nature biotechnology

Article

https://doi.org/10.1038/s41587-026-03142-z

Selection of human hematopoietic stem cells bearing the intended functional edit by transient AND-gate reporters

Received: 29 October 2024

Accepted: 21 April 2026

Published online: 01 June 2026

Check for updates

Daniele Canarutto[1,2,16], Martina Fiumara[1,2,16], Vigneshwaran Venkatesan[1], Chiara Gaddoni[1], Kohei Shiroshita[1], Angelica Varesi[1,4,5], Luisa Albano[1], Gabriele Pelosi[1], Alfredo Silva[1], Claudia Firrito[1,6], Giulia Schiroli[1,7], Deborah Cipria[1], Stefano Beretta[1], Angelo Amabile[1,8,9,10], Anna Villa[1,3], Erika Zonari[1], Bernhard Gentner[1,11,12], Angelo Lombardo[1,2], Pietro Genovese[1,13,14,15], Samuele Ferrari[1,2,17] & Luigi Naldini[1,2,17]

Targeted genomic integration of gene-sized cassettes into hematopoietic stem and progenitor cells (HSPCs) for genetic disease treatment is constrained by the low efficiency of homology-directed repair (HDR) and frequent unintended genetic changes at the editing site. Here, to overcome these challenges, we introduce selection by means of artificial transactivators (SMArT), which transiently implements AND reporter gates to achieve templated integration of a functional cassette at the target site. HDR-edited HSPCs were enriched to 80–100% purity through transient selector expression, whereas cells carrying undesired and potentially genotoxic on-target edits were preferentially depleted. Xenotransplantation of SMArT-enriched HSPCs in immunodeficient mice resulted in fully HDR-edited human grafts with the selector no longer detectable. SMArT strategies were implemented through clinically compliant manufacturing and selectors. They support both safe harbor integration and gene correction, can preserve physiological transcriptional regulation and are portable across loci also with polyfunctional editors. Overall, SMArT strategies may broaden the therapeutic applicability of gene-sized editing while reducing its genotoxic burden.

Gene editing mediated by homology-directed repair (HDR) allows for the insertion of gene-sized DNA sequences into targeted double-strand breaks (DSBs) generated with programmable nucleases, such as RNA-guided CRISPR–Cas systems[1]. The novel sequence is templated by exogenous DNA molecules homologous to the target site, often provided by viral vectors (for example, adeno-associated vector serotype 6, AAV6)[2–8]. HDR-mediated editing can fix or functionally override mutations at disease-causing loci or instruct novel cellular functions upon the insertion of gene expression cassettes in selected target sites or safe harbor loci[3,4,6–13]. However, HDR is confined to S/G2 cell-cycle phases[14], thus constraining HDR editing in slowly cycling and nonproliferating cells, such as long-term repopulating hematopoietic stem and progenitor cells (HSPCs)[15]. Moreover, divergent outcomes may stem from alternative repair pathways engaged at the DSB site. The nonhomologous or microhomology-mediated end joining (NHEJ/MMEJ) pathways can seal the break while inserting or deleting nucleotides (indels). Such an outcome has been mostly exploited to disrupt sequences (for example, enhancers and open reading frames[16–18]). Nuclease-based

A full list of affiliations appears at the end of the paper. e-mail: naldini.luigi@hsr.it

NHEJ/MMEJ-mediated editing has reached clinical approval in HSPC gene therapy, showing compelling efficacy for the treatment of hemoglobinopathies[19,20]. However, despite incremental improvements and proof of concept of therapeutic applicability[11,21–23], clinical applications of HDR editing lag considerably behind because of suboptimal efficiency and lower tolerability by long-term engrafting HSPCs[24,25].

Although nuclease-based gene editing enables precision targeting of the genome at variance with gene replacement strategies through semirandomly integrating vectors, it suffers from the potential genotoxic risk of unintended, disruptive and previously overlooked genomic rearrangements at the target sites triggered by DNA DSB repair[26]. A growing body of evidence supports the contention that large deletions, chromosomal translocations, chromothripsis and insertion of fragmented or full-length HDR template DNA at the nuclease on/off-target sites are relatively frequent outcomes of DNA repair in edited human HSPCs[24,27–32]. Consequently, contrary to its scope, HDR editing often yields genetically heterogeneous populations of treated cells within the final product. A fraction of these cells presumably lack therapeutic efficacy and can bear an undetermined genotoxic risk, which is likely specific to the target locus and the design of the corrective cassette.

To address the suboptimal efficiency of HDR editing and concomitantly mitigate the associated genotoxicity concerns, here, we couple HDR gene editing of human HSPCs with transient expression of a selector, in three SMArT (selection by means of artificial transcriptional activators) strategies with different purposes and stringency. These SMArT strategies enable ex vivo enrichment of cells edited at therapeutically relevant loci, purging of cells bearing unwanted and potentially genotoxic outcomes and/or avoiding permanent expression of the selector in HSPCs.

Results

SMArT-1 enriches HDR-edited cells

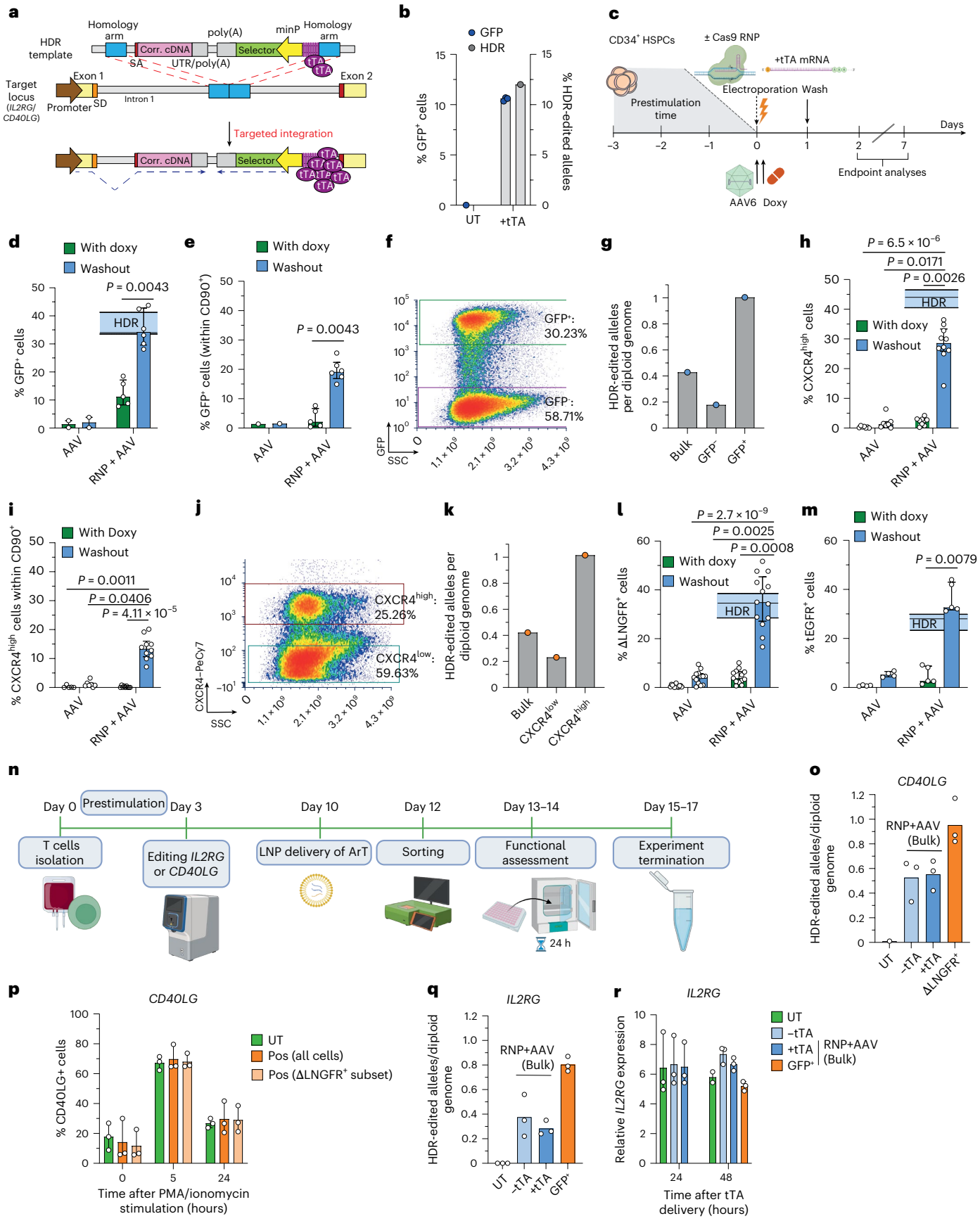
In the first configuration (SMArT-1), we design an AAV6-delivered HDR template as follows. The corrective complementary DNA (cDNA) for the gene of interest (GOI) is flanked by an independent expression cassette for the selector in reverse orientation, comprising seven repeats of the tetracycline operator (TetO7) next to a minimal promoter (minP). This promoter construct does not drive detectable expression unless bound by the tetracycline transactivator (tTA), which is transiently coexpressed with the editor by mRNA transfer. In this configuration, however, selector expression can be driven from both episomal and integrated donor templates. Because tTA activity can be tuned down by doxycycline administration, which inhibits its binding to TetO7, we suppressed selector expression until sufficient dilution of the non-integrated template DNA, thus allowing enrichment for HDR-edited cells (Fig. 1a).

We first applied SMArT-1 to the correction of the *IL2RG* gene[4], whose mutations cause X-linked severe combined immunodeficiency (SCID-X1). *IL2RG* encodes the common gamma chain (γc), which is a constituent of multiple interleukin receptors. *IL2RG* is expressed at high levels by lymphoid cells and at low levels by HSPCs. While successfully treated in clinical trials with lentiviral vector (LV)-mediated gene replacement therapy using the constitutive EF1α promoter[33,34], gene editing might provide the advantage of restoring function and its physiological regulation[4,10]. In K-562 cell lines, upon targeted integration of the SMArT-1 cassette into *IL2RG* and dilution of the non-integrated plasmid template, we found undetectable basal expression of the selector (*GFP*), which was conversely expressed at high levels upon delivery of tTA-encoding mRNA in nearly all HDR-edited cells (Fig. 1b and Extended Data Fig. 1a). We then tested SMArT-1 in male cord blood (CB)-derived HSPCs. Here, the tTA mRNA was codelivered by electroporation together with the *Streptococcus pyogenes* Cas9 ribonucleoprotein (RNP) targeting *IL2RG* after 3 days of prestimulation (Fig. 1c). Cells were then transduced with SMArT-1 AAV6 template, treated with doxycycline for increasing time after electroporation (Extended Data Fig. 1b–d) and compared for GFP expression. Doxycycline washout 24 h after editing resulted in high GFP expression in the edited samples compared to controls lacking RNP (Fig. 1d and Extended Data Fig. 1e). Copies of HDR-edited alleles per diploid genome, as measured by digital droplet PCR (ddPCR), were concordant with the percentage of GFP$^+$ cells. Selector transactivation was present throughout the bulk population, including the most primitive CD90$^+$ compartment, albeit to a lesser extent in this subset (Fig. 1e and Extended Data Fig. 1f). To confirm that GFP-expressing cells bore on-target integration of the cassette, we sorted GFP$^+$ and GFP$^-$ fractions and found nearly saturating HDR editing in the former (Fig. 1f,g). We then replicated these findings, targeting the same locus and substituting the *GFP* selector with *CXCR4*, a potentially clinically compatible selector that may be dually exploited for ex vivo enrichment and as an engraftment enhancer[35] (Fig. 1h–k and Extended Data Fig. 1g,h).

To demonstrate portability of the strategy, we applied SMArT-1 to mobilized peripheral blood (mPB) HSPCs and a different therapeutic model—X-linked hyper-IgM syndrome (HIGM1). HIGM1 is a combined immunodeficiency caused by mutations in the *CD40LG* gene[36], whose safe correction requires restoration of physiological control of expression[37,38]. Moreover, in the mouse model of the

Fig. 1 | SMArT-1 exploits tTA for transient selector expression upon integration. a, Schematics of the SMArT-1 selection strategy. A donor cassette consisting of (1) two homology arms (blue); (2) a splice acceptor sequence (red); (3) a corrective cDNA cassette (pink); (4) a selector (green) under the transcriptional regulation of minP (yellow) and (5) TetO7 (purple) is integrated into the target locus by HDR. Selector expression can be driven by the transiently delivered tTA, which can be repressed by concomitant doxycycline administration. **b**, Percentage of selector (GFP)$^+$ K-562 cells 9 days after editing at *IL2RG* locus upon delivery of the tTA mRNA and percentage of HDR-edited alleles (n = 1, 3 and 1). **c**, Experimental design. CB or mPB HSPCs were prestimulated, electroporated with editing reagents and the tTA mRNA and transduced with AAV6 encoding for the SMArT-1 donor cassette (either *IL2RG* or *CD40LG* loci). Doxycycline was added on the day of electroporation and washed out in the following days. **d**,**e**, Percentage of selector (GFP)$^+$ cells in bulk (**d**) or CD90$^+$ (**e**) CB HSPCs treated as detailed in **c** at *IL2RG*. Doxycycline was washed out 24 h after editing and selector expression was assessed 24 h thereafter by flow cytometry. HDR efficiency by ddPCR is reported in light blue (median and range, n = 2, 2, 5 and 6). Statistical analysis was conducted using a two-sided Mann–Whitney test. **f**, Gating strategy for sorting GFP$^+$ and GFP$^-$ cells. **g**, Fraction of HDR-edited alleles by ddPCR in the GFP$^+$, GFP$^-$ or unsorted bulk edited HSPCs from **f** (n = 1). **h**,**i**, Percentage of selector (CXCR4)^{high} cells in bulk (**h**) or CD90$^+$ (**i**) CB HSPCs treated as detailed in **c** (n = 7, 9, 7 and 10). Statistical analysis was conducted using a Kruskal–Wallis with Dunn's multiple-comparison test. The HDR efficiency range by ddPCR is reported in light blue (median and range). **j**, Gating strategy for sorting CXCR4^{high} and CXCR4^{low} cells. **k**, Fraction of HDR-edited alleles by ddPCR in the CXCR4^{high}, CXCR4^{low} or unsorted bulk edited HSPCs (n = 1). **l**,**m**, Percentage of selector (ΔLNGFR (**l**) tEGFR (**m**))$^+$ cells in bulk mPB HSPCs treated as detailed in **c** (n = 13). Statistical analysis was conducted using a Kruskal–Wallis with Dunn's multiple-comparison test. HDR efficiency range by ddPCR is reported in light blue (median and range). **n**, Experimental setup for SMArT-1 in T cells. T cells were isolated from PB, prestimulated with Transact and edited on day 3. The tTA ArT was delivered 7 days after editing by LNP. Cells were sorted 2 days after LNP delivery and *CD40LG* or *IL2RG* expression was evaluated 1–2 days thereafter. **o**, Fraction of HDR-edited alleles at the *CD40LG* locus in T cells in the absence of the tTA (−tTA) or with the tTA, before sorting (+tTA) and after sorting (ΔLNGFR$^+$) (n = 3). **p**, Expression of *CD40LG* in response to PMA/ionomycin stimulation in UT cells or sorted cells (Pos) and within the ΔLNGFR$^+$ sorted cells (mean and range, n = 3). **q**, Fraction of HDR-edited alleles at the *IL2RG* locus in T cells in the absence of the tTA (−tTA) or with the tTA, before sorting (+tTA) and after sorting (GFP$^+$) (n = 3). **r**, Sum of native and recoded *IL2RG* expression relative to *HPRT* upon editing (−tTA), tTA delivery (+tTA) and sorting (GFP$^+$). Each dot is a biological replicate (mean and range, n = 3). Schematic in **n** was created in BioRender; Ferrari, S. https://biorender.com/7f7mclp (2026).

https://doi.org/10.1038/s41587-026-03142-z



disease, we showed that administering a product made only of functional HSPCs outperforms those comprising the same or even higher doses of functional cells admixed with mutant ones in rescuing IgG production (Extended Data Fig. 1i–k)[8,9,32]. We used two different clinically compliant selectors, the truncated nerve growth factor receptor (ΔLNGFR; Fig. 1l and Extended Data Fig. 1l) and an improved version of the truncated epidermal growth factor receptor (*tEGFR* (ref. 9); Fig. 1m and Extended Data Fig. 1m), obtaining similar levels of editing and transactivation as shown above for *IL2RG*.

We then applied SMArT-1 to human T cells to assess functionality in the most relevant cell type that differentiates from *IL2RG*/*CD40LG*-edited and selected HSPCs. To avoid the need for transient repression of the tTA with doxycycline, we took advantage of longer cell culture for episomal vector dilution. We used our previously developed T cell HDR editing and expansion protocol[39] and coupled it with delayed lipid nanoparticle (LNP) delivery of the tTA mRNA, 7 days after editing[40] (Fig. 1n). Targeted integration in the *CD40LG* locus of the same construct used in HSPCs (Fig. 1l) resulted in high levels of enrichment (Fig. 1o) upon sorting and regulated *CD40LG* expression kinetics (Fig. 1p and Extended Data Fig. 1n)[9,32]. Delivery of the tTA did not demonstrably affect the expression of nearby genes *LNC00892* and *ARHGEF6*, which was maintained within the range of untreated (UT) cells (Extended Data Fig. 1o,p). Similarly, we could enrich T cells edited at the *IL2RG* locus to a high degree of purity (Fig. 1q). *IL2RG* expression was similar to the unedited control upon tTA delivery and sorting (Fig. 1r). In this context, we did observe transient and modest upregulation of the neighboring gene *FOXO4* upon sorting, which rapidly returned to basal levels (Extended Data Fig. 1q). Expression of the neighboring gene *MED12* remained instead at levels comparable to the UT control (Extended Data Fig. 1r).

Overall, these data show that the SMArT-1 strategy can be plugged into the gene-editing process of HSPCs and T cells without impacting its efficiency and recapitulating regulated expression of two paradigmatic disease-associated genes, without notable perturbation of nearby genes.

