

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

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

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

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

All relevant data are included in the manuscript.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

CD34+ HSPCs from mobilized peripheral blood of 8 different, deidentified healthy donors (male and female) were obtained from the Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829). Plerixafor-mobilized peripheral blood from patients with sickle cell disease was procured after Boston Children's Hospital institutional review board (IRB) approval and patient-informed consent.

Population characteristics

De-identified healthy donors and patients affected by sickle cell disease

Recruitment

Cells obtained for research purpose only through commercial sources (Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829)) and Boston Children's Hospital.

Ethics oversight

Fred Hutch Cooperative Center of Excellence in Hematology (U54 DK106829), Boston Children's Hospital institutional review board (IRB)

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

## Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size

Sample sizes were determined by the availability of primary human cells and, for in vivo studies, by engraftment capacity and ethical constraints on animal use. Within these limits, cells were allocated evenly across conditions, and we aimed for  $\geq 3-5$  independent biological replicates per group, consistent with standard practice in HSPC editing and xenotransplantation studies. The selected sample sizes enabled estimation of variance and application of appropriate statistical tests. Observed effect sizes were large relative to within-group variability and were reproducible across independent experiments, supporting the adequacy of the sample sizes to detect biologically meaningful differences. No samples were excluded unless predefined technical quality criteria were not met (see below). In some in vivo experiments, such as secondary transplantation, the total number of available cells was more constrained and limited to what could be harvested from the primary recipients.

Data exclusions

No data or sample were excluded from analysis, unless the data point was missing due to technical reasons (sample availability or instrumentation acquisition errors). Outlier exclusion was applied to the in vivo barcode library fully analyzed data (retrieved barcodes, Shannon H) and is fully described in the Figure legend/data description. N=1 outlier data point, likely due to sequencing issues. All these criteria were pre-established.

Replication

Number of biological replicates is specified for each experiment in figure legends. All attempts at replication were successful. Inferential techniques were applied in presence of adequate sample sizes ( $n \geq 3-5$ ), otherwise only descriptive statistics are reported.

Randomization

Mice were randomly distributed to each experimental group. For in vitro experiments, formal randomization was not applied, as experimental groups were defined by predetermined conditions (e.g., editing constructs, antibody treatments) rather than by assignment of interchangeable units. Cells from the same donor or culture were distributed evenly across all conditions and processed in parallel to minimize technical variability. Where applicable, treatments were applied simultaneously using identical reagents and conditions, and sample handling and readouts were performed in a consistent and unbiased manner. For experiments involving multiple donors or independent cultures, biological replicates were included and analyzed across conditions to control for inter-sample variability.

Blinding

Not relevant for objective measures.

# Reporting for specific materials, systems and methods

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

## Materials & experimental systems

| n/a                                 | Involved in the study                                           |
|-------------------------------------|-----------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |

## Methods

| n/a                                 | Involved in the study                              |
|-------------------------------------|----------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

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

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

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

## Animals and other research organisms

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

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

## Flow Cytometry

### Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation

Cell lines: harvested, washed, incubated with Fc-block reagent and stained at 4C for 15-30 min, then washed. Mouse peripheral blood: collected, lysed with ACK reagent for 10 minutes at room temperature, then washed (repeat 2x lysis and wash); incubated with human and murine Fc-block, stained at room temperature for 15 minutes, then washed. Bone marrow, spleen: collected, washed, lysed with ACK reagent for 5 minutes at room temperature and washed; incubated with human and murine Fc-block, stained on ice for 35 minutes, then washed. BD Brilliant Stain buffer was included when multiple Brilliant Violet or UV dyes were co-stained. Viability staining was performed using Live/Dead yellow, 7-AAD, Propidium Iodide and AnnexinV staining depending on the experiment. All wash steps were performed with PBS + 2% FBS or PBS + 2% FBS + 2mM EDTA.

Instrument

BD Fortessa flow cytometer, BD Fortessa X-20, BD FACS Melody sorter, BD Aria Fusion sorter.

Software

BD Diva (acquisition, sorting), BD Chorus (sorting), FCS Express v6 or v7 (analysis).

Cell population abundance

See Extended Data FIG.5 for gating strategy on mouse BM samples. We also directly report absolute counts of all hematopoietic populations from in vivo experiments in the Main Figures and Source Data.

Gating strategy

See Methods for details. For most analyses, gating strategy included FSC-A vs FSC-H to exclude doublets (diagonal gating), followed by viability through exclusion of PI/7AAD+ events, followed by physical parameters to detect cells with appropriate physical features according to the specific sample being analyzed (FSC-A vs SSC-A). Subsequent gating is experiment-specific and detailed in Supplementary Data Table 2, Extended Data Figure 1H, 7H and Supplementary Data FIG.6-7.

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