, cell proliferation could be observed in, and appropriate flow cytometric data could be obtained from, 46% of all sorted clones (261/570). Surface marker expression was heterogeneous, with 24% (63/261) containing at least 10% of CD201+CD150+KSL cells (Figure 4C). Fifty-one colonies were selected for transplantation into single CD45.2+ recipients. Donor chimerism of ≥5% was observed in 29/51 (57%) and 17/51 (33%) recipients 4 and 16 weeks after transplantation, respectively (Figures 4D and S4D). Among the recipients showing LT chimerism, multilineage reconstitution (myeloid, B cell, and T cell lineages ≥5% of donor hematopoiesis) was observed in 8 recipients (16% of recipients, Figure 4E), while the remaining mice showed biased donor
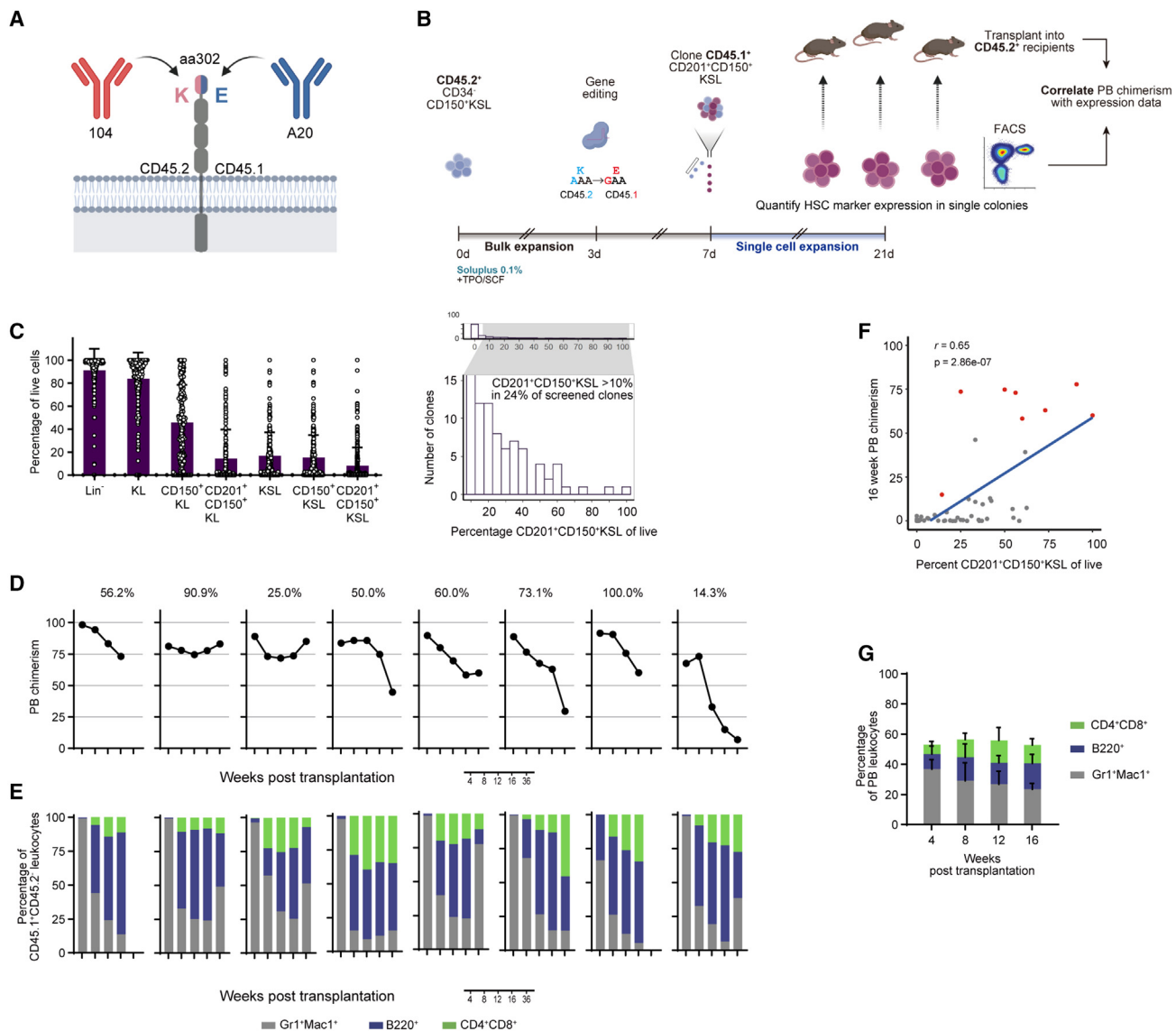
Figure 4. Single-cell cloning of gene-edited functional HSCs

(A) Schematic showing the extracellular domain of CD45 with allele-specific antibody clones 104 and A20 and the epitope-defining amino acid.

(B) Experimental setup of the single-cell editing and expansion experiment.

(C) Left: fractions of CD201+CD150+KSL cells in single-cell-derived cultures 14 days after cloning (n = 261 clones). Right: Histogram of CD201+CD150+KSL cell frequency. Zoomed-in region shows clones with >10% CD201+CD150+KSL cells.

(D and E) CD45.1+ donor PB chimerism (D) and lineage distribution (E) in single recipients with long-term (LT) engraftment ≥5% and multilineage reconstitution (n = 8). Numbers over graphs in (D) represent percentage of CD201+CD150+KSL cells in the transplanted clone (%).

(F) Linear correlation plots of CD201+CD150+KSL cell frequency and 16-week donor chimerism. Red dots indicate LT repopulating and multilineage clones. Pearson correlation.

(G) CD45.1+ PB chimerism and lineage distribution in secondary recipients (n = 5).

See also Figure S4 and Table S3.

Error bars represent SD.

hematopoiesis (Figure S4E). Although myeloid contribution seemed relatively low in some LT-engrafted recipients (Figure 4E), a comparison with the co-transplanted competitor graft revealed similar lineage contributions, suggesting an assay-specific phenomenon rather than a linage bias of the expanded HSC clones (Figure S4F). Linear correlation analysis of pre-SCT marker expression and 16-week chimerism revealed several parameters associated with LT engraftment, the strongest of which was the fraction of CD201+CD150+KSL cells in the transplanted graft (Figures 4F and S4G). Secondary transplantations were performed with whole bone marrow cells from a highly chimeric primary recipient. Analysis of bone marrow cells revealed high chimerism of 77% within the KSL population (Figure S4H). 16 weeks after secondary transplantation, multilineage PB donor

chimerism was observed in all secondary recipients (Figure 4G). WES analysis of gene-edited CD45.1$^+$CD45.2$^-$ cells obtained from the BM of these secondary recipients did not indicate non-synonymous variants in oncogenic driver genes (Table S3).

Together, these results established that HSCs can be gene edited and clonally expanded while maintaining their self-renewal properties using our expansion system. Our experiments also confirm the expression of CD201 and CD150 on expanded clones as predictive of LT engraftment. This approach therefore provided the framework for probing single HSC clones for on- and off-target edits prior to transplantation.

Single-cell expansion of edited HSCs permits the assembly of a genetically defined HSC graft

To explore this approach, we adopted our single-cell expansion platform to the SCID immunodeficiency model. Analogous to our previous experiments, we expanded SCID HSCs in Soluplus-supplemented expansion medium and edited SCID HSCs to correct the $Prkdc^{scid}$ mutation. After editing, CD201$^+$CD150$^+$CD48$^-$KL cells were cloned by flow cytometry and expanded for 14 days (Figure 5A). Since expression of CD201 and CD150 was predictive of LT engraftment, we first screened for colonies containing a CD201$^+$CD150$^+$KL population of over 10% and then checked for the presence of the corrected allele ($Prkdc^{HDR}$) and absence of off-target mutations. Candidate clones were then combined and administered to a SCID recipient (Figure 5A).

Phenotypic profiling data could be obtained from 19% (384) of sorted clones (Figure 5B). Of these, 26% (99/384) contained a population of CD201$^+$CD150$^+$KL HSCs $\geq$10%, which we classified as transplantable clones. Correlating HSC marker expression on the founder cell with the outcome of the expansion cultures, we observed that CD201^{hi}/CD150^{hi} cells were more likely to, and that CD105$^-$ cells did not, generate transplantable clones (Figure S5A). Sequencing of all intended loci ($Prkdc$, off-targets #1–7) could be achieved in most of these clones (96/384). We detected $Prkdc^{HDR}$ in 57% (55/96) of genotyped clones, and all corrected clones were free of off-target mutations at predicted sites (Figures 5C and S5B). As a result, an average of 18 HDR$^+$Off-target$^-$ colonies were selected for transplant per experiment. Due to this selection step, the combined allelic composition of the selected clones was dominated by $Prkdc^{HDR}$ alleles (67%, Figure 5D). This stands in contrast to our bulk-transplant approach, in which indel alleles were most abundant (Figure 1C). In PB samples from transplanted SCB17/SCID mice, we detected B and T lymphocytes from week 4 through week 20 post-transplant, confirming LT engraftment (Figures 5E and S5E). In contrast to our bulk-transplant experiments, the $Prkdc^{HDR}$ allele was highly prevalent not only in lymphoid, but also in myeloid cells (>60%, Figure 5F). Notably, $Prkdc^{indel}$ frequency was low in all PB lineages.

Having achieved robust reconstitution of lymphoid cells, we asked if correction of the $Prkdc^{scid}$ allele also led to development of a functional immune system. Double (CD4$^+$CD8$^+$) and single (CD4$^+$ and CD8$^+$) -positive cells were detected among thymocytes of SCID recipients (Figure S5C). Length diversity of the third complementary determining region (CDR3) in the T cell receptor gene is a direct function of $Prkdc$ activity. We measured CDR3 region length distributions within multiple T cell receptor beta chain ($Tcrb$) gene families of splenic CD4$^+$ T cells using

spectratype analysis.[26] Distribution profiles showed Gaussian distribution patterns without oligoclonal spikes (Figure S5D), suggesting that restored $Prkdc$ activity permitted the generation of unbiased CDR3 regions. To confirm the status of immune cell function *in vivo*, we immunized transplanted CB17/SCID mice with a T-dependent antigen, nitroiodophenyl (NIP)-conjugated OVA (NIPOVA), intraperitoneally (i.p.).[27] NIP$_{30}$-specific IgG and IgM levels were significantly elevated in the serum of transplanted CB17/SCID mice 19 days post-immunization (p.i.), indicative of a specific humoral immune response (Figure S5F). On the other hand, untreated CB17/SCID mice showed no specific response. To assess cellular immunity, we inoculated mice with the human lung cancer cell line A549, hypothesizing that reconstitution of immunity would trigger xenograft rejection. Indeed, rejection was observed in transplanted mice only, whereas progressive tumor growth was detected in non-transplanted SCID mice (Figure 5G). We therefore conclude that the molecular and functional hallmarks defining the $Prkdc^{scid}$ phenotype had been reversed. These results illustrate that a functional graft can be assembled from individually expanded and profiled gene-edited HSC clones.

As demonstrated above (Figure 1I), $Prkdc$ editing frequently results in the production of LDs around the cut site. At the single-clone level, these events lead to loss of heterogeneity (LOH) and may remain undetected upon genotyping if one of the priming sites is involved. Consequently, some HSC clones will appear to contain two HDR-corrected alleles (homozygous correction, $Prkdc^{HDR/HDR}$) but will actually harbor only one ($Prkdc^{HDR/\Delta}$, Figure S5G). To confirm that our system allows for the detection of these clones, we sequence-genotyped 173 Prkdc-edited colonies containing $\geq$10% CD201$^+$CD150$^+$KL HSCs. Of these, 29 (17%), produced a single, HDR-corrected sequencing trace (Figure 5H). Subsequent quantification of alleles using the Prkdc-0.0 kb probe revealed that 16 of the clones harbored two copies (HDR/HDR), whereas 13 clones were indeed hemizygous (HDR/Δ), indicative of LDs (Figures 5I and S5H). Thus, our approach enables the detection and exclusion of clones containing LDs in gene-edited HSC colonies.

Controlling genetic heterogeneity in gene-edited human HSCs

To demonstrate the potential of our clonal expansion concept in the human setting, we edited healthy cord blood (CB)-derived CD34$^+$CD45RA$^-$ human HSCs (huHSCs) at the *HBB* locus to knock in a two-base mutation, creating a sickle cell disease (SCD)-like genotype (Figure S6A). Analogous to our murine protocol, single cells were cloned and cultured for 2 weeks followed by genetic profiling, selection, and xenotransplantation (Figure 6A). Since *bona fide* HSCs cannot be expanded from single huHSCs using currently available culture protocols, we adopted a single-cell differentiation and proliferation assay to produce clonally derived HSPC colonies.[28] At the time of cloning, around one third of alleles (30% ± 7%) contained the HDR edit (HBB^{HDR}, Figure 6B). Though not to the extent seen in murine $Prkdc^{SCID}$ HSCs, we did detect copy-number losses at the HBB locus in edited bulk-expanded huHSCs, suggestive of LDs (Figures 6C and S6B). The SCD edit was detected in about half of the sequence-genotyped clones (49%, 65/133) (Figure 6D). Phenotypically, these colonies were dominated by erythro-myeloid
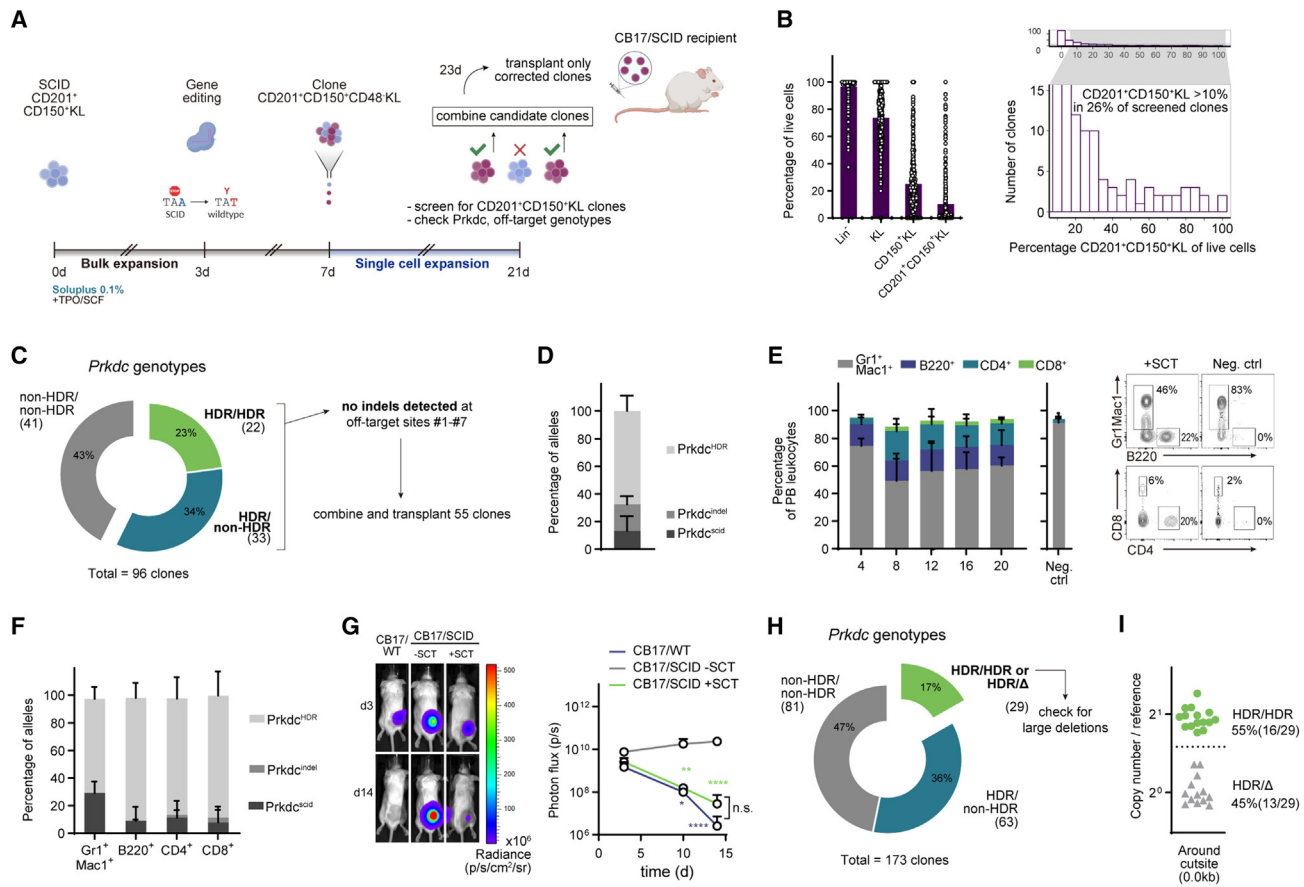
Figure 5. Autologous HSCT using gene-corrected HSC clones is curative in an immunodeficiency mouse model

(A) Schematic of the single clone *Prkdc*[scid] correction model.