SMArT-2 enriches functionally edited HSPCs, purging out unintended on-target events

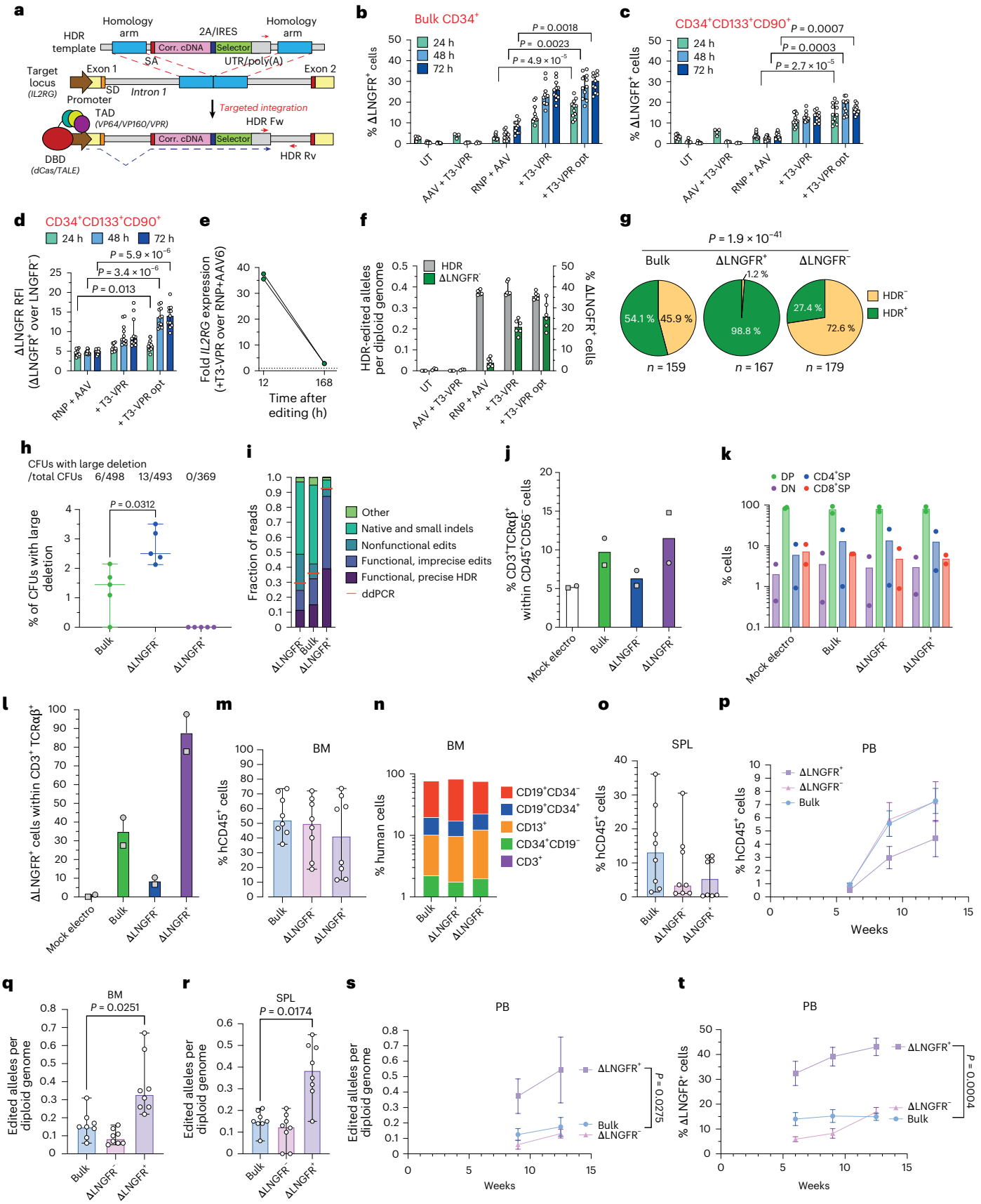
In a second SMArT configuration (SMArT-2), we aimed to subordinate selector expression to functional on-site integration of the HDR template, thus simplifying and increasing the stringency of the selection approach. In SMArT-2, the selector is linked to the corrective cDNA sequence by either self-cleaving peptides (2A) or internal ribosome entry sites (IRESs) and is under the transcriptional control of the endogenous enhancer and promoter sequences of the target gene (Fig. 2a). Transcriptional overexpression of the selector (and the edited gene) can be transiently achieved by delivering an ArT targeting the edited gene promoter, enabling enrichment of cells carrying targeted integration. Whereas residual selector expression will occur in the HSPCs and/or their progeny physiologically expressing the corrected gene, the selection only captures cells bearing the targeted integration. Moreover, if applied to the correction of an X-linked gene, it will concomitantly purge most if not all undesired on-target events.

Thus, we applied this strategy to gene correction of the X-linked *IL2RG* in male cells. Targeted integration of SMArT-2 *IL2RG*-corrective templates into intron 1 of the gene did not alter *IL2RG* expression in human T cells, as opposed to a corrective template carrying the selector (GFP) under the transcriptional control of the ubiquitous constitutive phosphoglycerate kinase promoter (*PGK–GFP*) (Extended Data Fig. 2a). This indicates that the absence of a constitutive promoter minimizes perturbations of physiological gene expression in the most relevant cell type, alleviating concerns stemming from overactive *IL2RG* signaling[41,42].

We then designed three transcriptional activator-like effector (TALE)-based DNA-binding domains (DBDs) targeting the *IL2RG* promoter and fused each of them with a VP160 transactivating domain (TAD), consisting of ten repeats of the VP16 protein (Extended Data Fig. 2b). TALE3 ArT (hereafter referred to as T3) allowed 11-fold *IL2RG* overexpression compared to UT control in K-562 cell line and was selected for further experiments (Extended Data Fig. 2c). We then tested SMArT-2 in male CB HSPCs, coelectroporating the *IL2RG* RNP and the mRNA encoding for the T3 ArT comprising either VP160 or VP64–p65–Rta (VPR) TAD[43] (Extended Data Fig. 2d). Cells were also transduced with an AAV6 carrying a simplified version of the SMArT-2 HDR template, comprising only the selector part of the cassette (splice acceptor (SA).2A.GFP) (Extended Data Fig. 2e). T3-VPR enabled higher overexpression compared to T3-VP160, reaching up to threefold higher relative fluorescence intensity (RFI) compared to the nontransactivated control and was, thus, selected for further studies (Extended Data Fig. 2f). We then edited male mPB HSPCs from different donors using the editing enhancers GSE56 (with or without the adenoviral protein Ad5-E4orf6/7)[5,24,25] and a 2A-based SMArT-2 corrective construct carrying the clinically compatible ΔLNGFR selector. In the absence of T3-VPR delivery, ΔLNGFR expression was low in bulk HSPCs (Fig. 2b) and nearly absent in the most primitive CD90$^+$ compartment (Fig. 2c,d), indicating that efficient enrichment of *IL2RG*-edited

Fig. 2 | SMArT-2 enriches for HSPCs edited by HDR using an ArT and a selector functionally coupled to the corrective cDNA. a, Schematics of the selection strategy. A donor cassette consisting of (1) two homology arms (blue); (2) a splice acceptor sequence (red); (3) a corrective cDNA cassette encoding for downstream exons (pink); (4) a 2A or IRES linker (dark blue); and (5) a selector (green). An ArT consisting of a DBD (dCas or TALE, in red) and a TAD binds the genome on the endogenous promoter of *IL2RG*. **b**–**d**, Expression of the ΔLNGFR selector in bulk (**b**; median and interquartile range (IQR), $n = 7, 4, 11, 11$ and 11) or CD90$^+$ (**c**; median and IQR, $n = 7, 4, 11, 11$ and 11) mPB HSPCs treated as illustrated in Extended Data Fig. 2d at different time points after editing and corresponding selector RFI in the CD90$^+$ population (**d**; median and IQR, $n = 11$). Statistical analysis was conducted using a Kruskal–Wallis with Dunn's multiple-comparison test. T3-VPR opt, optimized mRNA[35]. **e**, Fold *IL2RG* gene expression relative to the nontransactivated control after editing in the standard protocol ($n = 2$). **f**, Fraction of HDR-edited alleles by ddPCR and selector (ΔLNGFR)$^+$ mPB HSPCs after transactivation with T3-VPR or not ($n = 3, 3, 6, 6$ and 6). **g**, Proportion of HDR-positive HSPC-derived colonies in the bulk, selector-positive (ΔLNGFR$^+$) and selector-negative (ΔLNGFR$^-$) fraction ($n = 159$ bulk, 167 negative and 179 positive). A chi-square test for trend was used to compare selector-negative, bulk and selector-positive CFUs in this order ($P = 1.91 \times 10^{-41}$). The pool of colonies was derived from HSPCs edited with either SMArT-2 (**a**) or simplified SMArT-2 (Extended Data Fig. 2e) constructs. **h**, Fraction of colonies harboring deletions encompassing the *IL2RG* promoter, seeded from bulk, selector-negative (ΔLNGFR$^-$) or selector-positive (ΔLNGFR$^+$) HSPCs, assessed by ddPCR (median and range, $n = 498$ bulk, 493 negative and 369 positive). Statistical analysis was conducted using a one-tailed Wilcoxon *t*-test. Each dot represents a biological replicate. **i**, Targeted, native DNA, long-read sequencing of CFUs derived from bulk ($n = 1$ donor), selector-negative (ΔLNGFR$^-$; $n = 1$ donor) and selector-positive (ΔLNGFR$^+$; $n = 2$ donor) HSPCs edited and selected with SMArT-2; the red line indicates the HDR efficiency by ddPCR. **j**,**k**, Differentiation of CD34$^+$ cells in ATOs of electroporated cells (mock electro), bulk, selector-negative (ΔLNGFR$^-$) and selector-positive (ΔLNGFR$^+$) cells. **j**, Percentage of CD3$^+$TCRαβ$^+$ cells. **k**, Distribution of surface markers CD4 and CD8 within CD3$^+$TCRαβ$^+$ cells. DN, double negative (CD4$^-$CD8$^-$); DP, double positive (CD4$^+$CD8$^+$) ($n = 2$ biological replicates). **l**, ΔLNGFR expression in CD3$^+$TCRαβ$^+$ cells ($n = 2$). **m**–**t**, Xenotransplantation of SMArT-2 selector-positive (ΔLNGFR$^+$), bulk and selector-negative (ΔLNGFR$^-$) cells in NBSGW mice (pool of two independent experiments, $n = 8$ mice per group, two HSPC donors). **m**, Engraftment in the BM. **n**, BM composition. **o**, Engraftment in the SPL. **p**, Engraftment in PB. **q**–**s**, HDR efficiency in BM (**q**), SPL (**r**) and PB (**s**). **t**, ΔLNGFR expression in PB. Data are plotted as the median and range (**m**,**o**,**q**,**r**) or mean ± s.e.m. (**p**,**s**,**t**). Statistical analysis was conducted using a Kruskal–Wallis with Dunn's multiple-comparison test (at latest time point for PB). All cells were derived from male donors. Each dot is a biological replicate.

https://doi.org/10.1038/s41587-026-03142-z



HSPCs would not be feasible when using this clinically compliant selector without an ArT. Optimization of the T3-VPR mRNA structure and sequence, with previously described strategies[35], resulted in even greater expression of the ΔLNGFR selector (Fig. 2b,d), without impacting the phenotypic distribution of HSPC subpopulations compared to the no-ArT control (Extended Data Fig. 2g). Of note, *IL2RG* overexpression was transient, as expected, and normalized by day 7–11 after editing (Fig. 2e and Extended Data Fig. 2h); no upregulation of nearby genes *MED12* and *FOXO4* was observed after editing (Extended Data Fig. 2i).

We estimated the efficiency of selection by comparing the fraction of HDR-edited alleles by ddPCR with the percentage of cells expressing the selector and found selector expression with T3-VPR reaching an average of 75% of HDR-edited HSPCs (Fig. 2f). We next investigated at a clonal level the on-target genetic configuration of bulk and sorted HSPCs by screening hundreds of single mPB HSPC-derived colonies per condition with ddPCR assays probing for HDR and large deletions, as previously described[24]. Bulk SMArT-2 edited cells were enriched from 55% to 97% HDR editing in the positive fraction (Fig. 2g), confirming the high stringency of SMArT-2 in male HSPCs; HDR-edited HSPCs were scarce in colonies originating from selector-negative HSPCs. Concordantly with the fact that *IL2RG* is hemizygous in male cells, large deletions encompassing the *IL2RG* promoter were absent in colonies plated from ΔLNGFR$^+$ HSPCs, present in colonies originating from bulk unsorted HSPCs and enriched in those from ΔLNGFR$^-$ HSPCs compared to the former (Fig. 2h). Nanopore long-read sequencing of unamplified genomic DNA confirmed enrichment for predicted functional rearrangements, bearing precise (that is, HDR-mediated) or imprecise (conceivably mediated by NHEJ/MMEJ) transgene–genome junctions at either side (Fig. 2i).

To confirm the T cell differentiation potential of selected HSPCs, we seeded them in artificial thymic organoids (ATOs). Sorted ΔLNGFR$^+$ HSPCs gave rise to differentiated progeny up to the terminal CD4 or CD8 single-positive stage (Fig. 2j,k), showing a diverse T cell receptor (TCR) Vβ repertoire (Extended Data Fig. 2j) and retaining, as expected, the surface ΔLNGFR expression in nearly all cells (Fig. 2l). We then assessed the repopulation potential of SMArT-2 mPB HSPCs by transplanting equal numbers of CD90$^+$ cells from the positively and negatively enriched fractions and the bulk population in immunodeficient mice (Fig. 2m–p). HSPCs positively enriched by SMArT-2 generated robust grafts in the PB and bone marrow (BM), albeit with a trend for lower size in the former compared to bulk or negatively selected cells. Grafts from positively selected HSPCs displayed 2.2–3-fold higher editing efficiency, despite not reaching saturating levels (Fig. 2q–t). We speculate that this discrepancy from in vitro and ATO data, where selected cells comprise nearly 100% of the output, might be because of a more stringent B cell requirement for precise HDR rescue of *IL2RG* expression. Cells bearing imprecise reconstitution may fall short of full recapitulation of physiological control of expression (for instance, by interference of trapped vector sequences) and, thus, be outcompeted in part by residual unedited cells in the sample. While *IL2RG* surface expression appeared similar among all grafts (Extended Data Fig. 2k), this finding may still be consistent with an in vivo disadvantage of B cells with suboptimal *IL2RG* expression versus wild-type (WT) and precisely HDR-edited cells, not captured at this time point. Note, however, that competition with WT cells would not occur when dealing with *IL2RG*-deficient patient cells. Overall, these data highlight the potential of SMArT-2 in enriching for HDR-edited HSPCs and purging for those bearing unintended and potentially genotoxic on-target editing outcomes, while conserving HSPC engraftment and multilineage differentiation potential, including toward both T and B lymphoid lineages.

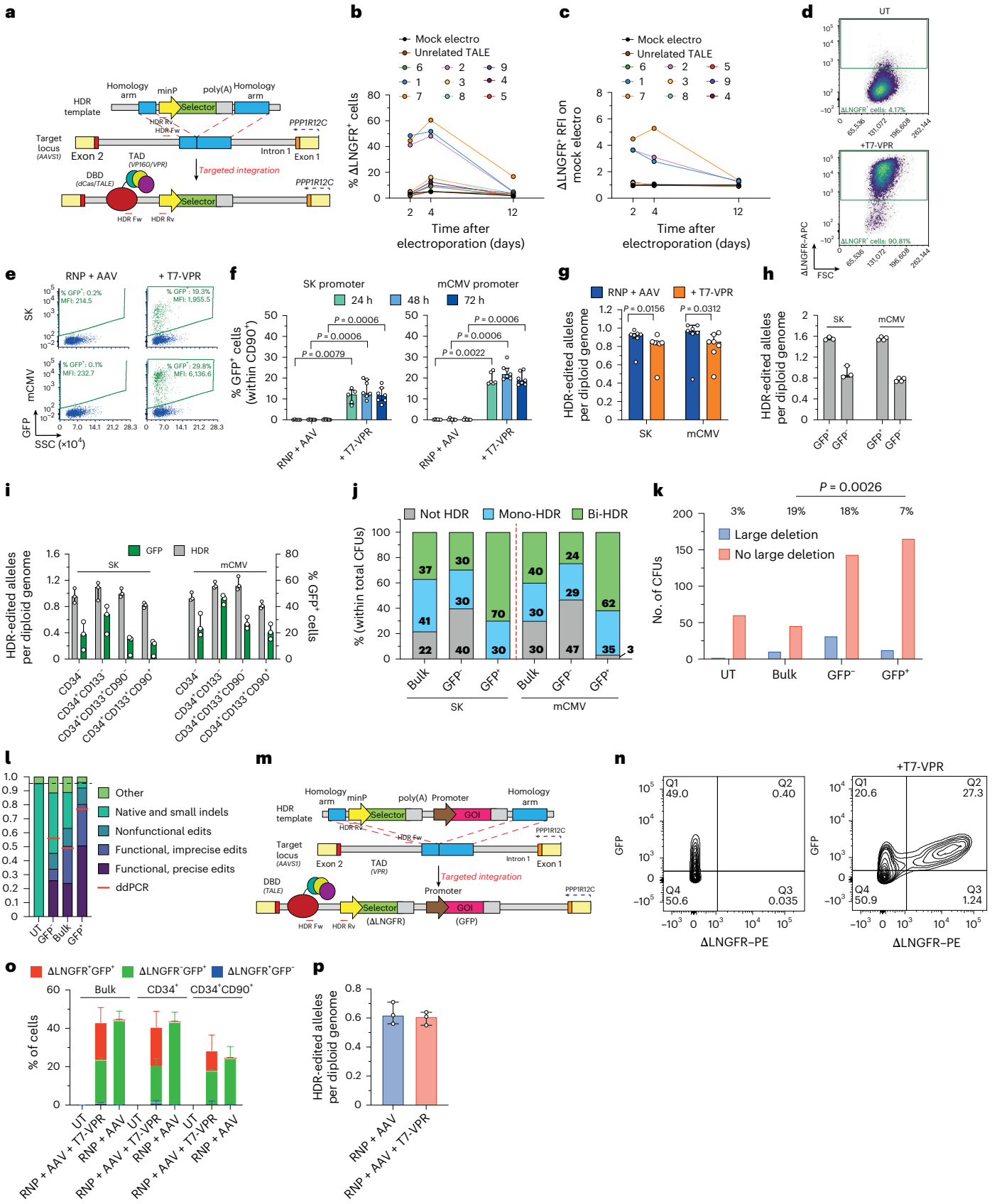
SMArT-3 enriches HDR-edited HSPCs at the *AAVS1* locus with transient selector expression

In some gene-editing applications, expression of the selector conditional to that of the edited gene or transient overexpression of the latter, as in SMArT-2, might be undesirable because of toxicity or immunogenicity concerns. Moreover, some genes may be refractory to transactivation, for example, in the presence of repressive chromatin or tight post-transcriptional regulation by 3′ untranslated region elements. Therefore, we designed another iteration of the SMArT strategy (SMArT-3) in which the HDR template comprises an independent selector cassette driven only by minP. Basal selector expression is, thus, expected to be nearly undetectable and induced to exploitable levels only upon transient delivery of an ArT, which is designed to bind the genome close to but outside of the sequence homologous to the template (Fig. 3a).

We first developed this strategy for the targeted integration into the *AAVS1* safe harbor locus[44] for broader applicability to therapeutic settings. To identify the optimal ArT for selection of *AAVS1*-edited cells, we designed a panel of TALE DBDs fused with VP160–TAD (Extended Data Fig. 3a, top) and a set of gRNAs to be coupled with either dCas9–VP160 or dCas9–VPR (Extended Data Fig. 3a, bottom) spanning the 645-bp genomic region upstream of the 5′ end of the homology sequence. All of them were electroporated as plasmids in a K-562 cell line clone (H1). This clone stably carries monoallelic HDR-mediated on-target integration of a minimal CMV (mCMV) promoter sequence

Fig. 3 | SMArT-3 enriches for HSPCs edited at the *AAVS1* safe harbor locus. **a**, Schematics of the selection strategy. A donor cassette consisting of (1) two homology arms for *AAVS1* locus (blue) and (2) minP (mCMV or SK, in yellow) driving the expression of a selector (green). An ArT consisting of a DBD (dCas or TALE, in red) and a TAD binds the genome upstream the left homology arm, driving the expression of the selector upon targeted integration. **b**,**c**, Percentage (**b**) and RFI (**c**) of the H1 reporter cell line expressing ΔLNGFR upon delivery of different TALE–VP160. **d**, ΔLNGFR expression in the H1 clone upon delivery of the T7-VPR. **e**, Gating strategy for GFP expression after transactivation with T7-VPR and mCMV or SK promoters. **f**, GFP expression in CD90$^+$ HSPCs upon delivery of the T7-VPR transactivator with the SK (n = 5, 7, 7, 5, 7 and 7) and mCMV (n = 6, 7, 7, 6, 7 and 7) promoters. Statistical analysis was conducted using a two-sided Mann–Whitney test (median and IQR). **g**, Fraction of HDR-edited cells by ddPCR upon editing of mPB HSPCs in presence of the T7-VPR (n = 7). Statistical analysis was conducted using a two-sided Wilcoxon matched-pairs signed-rank test (median and IQR). **h**, Fraction of HDR-edited alleles by ddPCR upon sorting of the GFP$^+$ and GFP$^-$ edited mPB HSPC fraction in the different SK (n = 3) and mCMV (n = 4) configurations (median and IQR). **i**, Fraction of HDR-edited alleles by ddPCR and GFP$^+$ cells in sorted mPB HSPC subpopulations with the different SK and mCMV configurations (median and IQR, n = 3). **j**, Monoallelic and biallelic editing efficiency in single colonies derived from bulk (n = 46 SK, n = 10 mCMV), GFP$^-$ (n = 98 SK, n = 77 mCMV) and GFP$^+$ (n = 83 SK, n = 95 mCMV) edited HSPCs with the different SK and mCMV promoters. **k**, Percentage and number of colonies harboring (blue) or not harboring (red) large deletions spanning the T7 DNA-binding site. A chi-square test for trend was used to compare GFP−, bulk and GFP$^+$ CFUs in this order. **l**, Targeted, native DNA, long-read sequencing of UT, bulk, GFP$^-$ and GFP$^+$ enriched CFUs derived from HSPCs edited with the SK vector construct. The horizontal line indicates the fraction of spurious events detected upon sequencing of the UT sample. ddPCR on the same samples is indicated in red. **m**, Schematics of the selection strategy with the inclusion of a model GOI (GFP). A donor cassette consisting of minP (SK, in yellow) driving the expression of the ΔLNGFR selector (green) is integrated by recombination of homology arms (blue) into the *AAVS1* target locus. An ArT consisting of a DBD (TALE7, in red) and a VPR TAD binds the genome upstream the left homology arm, driving the expression of the selector upon targeted integration. A PGK promoter placed downstream the selector drives the expression of the GOI (GFP). **n**, Expression of the GOI (GFP) and selector (ΔLNGFR) in bulk mPB HSPCs 24 h after editing in the absence (left) or presence (right) of the T7-VPR transactivator. **o**, Distribution of the GOI (GFP) and selector (ΔLNGFR)$^+$ cells in bulk edited, CD34$^+$ and CD90$^+$ mPB HSPCs 24 h after editing (median and range, n = 3 biological replicates). **p**, Editing efficiency by ddPCR of samples reported in **o** (median and range, n = 3).

https://doi.org/10.1038/s41587-026-03142-z



regulating ΔLNGFR selector expression, which was undetectable on cell surface at steady state (Extended Data Fig. 3b). As previously observed[43], dCas9–VPR performed better than dCas9–VP160 with most single guide RNAs, resulting in higher percentage and ΔLNGFR RFI at the peak over mock electroporated controls (Extended Data Fig. 3c,d). gRNA 10 (g10) was the most efficient one when coupled with dCas9-based ArTs. Combined delivery of two less effective gRNAs (g4 and g12) resulted in an additive transactivating effect, as previously shown in other contexts[43]. Among TALE–VP160 ArTs, TALE7 (hereafter referred to as T7) was the most efficient one and was selected for further experiments given the orthogonality with the Cas9 nuclease system here used for *AAVS1* editing (Fig. 3b,c). Notably, mRNA delivery of T7-VPR in the H1 clone resulted in potent ΔLNGFR expression in nearly 100% of the cells (Fig. 3d).

We then edited mPB HSPCs from six healthy donors by co-delivering *AAVS1* Cas9 RNP and GSE56/E4orf6/7 mRNA in the presence or absence of the optimized T7-VPR mRNA. We compared two GFP-based selector cassettes bearing the mCMV promoter or its improved version, 'T6-mCMV', which was reported to have even lower basal transcriptional activity (hereafter referred to as 'SK')[45]. Both SK and mCMV cassettes were delivered by AAV6. T7-VPR mediated potent GFP expression in a median of 25% and 35% of bulk-treated HSPCs with SK and mCMV promoters, respectively, at 48–72 h after editing (Fig. 3e and Extended Data Fig. 3e,f). mCMV outperformed SK in terms of percentage and extent of GFP expression upon ArT-mediated transactivation and did not show detectably higher basal activity. We observed a median of 15% and 25% GFP$^+$ cells in the most primitive CD90$^+$ HSPC compartment (Fig. 3f), with about half RFI of the bulk population (Extended Data Fig. 3f,g). These findings suggest that the CD90$^+$ cellular subset might be less permissive to ArT delivery and/or activity in these conditions. No changes in the HSPC phenotype were found (Extended Data Fig. 3h). Gene expression analyses showed normalization of GFP selector transcription in T7-VPR treated samples 7 days after the SMArT-3 procedure (Extended Data Fig. 3i) and no perturbation of expression of the *AAVS1* hosting gene (*PPP1R12C*) and the neighboring one (*TNNT1*) (Extended Data Fig. 3j,k). The percentage of *AAVS1* HDR-edited in the bulk population was 45–50% and similar between the SK and mCMV conditions while being slightly lower in the T7-VPR conditions (Fig. 3g), possibly because of competition between either the repair machinery and TALE–VPR for DNA binding or TALE–VPR and GSE56/E4orf6/7 mRNAs for delivery or translation. Sorting of GFP$^+$ cells greatly increased the fraction of HDR-edited alleles (Fig. 3h), providing the proof of concept of selection of HDR-edited cells by SMArT-3. Persistence of a considerable fraction of HDR-edited alleles in the sorted GFP$^-$ fraction suggested high stringency of the SMArT-3 strategy at the expense of the selection yield. Concordantly, comparison between HDR-edited alleles and percentage of GFP$^+$ cells across sorted HSPC subpopulations highlighted detectable GFP transactivation only in a fraction (estimated around 40–50%) of CD90$^+$ cells carrying targeted integration, as assessed by ddPCR (Fig. 3i). To evaluate the genotype composition and integrity of the target site, we analyzed hundreds of single colonies in sorted and bulk edited HSPCs and found that 100% and 97% of colonies from sorted GFP$^+$ HSPCs in the SK and mCMV conditions, respectively, carried at least one HDR-edited allele (Fig. 3j). Notably, SMArT-3 preferentially enriched for clonogenic HSPCs harboring biallelic HDR editing, particularly in the SK condition, possibly because of better transactivation when both alleles are corrected. Moreover, we found a reduction in the number of clones carrying large deletions spanning over the T7 DNA-binding site from the GFP$^+$ fraction (Fig. 3k) and a consistent increase in HDR events assessed by long-read sequencing (Fig. 3l and Extended Data Fig. 3l). These results highlight the potential of SMArT-3 to increase the purity of the product to at least monoallelic editing of the cells while reducing the carryover of unintended genomic outcomes at the target.

We then modeled the delivery of a therapeutic cassette that included GFP as GOI under the control of the ubiquitous, constitutive PGK promoter and ΔLNGFR as a clinically compliant selector under the SK promoter (Fig. 3m). This configuration was compatible with transactivation, as the inclusion of the PGK promoter did not interfere with transactivation (Fig. 3n,o) and the transactivation did not interfere with editing (Fig. 3p).

Xenografts of SMArT-3 selected HSPCs are fully edited without permanent selector expression

To assess the repopulation potential of SMArT-3-enriched mPB HSPCs, we transplanted an equivalent number of mCMV edited bulk, sorted GFP$^+$ or sorted GFP$^-$ mPB HSPCs from a pool of two human donors in immunodeficient mice (Fig. 4a,c). Strikingly, the grafts from the selected cells showed fivefold higher HDR editing than those from the bulk group, reaching 1–1.5 average copies of HDR-edited *AAVS1* alleles per diploid genome (Fig. 4d,e), indicating a fully edited graft. Moreover, GFP expression was undetectable in all the groups, confirming the lack of basal selector expression (Fig. 4f,g). Graft size, however, was 5–10-fold lower for the sorted GFP$^+$ group than the bulk and sorted GFP$^-$ counterparts in both PB and hematopoietic organs at the end of the study (Fig. 4h,i). This finding might be because of a lower HDR efficiency and, thus, GFP marking in long-term engrafting HSPCs, potentially leading to overrepresentation of more committed progenitors in the selected and infused cell product.

On the basis of this data, we pursued optimization of the HSPC culture and manipulation to overcome the relatively low process yield while achieving proof of translatability. We adapted a clinically compliant cell culture medium and cytokine supplementation, previously developed to obtain ex vivo expansion of repopulating mPB HSPCs[46], and used lower AAV6 multiplicity of infection (MOI) and GMP-compliant electroporation and sorting (Fig. 4j and Extended Data Fig. 4a–m). These optimizations allowed an average tenfold increase in editing and threefold increase in the size of the human xenografts, calculated on pooled experiments targeting the same locus and using different batches of reagents (Extended Data Fig. 4n,o). When applying these conditions to SMArT-3, we reached an average 29% GFP$^+$ cells within the CD90^{high} population, corresponding to a threefold increase from the previous method (Fig. 4b,k). Sorting GFP$^+$ HSPCs resulted in a population of cells with 1.4 HDR-edited *AAVS1* alleles per diploid genome from a value of 0.9 in the bulk (Fig. 4l). As before, xenotransplantation showed HDR levels averaging one edited allele per diploid genome in grafts from the selected cells (Fig. 4m,n). Again, no selector expression was present (Fig. 4o,p). Importantly, however, engraftment was comparable between positively selected and bulk cells, indicating better preservation of the repopulation capacity of treated HSPCs with the optimized process (Fig. 4q,r). A slightly higher graft size was observed in the PB for the negatively selected fraction, possibly suggesting better fitness of HSPCs that escaped one or more hits of the ex vivo editing and selection manipulation process (for example, AAV6 infection and/or DNA DSB). Overall, these results validate SMArT-3 as another strategy to obtain a fully HDR-edited HSPC product with alleviated genotoxic risk, no residual selector expression and preserved repopulating capacity in clinically compliant conditions.

SMArT-3 can be simplified, multiplexed and ported to other loci

To facilitate portability to different loci, we then set out to simplify SMArT-3 by using only RNA-guided DBDs and replaced the TALE–VPR with a dCas9–VPR. To avoid generation of DNA DSBs at the dCas9–VPR-binding sites, we exploited truncated gRNAs (tgRNAs), which allow DNA binding but prevent DSB when coupled with Cas9 (ref. 47). We applied this strategy at the *RAG1* locus, building upon our previous HDR gene correction approach[11], achieving transient selector expression (Extended Data Fig. 5a,b) while abolishing the burdensome requirement of designing a new TALE DBD for each locus.

https://doi.org/10.1038/s41587-026-03142-z

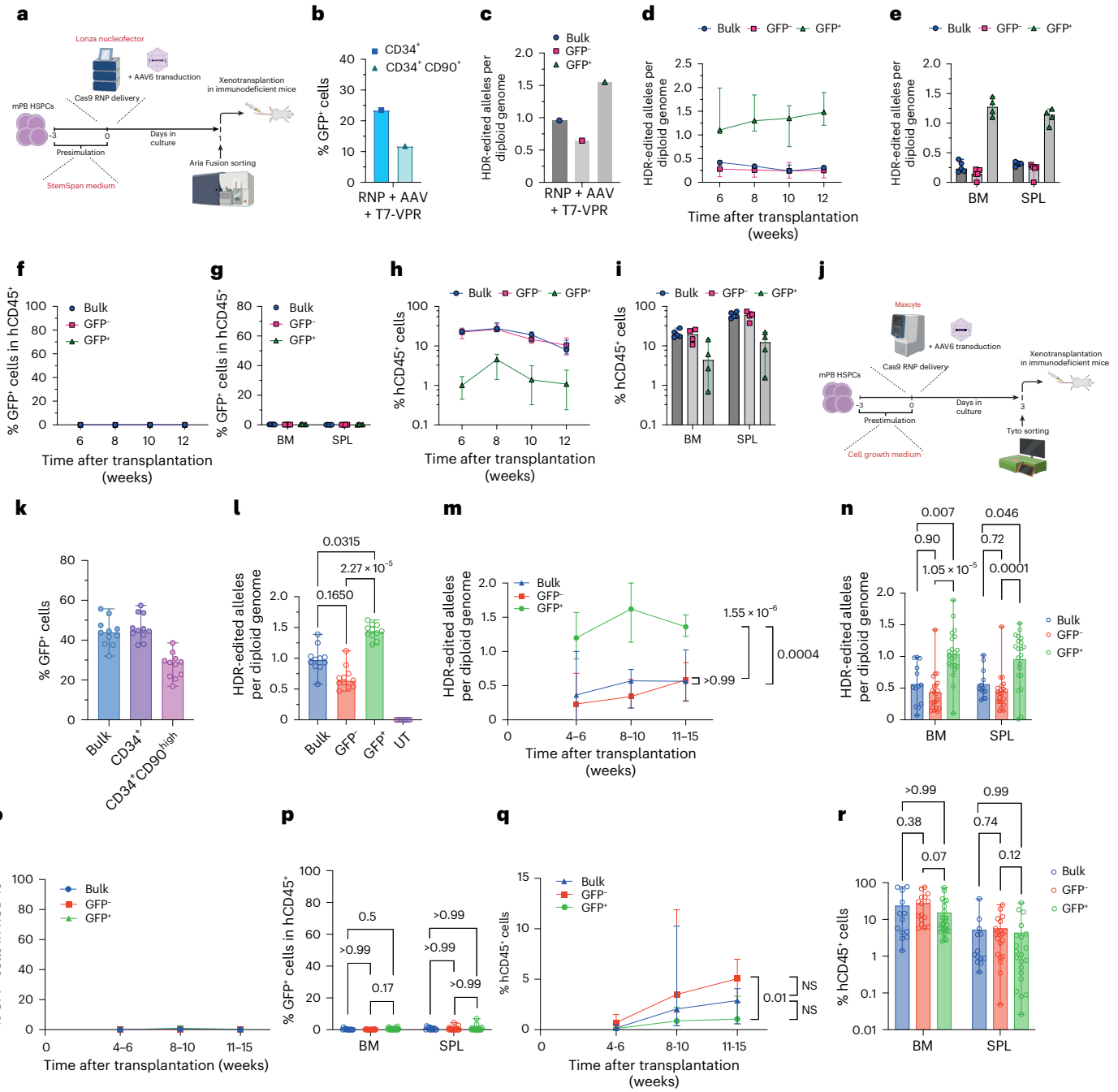


Fig. 4 | SMArT-3-enriched HSPCs engraft upon xenotransplantation. **a**–**i**, Xenotransplantation of edited mPB cells from two donors into a total of 13 NSG mice (median and 95% CI). **a**, Experimental design. mPB HSPCs were cultured for 3 days, electroporated with the Lonza 4D nucleofector and transduced with AAV6 encoding for the mCMV SMArT-3 donor. Then, 2 days after editing, cells were sorted with FACSAria Fusion and transplanted into mice. **b**, Fraction of $CD34^+$ and $CD90^+$ selector $(GFP)^+$ cells on the day of xenotransplant **c**, HDR-edited alleles per diploid genome by ddPCR 7 days after editing of bulk edited cells, GFP^- cells and GFP^+ cells. **d**,**e**, HDR-edited alleles per diploid genome by ddPCR in the human graft over time in the PB (**d**) or BM and SPL (**e**). **f**,**g**, Selector expression over time in the human graft in the PB (**f**) or BM and SPL (**g**). **h**,**i**, Engraftment of human $hCD45^+$ cells in the PB (**h**) or BM and SPL (**i**) of xenografts. Each dot is a different mouse. Data are plotted as the median and IQR (**d**,**f**,**h**) or mean and range (**e**,**g**,**i**). **j**–**r**, Xenotransplantation of SMART-3-selected cells in optimized conditions. **j**, Experimental design. mPB cells from two different donors edited were cultured separately in CG medium for 3 days, electroporated with the Maxcyte GTx electroporator and transduced with the AAV6 SK-GFP vector. Then, 3 days after electroporation, cells were sorted with either the FACSAria of the Milteny Tyto sorter and transplanted into mice. **k**,**l**, In vitro SMArT-3 efficiency with optimized conditions: fraction of bulk, $CD34^+$ and $CD90^{high}$ GFP^+ (**k**; median and range, $n = 11$ biological replicates) and fraction of HDR-edited alleles or diploid genomes by ddPCR in bulk edited, GFP^-, GFP^+ and UT cells (**l**; median and range). **m**,**n**, HDR-edited alleles or diploid genomes by ddPCR in the human graft over time in the PB (**m**) or BM and SPL (**n**) (pool of $n = 21$ positive, 21 negative and 13 bulk mice from different experiments). **o**,**p**, Selector expression over time in the human graft in the PB (**o**) or BM and SPL (**p**). **q**,**r**, Engraftment of human $hCD45^+$ cells in the PB (**q**) or BM and SPL (**r**) of xenografts. NS, not significant. Each dot is a different mouse. Data are plotted as the median and IQR (**m**,**o**,**q**) or mean and range (**n**,**p**,**r**). Statistical analysis was conducted using a Kruskal–Wallis with Dunn's multiple-comparison test at the latest time point for PB. Schematics created in BioRender; **a**, Ferrari, S. https://BioRender.com/fu6gmkk (2026); **j**, Canarutto, D. https://BioRender.com/zk3ltof (2026).