(B) Single-cell SCID HSC expansion outcomes. Left: frequencies of phenotypic HSC populations in screened colonies (n = 384 from 3 experiments). Right: histogram of CD201$^+$CD150$^+$KL cell frequency. Enlarged region shows clones with $\geq$10% CD201$^+$CD150$^+$KL cells.

(C) Genotyping of candidate clones (Sanger sequencing) (n = 96 clones, 3 experiments). Only clones with at least one HDR-corrected allele were sequenced at the off-target loci.

(D) Allelic composition of the combined cell mixture at the edited *Prkdc* locus (n = 3).

(E) Frequencies of PB leukocytes in CB17/SCID recipients. Left: lineage distributions in treated mice (n = 3) and in recipients receiving only 2 × 10^5 CB17/SCID whole bone marrow cells (neg. ctrl., n = 3). Right: representative FACS plots at 16 weeks post-SCT.

(F) Allele frequencies in sorted PB cells 20 weeks post-SCT (n = 3 mice from 3 experiments).

(G) Xenograft transplantation assay. A549 cells expressing the luminescent reporter Akaluc were injected subcutaneously (s.c.) and tumor growth was tracked by *in vivo* imaging. Left: representative images from CB17/WT, transplanted CB17/SCID, and untreated CB17/SCID mice 3 and 14 days after inoculation. Right: quantification of luminescence over a 14-day period (CB17/WT: n = 4, CB17/SCID − SCT: n = 3, CB17/SCID + SCT: n = 3). Two-way ANOVA and Tukey's multiple comparison test.

(H) Genotyping of candidate clones (Sanger sequencing) (n = 173 clones, 2 experiments). Clones producing a single, HDR-corrected sequencing trace were checked for LD (n = 29). *Prkdc*[HDR/HDR], homozygous correction; *Prkdc*[HDR/Δ], hemizygosity.

(I) *Prkdc* copy-number analysis using 0.0 kb *Prkdc* (around cutsite) probe, quantified against reference gene.

See also Figure S5.

Error bars represent SD. *p < 0.05, **p < 0.01, and ****p < 0.0001.

---

lineages (Figures S6C ans S6D). We separated these clones into 2 groups: those heterozygous for the SCD edit (*HBB*[HDR/non-HDR]) were selected for immediate xenotransplantation, while clones that produced a single, HDR-corrected sequencing trace (*HBB*[HDR/HDR] or *HBB*[HDR/Δ]) were subjected to copy-number profiling. We found that while the majority of these clones were homozygous, four out of 16 in fact only had one allele, indicating a *HBB*[HDR/Δ] genotype (Figure 6E). The *HBB*[HDR/HDR] colonies were selected for xenotransplantation. Human donor

(huCD45$^+$) splenic chimerism 7–10 days after transplantation was similar at 20%–30% in both groups (Figure 6F). We sorted huCD45$^+$ cells from recipient spleens and bone marrows and sequenced them at the HBB locus. As expected, cells in the heterozygous (*HBB*[HDR/non-HDR]) and homozygous (*HBB*[HDR/HDR]) transplant groups carried around 50% and 100% HDR alleles, respectively (Figures 6G and S6E). Therefore, our method has allowed us to select and transplant human graft cells with an entirely HDR-corrected genotype.
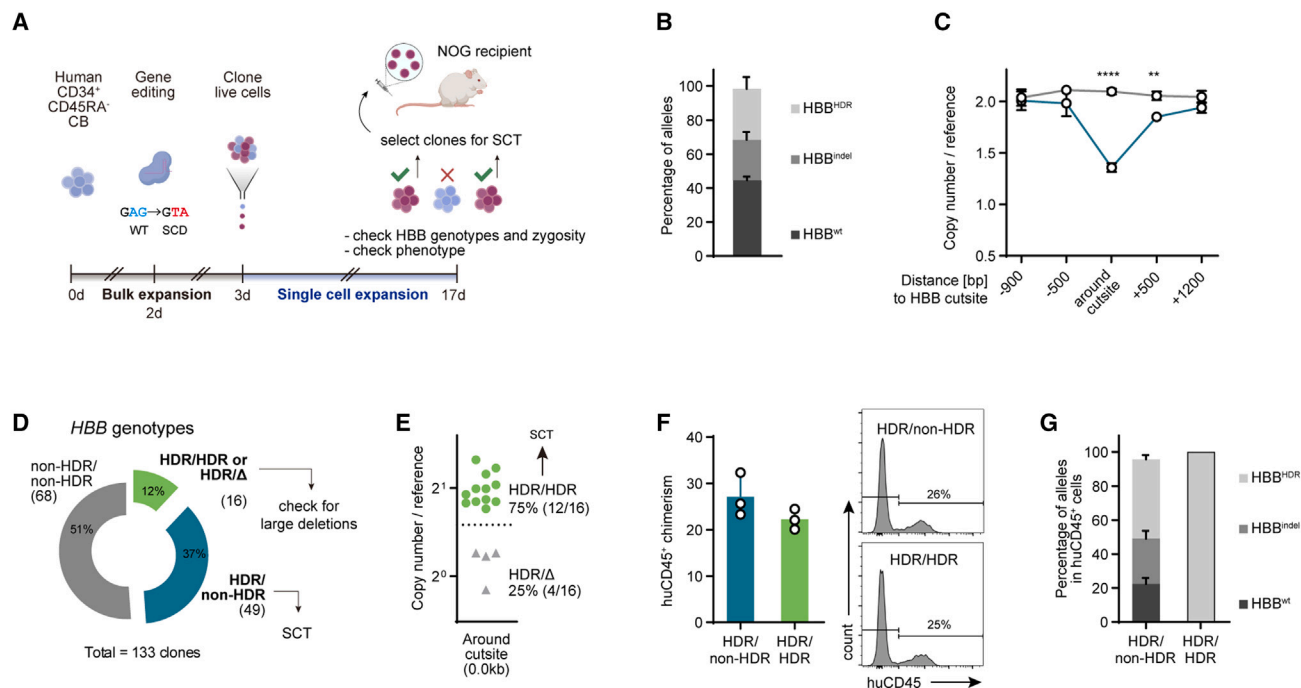
Figure 6. Controlling genetic heterogeneity in gene-edited human HSCs

(A) Scheme of experimental setup.

(B) Post-editing allele distribution at the *HBB* locus, assessed by ICE (n = 3 cultures).

(C) Copy-number analysis of *HBB* probes against reference gene (n = 3). Two-way ANOVA with Sidak's multiple comparison test.

(D) Genotyping of candidate clones (Sanger sequencing) (n = 133 clones, 2 experiments). Heterozygous clones were combined for SCT. Clones producing a single, HDR-edited sequencing trace were checked for LD (n = 16).

(E) Copy-number analysis of *HBB* probe 0.0 kb against reference gene.

(F) Left: spleen huCD45+ chimerism in *HBB*[HDR/non-HDR] and *HBB*[HDR/HDR] groups (n = 3 each). Right: representative FACS plots.

(G) Percentage of HDR alleles in huCD45+ cells sorted from bone marrow and spleens of *HBB*[HDR/non-HDR] and *HBB*[HDR/HDR] engrafted mice (n = 3 each).

**p < 0.01, ****p < 0.0001.

DISCUSSION

The platform established here permits clonal expansion, direct selection, and transplantation of a primary adult tissue stem cell population without the use of pluripotent intermediary cells. We believe our results not only have important implications for studying the genetics of hematopoiesis but also highlight the potential of this approach for gene therapy. With the widespread interest in gene editing as a therapeutic modality, concerns about hazardous mutations in edited cell products have become more visible.[29] Our expansion system addresses this concern in the murine model by enabling marker-free selection of edited LT-HSC clones with known mutational profiles. As shown here, this not only includes on- and off-target indels, but also long deletions, a recently identified by-product of Cas9 gene editing.[7] Importantly, our approach is not restricted to Cas9-induced aberrations but can be used to exclude from transplantation clones bearing any kind of collateral genetic lesion (e.g., adeno-associated virus [AAV]-associated insertions[30]).

The addition of Soluplus facilitated the expansion of HSC colonies to previously unattainable levels of over 30,000-fold. *In vivo*, this translated to high levels of chimerism in split-clone transplantations. The mechanism by which Soluplus, a biocompatible excipient for oral drug formulations, exerts its HSC-sup-

portive properties remain largely undefined, although we did find that it enhances the stability of TPO during culture. This phenomenon likely extends to other cytokines and essential factors in the medium as well.[20,31] Intriguingly, we have recently found that Soluplus supplementation also improved *in vitro* proliferation of human CB-derived CD34+ cells compared with PVA (55- versus 10-fold over 30 days) in a cytokine-free, small-molecule-based expansion system.[32]

Previous studies have achieved molecular reversion of the *Prkdc*[scid] mutation in Lin− bone marrow cells, but low efficiencies have obstructed effective functional correction of immunodeficiency *in vivo*.[33,34] This might be attributed to the higher occurrence of mutagenic repair events, such as microhomology-mediated end joining (MMEJ). Our demonstration that a sufficient corrected cell dose can be generated despite this challenging background highlights the utility of HSC expansion to propagate corrected cells *ex vivo*. Such an approach would also be highly relevant to settings where on-target mutations are detrimental and limit the curative potential of gene correction approaches, e.g., correction of the hemoglobin sickle allele.[17] To demonstrate this, we apply our expansion-selection strategy to a human *HBB* editing model and illustrate the feasibility to generate fully HDR-corrected cell grafts. In summary, we have developed an easily adoptable and powerful clonal expansion platform for precise

genetic and functional interrogation of murine HSCs at the single-cell level.

Although individual fold expansion rates were very high using our culture conditions, the overall yield of transplantable HSC clones was low, with only ~3% of sorted clones selected for transplantation. The fact that many HSCs did not produce colonies at all highlights the importance of identifying *bona fide* HSCs among precultured cells. CD201 is a useful marker indicating repopulation potential.[35] Indeed, we found that the CD201+CD150+CD48−KSL phenotype was most likely to contain LT-expanding clones, but additional markers are needed to further resolve LT-HSC activity in expansion cultures. Furthermore, in those HSC clones that do expand over 14 days, the majority still lose the CD201+CD150+KSL phenotype even with our optimized protocol, pointing to a need to further improve expansion conditions. Clonal yield might also be affected by the gene editing strategy. In our experiments, we used single-strand oligonucleotides (ssODNs) to deliver HDR templates, which preserve the clonogenicity of HSCs compared with viral vectors (AAV and integration-deficient lentiviral vector (IDLV)) by inducing a lower p53 response.[30,36] It is likely that using these more genotoxic delivery methods would further impact the clonal yield due to increased cellular stress responses. On the other hand, it would be interesting to investigate if non-DSB or template-free editing tools, such as base or prime editors, enhance colony-forming potential.[37,38] Likewise, the use long ssODNs of up to 2,000 nucleotides might enable the delivery of long templates without eliciting the stress response associated with viral vectors, although this has not been experimentally verified.

An important limitation of single-cell-expanded HSCs for SCT is the oligoclonal composition of the transplanted graft, which is a direct consequence of the low clonal yield frequencies discussed above and raises concerns regarding clonal dominance. While these concerns are warranted, it is important to consider the impact of Cas9 gene editing itself on the clonal composition of HSC pools *in vivo*. A recent study by Ferrari et al. has shown that the LT repopulating graft arising from a Cas9-edited huHSC pool is dominated by less than 10 clones after transplantation into NSG mice.[39] Similarly, Sharma et al. have reported that a median of 2 clones contribute to 50% of allele diversity in an *HBB* gene editing model.[40] Although these observations, which were obtained from xenograft models, may not accurately inform our understanding of clonal dynamics in the autologous setting, they do reflect the negative impact of Cas9 gene editing, mediated by DSB-induced p53 induction, on HSC clonal diversity. Thus, one can speculate that our screening process based on the CD201+CD150+CD48−KSL phenotype selects for clones that would have dominated LT hematopoiesis even if the bulk population had been transplanted.

Limitations of the study
Regarding applicability to huHSCs, the major obstacles to applying our approach currently are the limited expansion efficiency of huHSCs and the lack of reliable markers to identify and purify LT-expanding huHSC clones. Regarding the former, current culture protocols unfortunately do not support single-cell expansion of huHSCs. For this reason, we used a differentiation and proliferation protocol to demonstrate the advantages

of our expansion-selection platform in the human system. Furthermore, our work shows that sorting populations highly enriched for HSC activity (CD201+CD150+CD48−KSL) is critical in post-gene editing cloning and expansion. There is a pressing need to identify reliable markers of LT-HSC activity especially in the human system, where HSC frequencies among precultured cells is ~1:38 among the most stringently defined population (CD34+CD45RA−CD90+CD133+EPCR+ITGA3+).[41] The diverse genetic background in humans raises additional challenges to identifying stable markers and developing reliable enrichment strategies across individuals. To this end, techniques that link clonal phenotype with expansion outcome, like index sorting and single clone transplantation, will be essential.[35] Moreover, we recognize that our investigations concerning the function of Soluplus are mainly observational and acknowledge the need for further mechanistic studies.

With a view toward a hypothetical clinical application of our approach, the above issues of low clonal yield and graft oligoclonality are compounded by the high number of clones required for polyclonal reconstitution in humans and represent a major obstacle to translation. In our opinion, improving human expansion efficiencies to levels similar to the murine system, together with the identification of reliable HSC markers, will be key to apply our concept to the human system.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Primary cell cultures
  - Cell line cultures
- METHOD DETAILS
  - Murine HSC isolation
  - Murine CRISPR/Cas9 gene editing
  - Indel and HDR quantification in bulk-expanded cultures
  - ddPCR copy number quantification
  - Flow cytometric analysis of bulk expanded murine HSCs
  - Peripheral blood analyses
  - Fluorescence-activated single cell index sorting
  - RNA-seq analysis of expanded HSCs
  - Marker profiling of single cell expanded murine HSC clones
  - Murine TPO ELISA
  - Genotyping of Prkdc-edited single cell HSC clones
  - Murine HSC stem cell transplantation (SCT)
  - CDR3 spectratyping
  - Immunization and ELISA assays
  - Xenograft transplantation assay

- ○ Human HSC culture and HBB CRISPR/Cas9 gene editing
- ○ Human HSC single cell cloning and differentiation-expansion cultures
- ○ Genotyping of HBB-edited single cell human colonies
- ○ Human hematopoietic colony phenotyping
- ○ Human hematopoietic colony SCT and analysis
- ○ Whole exome sequencing (WES) of single cell-expanded clones
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

REFERENCES

1. Naldini, L. (2019). Genetic engineering of hematopoiesis: current stage of clinical translation and future perspectives. EMBO Mol. Med. *11*, 1–12. https://doi.org/10.15252/emmm.201809958.

2. Dever, D.P., Bak, R.O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C.E., Pavel-Dinu, M., Saxena, N., Wilkens, A.B., Mantri, S., et al. (2016). CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. Nature *539*, 384–389. https://doi.org/10.1038/nature20134.

3. Schiroli, G., Ferrari, S., Conway, A., Jacob, A., Capo, V., Albano, L., Plati, T., Castiello, M.C., Sanvito, F., Gennery, A.R., et al. (2017). Preclinical modeling highlights the therapeutic potential of hematopoietic stem cell gene editing for correction of SCID-X1. Sci. Transl. Med. *9*. https://doi.org/10.1126/scitranslmed.aan0820.

4. De Ravin, S.S., Li, L., Wu, X., Choi, U., Allen, C., Koontz, S., Lee, J., Theobald-Whiting, N., Chu, J., Garofalo, M., et al. (2017). CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. Sci. Transl. Med. *9*. https://doi.org/10.1126/scitranslmed.aah3480.

5. Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G., and Passegué, E. (2010). Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell *7*, 174–185. https://doi.org/10.1016/j.stem.2010.06.014.

6. Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T.D., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C., et al. (2014). Targeted genome editing in human repopulating hematopoietic stem cells. Nature *510*, 235–240. https://doi.org/10.1038/nature13420.

7. Adikusuma, F., Piltz, S., Corbett, M.A., Turvey, M., McColl, S.R., Helbig, K.J., Beard, M.R., Hughes, J., Pomerantz, R.T., and Thomas, P.Q. (2018). Large deletions induced by Cas9 cleavage. Nature *560*, E8–E9. https://doi.org/10.1038/s41586-018-0380-z.

8. Boutin, J., Rosier, J., Cappellen, D., Prat, F., Toutain, J., Pennamen, P., Bouron, J., Rooryck, C., Merlio, J.P., Lamrissi-Garcia, I., et al. (2021). CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. Nat. Commun. *12*, 4922. https://doi.org/10.1038/s41467-021-25190-6.

9. Leibowitz, M.L., Papathanasiou, S., Doerfler, P.A., Blaine, L.J., Sun, L., Yao, Y., Zhang, C.Z., Weiss, M.J., and Pellman, D. (2021). Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing. Nat. Genet. *53*, 895–905. https://doi.org/10.1038/s41588-021-00838-7.

10. Lattanzi, A., Camarena, J., Lahiri, P., Segal, H., Srifa, W., Vakulskas, C.A., Frock, R.L., Kenrick, J., Lee, C., Talbott, N., et al. (2021). Development of β-globin gene correction in human hematopoietic stem cells as a potential durable treatment for sickle cell disease. Sci. Transl. Med. *13*. https://doi.org/10.1126/scitranslmed.abf2444.

11. Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. Nat. Rev. Genet. *17*, 300–312. https://doi.org/10.1038/nrg.2016.28.

12. Wilkinson, A.C., Igarashi, K.J., and Nakauchi, H. (2020). Haematopoietic stem cell self-renewal in vivo and ex vivo. Nat. Rev. Genet. *21*, 541–554. https://doi.org/10.1038/s41576-020-0241-0.

13. Wilkinson, A.C., Ishida, R., Kikuchi, M., Sudo, K., Morita, M., Crisostomo, R.V., Yamamoto, R., Loh, K.M., Nakamura, Y., Watanabe, M., et al. (2019). Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. Nature *571*, 117–121. https://doi.org/10.1038/s41586-019-1244-x.

14. Araki, R., Fujimori, A., Hamatani, K., Mita, K., Saito, T., Mori, M., Fukumura, R., Morimyo, M., Muto, M., Itoh, M., et al. (1997). Nonsense mutation at Tyr-4046 in the DNA-dependent protein kinase catalytic subunit of severe combined immune deficiency mice. Proc. Natl. Acad. Sci. USA *94*, 2438–2443.

15. Vazquez, S.E., Inlay, M.A., and Serwold, T. (2015). CD201 and CD27 identify hematopoietic stem and progenitor cells across multiple murine strains independently of Kit and Sca-1. Exp. Hematol. *43*, 578–585. https://doi.org/10.1016/j.exphem.2015.04.001.

16. Owens, D.D.G., Caulder, A., Frontera, V., Harman, J.R., Allan, A.J., Bucakci, A., Greder, L., Codner, G.F., Hublitz, P., McHugh, P.J., et al. (2019). Microhomologies are prevalent at Cas9-induced larger deletions. Nucleic Acids Res. *47*, 7402–7417. https://doi.org/10.1093/nar/gkz459.

17. Wilkinson, A.C., Dever, D.P., Baik, R., Camarena, J., Hsu, I., Charlesworth, C.T., Morita, C., Nakauchi, H., and Porteus, M.H. (2021). Cas9-AAV6 gene correction of beta-globin in autologous HSCs improves sickle cell disease

erythropoiesis in mice. Nat. Commun. *12*, 686. https://doi.org/10.1038/s41467-021-20909-x.

18. Noda, S., Horiguchi, K., Ichikawa, H., and Miyoshi, H. (2008). Repopulating activity of ex vivo-expanded murine hematopoietic stem cells resides in the CD48−c-kit + Sca-1+ lineage marker− cell population. Stem Cells *26*, 646–655. https://doi.org/10.1634/stemcells.2007-0623.

19. Zhang, C.C., and Lodish, H.F. (2005). Murine hematopoietic stem cells change their surface phenotype during ex vivo expansion. Blood *105*, 4314–4320. https://doi.org/10.1182/blood-2004-11-4418.

20. Linn, M., Collnot, E.M., Djuric, D., Hempel, K., Fabian, E., Kolter, K., and Lehr, C.M. (2012). Soluplus® as an effective absorption enhancer of poorly soluble drugs in vitro and in vivo. Eur. J. Pharm. Sci. *45*, 336–343. https://doi.org/10.1016/j.ejps.2011.11.025.

21. Nishimura, T., Hsu, I., Martinez-Krams, D.C., Nakauchi, Y., Majeti, R., Yamazaki, S., Nakauchi, H., and Wilkinson, A.C. (2019). Use of polyvinyl alcohol for chimeric antigen receptor T-cell expansion. Exp. Hematol. *80*, 16–20. https://doi.org/10.1016/j.exphem.2019.11.007.

22. Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods *347*, 70–78. https://doi.org/10.1016/j.jim.2009.06.008.

23. Komuro, K., Itakura, K., Boyse, E.A., and John, M. (1974). Ly-5: a new T-lymphocyte antigen system. Immunogenetics *1*, 452–456. https://doi.org/10.1007/BF01564083.

24. Zebedee, S.L., Barritt, D.S., and Raschke, W.C. (1991). Comparison of mouse Ly5a and Ly5b leucocyte common antigen alleles. Dev. Immunol. *1*, 243–254. https://doi.org/10.1155/1991/52686.

25. Mercier, F.E., Sykes, D.B., and Scadden, D.T. (2016). Single targeted exon mutation creates a true congenic mouse for competitive hematopoietic stem cell transplantation: the C57BL/6-CD45.1 STEM mouse. Stem Cell Rep. *6*, 985–992. https://doi.org/10.1016/j.stemcr.2016.04.010.

26. Pannetier, C., Cochet, M., Darche, S., Casrouge, A., Zöller, M., and Kourilsky, P. (1993). The sizes of the CDR3 hypervariable regions of the murine T-cell receptor β chains vary as a function of the recombined germ-line segments. Proc. Natl. Acad. Sci. USA *90*, 4319–4323. https://doi.org/10.1073/pnas.90.9.4319.

27. Lee, M.S.J., Natsume-Kitatani, Y., Temizoz, B., Fujita, Y., Konishi, A., Matsuda, K., Igari, Y., Tsukui, T., Kobiyama, K., Kuroda, E., et al. (2019). B cell-intrinsic MyD88 signaling controls IFN-γ-mediated early IgG2c class switching in mice in response to a particulate adjuvant. Eur. J. Immunol. *49*, 1433–1440. https://doi.org/10.1002/eji.201848084.

28. Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science *351*, aab2116. https://doi.org/10.1126/science.aab2116.

29. Sheridan, C. (2021). CRISPR therapies march into clinic, but genotoxicity concerns linger. Nat. Biotechnol. *39*, 897–899. https://doi.org/10.1038/d41587-021-00017-3.

30. Ferrari, S., Jacob, A., Cesana, D., Laugel, M., Beretta, S., Varesi, A., Unali, G., Conti, A., Canarutto, D., Albano, L., et al. (2022). Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. Cell Stem Cell *29*. 1428.e9–1444.e9. https://doi.org/10.1016/J.STEM.2022.09.001.

31. Obata, T., Suzuki, Y., Ogawa, N., Kurimoto, I., Yamamoto, H., Furuno, T., Sasaki, T., and Tanaka, M. (2014). Improvement of the antitumor activity of poorly soluble sapacitabine (CS-682) by using Soluplus® as a surfactant. Biol. Pharm. Bull. *37*, 802–807.

32. Sakurai, M., Ishitsuka, K., Ito, R., Wilkinson, A.C., Kimura, T., Mizutani, E., Nishikii, H., Sudo, K., Becker, H.J., Takemoto, H., et al. (2023). Chemically defined cytokine-free expansion of human haematopoietic stem cells. Nature *615*, 127–133. https://doi.org/10.1038/s41586-023-05739-9.

33. Abdul-Razak, H.H., Rocca, C.J., Howe, S.J., Alonso-Ferrero, M.E., Wang, J., Gabriel, R., Bartholomae, C.C., Gan, C.H.V., Garín, M.I., Roberts, A.,

et al. (2018). Molecular evidence of genome editing in a mouse model of immunodeficiency. Sci. Rep. *8*, 8214. https://doi.org/10.1038/s41598-018-26439-9.

34. Rahman, S.H., Kuehle, J., Reimann, C., Mlambo, T., Alzubi, J., Maeder, M.L., Riedel, H., Fisch, P., Cantz, T., Rudolph, C., et al. (2015). Rescue of DNA-PK signaling and T-cell differentiation by targeted genome editing in a prkdc deficient iPSC disease model. PLoS Genet. *11*, e1005239. https://doi.org/10.1371/journal.pgen.1005239.

35. Che, J.L.C., Bode, D., Kucinski, I., Cull, A.H., Bain, F., Becker, H.J., Jassinskaja, M., Barile, M., Boyd, G., Belmonte, M., et al. (2022). Identification and characterization of *in vitro* expanded hematopoietic stem cells. EMBO Rep. *23*, e55502. https://doi.org/10.15252/embr.202255502.

36. Pattabhi, S., Lotti, S.N., Berger, M.P., Singh, S., Lux, C.T., Jacoby, K., Lee, C., Negre, O., Scharenberg, A.M., and Rawlings, D.J. (2019). In vivo outcome of homology-directed repair at the HBB gene in HSC using alternative donor template delivery methods. Mol. Ther. Nucleic Acids *17*, 277–288. https://doi.org/10.1016/j.omtn.2019.05.025.

37. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424. https://doi.org/10.1038/nature17946.

38. Nelson, J.W., Randolph, P.B., Shen, S.P., Everette, K.A., Chen, P.J., v Anzalone, A.V., An, M., Newby, G.A., Chen, J.C., Hsu, A., et al. (2022). Engineered pegRNAs improve prime editing efficiency. Nat. Biotechnol. *40*, 402–410. https://doi.org/10.1038/s41587-021-01039-7.

39. Ferrari, S., Jacob, A., Beretta, S., Unali, G., Albano, L., Vavassori, V., Cittaro, D., Lazarevic, D., Brombin, C., Cugnata, F., et al. (2020). Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. Nat. Biotechnol. *38*, 1298–1308. https://doi.org/10.1038/s41587-020-0551-y.

40. Sharma, R., Dever, D.P., Lee, C.M., Azizi, A., Pan, Y., Camarena, J., Köhnke, T., Bao, G., Porteus, M.H., and Majeti, R. (2021). The TRACE-Seq method tracks recombination alleles and identifies clonal reconstitution dynamics of gene targeted human hematopoietic stem cells. Nat. Commun. *12*, 472. https://doi.org/10.1038/s41467-020-20792-y.

41. Tomellini, E., Fares, I., Lehnertz, B., Chagraoui, J., Mayotte, N., MacRae, T., Bordeleau, M.È., Corneau, S., Bisaillon, R., and Sauvageau, G. (2019). Integrin-α3 is a functional marker of ex vivo expanded human long-term hematopoietic stem cells. Cell Rep. *28*. 1063.e5–1073.e5. https://doi.org/10.1016/j.celrep.2019.06.084.

42. Freier, C. (2020). indexSort: retrieve index sorted data and visually explore.fcs date files.. GitHub. https://github.com/Kawameicha/indexSort.

43. McInnes, L., Healy, J., and Melville, J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. https://doi.org/10.48550/arXiv.1802.03426.

44. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

45. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

46. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

47. Brinkman, E.K., Chen, T., Amendola, M., and Van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. *42*, e168. https://doi.org/10.1093/nar/gku936.

48. Igarashi, K.J., Kucinski, I., Chan, Y.Y., Tan, T.K., Khoo, H.M., Kealy, D., Bhadury, J., Hsu, I., Ho, P.Y., Niizuma, K., et al. (2023). Physioxia improves the selectivity of hematopoietic stem cell expansion cultures. Blood Adv. https://doi.org/10.1182/bloodadvances.2023009668.

49. Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J., and Bao, G. (2014). Cosmid: a web-based tool for identifying and validating CRISPR/Cas off-target sites. Mol. Ther. Nucleic Acids *3*, e214. https://doi.org/10.1038/mtna.2014.64.

50. Ahmed, M., Lanzer, K.G., Yager, E.J., Adams, P.S., Johnson, L.L., and Blackman, M.A. (2009). Clonal expansions and loss of receptor diversity in the naive CD8 T cell repertoire of aged mice. J. Immunol. *182*, 784–792. https://doi.org/10.4049/jimmunol.182.2.784.

51. Miqueu, P., Guillet, M., Degauque, N., Doré, J.C., Soulillou, J.P., and Brouard, S. (2007). Statistical analysis of CDR3 length distributions for the assessment of T and B cell repertoire biases. Mol. Immunol. *44*, 1057–1064. https://doi.org/10.1016/j.molimm.2006.06.026.

52. Iwano, S., Sugiyama, M., Hama, H., Watakabe, A., Hasegawa, N., Kuchimaru, T., Tanaka, K.Z., Takahashi, M., Ishida, Y., Hata, J., et al. (2018). Single-cell bioluminescence imaging of deep tissue in freely moving animals. Science *359*, 935–939. https://doi.org/10.1126/science.aaq1067.

53. DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.J., Mitros, T., Muñoz, D.P., Boffelli, D., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci. Transl. Med. *8*, 360ra134. https://doi.org/10.1126/scitranslmed.aaf9336.

54. Wagenblast, E., Araújo, J., Gan, O.I., Cutting, S.K., Murison, A., Krivdova, G., Azkanaz, M., McLeod, J.L., Smith, S.A., Gratton, B.A., et al. (2021). Mapping the cellular origin and early evolution of leukemia in Down syndrome. Science *373*. https://doi.org/10.1126/science.abf6202.

55. Ellis, P., Moore, L., Sanders, M.A., Butler, T.M., Brunner, S.F., Lee-Six, H., Osborne, R., Farr, B., Coorens, T.H.H., Lawson, A.R.J., et al. (2021). Reliable detection of somatic mutations in solid tissues by laser-capture microdissection and low-input DNA sequencing. Nat. Protoc. *16*, 841–871. https://doi.org/10.1038/s41596-020-00437-6.

56. Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J.E. (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science *333*, 218–221. https://doi.org/10.1126/science.1201219.

57. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-mouse Ckit-APC (2B8)	Thermo Fisher Scientific	Cat#17-1171-83; RRID:AB_469431
anti-mouse CD4-APC (RM4-5)	Thermo Fisher Scientific	Cat#17-0042-83; RRID:AB_469324
anti-mouse CD8a-APC (53-6.7)	Thermo Fisher Scientific	Cat#17-0081-83; RRID:AB_469336
anti-mouse CD48-APC (HM48-1)	Thermo Fisher Scientific	Cat#17-0481-82; RRID:AB_469408
anti-mouse CD135-APC (A2F10)	Thermo Fisher Scientific	Cat#17-1351-82; RRID:AB_10717261
anti-mouse CD45.1-APC/Cy7 (A20)	Tonbo Biosciences	Cat#25-0453; RRID:AB_2621629
anti-mouse Sca-1-APC/Cy7 (D7)	BioLegend	Cat#108126; RRID:AB_10645327
anti-mouse CD45R(B220)-APC/eFluor780 (RA3-6B2)	Thermo Fisher Scientific	Cat#47-0452-82; RRID:AB_1518810
anti-mouse Ckit-APC/H7 (2B8)	BD Biosciences	Cat#560185; RRID:AB_1645231
anti-mouse CD4-Biotin (RM4-5)	Thermo Fisher Scientific	Cat#13-0042-85; RRID:AB_466330
anti-mouse CD8a-Biotin (53-6.7)	Thermo Fisher Scientific	Cat#13-0081-85; RRID:AB_466347
anti-mouse CD45R(B220)-Biotin (RA3-6B2)	Thermo Fisher Scientific	Cat#13-0452-82; RRID:AB_466449
anti-mouse TER119-Biotin (TER119)	Thermo Fisher Scientific	Cat#13-5921-85; RRID:AB_466798
anti-mouse Gr1-Biotin (RB6-8C5)	Thermo Fisher Scientific	Cat#13-5931-85-85; RRID:AB_466801
anti-mouse CD11b-Biotin (M1/70)	Thermo Fisher Scientific	Cat#13-0112-85; RRID:AB_466360
anti-mouse CD127-Biotin (A7R34)	Thermo Fisher Scientific	Cat#13-1271-85; RRID:AB_466589
anti-mouse CD45.2-BV421 (104)	Thermo Fisher Scientific	Cat#48-0454-82; RRID:AB_11042125
anti-mouse CD34-FITC (RAM34)	Thermo Fisher Scientific	Cat#11-0341-85; RRID:AB_465022
anti-mouse CD45.2-FITC (104)	BioLegend	Cat#109806; RRID:AB_313443
anti-mouse CD48-FITC (HM48-1)	BioLegend	Cat#103404; RRID:AB_313019
anti-mouse CD4-FITC (RM4-5)	Tonbo Biosciences	Cat#35-0042; RRID:AB_2621666
anti-mouse CD8a-FITC (53-6.7)	Thermo Fisher Scientific	Cat#11-0081-85; RRID:AB_464916
anti-mouse CD45R(B220)-FITC (RA3-6B2)	Thermo Fisher Scientific	Cat#11-0452-85; RRID:AB_465055
anti-mouse TER119-FITC (TER119)	Thermo Fisher Scientific	Cat#11-5921-82; RRID:AB_465311
anti-mouse Gr1-FITC (RB6-8C5)	Thermo Fisher Scientific	Cat#11-5931-82; RRID:AB_465314
anti-mouse CD11b-FITC (M1/70)	Thermo Fisher Scientific	Cat# 11-0112-82; RRID:AB_464935
anti-mouse CD4-PB (RM4-5)	BioLegend	Cat#100531; RRID:AB_493374
anti-mouse CD8a-PB (53-6.7)	BD Biosciences	Cat#558106; RRID:AB_397029
anti-mouse CD45R(B220)-PB (RA3-6B2)	BD Biosciences	Cat#558108; RRID:AB_397031
anti-mouse TER119-PB (TER119)	BioLegend	Cat#116231; RRID:AB_2149212
anti-mouse Gr1-PB (RB6-8C5)	BioLegend	Cat#108430; RRID:AB_893556
anti-mouse CD11b-PB (M1/70)	BioLegend	Cat#101224; RRID:AB_755986
anti-mouse Sca-1-PE (D7)	BD Biosciences	Cat#12-5981-83; RRID:AB_466087
anti-mouse CD201-PE (eBio1560)	Thermo Fisher Scientific	Cat#12-2012-82; RRID:AB_914317
anti-mouse Gr1-PE (RB6-8C5)	Thermo Fisher Scientific	Cat#12-5931-83; RRID:AB_466046
anti-mouse CD11b-PE (M1/70)	BD Biosciences	Cat#557397; RRID:AB_396680
anti-mouse CD105-PE (MJ7/18)	Thermo Fisher Scientific	Cat#12-1051-82; RRID:AB_657524
anti-mouse CD150-PE/Cy7 (TC15-12F12.2)	BioLegend	Cat#115913; RRID:AB_439796
anti-mouse CD45.1-PE/Cy7 (A20)	Tombo Biosciences	Cat#60-0453; RRID:AB_2621850
anti-mouse CD8a-PE/Cy7 (53-6.7)	BioLegend	Cat#100722; RRID:AB_312761
anti-mouse Sca-1-BV605 (D7)	BioLegend	Cat#108133; RRID:AB_2562275
Streptavidin-APC/eFluor780	Thermo Fisher Scientific	Cat#47-4317-82; RRID:AB_10366688
Streptavidin-BV421	BioLegend	Cat#405225

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
anti-human CD41-BV421 (HIP8)	BioLegend	Cat#303730; RRID:AB_2629627
anti-human CD235a-FITC (HI264)	BioLegend	Cat#349104; RRID:AB_10613463
anti-human CD71-PE (CY1G4)	BioLegend	Cat#334106; RRID:AB_2201481
anti-human CD34-APC (581)	BioLegend	Cat#343510; RRID:AB_1877153
anti-human CD33-PE/Cy7 (WM53)	Thermo Fisher Scientific	Cat#25-0338-42; RRID:AB_1907380
anti-human CD45-APC/Cy7 (HI30)	BioLegend	Cat#304014; RRID:AB_314402
anti-mouse CD45.1-eFluor 450 (A20)	Thermo Fisher Scientific	Cat#48-0453-82; RRID:AB_1272189
Biological samples		
Human CD34$^+$ cord blood HSCs	StemExpress	Cat#CB3400.5C
Chemicals, Peptides, and Recombinant Proteins		
Polyvinyl alcohol (PVA), 87-90% hydrolyzed	Sigma	Cat#P8136; CAS 9002-89-5
Soluplus	BASF	CAS 402932-23-4
Recombinant Murine TPO	Peprotech	Cat#315-14; P40226
Recombinant Murine SCF	Peprotech	Cat#250-03; P20826
Insulin-Transferrin-Selenium	Thermo Fisher Scientific	Cat#41400045
Nitroiodophenyl (NIP)-conjugated OVA (NIPOVA)	Biosearch Technologies	Cat#N-5041-10
Deposited data		
RNA-Seq data	This paper	GEO: GSE232527
Whole exome sequencing data	This paper	BioProject: PRJNA974717
Raw figure data	This paper	Mendeley Data: https://doi.org/10.17632/br3y74vjh4.1
Experimental Models: Cell Lines		
Human: A549 (lung adenocarcinoma)	ATCC	Cat#CCL-185; RRID:CVCL_0023
Murine: MS-5 (bone marrow stroma)	Riken BRC	Cat#RCB4680; RRID:CVCL_2128
Experimental Models: Organisms/Strains		
Mouse: CB17/SCID: C.B-17/Icr-$^{scid/scid}$Jcl	Clea	RRID:IMSR_JCL:JCL:mID-0003
Mouse: CB17/WT: C.B-17/Icr-$^{+/+}$Jcl	Clea	RRID:IMSR_JCL:JCL:mID-0004
Mouse: CD45.2$^+$ C57BL/6: C57BL/6NCrSlc	SLC	RRID:MGI:5295404
Mouse: CD45.1$^+$ C57BL/6: B6.SJL-Ptprc^a Pepc^b/BoyJ	Sankyo Labo	RRID:IMSR_JAX:002014
Oligonucleotides		
	See Table S4 for list of oligonucleotides	N/A
Software and Algorithms		
FlowJo version 10	BD	https://www.flowjo.com; SCR_008520
IndexSort version 0.1.6 (FlowJo plugin)	Freier[42]	https://www.flowjo.com
UMAP version 3.1 (FlowJo plugin)	McInnes et al.[43]	https://arxiv.org/abs/1802.03426
Prism version 9.1	Graphpad	https://www.graphpad.com; SCR_002798
R version 4.0.0	R Foundation	https://www.r-project.org; SCR_001905
DESeq2 version 1.32.0	Love et al.[44]	SCR_015687
clusterProfiler version 4.0.2	Wu et al.[45]	SCR_016884
ComplexHeatmap version 2.8.0	Gu et al.[46]	SCR_017270
Inference of CRISPR edits (ICE)	Synthego	https://ice.synthego.com/
Tracking of Indels by Decomposition (TIDE)	Brinkman et al.[47]	http://shinyapps.datacurators.nl/tide/
QuantaSoft Analysis Pro 1.0	Bio-Rad	https://www.bio-rad.com/

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to Satoshi Yamazaki (y-sato4@md.tsukuba.ac.jp).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- The RNA-seq and whole exome sequencing data have been deposited at Gene Expression Omnibus (GEO) and Sequence Read Archive (SRA) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Data belonging to figures have been deposited at Mendeley Data and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice
Male C.B-17/Icr-$^{+/+}$JcI wildtype (CB17/WT) and male C.B-17/Icr-$^{scid/scid}$JcI (CB17/SCID) mice were obtained from Clea Inc., Japan. C57BL/6NCrSlc (Ly 5.2, CD45.2) mice were purchased from SLC Inc., Japan. C57BL/6-Ly5.1 (Ly 5.1, CD45.1) mice were purchased from Sankyo Labo, Japan. All mice were obtained at age 8-10 weeks and housed in specific-pathogen-free (SPF) conditions at up to 5 mice per cage, with free access to standard rodent feed and kept under a 12h light/12h dark cycle. All animal protocols were approved by the Animal Care and Use Committee of the Institute of Medical Science, University of Tokyo.

Primary cell cultures
All cell culture operations were conducted under sterile hoods. Cells were kept in an incubator (Panasonic) at 37°C and a constant CO_2 fraction of 5%. For gene editing experiments, HSCs were cultured in hypoxia (5% F_iO_2).[48] Cell concentrations were determined on a Countess II cytometer (Thermo Fisher Scientific) after staining with Turk's staining buffer (bone marrow cells) or trypan blue dead stain solution. Male murine HSCs were cultured in a Ham's F12 medium (Wako) supplemented with 10 mM HEPES (Thermo Fisher Scientific), recombinant cytokines murine TPO (100 ng/ml, Peprotech) and SCF (10 ng/ml, Peprotech), as well as insulin-transferrin-selenium (ITS, Thermo Fisher Scientific, 1:100 dilution) and 1% Penicillin-Streptomycin-L-Glutamine (PSG, Wako). Recombinant human albumin (Albumin Biosciences), polyvinyl alcohol (PVA, 84% hydrolyzed, Sigma), Kollidon 12 PF, Kollidon 17 PF, Kollidon 90 F, Poloxamer 188 Bio, Poloxamer 407 Geismar, povidone and Soluplus (all BASF) were added at a concentration of 0.1% v/v (except for Soluplus titration experiments). PVA and Soluplus-supplemented media are designated PVA-HSC and Soluplus-HSC expansion medium, respectively. Polymers were added from prepared stocks of 10% w/v in ddH$_2$O. Bulk murine HSC expansions were cultured on human fibronectin-coated 24-well dishes (Corning). Single cell expansions were cultured on untreated U-bottom 96-well plates (TPP, cultures starting with 1-50 cells).

Human CD34$^+$ cord blood HSCs (from a male infant born at 40 weeks, StemExpress) were cultured in human HSC bulk expansion medium, consisting of IMDM (Wako) supplemented with 10 mM HEPES, Soluplus 0.01%, and recombinant human cytokines (FLT3L 100ng/ml, G-CSF 10 ng/ml, SCF 100 ng/ml, IL-6 10 ng/ml (all from Peprotech), TPO 15 ng/ml (Shenandoah)). Recombinant cytokines, ITS, PSG and polymers were freshly added to base media before each application. Human bulk HSCs were cultured on CellBIND 24-well dishes (Corning). For human single cell expansions, CellBIND flat bottom 96-well dishes were used (Corning).

Cell line cultures
Human male epithelial lung cancer cell line A549 (ATCC) was cultured in DMEM (Wako) supplemented with 10% FBS (Thermo Fisher) and 1% Penicillin-Streptomycin (Wako). Murine stromal MS-5 cells were cultured in MEMα (Wako) supplemented with 10% FBS and 1% Penicillin-Streptomycin. Cells were passaged after reaching 70-80% confluency. Transduction procedure with Akaluc-expressing lentivirus is described in the method details section below.

METHOD DETAILS

Murine HSC isolation
Male 8-10 week-old mice were sacrificed by cervical dislocation after isoflurane anesthesia. Pelvic, femur and tibia bones were isolated and crushed, and the obtained cell solution was filtered through a 48 μm nylon mesh and whole bone marrow cells were counted. Positive selection of c-Kit$^+$ cells was performed with anti-APC magnetic-activated cell sorting (MACS, Miltenyi Biotec) antibodies after staining cells with c-Kit-APC antibody for 30 minutes. Enriched c-Kit$^+$ cells were incubated with anti-Lineage antibody cocktail (consisting of biotinylated Gr1[LY-6G/LY-6C], CD11b, CD4, CD8a, CD45R[B220], IL7-Rα, TER119) for

30 minutes. This was followed by staining with CD34-FITC, CD201-PE (for CB17 strains) or Sca-1-PE (for C57BL/6 strains), c-Kit-APC, streptavidin-APC/eFluor 780 and CD150-PE/Cy7 antibodies for 90 minutes. CD201$^+$CD150$^+$c-Kit$^+$Lin$^-$ (CD201$^+$CD150$^+$KL) cells and CD34$^-$CD150$^+$c-Kit$^+$Sca-1$^+$Lin$^-$ (CD34$^-$CD150$^+$KSL) cells from CB17 and C57BL/6 bone marrows, respectively, were sorted via fluorescence-activated cell sorting (FACS) on a Aria II cell sorter (BD) using a 100 μm nozzle and appropriate filters and settings. Propidium iodide (PI) was used to exclude dead cells. For bulk expansion of HSCs before gene editing, 5000 cells were sorted into 1 ml of HSC expansion medium per well. Medium changes were not performed until gene editing (day 3 of culture). For single cell expansion of freshly isolated HSCs, single HSCs were sorted into individual wells on a 96-well U-bottom plate (TPP) prefilled with 200 μl of culture medium. Culture medium was changed on day 7 post-sort, after which complete media changes were performed every 2-3 days.

Murine CRISPR/Cas9 gene editing

Seventeen micrograms of recombinant S. pyogenes Cas9 (S.p. Cas9 Nuclease V3, IDT) were complexed with single guide RNA (sgRNA, synthesized at IDT) at a molar ratio of 1:2.5 (104 pmol Cas9:260 pmol sgRNA) for 10 minutes at 25°C to form ribonucleo-protein (RNP) complexes. Sequences of sgRNA targeting *Prkdc*^{scid} (Prkdc_gRNA1) and *Ptprc*^{b} (CD45.2, Ptprc_gRNA1) are listed in Table S3. Expanded HSCs were washed twice with PBS, pelleted, and resuspended in 20 μl electroporation buffer P3 (Lonza). Cells were gently added to the RNP duplex. For knockin experiments, 200 pmol of single-strand oligonucleotide (ssODN) templates (synthesized at IDT, Table S4) were added to the cell-RNP suspension. The suspension was transferred to a single 20 μl electroporation cuvette on a 16-well strip (P3 Primary Cell 96-well-Nucleofector Kit, Lonza). Electroporation was carried out on a 4D nucleofector device (Lonza) using programs DI-100 (CB17 HSCs) and EO-100 (C57BL/6 HSCs). Cells were immediately recovered in pre-warmed medium and gently split-transferred into 3 wells on a human fibronectin-coated 24-well plate (Corning) at 1 ml per well. A medium change was performed one day after nucleofection, and further medium changes were performed every 2-3 days.

Indel and HDR quantification in bulk-expanded cultures

To quantify indel and HDR rates from bulk cultured cells, genomic DNA (gDNA) was extracted using NucleoSpin Tissue XS columns (Macherey-Nagel). DNA concentration was measured on a Nanodrop spectrophotometer (Thermo Fisher). For murine *Prkdc* editing experiments, 1-10 ng of gDNA was used for polymerase chain reactions (PCR), formulated as 0.5 μM forward and reverse primers (Prkdc_inner_F, Prkdc_inner_R), 10 μl 2X buffer, and 0.5 U of Gflex *Thermococcus* DNA polymerase (Takara) in a 20 μl reaction. The PCR reaction setup was as follows: initial denaturation 94°C, 60s; followed by 35 cycles of denaturation 98°C, 10s; annealing 60°C, 15s; extension 68°C, 45s; and final extension 68°C, 45s. For human *HBB* genotyping, a nested PCR design was devised. Two microliters of gDNA were used with KOD FX Neo polymerase (Toyobo) and primers: HBB_outer_F, HBB_outer_R according to the manufacturer's instructions. The PCR reaction setup was as follows: initial denaturation 94°C, 120s; followed by 35 cycles of denaturation 98°C, 10s; annealing 60°C, 30s; extension 68°C, 45s; and final extension 68°C, 120s. The second PCR was performed with two microliters of the first reaction and HBB_inner_F, HBB_inner_R, using identical cycle conditions. PCR products were separated on a 1.5% agarose gel via electrophoresis and fragments corresponding to the expected amplification target were cut and gel-purified using the Wizard SV gel and PCR clean-up system (Promega). Fourty nanograms of purified fragment was subjected to Sanger sequencing (FASMAC, Japan) using sequencing primers (Prkdc_inner_F, huHBB_seqprimer). For assessment of HDR rates in bulk cultures, we used the web-based tool Inference of CRISPR edits (ICE, Synthego, https://ice.synthego.com). Sequences from non-edited HSCs were provided as negative control samples. Potential off-target sites associated with the designed *Prkdc* gRNA were identified using COSMID (CRISPR Off-target Sites with Mismatches, Insertions, and Deletions)[49] with up to 3 mismatches in the absence of indels in the seed sequence and 2 mismatches in the presence of one insert or deletion. All targets showing a score <3 were amplified using the same cycling conditions outlined above. For bulk expansion cultures, only "inner" primer pairs were used for PCRs (see Table S4). Sequencing was performed with either forward or reverse PCR primers (specified with "SEQ" in Table S4), except off-target site #6, for which a dedicated sequencing primer was designed (OT_06_SEQ). On- and off target indel frequency was calculated with the Tracking of Indels by Decomposition (TIDE) algorithm.[47]

ddPCR copy number quantification

We used droplet digital PCR to quantify the frequency of large deletions after gene editing in both bulk and single cell murine and human editing models.[16] In both settings, five primer-probe sets were designed around the cutsite (Figures S1F and S6A; Table S4). Reactions were carried out with ddPCR Supermix for Probes (No dUTP)(Bio-Rad) using 2 μl of gDNA, 225 nM of each primer and 50 nM of each probe in a 20 μl reaction. Murine *Dot1l*[16] (SUN-labeled, Table S4) and human *RRP30* (Bio-Rad Copy Number Assay (HEX)) were used as internal copy number controls. ddPCR reactions were read out on a QX200 system (Bio-Rad) using the copy number variation method with the control gene copy numbers set to 2. Analysis was performed with QuantaSoft Analysis Pro software (V1.0, Bio-Rad).

Flow cytometric analysis of bulk expanded murine HSCs

Cell counting operations were performed on a Countess II cytometer (Thermo Fisher Scientific). For flow cytometric studies of bulk expansion cultures, i.e. HSPCs cultured in 1 ml of expansion media, a 100 μl aliquot was removed from the culture well, washed in PBS, and stained with lineage antibodies (PB- and BV421-conjugated against Gr1[LY-6G/LY-6C], CD11b, CD4, CD8, CD45R[B220],

TER119), CD34-FITC, CD201-PE, Sca-1-APC/Cy7, c-Kit-APC, CD150-PE/Cy7 antibodies for 45 minutes. After washing once with PBS, cells were analyzed on a FACSVerse or FACSAria II flow cytometer (BD).

Peripheral blood analyses

For chimerism and lineage analysis, peripheral blood was drawn from mice by retro-orbital sinus sampling under general anesthesia. Red blood cells (RBC) in a sample of 40 μl were lysed in 1 ml of Ammonium-Chloride-Potassium (ACK, 0.15 M NH4Cl, 0.01 M KHCO3, 0.1 mM Na2EDTA) buffer for 15 minutes at room temperature. RBC lysis was repeated 2 times. Lysed blood cells were stained with Gr1-PE, CD11b-PE, CD4-APC, CD45R[B220]-APC/eFluor 780, CD8-PE/Cy7 for SCID mouse samples and with Gr1-PE, CD11b-PE, CD4-APC, CD8a-APC, CD45R[B220]-APC/eFluor 780, CD45.1-PE/Cy7 and CD45.2-BV421 for C57BL/6 mice samples for 30 minutes at room temperature. Cells were resuspended in 200 μl PBS/PI before recording events on a FACSVerse (BD) analyzer using the appropriate filters and settings.

Fluorescence-activated single cell index sorting

CD34⁻CD150⁺KSL cells were isolated from C57BL/6 mice and cultured on a 96-well dish in PVA-HSC expansion medium at 50 cells per well. After 10 days, cells were stained with antibodies against KSL (biotinylated lineage-antibodies (same mixture as used in 'Murine HSC isolation'), followed by c-Kit-APC/H7, Sca-1-BV605, and streptavidin-BV421. Cells were then divided into two sets, and each set was stained with an antibody panel (panel 1: CD34-FITC, CD48-APC, CD105-PE; panel 2: CD135-APC, CD201-PE, CD150-PE/Cy7). We cloned single KSL cells using the index sorting function on a FACSAria II (BD). Well location and expression data of the sorted clones were extracted using the IndexSort plugin for FlowJo.[42] Dimensionality reduction was performed with the UMAP plugin for FlowJo.[43] Expression data of the sorted clones were log-transformed and normalized to mean.

RNA-seq analysis of expanded HSCs

Expanded cells were washed in PBS once and stained with the identical panel specified in 'Analysis of bulk expanded cells', except for CD48-FITC, which was used instead of CD34-FITC. Over 5000 cells per population were sorted into 1.5 ml tubes and subsequently lysed in 600 μl Trizol LS reagent (Thermo Fisher Scientific). RNA purification, library preparation and next-generation sequencing was performed by Tsukuba i-Laboratory, LLC. Libraries were prepared using the SMARTer cDNA synthesis kit (Takara) and the high-output kit v2 (Illumina), followed by sequencing on a NextSeq 500 sequencer (Illumina) at 2x 36 paired end reads. Data normalization and comparative analyses were performed with the DESeq2 package in R.[44] Genes with an adjusted $p < 0.05$ were considered differentially expressed. Enrichment analysis of differentially expressed genes was performed with the gene set enrichment analysis (GSEA) functions in the clusterProfiler package[45] using molecular signature database (MSigDB) gene ontology biological process (C5 GO:BP) as well as chemical and genetic perturbations (C2:CGP) gene set collections. Heatmaps were generated with the ComplexHeatmap package.[46]

Marker profiling of single cell expanded murine HSC clones

To measure HSC marker expression in single cell-derived clonal cultures, 30 μl aliquots of cells were recovered from HSC colonies (cultured in 200 μl), transferred to a 96-well staining plate, washed in PBS, and stained with either PB/BV421- or FITC-conjugated linage antibodies (Gr1[LY-6G/LY-6C], CD11b, CD4, CD8, CD45R[B220], TER119), CD201-PE, Sca-1-APC/Cy7, c-Kit-APC and CD150-PE/Cy7 for 45 minutes at room temperature. After washing with PBS on-plate, cells were resuspended in 200 μl PBS/PI and examined on a FACSVerse analyzer (BD) using appropriate filters and settings. Acquisition time was set to 20 seconds to ensure enough cells remained for genomic DNA extraction, if necessary.

Murine TPO ELISA

Ham's F12 with 100 ng/ml muTPO was supplemented with the indicated polymers and cultured for 3 days at 37°C. After the incubation period, 10 μl of the cultures were diluted tenfold and assayed for TPO concentrations using a commercial ELISA kit (Mouse Thrombopoietin Quantikine ELISA Kit, R&D systems) according to the manufacturer's instructions.

Genotyping of Prkdc-edited single cell HSC clones

To quantify HDR in single cell expanded clones of *Prkdc*-corrected HSCs (genotyping), cells left over from HSC marker profiling (see previous section) were subjected to gDNA extraction using NucleoSpin Tissue XS columns (Macherey-Nagel). gDNA was eluted in 18 μl of ddH₂O. Only clones containing CD201⁺CD150⁺KL cells were selected for genotyping. For genotyping of the *Prkdc* locus, a nested PCR strategy was employed. The outer PCR formulation was 5 μl of gDNA, 0.5 μM forward and reverse outer primers (Prkdc_outer_F, Prkdc_outer_R) and 12.5 μl of Q5 2X master mix (containing Q5 DNA polymerase, dNTPs and Mg²⁺) (New England Biosciences) in a 25 μl reaction. The PCR reaction setup was as follows: initial denaturation 98°C, 30s; followed by 35 cycles of denaturation 98°C, 10s; annealing 65°C, 15s; extension 72°C, 45s; and final extension 72°C, 120s. The PCR product was diluted 1:20 for the inner PCR reaction. For this reaction, 1 μl of diluted PCR product was combined with 0.5 μM forward and reverse nested primers (Prkdc_inner_F, Prkdc_inner_R) and 25 μl of Q5 2X master mix in a 50 μl reaction. 700 bp PCR products were purified and sequenced as outlined above ('indel and HDR quantification in bulk-expanded cultures'). A semi-nested PCR strategy was employed for sequencing of off-target edits, first amplifying all sites in a multiplex PCR reaction using outer and inner primers, followed by a second reaction for individual targets using inner primers only. The multiplex PCR reaction contained primers specific to all off-target loci. The

formulation was 5 μl of gDNA, 0.25 μM outer and inner primers, 25 μl 2X buffer, and 1.25 U of Gflex *Thermococcus* DNA polymerase (Takara) in a 50 μl reaction. The PCR reaction setup was as follows: initial denaturation 94°C, 60s; followed by 35 cycles of denaturation 98°C, 10s; annealing 60°C, 15s; extension 68°C, 120s; and final extension 68°C, 45s. The PCR product was diluted 1:20 and used for amplification of individual off-target sites. These reactions were formulated as follows: 1 μl of diluted PCR product, 0.25 μM inner primers, 12.5 μl 2X buffer, and 0.625 U of Gflex *Thermococcus* DNA polymerase (Takara) in a 25 μl reaction. The PCR reaction setup was as follows: initial denaturation 94°C, 60s; followed by 35 cycles of denaturation 98°C, 10s; annealing 60°C, 15s; extension 68°C, 120s; and final extension 68°C, 45s. After agarose gel separation, PCR products were purified and sequenced using the inner reverse primers. Sequencing was performed with either forward or reverse inner PCR primers (specified with "SEQ" in Table S4), except off-target site #6, for which a dedicated sequencing primer was used (OT_06_SEQ). Sequence traces were aligned to reference sequences to check for mutations.

Murine HSC stem cell transplantation (SCT)

Cells in *Prkdc*-edited bulk HSC cultures were washed, resuspended in 300 μl PBS and divided into 3 aliquots for transplantation into three recipients. Each recipient received 0.5 x10[6] cells. For experiments comparing PVA and Soluplus expansion conditions, single cell-derived clones were split into several aliquots for SCT into multiple recipients as stated in the main text. *Ptprc*-edited single cell clones were transplanted into a one recipient. For single cell *Prkdc*-corrected clones, candidate clones were selected based on HSC marker and genotyping and combined to a single dose for transplantation into one CB17/SCID recipient. For SCT with *Prkdc*-corrected cells, 0.2 x10[6] whole bone marrow (WBM) cells from 10 week old male CB17/SCID mice were added to the graft as support to ensure survival immediately after myeloablation. For non-edited and *Ptprc*-edited C57BL/6-derived HSCs, 0.2 x10[6] WBM competitor cells from C57BL/6 CD45.1[+]/CD45.2[+] F1 mice were added unless stated otherwise in the main text. CB17/SCID and C57BL/6 mice were lethally irradiated with 2.5 Gy and 9 Gy, respectively, immediately prior to transplantation. Cells were injected via tail vein injection. Secondary bone marrow transplantations were performed by extracting WBM cells from the primary recipient and transplanting 1.0 x10[6] cells into lethally irradiated secondary recipients.

CDR3 spectratyping

The spectratyping protocol originally published by Pannetrier et al. was followed with modifications by Ahmed et al.[26,50] Splenocytes were recovered by crushing freshly excised spleens between two glass slides (Matsunami). CD4[+] lymphocytes were enriched using CD4 magnetic-activated cell sorting (MACS) positive selection according to anufacturer's intructions (Miltenyi Biotec) and lysed in 300 μl Trizol reagent (Thermo Fisher) per 10[6] cells. RNA was purified using the Direct-zol RNA Microprep kit (Zymo) and eluted in 15 μl ddH$_2$O. 150 ng of RNA was subjected for cDNA synthesis using Superscript IV reverse transcriptase (Thermo Fisher) according to manufacturer's instructions. 2 μl of cDNA was used per Vβ PCR reaction. The PCR reaction was formulated as: 2 μl of cDNA, 20 pmol TCR constant region (TCR-Cb) and Vβ gene-specific primer (1 μM final concentration) (see Table S4), 10 μl 2X buffer (containing deoxynucleoside triphosphates (dNTPs) and Mg^{2+}), and 0.5 U of Gflex DNA polymerase (Takara) in a 20 μl reaction. The PCR reaction setup was as follows: initial denaturation 94°C, 120s; followed by 40 cycles of denaturation 98°C, 10s; annealing 62°C, 30s; extension 68°C, 90s; and final extension 68°C, 600s. 5 μl of PCR product was then used in a runoff reaction including a FAM-labeled TCR-Cb primer. The reaction mix was formulated as 5 μl PCR product, 4 pmol 5'-FAM-labeled TCR constant region primer (TCR-Cb-FAM, 0.2 μM final concentration), 10 μl 2X buffer (containing deoxynucleoside triphosphates (dNTPs) and Mg^{2+}), and 0.5 U of Gflex *Thermococcus* DNA polymerase (Takara) in a 20 μl reaction. The PCR reaction setup was as follows: initial denaturation 94°C, 120s; followed by 20 cycles of denaturation 98°C, 10s; annealing 62°C, 30s; extension 68°C, 90s; and final extension 68°C, 300s. Ten microliters of the reaction mix were used for fragment sizing (performed at FASMAC, Japan). Fragment size analysis was performed on the Thermo Fisher Connect platform (https://apps.thermofisher.com/) using the peak scanner application. Relevant peaks were filtered and imported into Prism software (GraphPad) for further analysis and visualization.[51] For Kolmogorov-Smirnov normality tests, a threshold level of 0.05 was selected to reject the hypothesis that data was normally distributed.

Immunization and ELISA assays

Mice at 20 weeks after SCT with *Prkdc*-corrected HSCs were immunized with 100 μg of nitroiodophenyl (NIP)-conjugated OVA (NIPOVA) (Biosearch Technologies) mixed 1:1 with aluminium hydroxide (Invivogen) intraperitoneally (i.p.). Blood samples were collected after 12 and 19 days post-immunization via retro-orbital sinus sampling. Serum was recovered by centrifugation of whole blood for 10 minutes at 5000g and stored at -20°C. For serum antibody detection, high binding 96-well microplates (Thermo Fisher) were pre-coated with NIP$_{30}$-BSA (Biosearch Technologies) at 2 μg/ml concentration overnight. After blocking wells with 1% BSA/PBS solution, 1:5000 dilutions of serum samples were applied to the wells and incubated overnight at 4°C. Wells were washed with 0.05% PBS/Tween-20 (PBS-T) followed by incubation with horse radish peroxidase (HRP)-conjugated secondary antibodies against murine IgG and IgM (1:5000 dilution, Southern Biotech) for 2 hours. Enzymatic reaction was initiated by adding 100 μl of 3,3',5,5'-Tetramethyl -benzidine (TMB) substrate solution (TCI) to each well, followed by termination of the reaction with 100 μl hydrochloric acid 1M (HCl, TCI). Absorbance readings were obtained on a microplate reader at 450 nm (Molecular Devices).

Xenograft transplantation assay

Human A549 cells were modified to constitutively express Akaluc, a firefly luciferase derivative with improved bioluminescent activity.[52] Cultured cells were transduced with a VSV-G pseudotyped lentiviral vector carrying an Akaluc-P2A-mNeonGreen transgene

under the control of the human ubiquitin C (UbC) promoter at an MOI of 10. After 14 days, stably transduced cells were selected by sorting mNeonGreen-positive cells on a FACSAria II (BD). Xenograft transplantations were performed by subcutaneously injecting 5 x10⁶ cells in 100 μl of PBS into the flanks of recipient mice. Prior to intravital imaging, the fur above of the injection site was removed with household depilatory cream. After anesthesia, 50 μl of Akalumine-HCl substrate (15 mM, Wako) were injected intraperitoneally and mice were placed in an IVIS in vivo imaging system (PerkinElmer). Images were acquired after 10-15 minutes using appropriate binning (1) and exposure settings.