We then further simplified the strategy by exploiting a single nuclease-competent Cas9–ArT editor for both HDR and transactivation at the respective target sites[48–50]. Notably, tgRNAs may also be optimized in length and combination or for maximum selector expression and multiplexed to achieve robust transactivation of endogenous genes of interest, such as *CXCR4* (refs. 35,51) (Fig. 5a). To this end, we first screened the previously identified gRNAs for *AAVS1* (Extended Data Fig. 3c,d) to establish the optimal truncation length that abolishes indel formation and maximizes selector transactivation. Electroporation of plasmids encoding gRNAs of 19 nt to 11 nt into the H1 reporter cell clone, together with Cas9–ArT mRNA, showed that progressive truncation of the gRNAs led to complete loss of nuclease activity along with increasingly stronger reporter transactivation (Fig. 5b). Additionally, given the multiplexing potential of polyfunctional editor toward transactivation of endogenous genes, we designed a panel of gRNAs spanning the *CXCR4* promoter (Extended Data Fig. 5c). We identified two gRNAs with the strongest transactivation activity upon electroporation of plasmids encoding full-length gRNA together with dCas9–ArT into K-562 cells (Extended Data Fig. 5d). As observed for *AAVS1*, gRNAs when used in combination produced a synergistic increase in CXCR4 RFI (Extended Data Fig. 5e). The 15-nt tgRNA variant of *CXCR4* g116 and the 14-nt variant of *CXCR4* g160 yielded the strongest transactivation when used with Cas9–ArT with no indels for tgRNAs shorter than 16 nt (Extended Data Fig. 5f). We then combined Cas9–ArT with *AAVS1* gRNA, *AAVS1* truncated g10 and CXCR4 truncated g116 and g160 for simultaneous transactivation of the ΔLNGFR reporter and *CXCR4* in HSPCs, resulting in a ~12-fold increase in CXCR4 RFI in the treated cells (Fig. 5c,d), among which 30% were also ΔLNGFR^{+} (Fig. 5e), allowing selection of a fully HDR-edited CXCR4^{high} cell product (Fig. 5f).

To demonstrate portability, we then applied the same workflow to *IL2RG* correction. Here, we modified the SMArT-2 donor construct to include the minimal SK promoter to drive selector expression in a SMArT-3 configuration (Fig. 5g). After screening for the two most efficient 15-nt tgRNAs in HSPCs (that is, *IL2RG* truncated g2 and g5; Extended Data Fig. 5g), we were able to transactivate edited HSPCs and enrich them to a very high degree of purity, with minimal impact of addition of tgRNAs on editing efficiency (Fig. 5h,i).

In summary, the Cas9–VPR polyfunctional editor simplifies SMArT-3, allowing for editing and selection of HSPCs with a single protein. Its versatility empowers portability and multiplexing, paving the way for the design of novel therapeutic strategies.

Discussion

Here, we implemented SMArT, transient synthetic AND gates selecting for integration of a gene-sized cassette into a genomic safe harbor or functional correction of disease-causing genes. SMArTs enrich HDR-edited HSPCs by making selector expression conditional to the intended functional genetic outcomes through ArT binding to the target locus and result in robust engraftment of a pure product with reduced or virtually abrogated genotoxic burden. Hereby, SMArTs tackle the suboptimal efficiency of integration of gene-sized DNA templates into HSPCs by HDR and the potential on-target genotoxicity of DNA DSBs, which constitute major concerns of an otherwise promising avenue for the treatment of genetic diseases.

We previously developed a strategy to enrich male T cells edited at the X-linked *CD40LG* locus with Cas9 and a donor template and demonstrated its positive effects in increasing purity and reducing the genotoxic burden of HDR gene editing[9,32], before pursuing its clinical translation. This is a rather privileged setting, where selector expression can be coupled with functional correction of T cells without the need for an ArT. Here, instead, we transiently instruct a synthetic transactivation circuit composed of ArT binding to genomic sequences repurposed as enhancer-binding sites and coupled with either the target gene promoter or an exogenous minP embedded in the HDR template.

The first iteration, SMArT-1, is inherently portable to virtually any site without the need for customization and selector expression wanes as the ArT disappears. However, in principle, it is blind to whether integration occurred at the target site, in the correct orientation or not, or off-target. Compared to the previously reported CRISPRa enrichment[52], SMArT-1 may be implemented by a single delivery of nucleic acids, thus being gentler on HSPCs; moreover, is more efficient and is compatible with previously validated, clinically compliant selectors. Because of the easier ex vivo manipulation and lower risk of transformation of T cells, SMArT-1 may find application in the field of inborn errors of immunity or cancer gene therapy.

SMArT-2 builds upon our previously published T cell enrichment strategy; it is effective in selected settings, such as *IL2RG*, which is expressed at low levels in HSPCs. As male cells bear a single copy of *IL2RG*, selection results in a pure, fully edited product, with exhaustive purging of large deletions encompassing the ArT-binding domain and, thus, by extension the *MED12* gene, located immediately upstream of *IL2RG*. Indeed, loss of function of *MED12* is implicated in the development of T lymphoblastic leukemia[53]. SMArT-2, thus, constitutes an alternative strategy for the genetic correction of *IL2RG* deficiency, which couples in situ restoration of physiological gene regulation with great abatement the genotoxic risk of gene editing[27] by purging out cells bearing large deletions potentially encompassing tumor suppressor genes. SMArT-2-selected cells preserve their robust multilineage differentiation potential, including toward functional T and B lymphoid lineages; as such, SMArT-2 could become a better and potentially safer therapeutic strategy than LV-mediated gene therapy, amenable also to the treatment of dominant negative mutations where LV gene addition is ineffective[54].

SMArT-3 is arguably the most advanced strategy, which bypasses the requirement for an actionable endogenous promoter by SMArT-2 by exploiting an exogenous minP embedded in the HDR template. Moreover, conversely to SMArT-2, selector expression extinguishes shortly after the process. While SMArT-3 also reduces the genotoxic burden of large deletions, full purge out cannot be expected when editing autosomal loci or the X chromosome in females. Of note, using very strong promoters in the context of SMArT-3, even when targeting a safe genomic harbor, might not be advisable because of the likelihood of constitutively transactivating the selector. In its latest Cas9–VPR-based iteration, tgRNAs replace the TALE DBD and SMArT-3 becomes easy to adopt, multiplexable and portable across safe harbor and gene correction approaches, requiring a single universal polyfunctional editor. In future developments, this polyfunctional editor may allow to couple selection with additional, transient endowment of increased cell fitness and expansion, further improving engraftment and graft function through the upregulation of endogenous genes, as shown here by concomitant overexpression of the selector and *CXCR4*.

Selector expression abates nearly completely after selection when SMArT drives it independently from the endogenous gene promoter. This allows for stealthier reconstitution of hematopoiesis. While selector expression allows tracking the gene-edited cells and, potentially, their depletion in the host, biological selectors such as ΔLNGFR may sequester its bioactive endogenous ligand with potentially adverse consequences. This concern most applies when the selector is widely expressed throughout hematopoiesis, as in previously reported selection strategies based on constitutive selector expression[3].

Another study reported a strategy to enrich for cells harboring targeted integration in essential genes in lymphoid cells (SLEEK—selection by essential-gene exon knock-in)[55]. This strategy, however, constrains the choice of target sites, while SMArT-3 is more flexible, is portable to different loci of interest thanks to its simplified design leveraging on polyfunctional editors and can be exploited for gene correction. Another reported approach exploits simultaneous editing of two different loci with orthogonal platforms, with editing outcome at the selector locus conferring resistance to a cytotoxic drug[56]. In addition

https://doi.org/10.1038/s41587-026-03142-z

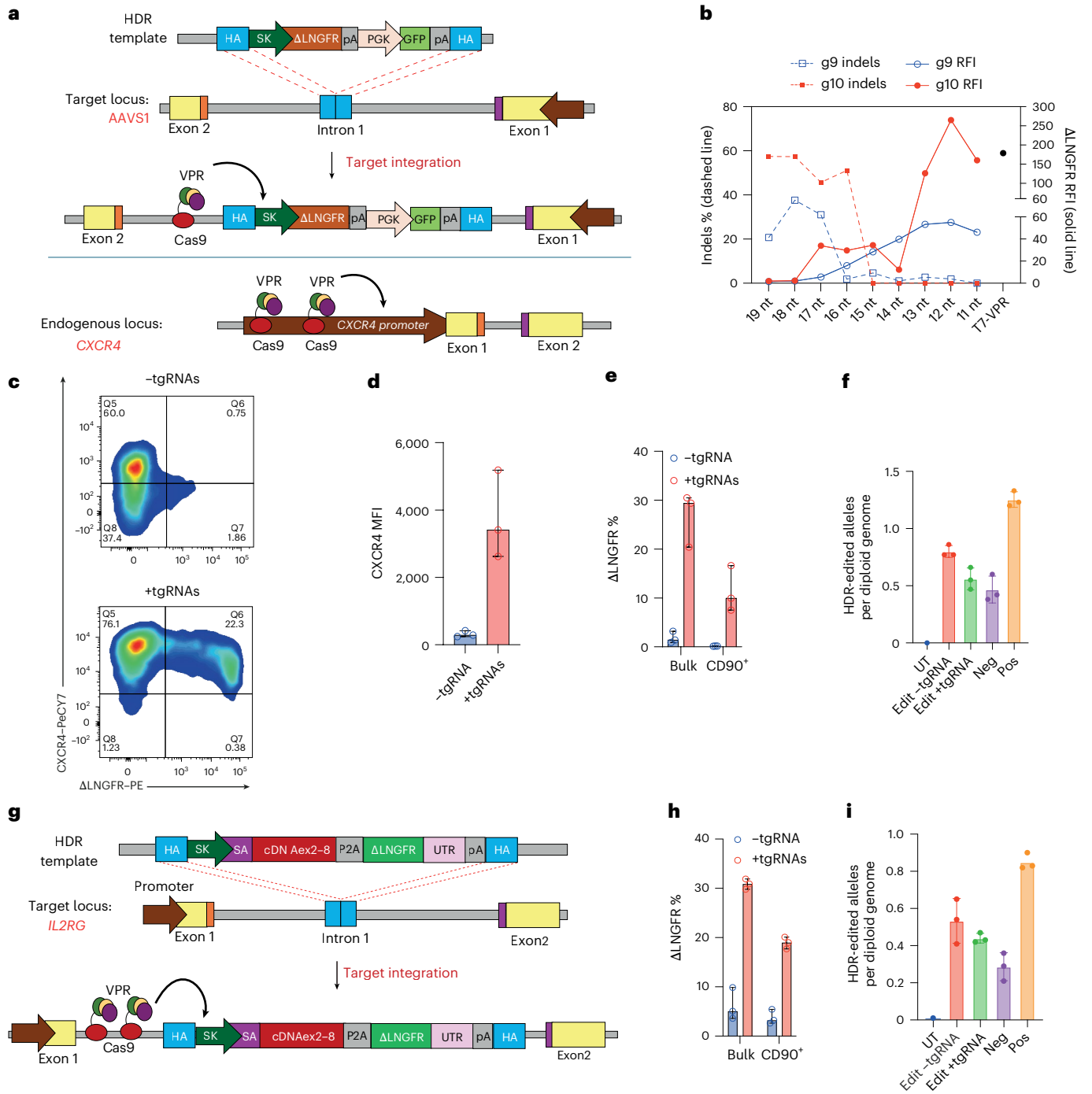


Fig. 5 | SMArT-3 editing and selection using Cas9–VPR polyfunctional editor. a, Schematic of SMART-3 with Cas9–VPR: (1) the DNA DSB is delivered by the Cas9–VPR coupled with a full-length gRNA; (2) the template (same construct of Fig. 3m) is integrated by HDR; (3) transactivation of the selector is achieved by Cas9–VPR coupled to tgRNAs binding outside the homology arms. Simultaneously, tgRNAs allow for transient upregulation of endogenous genes. **b**, Percentage of indels (dashed line) and ΔLNGFR RFI (solid line) in the H1 clone with progressive truncation of g9 and g10 reported in Extended Data Fig. 3c,d ($n = 1$). **c**–**f**, Polyfunctional editing at the *AAVS1* locus. Flow cytometric analysis of the ΔLNGFR selector and endogenous *CXCR4* transactivation analyzed 72 h after editing from the construct integrated at *AAVS1* locus in mPB HSPCs (n = 3 biological replicates): representative FACS plot (**c**), CXCR4 MFI (**d**; median and range) and ΔLNGFR expression (**e**; median and range) in HSPCs in the presence and absence of tgRNAs. **f**, Fraction of HDR-edited cells at *AAVS1* locus in edited mPB-derived HSPCs, quantified by ddPCR (mean and s.d.). **g**–**i**, Polyfuncional editing at the *IL2RG* locus. **g**, Schematic representation of the SMArT-3 construct used for *IL2RG* correction (same construct as Fig. 2a, with the addition of a minimal SK promoter) and location of tgRNAs targeted to the endogenous site upstream to the homology arm for transactivation. **h**, Flow cytometric analysis of the ΔLNGFR transactivation analyzed 72 h after editing from the construct integrated at *IL2RG* locus, reported in **g** (median and range; $n = 3$ biological replicates). **i**, Fraction of HDR-corrected cells at the *IL2RG* locus in edited mPB-derived HSPCs, quantified by ddPCR (mean and s.d.). Each dot is a biological replicate.

to the complexity of the procedure and the requirement for drug selection, this approach cannot ensure coselection of the intended outcome at the GOI, as it only relies on the likelihood of co-occurrence of both editing events.

In our study, the fraction of selector-expressing cells upon transactivation was generally less than that estimated to undergo HDR by ddPCR. This discrepancy can be explained by a combination of factors. As we and others have shown[32,57], targeted ddPCR interrogating the genome–template junction, while representing a good surrogate for HDR, is blind to more complex genomic rearrangements and imprecise targeted integrations, such as NHEJ-mediated trapping of vector and its concatemers. Notably, some of these outcomes may still be functional and responsive to transactivation, depending on vector design and locus expression features. A more comprehensive representation of gene-editing outcome is provided by native DNA long-read sequencing of the target loci, which does not suffer from PCR biases, such as loss of primer-binding sites, sequence variations and preferential amplification of short amplicons. Another mechanism potentially limiting reporter expression shortly after editing is the transiently reduced transcriptional accessibility of the locus targeted by DSB[25,58].

An intrinsic limitation of any selection approach is the expected reduction of clonal diversity in the cell product. This bottleneck effect depends on the efficiency of HDR and the efficiency of selection. Grounded on our previous work on enhancing HDR editing[24,25], we further optimized culture conditions and the cell manufacturing process, leading to higher HDR editing and selection efficiency. These improvements might be because of several factors, whose relative contribution remains to be investigated but likely comprise (1) better preservation of the long-term repopulation potential of HSC that engage in cell replication and become permissive to HDR editing; (2) improved transactivation of the more primitive HSPC subset, which mitigates overrepresentation of the more committed progenitors in the selected and infused cell product; and (3) lower cytotoxicity of the procedure by gentler electroporation, sorting and lower AAV6 dose. Future optimization of viral vector preparations or the use of alternative viral vectors or nonviral templates, possibly delivered by LNPs instead of electroporation, may further increase the process yield and product quality[12,24,59–61]. Ultimately, clinical application requires addressing the concern of shrinkage in clonal diversity and effective cell yield in the specific therapeutic context, which can be addressed preclinically by incorporating degenerated molecular barcodes in the DNA template to track HDR-edited clones and quantify clonal diversity[5,62,63]. Prospectively, intensive HSPC mobilization[64] to increase the starting number of HSPCs may further mitigate this concern, resulting in the infusion of a sufficiently high number of corrected HSPCs. Moreover, expansion of HSPCs, if robustly achieved[65], may eventually serve the same purpose, also allowing delayed ArT delivery for greater or more convenient transactivation[40], as shown here for T cells.

AAV6 cargo size can be a theoretical constrain for clinical application of SMArTs. The ΔLNGFR selector cDNA is 843 bp long, leaving around 3.9 kb for the remaining HDR donor template features (transgene, homology arms, etc.). Considering that the median length of human protein coding transcripts is 2.8 kb (ref. 66), cargo space appears sufficient to accommodate a substantial number of therapeutic corrective cDNAs, beyond applications already illustrated in this manuscript. Nevertheless, SMArT strategies could also be developed using integrase defective LVs in lieu of AAV6 for HDR donor template delivery, should a larger cargo size be required for specific applications[24]. SMArTs might also be applied beyond HDR-mediated editing to other technologies aiming to targeted integration of gene-sized DNA, such as PASSIGE[67], PASTE[68] or CAST[69]. However, these strategies currently require serial lengthy and even more complex and less efficient editing procedures than those used in our work, constraining feasible application for translational purposes in primary human cells with limited life in culture.

SMArT strategies may be coupled with novel biological conditioning regimens[35,70,71], which have limited toxicity and are not myeloablative. Accordingly, infusion of a limited number of pure edited HSPCs may suffice to achieve clinical benefit either in the context of inborn errors of immunity or in cancer gene therapy[72]. We expect that the possibility to select cells bearing the desired editing outcome with a reduced genotoxic burden will broaden translational applications of gene editing by improving their risk–benefit profile.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-026-03142-z.

References

1. Porteus, M. H. A new class of medicines through DNA editing. *N. Engl. J. Med.* **380**, 947–959 (2019).
2. Wang, J. et al. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**, 1256–1263 (2015).
3. Dever, D. P. et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384–389 (2016).
4. Schiroli, G. et al. Preclinical modeling highlights the therapeutic potential of hematopoietic stem cell gene editing for correction of SCID-X1. *Sci. Transl. Med.* **9**, eaan0820 (2017).
5. Ferrari, S. et al. Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nat. Biotechnol.* **38**, 1298–1308 (2020).
6. Rai, R. et al. Targeted gene correction of human hematopoietic stem cells for the treatment of Wiskott–Aldrich syndrome. *Nat. Commun.* **11**, 4034 (2020).
7. Eyquem, J. et al. Targeting a CAR to the *TRAC* locus with CRISPR/Cas9 enhances tumour rejection. *Nature* **543**, 113–117 (2017).
8. Hubbard, N. et al. Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. *Blood* **127**, 2513–2522 (2016).
9. Vavassori, V. et al. Modeling, optimization, and comparable efficacy of T cell and hematopoietic stem cell gene editing for treating hyper-IgM syndrome. *EMBO Mol. Med.* **13**, e13545 (2021).
10. Pavel-Dinu, M. et al. Gene correction for SCID-X1 in long-term hematopoietic stem cells. *Nat. Commun.* **10**, 1634 (2019).
11. Castiello, M. C. et al. Exonic knockout and knockin gene editing in hematopoietic stem and progenitor cells rescues RAG1 immunodeficiency. *Sci. Transl. Med.* **16**, eadh8162 (2024).
12. Roth, T. L. et al. Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405–409 (2018).
13. Lombardo, A. et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869 (2011).
14. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
15. Genovese, P. et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014).
16. Xu, L. et al. CRISPR-edited stem cells in a patient with HIV and acute lymphocytic Leukemia. *N. Engl. J. Med.* **381**, 1240–1247 (2019).
17. Stadtmauer, E. A. et al. CRISPR-engineered T cells in patients with refractory cancer. *Science* https://doi.org/10.1126/science.aba7365 (2020).
18. Wu, Y. et al. Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* **25**, 776–783 (2019).