Human HSC culture and HBB CRISPR/Cas9 gene editing

CD34⁺ cord blood HSCs (0.5-1 x10⁶) were thawed and cultured in human HSC bulk expansion medium (see above) at a density of 0.3-0.5 x10⁶ cells/ml before gene editing on day 2. The guide RNA design targeting the human hemoglobin beta gene (HBB_gRNA10) was derived from DeWitt et al. (Table S4).[53] Editing was performed with the DZ-100 nucleofection program on a 4D nucleofector (Lonza) using the P3 primary cell buffer kit and 20 μl cuvettes (Lonza) analogous to our murine HSC editing protocol with the same amounts of Cas9 protein, sgRNA and ssODN template (synthesized at IDT, Table S4). After editing, cells were allowed to recover in fresh bulk expansion medium overnight before viable cells were cloned on day 3.

Human HSC single cell cloning and differentiation-expansion cultures

For single cell HSC expansion, we adopted a previously established protocol for the generation of erythro-myeloid-megakaryocytic colonies from single HSC clones.[54] Briefly, two days prior to HSC cloning, flat-bottom 96-well plates (Corning) were treated with 50 μl gelatin (0.2%) (Sigma-Aldrich) for 1 hour. After removing the gelatin solution, 1.5 x10³ MS-5 murine bone marrow stroma cells in 100 μl HS5100 medium (Stemcell Technologies) were seeded into each well and allowed to become adherent. Two days later, immediately prior to cloning, H5100 medium was replaced with 100 μl of HSC differentiation-expansion medium (StemPro34 (Thermo Fisher) supplemented with 1% PSG and human cytokines FLT3L (20 ng/ml), GM-CSF (20 ng/ml), SCF (100 ng/ml), TPO (100 ng/ml), EPO (3 ng/ml), IL-2 (10 ng/ml), IL-3 (10 ng/ml), IL-6 (50 ng/ml), IL-7 (20 ng/ml) and IL-11 (50 ng/ml); all from Peprotech). Viable (PI negative) HSCs were cloned by fluorescence-activated cell sorting (FACSFusion, (BD)) at a density of one cell per well. Single clones were cultured for 14 days, with the addition of 100 μl fresh medium after seven days. At the end of the culture period, wells containing colonies were marked for downstream analyses (genotyping, ddPCR and transplantation assays).

Genotyping of HBB-edited single cell human colonies

Fifty microliters were recovered from each well and transferred to a fresh 96-well plate. Cells were spun down on the plate at 400g for 5 minutes. Using a multichannel pipette, as much supernatant as possible was removed without disturbing the cell pellet. Adopting a protocol outlined in a previous report,[55] 20 μl of Arcturus PicoPure DNA Extra (Thermo Fisher) extraction solution was added to each well and the crude lysate was incubated at 60°C for 3 hours, followed by 75°C for 30 minutes. As for the Prkdc genotyping protocol, a nested PCR design was devised. For the first genotyping PCR, two microliters of the lysate were used with KOD FX Neo polymerase (Toyobo) and primers: HBB_outer_F, HBB_outer_R according to the manufacturer's instructions. The PCR reaction setup was as follows: initial denaturation 94°C, 120s; followed by 35 cycles of denaturation 98°C, 10s; annealing 60°C, 30s; extension 68°C, 45s; and final extension 68°C, 120s. The second PCR was performed with two microliters of the first reaction and HBB_inner_F, HBB_inner_R, using identical cycle times. Analogous to our murine Prkdc genotyping workflow, PCR products were gel purified and Sanger sequenced (huHBB_seqprimer), and editing outcomes were analyzed with TIDE.

Human hematopoietic colony phenotyping

For phenotypic analysis of generated hematopoietic colonies, we followed a previously published protocol by Notta et al.[56] Briefly, 30 μl of cell culture were transferred from each well onto a fresh 96-well plate. Cells were washed on plate once (400 g, 5 min.) and stained with a lineage marker mix (huCD41-BV421, huCD235a-FITC, huCD71-PE, huCD34-APC, huCD33-PE/Cy7, huCD45-APC/Cy7). After washing once, cells were read out on a FACSVerse analyzer (BD) equipped with a high throughput sampler using the appropriate filters and settings. CD235a expression marked erythroid cells, while CD71 expression and CD33 indicated megakaryocytic and myeloid lineages, respectively. Ten or more positive events were required to consider a colony lineage positive (Figure S6C).

Human hematopoietic colony SCT and analysis

Selected human hematopoietic clones were combined and transplanted into 8-10 week-old male NOG mice (In-Vivo Science Inc.) via intrasplenic injection. Mice were irradiated with 1.5 Gy immediately prior to transplantation. Seven to ten days post transplantation, the recipients were sacrificed, and cells were recovered from the spleen and bone marrow. Cells were stained with huCD45-APC/Cy7 and muCD45.1-eFluor 450, and huCD45⁺ cells were sorted on a FACSFusion (BD) cell sorter. Genomic DNA was isolated and HBB editing outcomes quantified as described above (Indel and HDR quantification in bulk-expanded cultures).

Whole exome sequencing (WES) of single cell-expanded clones

To identify mutations acquired during single cell expansion, we isolated c-Kit⁺ cells from the bone marrow of one C57BL/6 mouse. Genomic DNA was isolated from an aliquot of this population (NucleoSpin Tissue column, Macherey-Nagel). From this parental c-Kit⁺ population, CD34⁻CD150⁺KSL cells were cloned and cultured in HSC expansion medium for 28 days, followed by gDNA extraction.

WES Library generation and sequencing was performed at Macrogen Japan Corp using on a NovaSeq sequencer (Illumina). Target enrichment by was performed using SureSelect (Agilent). FASTQ files were mapped with Burrows-Wheeler Aligner (BWA). PCR duplicates were flagged with Picard. Base Quality Score Recalibration and variant calls were performed with GATK. Variant annotation was performed with SnpEff. Sequence data was annotated using the mouse reference genome (GRCm39) and filtered for mutations that were present only in the expanded clone and not in the parental population. We only considered nonsynonymous mutations with a coverage ≥ 10.

To detect sequence divergence in a gene-edited, expanded and transplanted clone, CD45.1$^+$c-Kit$^+$Lin$^-$ cells were isolated from secondary recipients and gDNA was extracted for WES analysis. Since the parental population's data was not available, we compared the exome sequences to the mouse reference genome and searched for divergence in the genes specified in the text.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details regarding employed statistical tests as well as number of subjects and groups are stated in the figure legends. Student's t-tests, one- and two-way analysis of variance (ANOVA) were performed in Prism (version 9.1, Graphpad). Error bars denote standard deviations, unless otherwise stated in the figure legends. Statistical evaluation surrounding RNA-seq analysis, correlation calculations between chimerism and CD201$^+$CD150$^+$KSL marker expression, as well as select visualizations of peripheral blood lineage distributions were performed in R version 4[57] with the appropriate packages (outlined in key resources table). Pictograms and illustrations were generated with BioRender (https://www.biorender.com/).

Supplemental Information

Controlling genetic heterogeneity in gene-edited

hematopoietic stem cells by single-cell expansion

Hans Jiro Becker, Reiko Ishida, Adam C. Wilkinson, Takaharu Kimura, Michelle Sue Jann Lee, Cevayir Coban, Yasunori Ota, Yosuke Tanaka, Meike Roskamp, Tsubasa Sano, Arinobu Tojo, David G. Kent, and Satoshi Yamazaki

A

All events

Lymphocytes 76%

SSC-A / FSC-A

Lymphocytes

Singlets 95.9%

SSC-H / FSC-H

Singlets

Live cells 93.1%

PI / CD201-PE

Live cells

Ckit⁺Lin⁻ (KL) 56.6%

Lin-APC-Cy7 / c-Kit-APC

KL

CD201⁺KL 4.3%

CD201-PE / c-Kit-APC

CD201⁺KL

CD201⁺CD150⁺KL 25.4%

c-Kit-APC / CD150-PE-Cy7

B

Percentage of HDR-corrected alleles

KL / CD201⁻CD150⁻KL / CD201⁻CD150⁺KL / CD201⁺CD150⁺KL

* ** ***

C

Gr1⁺Mac1⁺ ■ B220⁺ ■ CD4⁺ ■ CD8⁺

Percentage of splenocytes

CB17/SCID +SCT / CB17/WT / CB17/SCID

** **** *

D

■ CD4⁺ ■ CD8⁺ ■ CD4⁺CD8⁺

Percentage of thymocytes

CB17/SCID +SCT / CB17/WT / CB17/SCID

**** ***

E

CB17/SCID +SCT

HE

CK

F

Murine *Prkdc*

-1.1 kb (5') / -0.5 kb (5') / 0.0 kb / +0.6 kb (3') / +1.1 kb (3')

Exon 84   TAA>TAT   Exon 85

-1000   -500   0   +500   +1000

position relative to cutsite [bp]

Fig. S1, Becker et al.

Figure S1. Functional correction of *Prkdc*scid HSCs. Related to Fig. 1.

(A) Gating strategy for the isolation of CD201+CD150+KL cells from CB17/SCID mouse BM. **(B)** Frequency of HDR+ alleles within phenotypically defined HSPC populations 7 days post gene editing (n=3 cultures). **(C)** Immunophenotype of splenocytes 20 weeks post-SCT (n=3 mice per group). **(D)** Frequencies of double (CD4+CD8+) and single positive (CD4+ and CD8+) thymocytes 20 weeks post-SCT. Data points represent individual mice. **(E)** Sections of thymi isolated from a CB17/SCID recipient transplanted with gene edited HSCs 20 weeks post-SCT. Upper panel: hematoxylin-eosin (HE), lower panel: cytokeratin (CK) stains; Magnified section of left panels (scale bar: 500 µm) shown on right (scale bar: 300 µm). **(F)** Design of ddPCR probes for copy number assays. Blue boxes: probes; red line: cut site; blue line: targeted mutation. Black arrows: PCR primers.

One-(B) and two-way(C, D) ANOVA with Tukey's multiple comparison test.
Error bars represent SD. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

A

All events
Lymphocytes 83.9%

Lymphocytes
Singlets 99.6%

Singlets
live 99.8%

live
KL 97.4%

KL
KSL 68.4%

KSL
CD201⁻CD150⁺ KSL 14.8%
CD201⁺CD150⁺ KSL 53.8%

CD201⁺CD150⁺KSL
CD201⁺CD150⁺ CD48⁺KSL 74.2%
CD201⁺CD150⁺ CD48⁻KSL 20.6%

B

CD201⁺CD150⁺ CD48⁺KSL

CD201⁺CD150⁺ CD48⁻KSL

CD201⁻ CD150⁺KSL

C

D

enriched in CD201⁻CD150⁺ KSL

enriched in CD201⁺CD150⁺ CD48⁻KSL

positive regulation of cell migration
positive regulation of locomotion
ameboidal-type cell migration
positive regulation of cell motility
response to wounding
positive regulation of cellular component movement
cell junction organization
vasculature development
blood vessel development
sensory perception of chemical stimulus
RNA processing
ncRNA metabolic process
ncRNA processing
rRNA metabolic process
ribonucleoprotein complex biogenesis
DNA-dependent DNA replication
ribosome biogenesis
rRNA processing

NES

E

HSC marker expression profile PVA expansion d14

Fig. S2, Becker et al.

Figure S2. Transcriptional and immune phenotype of cultured, gene-edited HSCs. Related to Fig. 2. **(A)** Gating strategy applied for the isolation of CD201$^+$CD150$^+$CD48$^-$KSL, CD201$^+$CD150$^+$CD48$^+$KSL and CD201$^-$CD150$^+$KSL cells for RNA-seq analysis on day 10 of bulk expansion. **(B)** Principal component analysis (PCA) of CD201$^+$CD150$^+$CD48$^-$KSL, CD201$^+$CD150$^+$CD48$^+$KSL and CD201$^-$CD150$^+$KSL cells (n=3 replicates). **(C)** Heatmap showing the top 40 differentially regulated (up- and down-regulated) genes in CD201$^+$CD150$^+$CD48$^-$KSL (1) and CD201$^-$CD150$^+$KSL (3) populations, with expression data of CD201$^+$CD150$^+$CD48$^+$KSL (2) cells (n=3 replicates). **(D)** Nine most significantly enriched GO terms (biological process) in CD201$^+$CD150$^+$CD48$^-$ and CD201$^-$CD150$^+$KSL populations. **(E)** Expansion of precultured and cloned CD201$^+$CD150$^+$CD48$^-$KSL cells. CD34$^-$CD150$^+$KSL cells were isolated and expanded for 10 days before cloning. The colonies generated from these clones in PVA-based expansion medium were analyzed 14 days post-sort. Left: Representative images of single clone-derived HSC colonies 13 days post-sort. Right: Phenotypic fractions within expanded colonies, as a percentage of live cells (n=14).
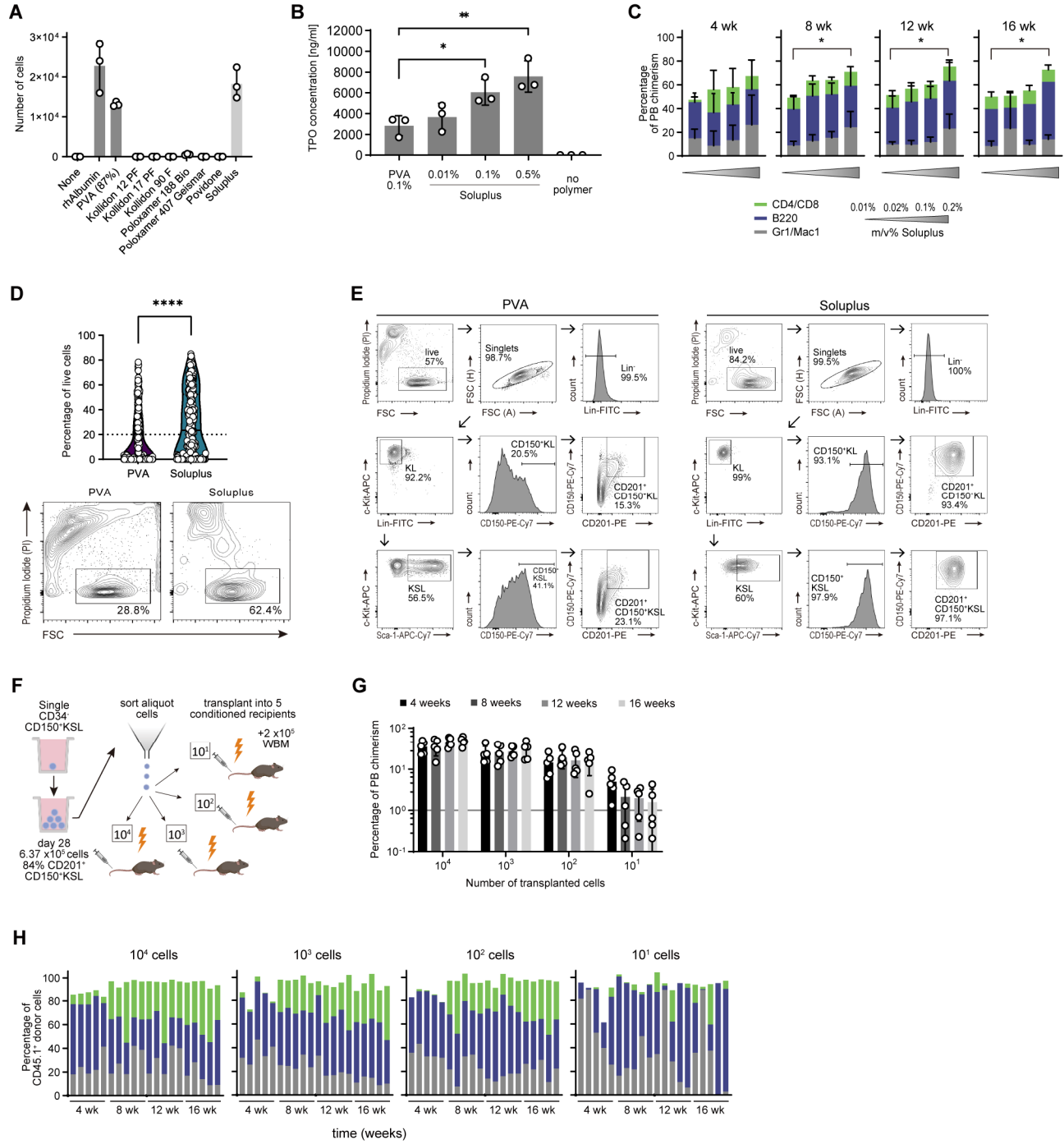
Fig. S3, Becker et al.