19. Locatelli, F. et al. Exagamglogene autotemcel for transfusion-dependent β-thalassemia. *N. Engl. J. Med.* **390**, 1663–1676 (2024).
20. Frangoul, H. et al. Exagamglogene autotemcel for severe sickle cell disease. *N. Engl. J. Med.* **390**, 1649–1662 (2024).
21. Gardner, C. L. et al. Gene editing rescues in vitro T cell development of RAG2-deficient induced pluripotent stem cells in an artificial thymic organoid system. *J. Clin. Immunol.* **41**, 852–862 (2021).
22. De Ravin, S. S. Targeted gene addition in human $CD34^+$ hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.* **34**, 424–429 (2016).
23. Cromer, M. K. et al. Gene replacement of α-globin with β-globin restores hemoglobin balance in β-thalassemia-derived hematopoietic stem and progenitor cells. *Nat. Med.* **27**, 677–687 (2021).
24. Ferrari, S. et al. Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. *Cell Stem Cell* **29**, 1428–1444 (2022).
25. Schiroli, G. et al. Precise gene editing preserves hematopoietic stem cell function following transient p53-mediated DNA damage response. *Cell Stem Cell* **24**, 551–565 (2019).
26. Ferrari, S. et al. Genetic engineering meets hematopoietic stem cell biology for next-generation gene therapy. *Cell Stem Cell* **30**, 549–570 (2023).
27. Leibowitz, M. L. et al. Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing. *Nat. Genet.* **53**, 895–905 (2021).
28. Cancellieri, S. et al. Human genetic diversity alters off-target outcomes of therapeutic gene editing. *Nat. Genet.* **55**, 34–43 (2023).
29. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
30. Adikusuma, F. et al. Large deletions induced by Cas9 cleavage. *Nature* **560**, E8–E9 (2018).
31. Turchiano, G. et al. Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. *Cell Stem Cell* **28**, 1136–1147 (2021).
32. Canarutto, D. et al. Unbiased assessment of genome integrity and purging of adverse outcomes at the target locus upon editing of $CD4^+$ T-cells for the treatment of hyper IgM1. *EMBO J.* **42**, e114188 (2023).
33. Mamcarz, E. et al. Lentiviral gene therapy combined with low-dose busulfan in infants with SCID-X1. *N. Engl. J. Med.* **380**, 1525–1534 (2019).
34. Ginn, S. L. et al. Lymphomagenesis in SCID-X1 mice following lentivirus-mediated phenotype correction independent of insertional mutagenesis and γc overexpression. *Mol. Ther.* **18**, 965–976 (2010).
35. Omer-Javed, A. et al. Mobilization-based chemotherapy-free engraftment of gene-edited human hematopoietic stem cells. *Cell* **185**, 2248–2264 (2022).
36. Yazdani, R. et al. The hyper IgM syndromes: epidemiology, pathogenesis, clinical manifestations, diagnosis and management. *Clin. Immunol.* **198**, 19–30 (2019).
37. Brown, M. P. et al. Thymic lymphoproliferative disease after successful correction of CD40 ligand deficiency by gene transfer in mice. *Nat. Med.* **4**, 1253–1260 (1998).
38. Sacco, M. G. et al. Lymphoid abnormalities in CD40 ligand transgenic mice suggest the need for tight regulation in gene therapy approaches to hyper immunoglobulin M (IgM) syndrome. *Cancer Gene Ther.* **7**, 1299–1306 (2000).
39. Asperti, C. et al. Scalable GMP-compliant gene correction of $CD4^+$ T cells with IDLV template functionally validated in vitro and in vivo. *Mol. Ther. Methods Clin. Dev.* **30**, 546–557 (2023).
40. Vavassori, V. et al. Lipid nanoparticles allow efficient and harmless ex vivo gene editing of human hematopoietic cells. *Blood* **142**, 812–826 (2023).
41. Woods, N.-B., Bottero, V., Schmidt, M., von Kalle, C. & Verma, I. M. Therapeutic gene causing lymphoma. *Nature* **440**, 1123–1123 (2006).
42. Pike-Overzet, K. et al. Is *IL2RG* oncogenic in T-cell development? *Nature* **443**, E5–E5 (2006).
43. Chavez, A. et al. Comparison of Cas9 activators in multiple species. *Nat. Methods* **13**, 563–567 (2016).
44. Hockemeyer, D. et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* **27**, 851–857 (2009).
45. Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol.* **10**, 81 (2010).
46. Zonari, E. et al. Ex vivo expansion of hematopoietic stem and progenitor cells from human mobilized peripheral blood for gene therapy applications. Preprint at *bioRxiv* https://doi.org/10.64898/2026.04.08.716064 (2026).
47. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR–Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
48. Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
49. Kiani, S. et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* **12**, 1051–1054 (2015).
50. Cipria, D. et al. Simultaneous orthogonal cell engineering by a single CRISPR–Cas9 polyfunctional editor. *Nat. Commun.* https://doi.org/10.1038/s41467-026-72846-2 (2026).
51. Canarutto, D., Omer Javed, A., Pedrazzani, G., Ferrari, S. & Naldini, L. Mobilization-based engraftment of haematopoietic stem cells: a new perspective for chemotherapy-free gene therapy and transplantation. *Br. Med. Bull.* **147**, 108–120 (2023).
52. Mikkelsen, N. S., Hernandez, S. S., Jensen, T. I., Schneller, J. L. & Bak, R. O. Enrichment of transgene integrations by transient CRISPR activation of a silent reporter gene. *Mol. Ther. Methods Clin. Dev.* **29**, 1–16 (2023).
53. Pölönen, P. et al. The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. *Nature* **632**, 1082–1091 (2024).
54. Hussain, M. et al. A novel dominant negative variant in *IL2RG* gene that failed to be corrected by lentiviral gene therapy. *J. Hum. Immun.* https://doi.org/10.70962/CIS2025abstract.177 (2025).
55. Allen, A. G. et al. A highly efficient transgene knock-in technology in clinically relevant cell types. *Nat. Biotechnol.* **42**, 458–469 (2024).
56. Agudelo, D. et al. Marker-free coselection for CRISPR-driven genome editing in human cells. *Nat. Methods* **14**, 615–620 (2017).
57. Suchy, F. P. et al. Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. *Nat. Biotechnol.* **43**, 204–213 (2024).
58. Iannelli, F. et al. A damaged genome's transcriptional landscape through multilayered expression profiling around in situ-mapped DNA double-strand breaks. *Nat. Commun.* **8**, 15656 (2017).
59. Nguyen, D. N. et al. Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat. Biotechnol.* **38**, 44–49 (2020).
60. Lin-Shiao, E. et al. CRISPR–Cas9-mediated nuclear transport and genomic integration of nanostructured genes in human primary cells. *Nucleic Acids Res.* **50**, 1256–1268 (2022).
61. Shy, B. R. et al. High-yield genome engineering in primary cells using a hybrid ssDNA repair template and small-molecule cocktails. *Nat. Biotechnol.* **41**, 521–531 (2023).

62. Sharma, R. et al. The TRACE-Seq method tracks recombination alleles and identifies clonal reconstitution dynamics of gene targeted human hematopoietic stem cells. *Nat. Commun.* **12**, 472 (2021).
63. Ferrari, S. et al. BAR-Seq clonal tracking of gene-edited cells. *Nat. Protoc.* **16**, 2991–3025 (2021).
64. Canarutto, D. et al. Peripheral blood stem and progenitor cell collection in pediatric candidates for ex vivo gene therapy: a 10-year series. *Mol. Ther. Methods Clin. Dev.* **22**, 76–83 (2021).
65. Sakurai, M. et al. Chemically defined cytokine-free expansion of human haematopoietic stem cells. *Nature* **615**, 127–133 (2023).
66. Piovesan, A., Caracausi, M., Antonaros, F., Pelleri, M. C. & Vitale, L. GeneBase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics. *Database (Oxford).* **2016**, 331–345 (2016).
67. Pandey, S. et al. Efficient site-specific integration of large genes in mammalian cells via continuously evolved recombinases and prime editing. *Nat. Biomed. Eng.* **9**, 22–39 (2024).
68. Yarnall, M. T. N. et al. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat. Biotechnol.* **41**, 500–512 (2023).
69. Witte, I. P. et al. Programmable gene insertion in human cells with a laboratory-evolved CRISPR-associated transposase. *Science* **388**, eadt5199 (2025).
70. Casirati, G. et al. Epitope editing enables targeted immunotherapy of acute myeloid leukaemia. *Nature* **621**, 404–414 (2023).
71. Garaudé, S. et al. Selective haematological cancer eradication with preserved haematopoiesis. *Nature* **630**, 728–735 (2024).
72. Birocchi, F. et al. Targeted inducible delivery of immunoactivating cytokines reprograms glioblastoma microenvironment and inhibits growth in mouse models. *Sci. Transl. Med.* **14**, eabl4106 (2022).

[1]San Raffaele Telethon Institute for Gene Therapy, IRCCS San Raffaele Scientific Institute, Milan, Italy. [2]Vita-Salute San Raffaele University, Milan, Italy. [3]Milan Unit, Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Milan, Italy. [4]Present address: Princess Margaret Cancer Centre, Toronto, Ontario, Canada. [5]Present address: Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [6]Present address: Fondazione Telethon ETS, Rome, Italy. [7]Present address: Tessera Therapeutics, Sommerville, MA, USA. [8]Present address: Icahn Genomics Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [9]Present address: RNA Nanoparticle Core, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [10]Present address: Department of Immunology and Immunotherapy, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [11]Present address: Ludwig Institute for Cancer Research and Department of Oncology, University of Lausanne (UNIL), Lausanne, Switzerland. [12]Present address: Lausanne University Hospital (CHUV), Lausanne, Switzerland. [13]Present address: Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA. [14]Present address: Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [15]Present address: Harvard Medical School, Boston, MA, USA. [16]These authors contributed equally: Daniele Canarutto, Martina Fiumara. [17]These authors jointly supervised this work: Samuele Ferrari, Luigi Naldini. ✉e-mail: naldini.luigi@hsr.it

Methods

Nucleases

Previously reported[4,9,15,25,73] and newly designed gRNA spacers are listed in Supplementary Table 1. RNP complexes were assembled by incubating at 1:1.5 (*IL2RG* and *AAVS1*) or 1:2 (*CD40LG* and *B2M*) molar ratio with *Sp*Cas9 protein (Aldevron) with either preannealed synthetic Alt-R crRNA:tracrRNA or sgRNA (Synthego) for 10 min at 25 °C. The H1 clone was generated using zinc finger nucleases targeting *AAVS1* (ref. 13).

Plasmids and vectors

AAV transfer vector constructs for *IL2RG*, *CD40L*, *AAVS1* and *RAG1* editing are described in Figs. 1–3, Extended Data Fig. 5a and Supplementary Fig. 1. All cargo cassettes are flanked by AAV serotype 2 inverted terminal repeats (ITRs)[74].

For screening purposes, full-length gRNAs and tgRNAs targeting *AAVS1* and *CXCR4*, to be coupled with dCas9–TAD or Cas9–TAD were cloned in a pUC.U6 plasmid with BsaI (New England Biolabs (NEB)). gRNA spacers are listed in Supplementary Table 1. The Sp.dCas9-VPR plasmid was bought from Addgene (63789). The Sp.dCas9-VP160 plasmid was cloned by swapping VPR with VP160 DNA sequence. Cas9–VPR was derived from the former by restoring the Cas9 sequence. TALE DBDs were cloned by Golden Gate strategy (Addgene) in a pUC plasmid and fused in frame at the C terminus with either VP160–TAD or VPR–TAD. ArT sequences are listed in Supplementary Table 2. Vector maps were designed with SnapGene (GSL Biotech, https://www.snapgene.com/).

mRNA in vitro transcription (IVT)

GSE56, GSE56/E4orf6/7, TALE–TA, dCas9–VPR and Cas9–VPR were subcloned in a pVax plasmid. Specifically, ArT and GSE56 constructs were cloned into optimized 'pVax' plasmids under the control of the following 5′ aptamer sequence: CapAG–eIF4G aptamer (GACTCACTATTTGTTTTCGCGCCCAGTTGCAAAAAGTGTCG)–Kozak sequence (CCACC)–start codon (ATG). Downstream, the codon-optimized sequence follows a woodchuck hepatitis virus post-transcriptional regulatory element and a 64-bp poly(A) sequence. Methods for mRNA IVT, quantification and quality assessment were previously described[5,25,35]. Briefly, pVax plasmids were linearized with SpeI (NEB) and purified by phenol–chloroform extraction. mRNA IVT was performed with THE 5× MEGAscript T7 kit (Thermo Fisher) and capped with 4.5 mM anti-reverse cap analog 3′-*O*-methyl-mG(5′) ppp(5′)G (New England Biolabs) mixed in a 1:5 ratio with dGTP nucleotides or 5 mM CleanCapAG (Trilink). mRNAs were then purified using an RNeasy Plus mini kit (Qiagen) followed by high-performance liquid chromatography (ADS BIOTEC WAVE System) and Amicon Ultra-15 (30 kDa) tube (Millipore) concentration. mRNAs were aliquoted and stored at −80 °C.

Cell lines and primary cell culture

K-562 cells and the H1 clone were cultured in Iscove's modified Dulbecco's medium (Corning) supplemented with 10% heat-inactivated FBS (Euroclone), 100 IU per ml penicillin, 100 µg ml^{-1} streptomycin and 2% glutamine.

For T cells, buffy coats were obtained as anonymized residues of blood donations, used upon signature of specific institutional informed consent for blood product donation by healthy blood donors in accordance with OSR protocols. CD4 T cells isolation was performed using the CD4$^+$ T cell isolation kit (Miltenyi Biotec) and cultured in X-Vivo 15 (Lonza) supplemented with 0.5% HSA (Baxalta), 1% penicillin–streptomycin (Lonza), 100 IU per ml IL-7 and 200 IU per ml IL-15 (both from Miltenyi Biotec).

CB CD34$^+$ HSPCs were purchased from Lonza according to the TIGET-HPCT protocol approved by OSR Ethical Committee. Cells were seeded at 5 × 10^5 cells per ml in serum-free StemSpan medium (StemCell Technologies) supplemented with 100 IU per ml penicillin, 100 µg ml^{-1} streptomycin, 2% glutamine, 100 ng ml^{-1} hSCF (PeproTech), 100 ng ml^{-1} hFlt3-L (PeproTech), 20 ng ml^{-1} hTPO (PeproTech), 20 ng ml^{-1} hIL-6 (PeproTech), 10 µM PGE2 (at the beginning of the culture, Cayman), 1 µM SR1 (Biovision) and 50 nM UM171 (StemCell Technologies).

G-CSF or G-CSF + Plerixafor mPB CD34$^+$ HSPCs were purified with the CliniMACS CD34 reagent system (Miltenyi Biotec) from Mobilized Leukopak (AllCells) according to the TIGET-HPCT protocol, following the manufacturer's instructions. HSPCs were seeded at 5 × 10^5 cells per ml in serum-free StemSpan medium (StemCell Technologies) supplemented with 100 IU per ml penicillin, 100 µg ml^{-1} streptomycin, 2% glutamine, 300 ng ml^{-1} hSCF (PeproTech), 300 ng ml^{-1} hFlt3-L (PeproTech), 100 ng ml^{-1} hTPO (PeproTech), 1 µM SR1 (Biovision) and 35 nM UM171 (StemCell Technologies) (referred to as 'standard medium'). Alternatively, mPB HSPCs were seeded in GMP stem cell growth medium (CellGenix, with or without phenol red) supplemented with 100 IU per ml penicillin, 100 µg ml^{-1} streptomycin, 100 ng ml^{-1} hSCG, 100 ng ml^{-1} hFLT3, 50 ng ml^{-1} hIL-3, 50 ng ml^{-1} hIL-6, 50 ng ml^{-1} hTPO (all from Sartorius CellGenix), 1 µM SR1 and 35 nM UM171 (referred to as 'optimized medium').

All cells were cultured in a 5% CO_2 humidified atmosphere at 37 °C.

Gene editing of human HSPCs, T cells and K-562 cell line

Gene-editing protocols for human HSPCs and T cells were previously described[9,39,63] and are shown in Figs. 1c,n and 4a,j and Extended Data Fig. 2d.

HSPCs. After 3 days of stimulation, cells were washed with ten volumes of Dulbecco's PBS without calcium and magnesium (DPBS) and electroporated using the P3 primary cell 4D-Nucleofector X kit and the Nucleofector 4D device (program EO-100, Lonza; 1 × 10^5–9 × 10^5 HSPCs per well; Figs. 1b–m, 2b–g, 3a–k,n–p and 4b–i), or Opti-MEM (Gibco) with MaxCyte GTx (program HSC-3, MaxCyte, 1 × 10^6–20 × 10^6 HSPCs per well, OC-25×3 or OC100x2 cuvettes; Figs. 1o–r, 2h–t, 3l and 4k–r). Cells were electroporated with Cas9 RNPs at a final concentration of 1.25–2.5 µM. In some experiments, 0.1 nmol of Alt-R Cas9 electroporation enhancer (Integrated DNA Technologies) was added to the mix. AAV6 transduction was performed at a dose of 2 × 10^4 vector genome copies (vg) per cell (standard medium) or 5 × 10^3 vg per cell (optimized medium; Figs. 2h–t, 3n–p and 4j–r) 15 min after electroporation, unless otherwise specified. When indicated, mRNAs were added to the electroporation mix at the following final concentrations: 150 µg ml^{-1} tTA, 150 µg ml^{-1} T3-VPR, 250 µg ml^{-1} T7-VPR, 150 µg ml^{-1} GSE56 and 250 µg ml^{-1} GSE56/E4orf6/7.

For SMArT-3 Cas9 RNP and dCas9–VPR in Extended Data Fig. 5a,b, cells were electroporated with *RAG1* RNP, dCas9–VPR mRNA, 72.5 pM each of *RAG1* 15-nt truncated g2 and g3 or g4 and g5 and GSE56 mRNA, using the Nucleofector 4D device as described above. For the SMArT-3 Cas9–VPR experiments at the *AAVS1* locus in OC-25×3 (Fig. 5c–f), cells were electroporated with 38 pM *AAVS1* gRNA, Cas9–VPR mRNA, 114 pM *AAVS1* truncated g10 (12 nt), 57 pM each of the *CXCR4* truncated g116 (15 nt) and g160 (14 nt) and GSE56 mRNA. For *IL2RG* (Fig. 5g,h), cells were electroporated with Cas9–VPR mRNA, 38 pM *IL2RG* gRNA and 57 pM each of the *IL2RG* truncated g2 (15 nt), *IL2RG* truncated g5 (15 nt) and GSE56 mRNA. For Extended Data Fig. 5f, reagent concentrations were 114 pM for individual tgRNAs or 57 pM each when tgRNAs were combined. Cells were electroporated with the Nucleofector 4D device as described above. AAV6 transduction was performed at a dose of 2 × 10^3–2.5 × 10^3 vg per cell (optimized medium) 15 min after electroporation. When indicated, mRNAs were added to the electroporation mix at final concentrations of 200 µg ml^{-1} dCas9–VPR, 200 µg ml^{-1} Cas9–VPR and 120 µg ml^{-1} GSE56.