Figure S3. Titration of Soluplus supplementation and immune phenotype of Soluplus-expanded HSC clones. Related to Fig. 3.

(A) Albumin replacement polymer screening for *ex vivo* expansion of murine HSCs. 50 freshly isolated C57BL/6 CD34⁻KSL cells were cultured in media supplemented with the indicated polymers for 7 days. Recombinant human albumin and PVA (87% hydrolyzed) served as positive controls. Concentration of all polymers was 0.1% (m/v). Total cells were counted to assess growth support (n=3 cultures). **(B)** Murine TPO ELISA. Ham's F12 with 100 ng/ml TPO was supplemented with the indicated polymers and cultured for 3 days. ANOVA with Tukey's multiple comparison test. (n=3). **(C)** Soluplus supplementation titration assay. Fifty C57BL/6-Ly5.1 (CD45.1⁺) CD34⁻CD150⁺KSL cells grown in titrated concentrations of Soluplus (0.01%, 0.02%, 0.1% and 0.2%) were expanded for 14 days and split-transplanted into CD45.2⁺ recipients (n=4 to 9 per group) against 2 x10⁵ CD45.1⁺/CD45.2⁺ whole bone marrow competitor cells. Peripheral blood (PB) chimerism and lineage distribution is shown. Supplementation with 0.2% Soluplus produced mild micelle formation in cultures, which was not toxic, but occasionally obstructed the visibility of cells. Two-way ANOVA with Tukey's multiple comparison test. **(D)** Left panel: Percentage of viable cells in HSC colonies grown from freshly isolated single CD34⁻CD150⁺KSL cells after 19 days of culture in PVA (n=288)- and Soluplus (n=290)-based media, as evaluated by flow cytometry (%PI⁻ of all events). Right panel: Representative FACS plots of single-cell derived colonies (day 19). Two-tailed Mann-Whitney test. **(E)** Representative FACS plots of individual clones expanded in PVA- and Soluplus-supplemented media. **(F)** Schematic of limiting dilution assay (LDA). **(G)** Donor PB chimerism four to 16 weeks post-SCT (5 mice per group). Cutoff level (1%) denoted with gray line. **(H)** PB lineage distribution of CD45.1⁺ donor cells. Each bar represents an individual recipient.

Error bars represent SD. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.
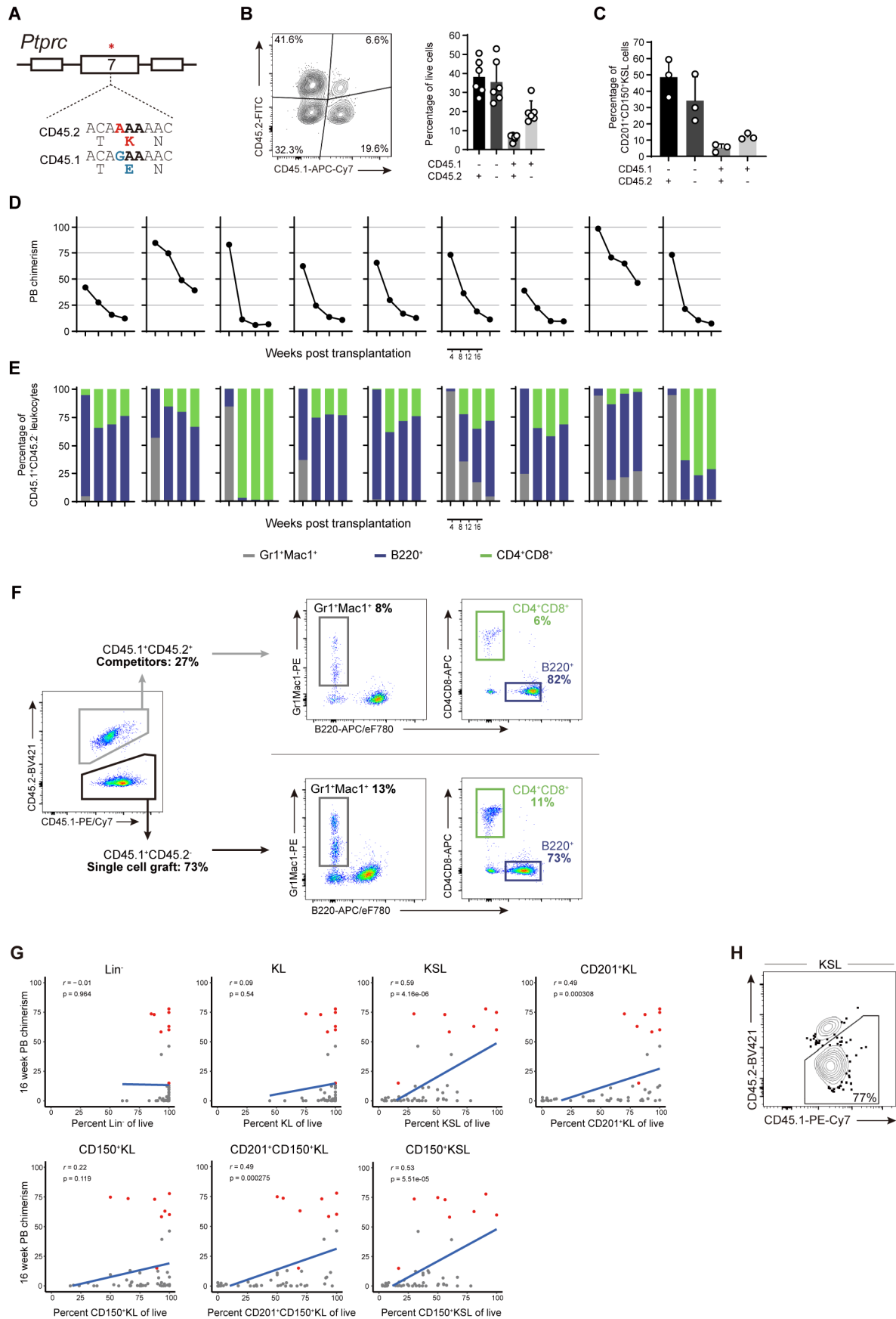
Fig. S4, Becker et al.

Figure S4. *Ptprc* allele conversion efficiency and single clone transplantation assays. Related to Fig. 4. **(A)** Genomic context of the *Ptprc*ᵃ and *Ptprc*ᵇ alleles. White boxes: exons. *denotes site of SNP. **(B)** Distribution of CD45 phenotypes among live cells 4 days post-editing. Representative FACS plot (left) and summary data (right, n=6 cultures). **(C)** Distribution of CD45 phenotypes among CD201⁺CD150⁺KSL cells 4 days post-editing (n=3 cultures).

(D-E) CD45.1⁺ chimerism **(D)** and lineage distribution **(E)** in single recipients that did show LT chimerism ≥5% but did not display multilineage distribution (defined as each lineage ≥5%). Four to 16 weeks post-SCT. Related to Fig. 4d-e. **(F)** Peripheral blood analysis of one of the transplanted mice from Fig. 4D at 16 weeks post-SCT. The single cell graft (CD45.1⁺CD45.2⁻) as well as the whole bone marrow competitor graft (CD45.1⁺CD45.2⁺) contribute similar amounts to the three lineages examined, suggesting that there is no linage bias in the expanded HSC graft. **(G)** Correlation plots of different expansion culture phenotypes versus LT donor chimerism (16 weeks). Red dots indicate multilineage and LT repopulating clones. Pearson correlation. **(H)** Bone marrow CD45.1⁺ donor chimerism within the KSL compartment of a representative primary recipient.

Error bars represent SD.

A

cKit
Sca-1
Lin

CD150
CD201
CD105

normalized log(FI)

all sorted cells

CD201hi, Sca-1hi

Sca-1-BV605

CD201-PE

not transplanted
selected for transplant

B

Prkdc   Off-target
#1 #2 #3 #4 #5 #6 #7

clone #

10

20

30

40

50

60

90

use for SCT

not selected

Prkdc^HDR/HDR
Prkdc^HDR/indel
Prkdc^HDR/scid
Prkdc^noHDR

no off-target mutation

not determined

C

live cells

5.3%   33.5%

CD8-PE-Cy7

31.7%   29.5%

CD4-APC

D

$V_{\beta1}$
$V_{\beta8.1}$
$V_{\beta9}$

Relative frequency

amino acids

E

Gr1+Mac1+
B220+
CD4+
CD8+

PB cell frequency [µl⁻¹]

weeks post SCT

F

IgM
19d p.i.

IgG
19d p.i.

Optical density (450nm)

**   *

****   ****

SCID+SCT   WT   SCID-SCT

G

HDR/HDR: 2 copies

TAT

TAT

HDR/Δ: 1 copy

TAT

deletion

Sanger sequence

TACGCTATGCT

Homozygous gene correction not distinguishable from hemizygosity

H

Prkdc 0.0 kb probe

single cell clones   NTC

#1 #2

Probe Amplitude

Amplitude

Internal control Amplitude

Amplitude

copy number: 2.0   1.0   -

Fig. S5, Becker et al.

Figure S5. Functional rescue of SCID immunodeficiency with edited and single- cell expanded HSC graft. Related to Fig. 5.

(A) Quantification of markers associated with HSC expansion. Marker expression at sorting was compared between clones that were transplanted after the expansion period (CD201$^+$CD150$^+$KL$\geq$10%, HDR$^+$) and those that were not (remaining clones). Left: Fluorescence intensity (FI) measured at sorting. Data presented as log-transformed and normalized to mean. Box plots with whiskers showing minimum and maximum. Right: Representative FACS sorting plot, transplanted clones indicated in red. **(B)** Genotypes of all profiled candidate clones (n=96). **(C)** Representative FACS plot showing frequencies of double (CD4$^+$CD8$^+$) and single positive (CD4$^+$ and CD8$^+$) cells from the thymus of a CB17/SCID recipient 20 weeks after transplantation. **(D)** CDR3 length spectratype analysis of the Tcrb-V1, -V8.1 and -V9 genes in splenic CD4$^+$ cells. Each bar represents the relative frequency of a CDR3 length species (n=1 mouse). **(E)** Absolute PB cell frequencies in CB17/SCID recipients. (F) Serum levels of NIP$_{30}$-specific IgG and IgM 19 days after immunization (CB17/SCID+SCT: n=4 mice; CB17/WT and treatment-naïve CB17/SCID: n=3). (G) LD schematic **(H)** Example data showing ddPCR quantification of *Prkdc* alleles (0.0kb probe) in two clones.

One-(E) and two-way (A, B) ANOVA with Tukey's multiple comparison test.
Error bars represent SD. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

A

Human *HBB*

-0.9 kb (5') -0.4 kb (5') 0.0 kb +0.5 kb (3') +1.2 kb (3')

GAG>GTA   Exon 1   Exon 2   Exon 3

position relative to cutsite

B

Gene edited   neg. control

HBB probes Amplitude

Internal control Amplitude

Copy number: 1.9 1.9 1.4 1.9 1.9   2.1 2.1 2.1 2.1 2.0

C

E, M, Meg   E, M   M, Meg   M only   E only

CD235a-FITC

CD41-BV421

CD33-PE-Cy7

huCD45-APC-Cy7

D

E only   18.5%
M only   1.5%
Meg only   1.5%

E, M   56.7%
E, Meg   0%
M, Meg   4.6%

E, M, Meg   9.2%

Percentage of clones

E

sorted huCD45⁺ cells   Wildtype ctrl.

C C T G A A G A G   C C T G T A G A G   C C T G A G G A G

HDR/non-HDR   HDR/HDR
(49%)   (100%)

Fig. S6, Becker et al.

Figure S6. Expansion and profiling of single, *HBB* gene-edited human HSCs. Related to Fig. 6.

(A) Editing strategy of the *HBB* locus with ddPCR probes for copy number assays. Blue boxes: ddPCR probes; red line: cut site; blue line: targeted mutation. **(B)** Example ddPCR quantification data of bulk gene edited CB HSCs (left) and negative control (unedited) cells. **(C)** Example FACS plots showing colony differentiation outcomes for erythroid (E, CD235a$^+$), myeloid (M, CD33$^+$) and megakaryocytic (Meg,CD41$^+$)-containing colonies. A lineage was defined positive if more than ten events were detected in the in the relevant gate. **(D)** Summary of single huHSPC differentiation outcomes. E, erythroid; M, Myeloid; Meg, megakaryocytic. (n=65). **(E)** Example sequencing traces from sorted huCD45$^+$ cells isolate from NOG recipients transplanted with HPSCs heterozygous (*HBB*$^{\text{HDR/non-HDR}}$) or homozygous (*HBB*$^{\text{HDR/HDR}}$) for the SCD mutation. Healthy control showing wild-type sequence (right).

#	Hit	Permutation	Mismatches	Position (GRCm38)	strand	Gene
	CTTACCAAGTTATAACAGCTNGG	-	-	-	-	Prkdc_gRNA1
1	**A**TTAC**TT**AGTTATAACAGCTGGG	No indel	3	Chr5:57647418-57647440	-	4932441J04Rik
2	**C**TTA**A**AAAGT**C**ATAACAGCTGGG	No indel	3	Chr5:21558299-21558321	-	Fbxl13; Lrrc17
3	CTTA**−TT**AA**C**TTATAACAGCTAGG	Del 15, or Del 16	2	Chr14:96347342-96347363	-	Klhl1
4	CTTA**A**CAA**−TT**C**TAACAGCTGGG	Del 12	2	Chr8:73726599-73726620	+	No known gene
5	CTTA**G**CAA**−TT**T**TAACAGCTCGG	Del 12	2	Chr1:24130829-24130850	-	No known gene
6	CTT**C**CCAAG**−TA**G**AACAGCTGGG	Del 10, or Del 11	2	Chr4:54112620-54112641	+	No known gene
7	CTT**C**CCAAG**−TA**T**T**ACAGCTTGG	Del 10, or Del 11	2	Chr2:140296967-140296988	+	Sel1l2

Supplemental Table S1. Candidate off-target sites. Related to Fig. 1.

Candidate off-target sites of the guide RNA targeting *Prkdc* were identified using the COSMID algorithm. Mismatches and permutations in the seed region are indicated in red.

The first row shows the seed sequence and PAM of the sgRNA used (Prkdc_gRNA1).