T cells. First, 3 days after stimulation with T cell TransAct 60 µl ml^{-1} (Miltenyi Biotec), cells were edited at the *IL2RG* locus using the P3 primary cell 4D-Nucleofector X kit and Nucleofector 4D device with RNPs

at a final concentration of 2.5 µM. AAV6 transduction was performed at a dose of 5×10^4 vg per cell 15 min after electroporation.

For Fig. 1n–r, CD4 T cells were processed as above until day 3, stimulated with Transact 10 µl ml^{-1} and edited as reported previously[39]. Then, 7 days after editing, 0.5 µg of tTA mRNA was delivered to 3×10^5 T cells by LNPs (T cell kit; 1000701, Cytiva). For LNP preparation, a single reaction of mRNA working solution was prepared by mixing 11 µg of tTA mRNA with 3.52 µl of 10× formulation buffer and filled up to 35.2 µl with sterile water. Quantification of encapsulated mRNA was performed using the Quant-it RiboGreen RNA assay kit (Thermo Fisher Scientific), following the manufacturer's instructions. Cells were washed 24 h after LNP delivery. For evaluation of *CD40LG* expression, T cells were stimulated with phorbol 12-myristate 13-acetate (PMA; 10 ng ml^{-1}) and ionomycin (0.5 µg ml^{-1}) as previously described[39].

Cell lines. K-562 cells were purchased from the American Type Culture Collection. For Fig. 5b, the H1 K-562 cell clone was electroporated with the 0.5 µg of plasmid expressing gRNA along with 2 µg of Cas9–VPR mRNA using the P3 primary cell 4D-Nucleofector X kit and Nucleofector 4D device with the nucleofection code FF120. For all other experiments, the K-562 cell line or clone H1 was nucleofected with 1.5 µg of the TALE–ArT-expressing plasmid or mRNA or 1.5 µg and 0.5 µg of the plasmids expressing dCas9–ArT and gRNA, respectively.

Flow cytometry

Immunophenotypic analyses were performed on a fluorescence-activated cell sorting (FACS) Canto II (BD Pharmingen). A total of 0.5×10^5–2×10^5 cells from culture or mouse samples were stained for 15 min at 4 °C with antibodies listed in Supplementary Table 3 in a final volume of 100 µl and then washed with DPBS + 2% heat-inactivated FBS. Single stained samples or compensation beads were used as controls. The live/dead fixable dead cell stain kit (Thermo Fisher) or 7-aminoactinomycin D (Sigma Aldrich or Biolegend) was included during sample preparation according to the manufacturers' instructions.

Cell sorting was performed on a BD FACSAria Fusion (BD Biosciences) equipped with four lasers: blue (488 nm), yellow/green (561 nm), red (640 nm) and violet (405 nm). Cells were sorted with an 85-mm nozzle. Sheath fluid pressure was set at 45 psi. A highly pure sorting modality (four-way purity sorting) was chosen. Sorted cells were collected in 1.5-ml Eppendorf tubes containing 500 µl of DPBS or HSPC medium. Cell sorting in Figs. 1n, 2h–t, 4j–r and 5 and corresponding Extended Data figures were performed on a MACSQuant Tyto Cell Sorter (Miltenyi Biotec) using the MACSQuantify Tyto software and equipped with blue (488 nm), red (640 nm) and violet (405 nm) lasers. Gating strategies for flow cytometry analyses are provided in Supplementary Fig. 2. Data were analyzed with FCS Express 7 Flow or FlowJo 10.

Clonogenic assay

A colony-forming-unit (CFU) cell assay was performed 1 day after editing by plating 800 mPB-derived HSPCs in methylcellulose-based medium (MethoCult H4434, StemCell Technologies) supplemented with 100 IU per ml penicillin and 100 µg ml^{-1} streptomycin. Then, 2 weeks after plating, CFUs were counted and manually picked.

Quantification of NHEJ, HDR editing efficiency and large deletions

gDNA was isolated with QIAamp DNA micro kit (Qiagen), according to the manufacturers' instructions. NHEJ frequency was assessed by T7 surveyor assays. The target region was PCR-amplified using the primers listed in Supplementary Table 3. The PCR products were then denaturated and reannealed to allow the formation of heteroduplex DNA, which was subsequently cleaved by T7 endonuclease I (NEB), as previously described[63]. The digested DNA fragments were quantified using the 4200 TapeStation system (Agilent). NHEJ frequency was calculated as follows:

$$\frac{\sum(\text{edited band intensities})}{\sum(\text{edited band intensities}) + \text{parent band intensity}} \times 100$$

For HDR ddPCR analysis, 5–50 ng of gDNA was analyzed using the QX200 or QX600 ddPCR System (Bio-Rad) according to the manufacturer's instructions. HDR ddPCR primers and probes were designed to span the vector-genome junction (Figs. 2a and 3a)[5,9,63]. Human *TTC5* (Bio-Rad) or *GAPDH* (Bio-Rad) was used as reference. Large deletions were assessed on HSPC-derived colonies as previously described[24]. Briefly, DNA from colonies was extracted with the QuickExtract DNA extraction solution (Lucigen) according to manufacturer instructions; 5 µl of the DNA was used for ddPCR assays. The large deletion assays amplify a sequence nearby the gRNA target site but outside homology arms. Single colonies showing amplification of the reference genomic amplicon but not of the large deletion assay were considered as bearing a large deletion. Colonies with an insufficient number of positive droplets were excluded. The deletion frequency was calculated as number of colonies bearing a deletion divided by the total number of colonies analyzed. Primers and probes are listed in Supplementary Table 4.

Long-read sequencing of unamplified genomic DNA

For long-read sequencing, DNA from CFUs was extracted using the Monarch HMW DNA extraction kit for cells and blood (NEB). No additional fragmentation was performed. First, 3–5 µg of DNA was processed with the ligation sequencing kit v14 (Oxford Nanopore) and acquired on a P2 Solo R10 flow cell (Oxford Nanopore) using adaptive sampling enrichment for the target locus. Reads were basecalled with dorado dna_r10.4.1_e8.2_400bps_sup@v5.2.0 for maximum quality, aligned to hg38 with dorado version 0.7 (later updated to 1.1.1) and classified manually. Because of limited software accuracy at the single-base level, WT reads and those containing indels or deletions < 50 bp were grouped together. Deletions > 50 bp were classified as deletions. Non-precise HDR events were classified as either presumably functional and imprecise (for example, donor trapping) or non-functional (for example, trapping of genomic sequences). Representative alignments are reported in Supplementary Fig. 3. Locus coverage was 54.5–104.5 for *AAVS1* and 40–65 for *IL2RG* (for coverage estimation, reads not covering the entire locus but only a vector–genome junction were counted half). The frequency of false-positive translocations (that is, technical artifacts) was established by performing long-read sequencing on a UT sample.

Long-read sequencing of PCR-amplified DNA

For amplicon sequencing of the target locus, we designed primers binding outside the homology arms (Supplementary Table 5). First, 150–400 ng of DNA was amplified with the LongAmp Taq 2× master mix kit (NEB; thermal protocol in Supplementary Table 5). Subsequently, 2–3 µg of amplified DNA was barcoded and sequenced with the SQK-NBD114.24 kit (Oxford Nanopore) according to the manufacturer's instructions. Libraries were sequenced on a P2 Solo R10 flow cell (Oxford Nanopore) and basecalled with model dna_r10.4.1_e8.2_400bps_sup@v5.2.0.

FUSiLLi

The FUSiLLi pipeline for data analysis was built upon the RAAVioli pipeline[63,75,76]. Briefly, the approach of the pipeline is to filter informative reads, reconstruct vector–genome junctions and feature sequences and compare all-versus-all reads for unsupervised clustering. Specific details of the pipelines will be reported in a follow-up methodological paper. The main steps of the pipeline are as follows: (1) filtering of reads mapping onto the target locus, beyond the homology arms. whereby reads are mapped with minimap2 onto a custom

human genome, where the sequence of the homology arms has been removed and only those within a set interval from the DNA DSB are retained; (2) quality filtering; (3) Alignment of reads with minimap2 to a custom genome consisting of the donor vector genome and the human genome; (4) downsampling across samples to the same number of reads; (5) processing of the alignments with a custom script to identify integration loci (homology arms, ITRs, promoters, etc.) and vector rearrangements using the CIGAR (concise idiosyncratic gapped alignment report) string, akin to the RAAVioli pipeline; (6) recoding of each read as a synthetic feature chain; and (7) all-versus-all comparison and unsupervised clustering of synthetic feature chains, accounting for potential Nanopore read truncations. Editing events are reported as the abundance of feature chains.

Gene expression analyses

Total RNA was extracted using RNeasy Plus micro kit or RNeasy micro kit (Qiagen), according to the manufacturer's instructions and DNAse treatment was performed using RNase-free DNAse Set (Qiagen). cDNA was synthesized with SuperScript VILO IV cDNA synthesis kit (Thermo Fisher) with EzDNAse treatment. cDNA was then used for qPCR in a Viia7 real-time PCR thermal cycler using TaqMan gene expression assays (Applied Biosystems) mapping to genes listed in Supplementary Table 4. Data were analyzed with QuantStudio real-time PCR software version 1.1 (Applied Biosystem). Alternatively, 2 ng of cDNA was used for gene expression analysis by ddPCR. ddPCR data were analyzed with QX Manager standard edition 2.0.0. Expression of each target gene was normalized to *HPRT* and, when indicated, represented as fold changes relative to the UT cells. Gene expression assays are listed in Supplementary Table 4.

ATOs

For preparation of ATO medium, RPMI-1640 medium was supplemented with 4% B-27 supplement (Thermo Fisher), 1% GlutaMAX (Thermo Fisher), 0.2% Primocine (InvivoGen), 30 µM ascorbic acid (Sigma Aldrich), 10 ng ml^{-1} rhIL-7 and 10 ng ml^{-1} rhFLT3 (both from Peprotech). For ATO seeding, a 0.4-µm Millicell transwell insert (EMD Millipore) was placed in a six-well plate containing 1.1 ml of ATO medium per well. On day 4 after editing, 1.0×10^5–1.2×10^5 HSPCs were mixed with 2×10^5 MS5-hDLL4, the cell aggregates were seeded onto the transwell insert and the number of ATOs per insert was set to 3. ATOs were cultured in 5% CO_2 humidified atmosphere at 37 °C and the medium was changed twice per week by aspirating the medium from around the insert and replacing it with 1 ml of fresh ATO medium. ATOs were isolated, disaggregated and passed through a 50-µm nylon strainer. A previously described staining panel was used[11], except for anti-CD8b–PE that was replaced by anti-NGFR–PE (Supplementary Table 3). The TCR repertoire was analyzed by the IOTest Beta mark kit (Beckman Coulter) following the manufacturer's instructions. Flow cytometry and analyses were performed as described above.

Mice

Experiments and procedures involving animals were performed with the approval of the Animal Care and Use Committee of the San Raffaele Hospital (approval nos. 818, 1206 and 1358) and authorized by the Italian Ministry of Health and local authorities. C57Bl/6 Ly45.1 mice were purchased from Charles River Laboratory, *Cd40lg*$^{-/-}$ (B6.129S2-*Cd40lg*^{tm1Imx}/J), NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ (NSG) and NOD,B6.*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ *Kit*$^{W41/W41}$ (NBSGW) female mice (The Jackson Laboratory) were held in specific pathogen-free conditions.

For *Cd40lg*$^{-/-}$ experiments, donor mice between 6 and 10 weeks of age were killed and BM cells were retrieved from femurs and tibias. HSPCs were purified by Lin$^-$ selection using the mouse lineage cell depletion kit (Miltenyi Biotec) according to the manufacturer's instructions. Cells were then cultured in serum-free StemSpan medium (StemCell Technologies) containing penicillin, streptomycin, glutamine and a combination of mouse cytokines (20 ng ml^{-1} IL-3, 100 ng ml^{-1} SCF, 100 ng ml^{-1} Flt-3L, 50 ng ml^{-1} TPO; PeproTech), at a concentration of 10^6 cells per ml. C57BL/6-Ly5.1 and *Cd40lg*$^{-/-}$ (CD45.2) Lin$^-$ cells were cultured for 16 h in the medium described above, mixed at the indicated ratios and transplanted at a total dose of 10^6 cells per mouse unless otherwise specified into 8-week-old lethally irradiated *Cd40lg*$^{-/-}$ CD45.2 mice. Blood collection and analysis was performed to monitor donor cell engraftment. Mice were immunized intraperitoneally with 100 µg of TNP-KLH (Lgc Biosearch Technologies) in Imject Alum Adjuvant (1:2 ratio) (Thermo Fisher Scientific). Serum was collected at the indicated time points after immunization. Immunization was boosted as described above and the serum was collected after rechallenge. The concentration of antigen-specific IgGs in mouse sera was determined by ELISA, as previously described[9].

For transplantation of mPB CD34$^+$ HSPCs, the outgrowths of 1×10^6–2×10^6 culture-initiating HSPCs were injected intravenously into sublethally irradiated NSG mice (150–180 cGy) or nonirradiated NBSGW mice. Sample size for each experiment was determined by the total number of available treated cells. Mice were randomly distributed to each experimental group. Human CD45$^+$ cell engraftment and the presence of edited cells were monitored by serial blood collection. At the endpoint, the BM and spleen (SPL) were collected for analyses.

Quantifications and statistical analyses

The number of biologically independent samples, animals, colonies or experiments is indicated by dots or *n* in figures. For some experiments, different HSPC donors were pooled to account for donor-related variability and reach the number of cells needed for the analyses. Data were summarized as the median and 95% CI, median and range or mean ± s.e.m. depending on data distribution. Inferential two-tailed techniques were applied in the presence of adequate sample sizes ($n \geq 5$); otherwise, only descriptive statistics are reported. Analyses were performed using GraphPad Prism version 10.6.1. Statistical tests are indicated in the figure legends.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All relevant data are included in the manuscript. Long-read sequencing data are available from the European Nucleotide Archive under accession number PRJEB111192. The reagents described in this manuscript are available under a material transfer agreement with IRCCS Ospedale San Raffaele and Fondazione Telethon; requests for materials should be addressed to L.N. All other raw data from the main figures were deposited to Mendeley and are publicly available as of the date of publication (https://doi.org/10.17632/btrk86ygj6.1)[77]. Source data are provided with this paper.

Code availability

The FUSiLLi pipeline is provided as Supplementary Code.

References

73. Gaudelli, N. M. et al. Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.* **38**, 892–900 (2020).
74. Samulski, R. J., Srivastava, A., Berns, K. I. & Muzyczka, N. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. *Cell* **33**, 135–143 (1983).
75. Calabria, A. et al. Intrathymic AAV delivery results in therapeutic site-specific integration at TCR loci in mice. *Blood* **141**, 2316–2329 (2023).

76. Cipriani, C. et al. RAAVioli: a comprehensive approach to characterizing AAV vector integrations and rearrangements. *Mol. Ther. Adv.* **34**, 201659 (2026).
77. Canarutto, D. et al. Selection of human hematopoietic stem cells bearing the intended functional edit by transient 'AND-gate' reporters. *Mendeley* https://doi.org/10.17632/btrk86ygj6.1 (2026).

Acknowledgements

We thank all members of the L.N. laboratory for discussion, the IRCCS San Raffaele Hospital Flow Cytometry facility, A. Auricchio (Telethon Institute of GEnetics and Medicine Vector Core), M. Doria (InnovaVector), V. Blouin and F. Balter (Vectors Production Center, INSERM UMR1089) for producing and supplying some recombinant AAVs used in this study. We thank I. Cuccovillo, C. Asperti, G. Desantis, C. Rossi, K. Giannetti, E. Rovelli, E. Pettinato, T. Plati and T. Di Tomaso for purification of mPB HSPCs and E. Mercuri for contributing to the murine *Cd40lg*$^{-/-}$ experiment. This work was supported by grants to L.N. from Fondazione Telethon (SR-Tiget Core Grant), the EU's Horizon Europe Program (UPGRADE no. 825825; X-PAND no. 101070950), the Italian Ministry of University and Research (PRIN 2017 no. 20175XHBPN; FIS 2021 Advanced Grant no. FIS00002157) and the Louis-Jeantet Foundation through the 2019 Jeantet-Collen Prize for Translational Medicine, to S.F. from the Italian Ministry of Health (Ricerca Finalizzata SG-2021-12374529), to P.G. from the National Institutes of Health (R01AI155796 and R01HL170629), to B.G. from the EU's Horizon Europe Program (X-PAND no. 101070950), to D.C. from the 2023 Hyper-IgM Foundation grant, to K.S. from the Japan Society for the Promotion of Science Overseas Research Fellowship, Mochida Memorial Foundation for Medical and Pharmaceutical Research Study Abroad Assistance Grant and Kagawa University Faculty of Medicine School of Medicine Sanju Alumni Association Sanjukai Overseas Fellowship and to V.V. from the Marie Skłodowska-Curie Postdoctoral Fellowship funded by the European Union's Horizon Europe research and innovation program (grant agreement no. ENRICHING-HDR-101210738).

Author contributions

D. Canarutto and M.F. performed the research and interpreted the data. C.G. and L.A. provided technical support. A. Varesi, C.F., K.S., V.V. and G.S. performed some experiments. G.P. and A.S. performed the long-read sequencing experiments. A.A. designed the TALE-based ArT. E.Z. developed the improved HSPC expansion protocol under the supervision of B.G. A.L., P.G., S.F. and L.N. conceptualized the SMArT strategies. D. Cipria contributed to the conceptualization of the polyfunctional editor in A.L. laboratory. D. Canarutto developed the FUSiLLi pipeline. S.B. contributed to the development the FUSiLLi pipeline. P.G., S.F., A. Villa and L.N. supervised the research. L.N. coordinated the work. D. Canarutto, M.F., S.F. and L.N. wrote the manuscript. All authors proofread the manuscript.

Competing interests

D. Canarutto, M.F., G.S., A.A., E.Z., B.G., A.L., P.G., S.F., A. Villa, V.V. and L.N. are inventors of patents on gene therapy owned and managed by IRCCS Ospedale San Raffaele and Fondazione Telethon ETS. The other authors declare no competing interests.

Additional information

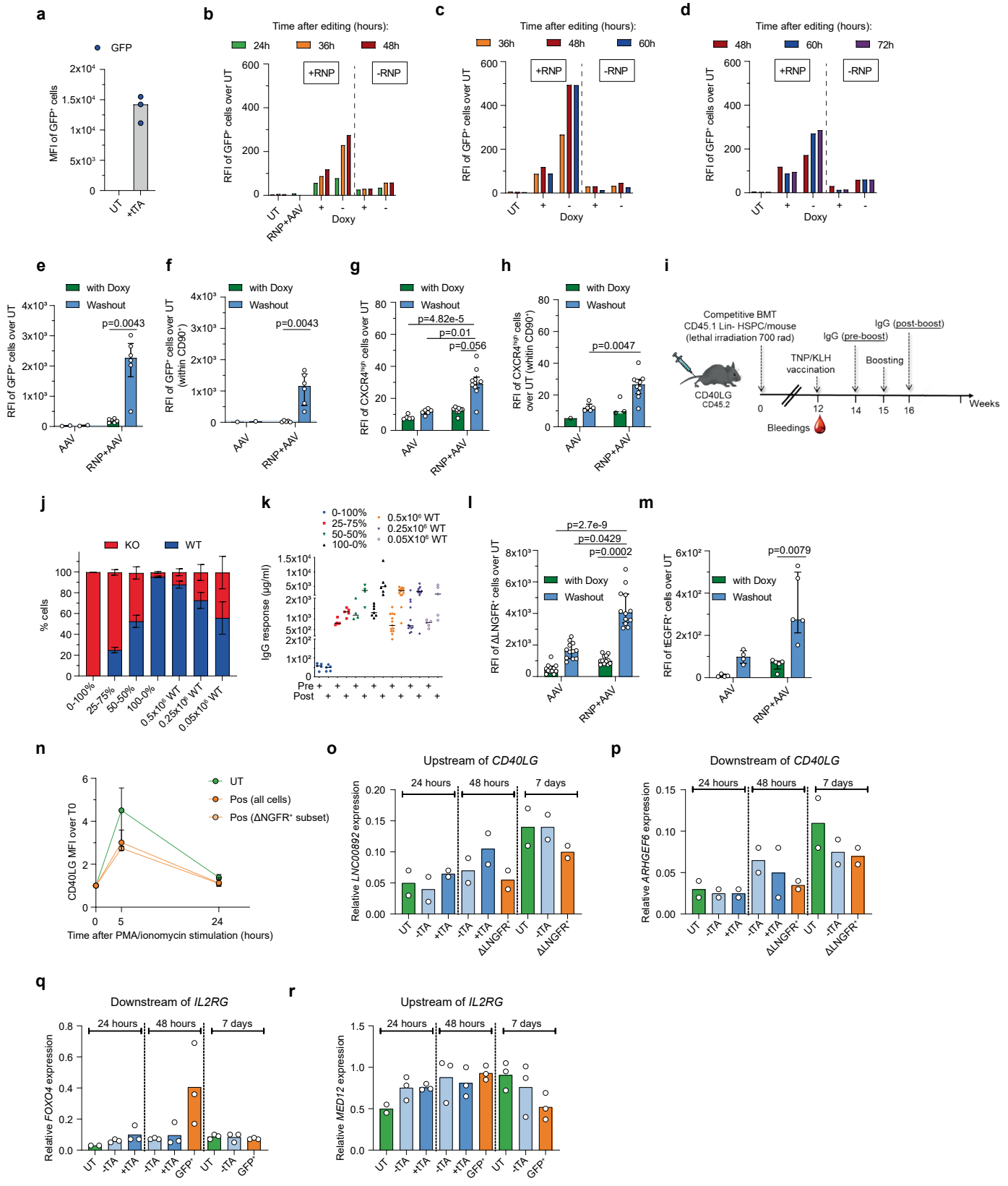
Extended data is available for this paper at https://doi.org/10.1038/s41587-026-03142-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-026-03142-z.

Correspondence and requests for materials should be addressed to Luigi Naldini.

Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41587-026-03142-z

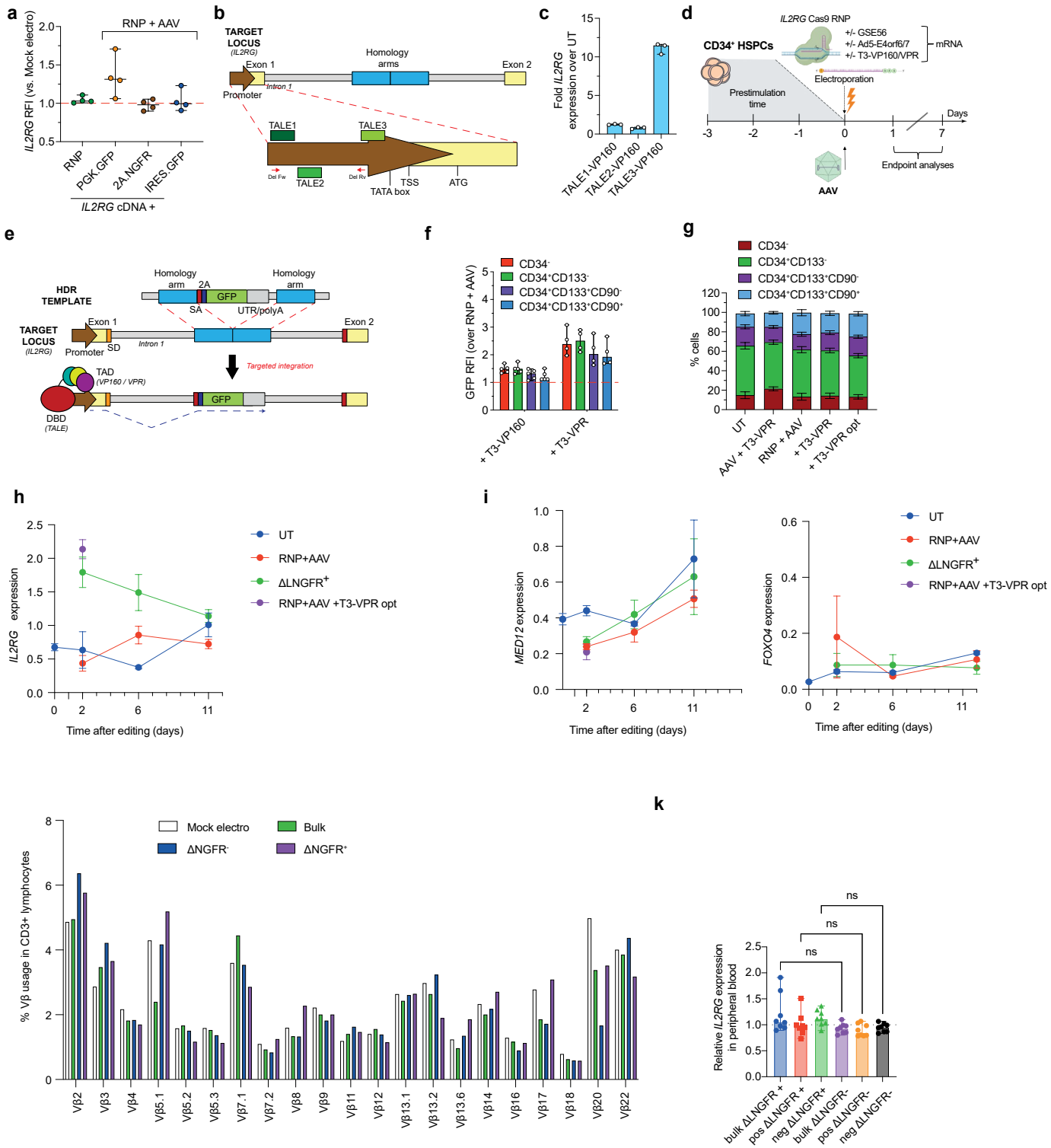


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41587-026-03142-z

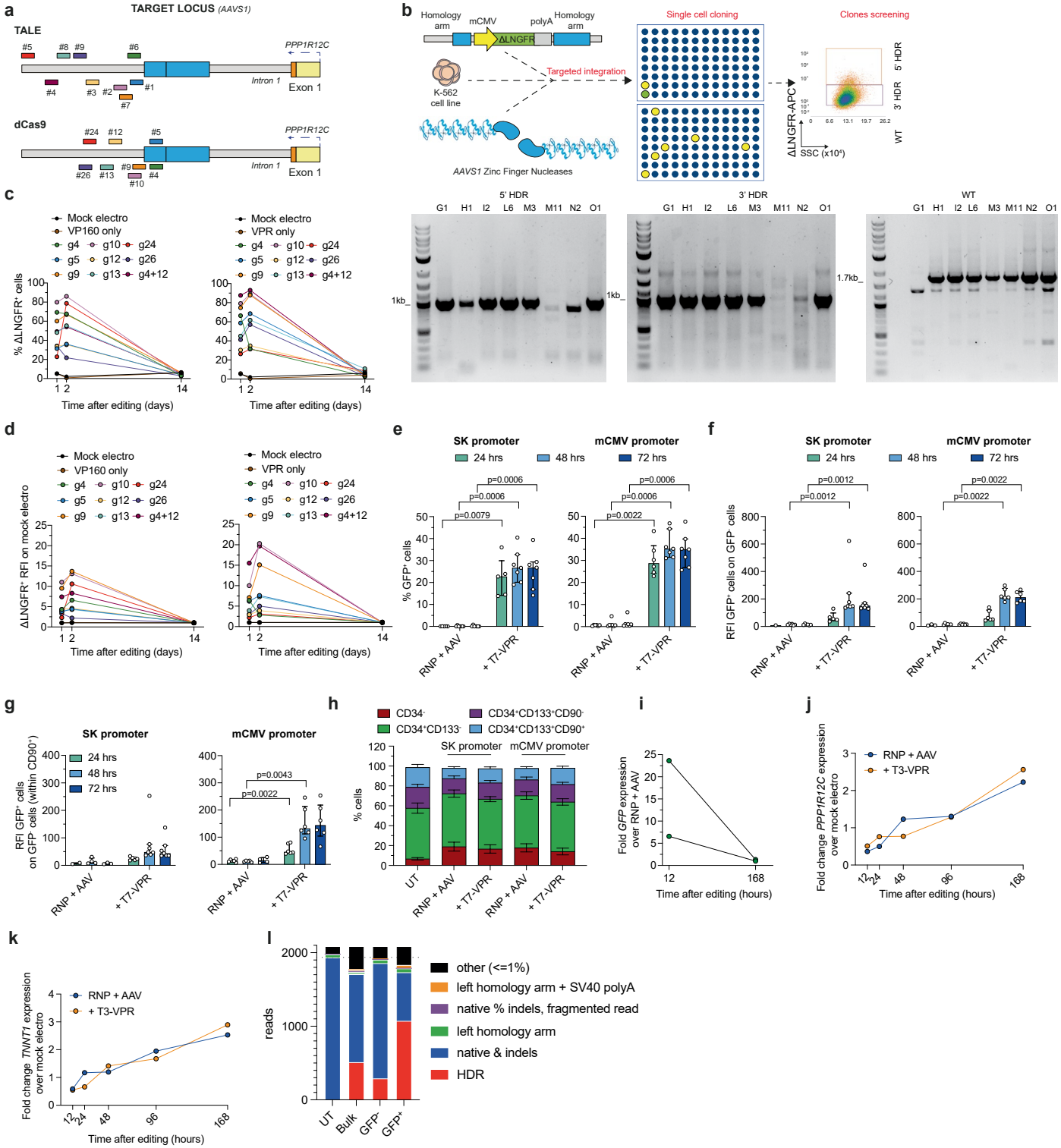
Extended Data Fig. 1 | SMArT-1 enables enrichment of cells with integration of the corrective payload via tTA-mediated transactivation of the selector. **a**, Mean fluorescent intensity (MFI) of selector (GFP)$^+$ K-562 cells 9 days after editing upon delivery of the tTA mRNA. n = 1.3. **b-d**, Relative fluorescence intensity (RFI) of the GFP selector in CB HSPCs edited at the *IL2RG* locus at different timepoints after doxycycline washout. Doxycycline was washed out 12 hours (**b**), 24 hours (**c**) or 36 hours (**d**) after editing; - RNP controls serve as measure of transactivation of episomal AAV6. **e,f**, RFI of selector (GFP)$^+$ cells in bulk (**e**) or CD90$^+$ (**f**) CB HSPCs reported in Fig. 1d, e. Median with IQR. n = 2,2,5,6. Two-sided Mann–Whitney test. **g,h**, RFI of selector (CXCR4)^{high} cells in bulk (**g**; n = 5,5,9,10 Kruskal-Wallis test with Dunn's multiple comparisons) or CD90$^+$ (**h;** n = 1,5,3,10 Two-sided Mann–Whitney test) CB HSPCs reported in Fig. 1h, i. Median with IQR. **i**, Schematic representation of competitive transplantation in *Cd40lg-/-* mice. **j**, Peripheral blood (PB) cell chimerism 26 weeks after transplant of mice transplanted with different numbers and ratios of WT and *Cd40lg-/-* Lin$^-$HSPCs (n = 5, 5, 5, 5, 15, 12, 6). Mean and SEM. **k**, Concentration of IgG anti-2,4,6-tri-nitrophenil conjugated keyhole limpet hemocyanin (TNP-KLH) in mice from the indicated experimental groups before (PRE) and after (POST) boosting of the TNP-KLH response (n = 5, 5, 5, 5, 4, 4, 7, 6, 13, 14, 12, 12, 5, 5). **l,m**, RFI of selector-positive (ΔLNGFR (**l**) n = 12, Kruskal-Wallis test with Dunn's multiple comparisons; tEGFR (**m**) n = 4,4,5,5 Two-sided Mann–Whitney test) cells in bulk mPB HSPCs reported in Fig. 1l–m. Median with IQR. **n**, Fold CD40LG surface expression (MFI) during PMA-ionomycin stimulation in T cells from Fig. 1p. n = 3, Median and range. **o,p**, Expression of genes neighboring *CD40LG* (*LNC00892* and *ARHGEF6)* upon culture in edited T cells (-tTA), after exposure to tTA (+tTA), and after sorting (ΔLNGFR$^+$), over time in culture. n = 2. **q,r**, Expression of genes neighboring *IL2RG (FOXO4* and *MED12)* upon culture in edited T cells (-tTA), after exposure to tTA (+tTA), and after sorting (GFP$^+$), over time in culture. n = 3. All gene expressions are relative to *HPRT*.

https://doi.org/10.1038/s41587-026-03142-z



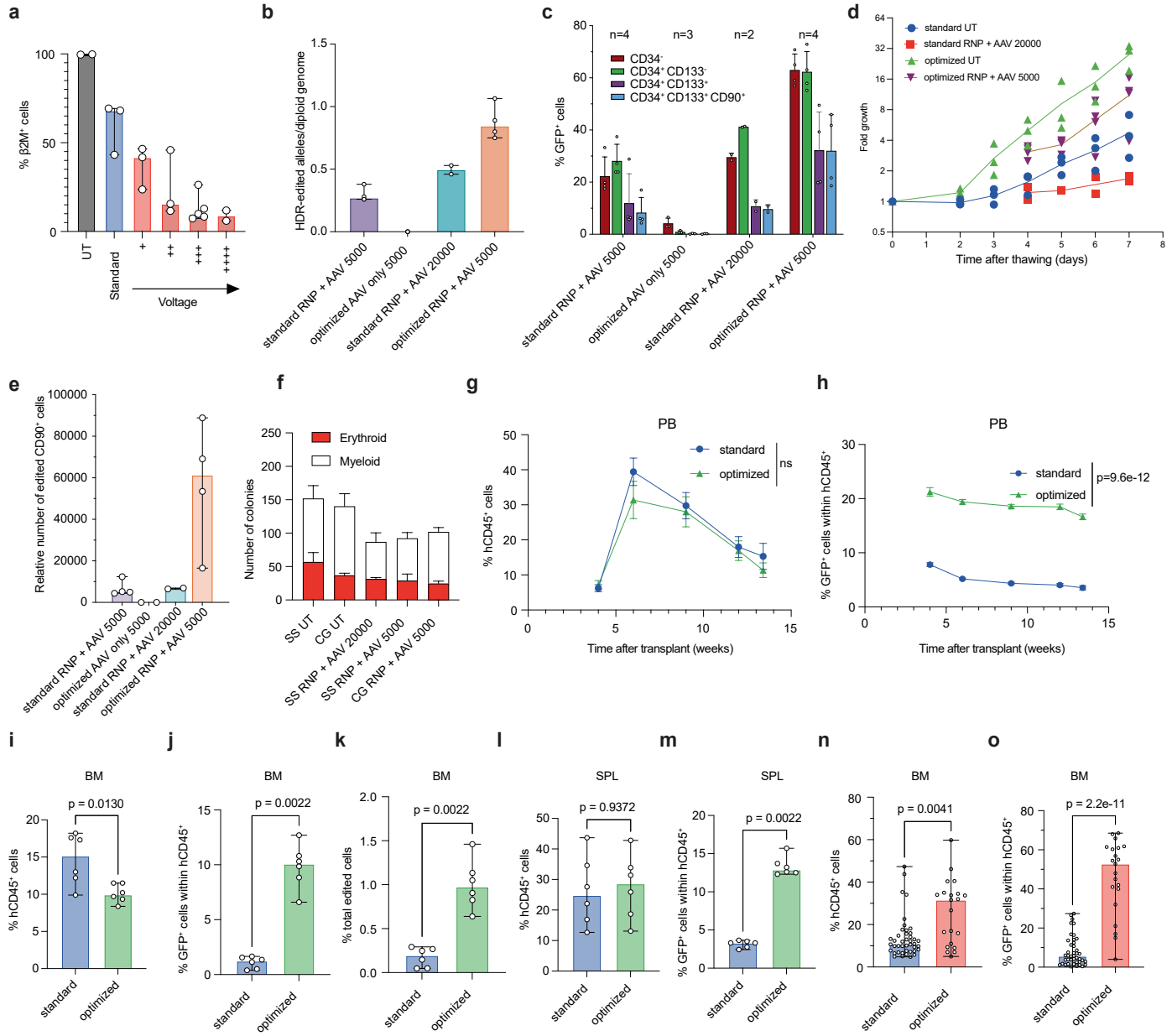
Extended Data Fig. 2 | SMArT-2 enriches HSPCs bearing HDR-mediated targeted integration in the context of IL2RG gene correction. a, RFI of *IL2RG* in selector positive male CD4 T cells edited at the *IL2RG* locus as illustrated in Fig. 2a with different SMArT-2 donor configurations, in absence of ArT, 3 days after editing, compared to the mock electroporated condition (dotted red line). n = 4, median and range. p = 0.0517. Friedman test. **b**, Location of candidate DNA binding sites for alternative TALE configurations at the *IL2RG* promoter. **c**, Fold *IL2RG* expression in K-562 with different TALE-VP160. n = 3, median and range. **d**, Experimental design. HSPCs were prestimulated, electroporated with editing reagents and the T3 mRNA, and transduced with AAV6 encoding for the SMArT-2 donor cassette. **e**, Simplified configuration of the SMArT-2 strategy illustrated in Fig. 2a, comprising only the selector linked by a 2 A and a splice acceptor to the first exon of *IL2RG*. **f**, RFI of selector (GFP)$^+$ HSPCs subpopulations transactivated with either T3-VP160 or T3-VPR over the non-transactivated control 24 hours after editing. n = 4, Median and 95% CI. **g**, Phenotypic composition of HSPCs from Fig. 2b. n = 4, mean and SEM. **h,i**, Relative expression of native and recoded *IL2RG* (**h**) and neighboring genes *MED12* and *FOXO4* (**i**) in SMArT-2 edited HSPCs over time in culture in the optimized protocol and selection. n = 3 biological replicates, mean and SEM. **j**, Percentage of CD3$^+$ cells expressing the listed Vβ families in mock edited, bulk edited, ΔLNGFR$^+$ sorted cells, and ΔLNGFR$^-$ sorted cells. n = 1, also included in Fig. 2j. **k**, Comparison of relative *IL2RG* expression by flow cytometry in ΔLNGFR$^+$ and ΔLNGFR$^-$ fractions in the PB, at the last timepoint. n = 8. Kruskal-Wallis test (p = 0.02) with Dunn's multiple comparisons. Median and range.