Chr	Position	Reference	Allele	Count	Coverage	Freq.	Control count	Control coverage	Transcript affected	AA change in longest transcript	Gene
Indels											
chr1	85504851	CACACACACACA CACACACACACA CACAG	C	11	39	28.2%	0	13	ENSMUST00000178024.2; c.257_284delCACACACACACACACACACAC ACACAGA	p.Thr86fs	ENSMUSG00000094127; G530012D18Rik
chr1	85504943	GAC	G	34	101	33.7%	0	39	ENSMUST00000178024.2; c.349_350delCA	p.Gln117fs	ENSMUSG00000094127; G530012D18Rik
chr2	98497361	CAAGAAAACTGA AAATCA	C	75	111	67.6%	0	449	ENSMUST00000099683.2; c.239_255delTGATTTTCAGTTTTCTT	p.Leu80fs	ENSMUSG00000075014; Gm10800
chr4	126059360	TC	T	6	34	17.6%	0	20	ENSMUST00000080919.12; c.2684delG	p.Arg895fs	ENSMUSG00000043962; Thrap3
chr9	3002239	A	ATTTCTCAT TTTTCACAT	2	16	12.5%	0	18	ENSMUST00000151376.3; c.575_576insATTTTCTCATTTTTCAC	p.Phe194fs	ENSMUSG00000091028; Gm10722
chr9	3025119	C	CA	2	15	13.3%	0	11	ENSMUST00000099046.4; c.583dupA	p.Ser195fs	ENSMUSG00000095186; Gm10718
chr17	23530080	TC	T	56	355	15.8%	0	93	ENSMUST00000168033.3; c.325delG	p.Glu109fs	ENSMUSG00000091945; Vmn2r114
chr17	23530164	GA	G	184	300	61.3%	0	93	ENSMUST00000168033.3; c.241delT	p.Ser81fs	ENSMUSG00000091945; Vmn2r114
SNPs											
chr1	85504886	A	G	15	46	32.6%	0	18	ENSMUST00000178024.2; c.290A>G	p.Glu97Gly	ENSMUSG00000094127; G530012D18Rik
chr1	85504923	G	C	33	78	42.3%	0	36	ENSMUST00000178024.2; c.327G>C	p.Glu109Asp	ENSMUSG00000094127; G530012D18Rik
chr1	85504941	G	C	34	93	36.6%	0	39	ENSMUST00000178024.2; c.345G>C	p.Glu115Asp	ENSMUSG00000094127; G530012D18Rik
chr1	85504970	C	T	43	124	34.7%	0	39	ENSMUST00000178024.2; c.374C>T	p.Ala125Val	ENSMUSG00000094127; G530012D18Rik
chr1	85504990	A	G	42	125	33.6%	0	39	ENSMUST00000178024.2; c.394A>G	p.Ile132Val	ENSMUSG00000094127; G530012D18Rik
chr1	133284294	G	A	23	91	25.3%	0	63	ENSMUST00000094556.3; c.658G>A	p.Glu220Lys	ENSMUSG00000070645; Ren1
chr2	98492825	T	G	11	52	21.2%	0	23	ENSMUST00000099684.4; c.244T>G	p.Cys82Gly	ENSMUSG00000075015; Gm10801
chr2	98497422	T	G	1009	1036	97.4%	0	2036	ENSMUST00000099683.2; c.220A>C	p.Ser74Arg	ENSMUSG00000075014; Gm10800
chr3	98334623	G	T	2	11	18.2%	0	12	ENSMUST00000056096.15; c.387G>T	p.Glu129Asp	ENSMUSG00000050064; Zfp697
chr4	126059371	C	T	6	34	17.6%	0	27	ENSMUST00000080919.12; c.2674G>A	p.Gly892Ser	ENSMUSG00000043962; Thrap3
chr4	126059400	C	T	4	23	17.4%	0	23	ENSMUST00000080919.12; c.2645G>A	p.Arg882Gln	ENSMUSG00000043962; Thrap3
chr7	11775508	T	G	4	18	22.2%	0	12	ENSMUST00000227320.2; c.283T>G	p.Cys95Gly	ENSMUSG00000095864; Vmn1r77
chr7	11775524	C	T	7	18	38.9%	0	13	ENSMUST00000227320.2; c.299C>T	p.Ala100Val	ENSMUSG00000095864; Vmn1r77
chr7	11775731	A	T	4	23	17.4%	0	23	ENSMUST00000227320.2; c.506A>T	p.Tyr169Phe	ENSMUSG00000095864; Vmn1r77
chr9	3000927	C	A	2	11	18.2%	0	12	ENSMUST00000151376.3; c.6C>A	p.Cys2*	ENSMUSG00000091028; Gm10722
chr9	3006907	A	C	3	15	20.0%	0	11	ENSMUST00000179881.2; c.631A>C	p.Asn211His	ENSMUSG00000096385; Gm11168
chr9	3025133	C	T	2	14	14.3%	0	11	ENSMUST00000099046.4; c.596C>T	p.Ser199Leu	ENSMUSG00000095186; Gm10718

chr11	115895814	C	A	4	16	25.0%	0	11	ENSMUST00000106458.2; c.4199C>A	p.Pro1400Gln	ENSMUSG00000020758; Itgb4
chr15	71335125	G	T	16	97	16.5%	0	37	ENSMUST00000022953.10; c.2068C>A	p.Pro690Thr	ENSMUSG00000036800; Fam135b
chr17	23530099	T	C	236	352	67.0%	0	93	ENSMUST00000168033.3; c.307A>G	p.Lys103Glu	ENSMUSG00000091945; Vmn2r114
chr17	23530105	T	G	230	346	66.5%	0	93	ENSMUST00000168033.3; c.301A>C	p.Ile101Leu	ENSMUSG00000091945; Vmn2r114
chr17	23530113	C	A	226	348	64.9%	0	93	ENSMUST00000168033.3; c.293G>T	p.Arg98Ile	ENSMUSG00000091945; Vmn2r114
chr17	23530117	C	G	221	344	64.2%	0	93	ENSMUST00000168033.3; c.289G>C	p.Val97Leu	ENSMUSG00000091945; Vmn2r114
chr17	23530182	A	G	155	269	57.6%	0	93	ENSMUST00000168033.3; c.224T>C	p.Val75Ala	ENSMUSG00000091945; Vmn2r114
chr17	23530186	T	A	153	264	58.0%	0	93	ENSMUST00000168033.3; c.220A>T	p.Asn74Tyr	ENSMUSG00000091945; Vmn2r114
chr17	23530192	A	G	142	249	57.0%	0	93	ENSMUST00000168033.3; c.214T>C	p.Tyr72His	ENSMUSG00000091945; Vmn2r114
chr17	45879194	G	A	7	61	11.5%	0	47	ENSMUST00000024739.14; c.2026C>T	p.His676Tyr	ENSMUSG00000023944; Hsp90ab1
chr17	45879196	G	A	7	59	11.9%	0	48	ENSMUST00000024739.14; c.2024C>T	p.Thr675Ile	ENSMUSG00000023944; Hsp90ab1
chr17	45879205	T	A	7	56	12.5%	0	46	ENSMUST00000024739.14; c.2015A>T	p.Asp672Val	ENSMUSG00000023944; Hsp90ab1

Supplemental Table S2. Mutations in a single cell-derived HSC colony. Related to Fig. 3.

Nonsynonymous mutations (indels and single nucleotide variations (SNPs)) detected in an HSC clone expanded over a period of 28 days. Whole exome sequencing (WES) was

performed on the expanded clone and the parent population from which the clone was derived. Both samples were compared to the murine reference genome, and only mutations

detected in the expanded clone and not in the parent population were considered.

Chromosome	Position	Reference	Allele	Count	Coverage	Frequency	Transcript affected	AA change in longest transcript	Gene
chr5	147293801	AT	A	38	40	95.0%	ENSMUST00000049324.13; c.1313-40delA	None (intron_variant)	ENSMUSG00000042817; FLT3
chr5	147294645	CT	C	6	18	33.3%	ENSMUST00000049324.13; c.1312+112delA	None (intron_variant)	ENSMUSG00000042817; FLT3
chr1	65198550	TA	T	16	36	44.4%	ENSMUST00000097709.11; c.*73delT	None (3_prime_UTR_variant)	ENSMUSG00000025950; IDH1
chr7	79745297	TG	T	6	34	17.6%	ENSMUST00000107384.10; c.1271+71delC	None (intron_variant)	ENSMUSG00000030541; IDH2
chr7	79748080	CCCAGGG	C	24	24	100.0%	ENSMUST00000107384.10; c.679-66_679-61delCCCTGG	None (intron_variant)	ENSMUSG00000030541; IDH2

Supplemental Table S3. Screen for sequence variants in critical genes in a gene edited and transplanted HSC clone. Related to Fig. 4.

Gene-edited CD45.1$^+$cKit$^+$Lin$^-$ cells were recovered from the bone marrow of CD45.2$^+$ secondary recipients of a single cell-derived HSC graft, followed by WES analysis. We checked several known tumor-associated genes for sequence variants: TP53, FLT3, NPM1, IDH1, IDH2, EVI1, TET2, RUNX1, EZH2 and DNMT3A. Since the parent population was not available, sequences were compared to the mouse reference genome. Therefore, it is possible that some of these variants are part of the genomic background of the mouse strain and have not been acquired during the editing and single cell expansion process.

Supplemental Table 4. Oligonucleotides used in this study. Related to STAR Methods.

Sequences of DNA and RNA oligos used in this work.

Sequence	Source	Identifier
Synthetic guide RNA (sgRNA) (mN*: Phosphorothioated 2'-O-methyl RNA base)		
mC*mU*mU*ACCAAGUUAUAACAGCUGUUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U	This study	Prkdc_gRNA1
mA*mU*mA*CUUCAAUUUGUUUGGAGGUUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U	This study	Ptprc_gRNA1
mU*mG*mC*CCCACAGGGCAGUAAGUUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U	DeWitt et al., 2016 [S1]	HBB_gRNA10
Single-strand oligodeoxynucleotide (ssODN) HDR templates (N*: Phosphorothioated DNA base)		
G*G*A*TTCAAGAAATAAATGTAACGGAAAAGAATTGGTATCCACAACATAAAATACGCTATGCTAAGAGGAAGTTAGCAGGTGCCAATCCAGCTGTTATAACTTGGTAAGACTTGTGAATGCAGAA*T*C*A	This study	Prkdc_HDR_ssODN_asym
T*G*C*CCAGCATCGTACCTGGCTCACAGTGGAGTACATATGAAATATTGTCACTGTTGCATTTTCTGAAATCAAGGTTTTCTGTCTTCCATTCCAAACAAATGGAAGTATTAGCCTTTTCTTTTGG*T*G*T	This study	Ptprc_HDR_ssODN_asym
T*C*AGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGACTCCTGTAGAGAAGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAG*G*T	This study	HBB_HDR_ssODN_asym
PCR and sequencing primers (FAM-: 5' 6-carboxyfluorescein label)		
CTTTGTTTTAGGGTCATTACTTGGT	This study	Prkdc_inner_F
TGCTCAGAACTGAAGTCTAAGGT	This study	Prkdc_inner_R
CAATTATCCAGACTATCCCCGAAA	This study	Prkdc_outer_F
TCTTGCCTACACCCTGTAAAGC	This study	Prkdc_outer_R
TCCCTCCTAGAAGCACTTGTT	This study	Ptprc-e7-F
CCCCTAGCGAAATCTCCTGC	This study	Ptprc-e7-R
AGCGGAAGGGCAACTTTACTA	This study	OT_01_outer_F
GATCTCATACACAGACAGGGAAG	This study	OT_01_inner_F
GCATACGCCATTCTGCTCAC	This study	OT_01_inner_R_SEQ[a]
TGCCTGTATGGTAAGCACCC	This study	OT_02_inner_F_SEQ[a]
TTAGTGGCAGGCATGCTTCA	This study	OT_02_outer_R
AGTCACCATTCGTGTGTCCC	This study	OT_02_inner_R
TGGATCTTGTTCATCTGGGGC	This study	OT_03_inner_F_SEQ[a]
GCCTGAGGGACTCAGTATTGT	This study	OT_03_outer_R
AGATTTCAAGATGTCCTTACGA	This study	OT_03_inner_R
TATTAGCACCTACACCAATGCT	This study	OT_04_outer_F
GTGGCACAAAGAAAGATGTATGG	This study	OT_04_inner_F
TGCAGTCCTTGTAAGAGGGT	This study	OT_04_inner_R_SEQ[a]
GATGCAGCTGAGAGACTCGT	This study	OT_05_inner_F_SEQ[a]

CACTCCCTGTGTGTTTGTTTC	This study	OT_05_inner_R
AGTAGGTCTTTGTAGGCACGC	This study	OT_05_outer_R
CGACAACACTCTGACTCCCATA	This study	OT_06_outer_F
CCTTTCCCTTGGGTACTTCTTG	This study	OT_06_inner_F
TGCAAATGACCGGAAATCTGTAAA	This study	OT_06_inner_R
GCCCTCAAGATTTTGCTGTCAAG	This study	OT_06_SEQ[a]
TACCTTCTCACAAGCAGGGAGG	This study	OT_07_outer_F
GAGGAGCCTTATGGAAGAGTTG	This study	OT_07_inner_F
TTAAGGCATTCCTGTCTGCCA	This study	OT_07_inner_R_SEQ[a]
CTGAATGCCCAGACAGCTCCAAGC	Ahmed et al., 2009 [S2]	TCR-Vb1
CATTACTCATATGTCGCTGAC	Ahmed et al., 2009 [S2]	TCR-Vb8.1
TGCTGGCAACCTTCGAATAGGA	Ahmed et al., 2009 [S2]	TCR-Vb8.3
TCTCTCTACATTGGCTCTGCAGGC	Ahmed et al., 2009 [S2]	TCR-Vb9
CTTGGGTGGAGTCACATTTCT	Ahmed et al., 2009 [S2]	TCR-Cb
FAM−CTTGGGTGGAGTCACATTTCT	Ahmed et al., 2009 [S2]	TCR-Cb-FAM
AACTTGGTAAGACTTGTGAATGC	This study	Prkdc_dd_00-2_F
ACACAGTGAAGTGCCATACT	This study	Prkdc_dd_00-2_R
GAAGAGGGAGCACCTGAATTA	This study	Prkdc_dd_5p-500-1_F
TGTTCTTACAGAGGACCAACC	This study	Prkdc_dd_5p-500-1_R
TCAGTTCTGAGCAGTAGTAAGATG	This study	Prkdc_dd_3p-500-1_F
ATTGATTTCTGCAGCTCATTCTC	This study	Prkdc_dd_3p-500-1_R
GCTCACTGACACCATGGAAA	This study	Prkdc_dd_5p-1000-1_F
TCACAATCCAAAGGAGACATGA	This study	Prkdc_dd_5p-1000-1_R
TGGTTCTTCCACCTCCAATC	This study	Prkdc_dd_3p-1000-3_F
CATGATGCCTATGTCTGAGGAG	This study	Prkdc_dd_3p-1000-3_R
GCCCCAGCACGACCATT	This study	Dot1l_dd_1_F
TAGTTGGCATCCTTATGCTTCATC	This study	Dot1l_dd_1_R
AAGTAACTAATGCACAGAGCACAT	This study	huHBB_outer_F
AATGTACTAGGCAGACTGTGTAA	This study	huHBB_outer_R
ATGCTTAGAACCGAGGTAGAGTT	This study	huHBB_inner_F
CCTGAGACTTCCACACTGATGC	This study	huHBB_inner_R
ACCATGGTGCATCTGACTC	This study	huHBB_00-3_F
GGTCTCCTTAAACCTGTCTTGT	This study	huHBB_00-3_R
AACGATCTTCAATATGCTTACCAAG	This study	huHBB_5p-500-3_F
CCCATACCATCAGTACAAATTGC	This study	huHBB_5p-500-3_R
CACGTGGATCCTGAGAACTT	This study	huHBB_3p-500-1_F
CTTCCTATGACATGAACTTAACCATAG	This study	huHBB_3p-500-1_R
GTCCAGGCAGAAACAGTTAGA	This study	huHBB_5p-1000-3_F
CAACCCAAAGTGTGACTATCAATG	This study	huHBB_5p-1000-3_R

ACTGATGTAAGAGGTTTCATATTGC	This study	huHBB_3p-1000-1_F
GCCTAGCTTGGACTCAGAATAA	This study	huHBB_3p-1000-1_R
CTAGGGTTGGCCAATCTACTCC	This study	huHBB_seqprimer
aSEQ denoted off-target PCR primers were also used as sequencing primers		
ddPCR probes (synthesized at IDT)		
CCAGCTCTCAAGTCG	This study	Dot1l-1_probe (SUN)
TGCAAAGCACTTCACACACTTCTGAGC	This study	Prkdc_00-2_probe (FAM)
ACCAATGTTTCTGGGACCTGAATGCT	This study	Prkdc_5p-500-1_probe (FAM)
AGTGTCAGATCTATACAGTGACTGGTGCT	This study	Prkdc_3p-500-1_probe (FAM)
TGTGAAGGAACCCTCCTTTGATTGGAAGG	This study	Prkdc_5p-1000-1_probe (FAM)
TCTGTACCTCCGGTTCTAGGGAACCT	This study	Prkdc_3p-1000-3_probe (FAM)
CAAGGTGAACGTGGATGAAGTTGG	This study	huHBB_00-3_probe (FAM)
CGTAAATACACTTGCAAAGGAGGATGT	This study	huHBB_5p-500-3_probe (FAM)
AGGGTGAGTCTATGGGACGCTTGA	This study	huHBB_3p-500-1_probe (FAM)
ACCTCCTATTTGACACCACTGATTACCC	This study	huHBB_5p-1000-3_probe (FAM)
CAGCTACAATCCAGCTACCATTCTGCT	This study	huHBB_3p-1000-1_probe (FAM)

SUPPLEMENTAL REFERENCES

S1.   DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.-J., Mitros, T., Munoz, D.P., Boffelli, D., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci Transl Med *8*, 360ra134-360ra134. 10.1126/scitranslmed.aaf9336.

S2.   Ahmed, M., Lanzer, K.G., Yager, E.J., Adams, P.S., Johnson, L.L., and Blackman, M.A. (2009). Clonal Expansions and Loss of Receptor Diversity in the Naive CD8 T Cell Repertoire of Aged Mice. The Journal of Immunology *182*, 784–792. 10.4049/jimmunol.182.2.784.