https://doi.org/10.1038/s41587-026-03142-z



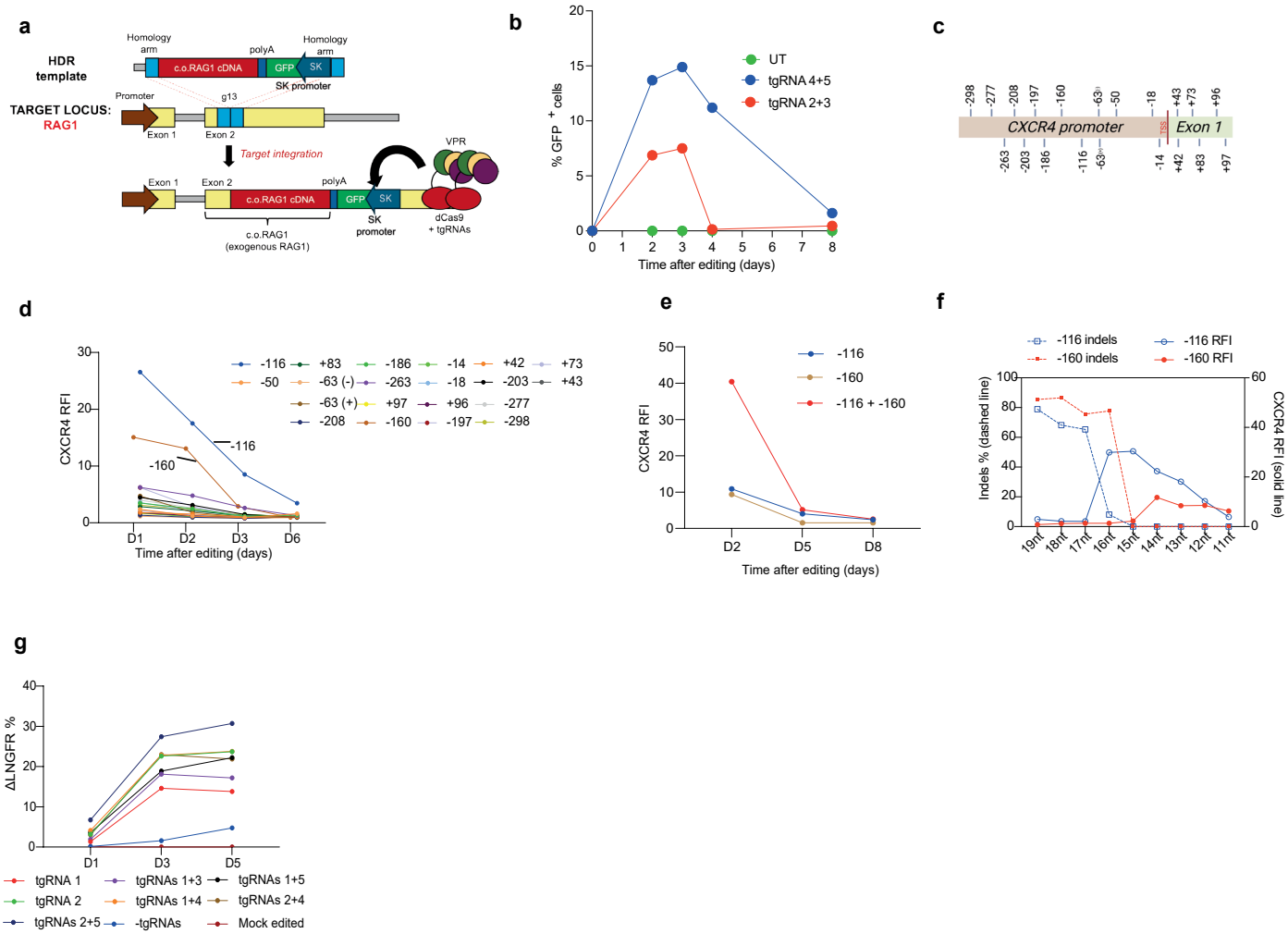
Extended Data Fig. 3 | SMArT-3 enable the enrichment of HDR-edited HSPCs carrying targeted integration of transgene cassettes at the *AAVS1* locus. **a**, Location of candidate DNA binding sites for alternative TALE (top) and dCas9 (bottom) configurations. The homology arms are indicated in blue. **b**, Generation and validation of single cell-derived K-562 clones harboring the desired cassette integration into the *AAVS1* locus. A reporter cassette consisting of the mCMV promoter (in yellow) driving the expression of the ΔLNGFR selector was integrated into the *AAVS1* locus using previously reported zinc-finger nucleases[13]; correct HDR junctions were confirmed by PCR (5′ HDR, 3′ HDR, and wild type (WT) allele). Each lane corresponds to a different clone. The H1 clone was selected for further experiments. n = 1 replicate. **c,d**, Percentage (**c**) and RFI (**d**) of selector (ΔLNGFR)$^+$ cells upon delivery of different gRNAs in the H1 clone using dCas9-VP160 (left) or dCas9-VPR (right). **e,f**, Percentage (**e**) and RFI of GFP$^+$/GFP$^-$ (**f**) bulk HSPCs after editing in presence or absence of T7-VPR with either the SK (n = 5,7,7,5,7,7) or the mCMV (n = 6,7,7,6,7,7) promoters. Two-sided mann–Whitney test. **g**, RFI of GFP$^+$ CD90 cells / GFP$^-$ CD90 cells with either the SK (n = 5,7,7,5,7,7) or the mCMV (n = 6,7,7,6,7,7) promoter. Two-sided Mann–Whitney test. **h**, Phenotypical distribution of HSPCs in presence or absence of the T7-VPR transactivator with either the SK or the mCMV. n = 6 **i**, Fold expression of GFP in transactivated cells over edited non-transactivated HSPCs. n = 2. **j,k**, Fold *PPP1R12C* (**j**) and *TNNT1* (**k**) gene expression over time in transactivated (T7-VPR) and non-transactivated (RNP + AAV) over mock-electroporated HSPCs. **l**, Long-read sequencing of the same samples presented in Fig. 3l after long range PCR amplification and analyzed with FUSiLLi.

https://doi.org/10.1038/s41587-026-03142-z



Extended Data Fig. 4 | Development of an optimized HDR editing protocol for human HSPCs using GMP-ready editing conditions. a, B2M knockout efficiency in HSPCs by flow cytometry using the Lonza electroporator (blue) or increasing voltages with the Maxcyte electroporator OC-25×3 (HSC1-4) n = 2,3,3,3,5,2. Median with range. **b,c**, Fraction of HDR-edited alleles by ddPCR (**b**, n = 4,3,2,4 biological replicates, median and range) and percentage of GFP⁺ cells in different subpopulations (**c**, mean and SD) with Stemspan (standard) or GMP Stem Cell Growth Medium (optimized) at *AAVS1* locus with an AAV6 encoding for a PGK-GFP reporter and two homology arms, at reported MOIs (5000 or 20000). n indicates biological replicates **d**, Fold growth of unedited and HDR-edited mPB HSPCs with different media and MOIs (n = 3 biological replicates standard UT and optimized UT, n = 2 biological replicates standard RNP + AAV 20k, n = 4 biological replicates optimized RNP + AAV 5k, mean and SEM). Cells were edited on day 3. Lines indicate mean. **e**, Relative number of edited CD90⁺ cells by flow cytometry with different media and MOIs. n = 4,2,2,4. Median with range. **f**, Clonogenic potential of HSPCs edited with the different media and MOIs, n = 3 independent donors from 2 experiments. Mean and SD. **g-m**, Side-by-side comparison of *AAVS1* HDR editing in standard (Stempsan medium, MOI 20k, Lonza electroporator) and optimized conditions (GMP Stem Cell Growth Medium medium, MOI 5k, Maxcyte electroporation) upon xenotransplantation into NSG mice (n = 6 mice/group, 1 HSPC donor), using an AAV6 encoding for PGK-GFP and two homology arms. **g,h**, engraftment (**g**) and editing efficiency (**h**) by GFP expression in PB. 2-way ANOVA. Mean and SEM **i-k**, Human engraftment (**i**), HDR efficiency by GFP (**j**) and total number of edited cells (**k**) in the BM after xenotransplantation. Median and range. **l,m**, Engraftment (**l**) and editing efficiency by GFP (**m**) in the SPL after xenotransplantation. Median and range. **n,o**, Comparison of engraftment (**n**) and HDR efficiency by GFP expression (**o**) in the BM of NBSGW mice (pool of n = 42 mice in standard conditions, n = 21 mice in optimized conditions from different experiments). Median and range. Two-tailed Mann-Whitney tests in panels **i-o**. Each dot is a biological replicate.

https://doi.org/10.1038/s41587-026-03142-z



Extended Data Fig. 5 | Polyfunctional editors improve portability and broaden applications of SMArT. a, Schematic representation of the SMArT-3 construct used for *RAG1* correction. **b**, Percentage of GFP$^+$ CD34$^+$ HSPCs after editing in presence dCas9-VPR with the construct shown in panel **a** with two different pairs of truncated gRNAs (tgRNAs) (n = 1). **c-f**, Screening of gRNAs for CXCR4 transactivation. **c**, Location of gRNAs with respect to *CXCR4* transcription start site (TSS). **d**, Relative expression of *CXCR4* by flow cytometry with gRNAs illustrated in **c**. **e**, Synergistic CXCR4 transactivation using the two best performing gRNAs in **d. f**, Transactivation and cutting efficiency by surveyor assay of CXCR4 using Cas9-VPR gRNAs -116 and -160 of progressively shorter length in K562 cell lines. **g**, ΔLNGFR expression in HDR edited mPB HSPCs using different combinations of tgRNAs illustrated in Fig. 5g (n = 1).

nature portfolio

Corresponding author(s): Luigi Naldini

Last updated by author(s): Apr 3, 2026

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	Immunophenotypic analysis was performed on FACS Canto II (BD Pharmingen). ddPCR data was acquired with Biorad QX200 or QX600. Long read sequencing data was acquired with Oxford Nanopore Promethion P2 solo.
Data analysis	Flow citometry was analyzed with FlowJo of FCS Express. ddPCR data was analyzed with QXManager. Long read sequencing data was basecalled with dorado, aligned with minimap2, and analysed with IGV and SnapGene.

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

Data

Policy information about availability of data

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

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

All relevant data are included in the manuscript. The reagents described in this manuscript are available under a material transfer agreement with IRCCS Ospedale San Raffaele and Fondazione Telethon; requests for materials should be addressed to L.N.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Male cells were used for IL2RG and CD40G experiments. Gender was not considered.
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	Human T-cells and hematopoietic stem cells were sourced from anonymized healthy donors.
Recruitment	Primary T cells were isolated from buffy coats from healthy donors as results of residues of blood product donations. mPB HSPCs were purified from commercial aphereses according to the TIGET-HPCT protocol.
Ethics oversight	Buffy coats were obtained in accordance with the Declaration of Helsinki, as anonymized residues of blood donations, used upon signature of specific institutional informed consent for blood product donation by healthy blood donors. HSPCs were obtained in accordance with the Declaration of Helsinki. T and HSPCs collection were approved by the Ospedale San Raffaele Scientific Institute bioethical committee (TIGET-HPCT).

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

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	Sample size for each experiment was determined by the total number of available cells, which is constrained by the human source of the material, to be split among each experimental conditions. Whenever possible we aimed to reach at least 3 replicates per group.
Data exclusions	For in vivo experiments failure during injection of treated HSPCs, confirmed by graft failure, in recipient animals led to exclusion of that mouse from the experimental group. No other data or sample were excluded from analysis. All these criteria were pre-established.
Replication	Number of biological replicates is specified for each experiment in figure legends or methods. All attempts at replication were successful.
Randomization	For tissue culture experiments, conditions were assigned to wells in 96- or 48- or 24-well or 6-well plates; plate position is not expected to affect editing efficiency. Mice were randomly distributed to each experimental group.
Blinding	Experiments were not performed in blind-fashion. Blinding during data collection and analysis was not feasible due to the nature of the experimental procedures, which required awareness of treatment conditions for correct execution and data interpretation. This limitation is not expected to have introduced bias, as objective quantitative measurements were used for all primary outcomes. Blood samples ware collected by a blind operator and mice were identified by a code not reflecting the treatment group.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

Anti-human CD34 VioBlue AC136 Miltenyi Biotec Catalog n°130-113-182 1:50
Anti-human CD34 PE-Vio770 8G12 BD Biosciences Catalog n° 348811 1:20
Anti-human CD133/2 PE REA816 Miltenyi Biotec Catalog n°130-112-157 1:50
Anti-human CD90 APC 5E10 BD Biosciences Catalog n°559869 1:33
Anti-human CXCR4 PE-Vio770 REA649 Miltenyi Biotec Catalog n°130-109-845 1:50
Anti-human NGFR PE ME20.4-1.H4 Miltenyi Biotec Catalog n°130-113-421 1:50
Anti-human NGFR APC ME20.4-1.H4 Miltenyi Biotec Catalog n°130-113-418 1:50
Anti-human EGFR APC AY13 BioLegend Catalog n°352906 1:50
Annexin V VioBlue BioLegend Catalong n° 640918 1:50
7AAD BioLegend Catalong n° 420404 1:33
Anti-human CD45 VioBlue HI30 BioLegend Catalog n°304029 1:50
Anti-human CD45 APC 2D1 BD Biosciences Catalog n°340910 1:50
Anti-human CD45 APC-Vio770 2D1 BD Biosciences Catalog n°348815 1:50
human Fc blocking Miltenyi Biotec Catalog n°130-059-901 1:50
mouse Fc blocking BD Biosciences Catalog n°553142 1:100
Anti-human CD132 APC TUGh4 BioLegend Catalog n°3386081:50
Anti-human CD132 PE AG184 BD Biosciences Catalog n°5616991:50
Anti-human CD90 BV421 5E10 BioLegend Catalog n°3281221:50
Anti-human CD3 Pecy7 HIT3a BioLegend Catalog n°3003161:50
Anti-human CD40LG PE 24-31 Invitrogen Catalog n°12-1548-421:50
Anti-human CD45 PO 5B1 Miltenyi Biotec Catalog n°130-113-1241:50
Anti-human CD19 Pecy7 HIB19 BioLegend Catalog n°302-2161:50
Anti-human CD3 FITC SK7 BD Biosciences Catalog n°3457631:50
Anti-human CD13 APC-H7 REA263 Miltenyi Biotec Catalog n°130-103-6711:50
Anti-human CD13 BV421 WM15 BD Biosciences Catalog n°5625961:50
Anti-human CD90 PE 5E10 BD Biosciences Catalog n°5555961:50
Anti-human CD40LG APC 24-31 Invitrogen Catalog n° 17-1548-425:100

Validation

Antibodies were used according to manufacturers' instructions and validation.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

K-562 were purchased by ATCC

Authentication

Not authenticated

Mycoplasma contamination

Cell lines were routinely checked for mycoplasma contatmination

Commonly misidentified lines
(See ICLAC register)

NA

Animals and other research organisms

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

Laboratory animals

C57Bl/6 Ly45.1 mice were purchased from Charles River Laboratory, Cd40lg−/− (B6.129S2-Cd40lgtm1Imx/J), NOD-SCID-IL2Rg−/− (NSG) and NOD,B6.SCID Il2rgamma-/- Kit(W41/W41) (NBSGW)mice (7-10 weeks of age) were obtained from Jackson Laboratory. Animals were maintained in Specific pathogen-free (SPF) animal research facilities with a 12h/12h dark/light cycle and standardized temperature (22 +/- 2°C) and humidity (55 +/- 5%).

Wild animals

NA

Reporting on sex	The study involves only female animals as they allow better human engraftment
Field-collected samples	NA
Ethics oversight	IACUC no. 818, 1206 and 1358

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

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA

Flow Cytometry

Plots

Confirm that:

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

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

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

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

Methodology

Sample preparation	Immunophenotypic analyses were performed on the fluorescence activated cell sorting Canto II (BD Pharmingen). From 0.5x10E5 to 2x10E5 cells (either from culture or mouse samples) were stained for 15 min at 4 °C with antibodies listed in Supplementary Table 2 in a final volume of 100 µl and then washed with DPBS + 2% heat inactivated FBS. Single stained samples or compensation beads were used as controls. The Live/Dead Fixable Dead Cell Stain Kit (Thermo Fisher) or 7-aminoactinomycin D (Sigma Aldrich or Biolegend) were included during sample preparation according to the manufacturer's instructions to identify dead cells.
Instrument	Canto II (BD Pharmingen); Cell sorting was performed on a BD FACSAria Fusion (BD Biosciences) using BDFACS Diva software and equipped with four lasers: blue (488 nm), yellow/green (561 nm), red (640 nm) and violet (405 nm). Cells were sorted with an 85 mm nozzle. Sheath fluid pressure was set at 45 psi. A highly pure sorting modality (four-way purity sorting) was chosen. Sorted cells were collected in 1.5 ml Eppendorf tubes containing 500 µl of DPBS or HSPC medium. Sheath fluid pressure was set at 17-18psi. Cell sorting in Figure 4j was performed on a MACSQuant Tyto Cell Sorter (Miltenyi Biotec) using the MACSQuantify™ Tyto Software and equipped with blue (488nm), red (640nm) and violet (405nm).
Software	FCS Express or Flow-jo
Cell population abundance	The abundance of population within post-sort fractions was >800 cells
Gating strategy	Gating Stategies are illustrated in supplementary figure 2.

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