



Universiteit  
Leiden  
The Netherlands

## **Decoding telomere gene variation through saturation genome editing: from functional genomics to clinical interpretation**

Obolenski, S.

### **Citation**

Obolenski, S. (2026, September 24). *Decoding telomere gene variation through saturation genome editing: from functional genomics to clinical interpretation*. Retrieved from <https://hdl.handle.net/1887/4310110>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4310110>

**Note:** To cite this publication please use the final published version (if applicable).

# Protocol for the Functional Evaluation of Genetic Variants using Saturation Genome Editing

*STAR Protocols* 2025;6:103710

Sofia Obolenski<sup>1,2,3</sup>, Rebeca Olvera-León<sup>1</sup>, Dijue Sun<sup>1</sup>, David J. Adams<sup>1</sup>, Andrew J. Waters<sup>1,3,4</sup>

## Affiliations

<sup>1</sup>Experimental Cancer Genetics, Wellcome Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SA Cambridge, UK

<sup>2</sup>Department of Dermatology, Leiden University Medical Centre, 2333 ZA Leiden, the Netherlands

<sup>3</sup>Technical contact

<sup>4</sup>Lead contact

Corresponding authors: [da1@sanger.ac.uk](mailto:da1@sanger.ac.uk) and [aw28@sanger.ac.uk](mailto:aw28@sanger.ac.uk)

## SUMMARY

Saturation genome editing (SGE) employs CRISPR-Cas9 and homology-directed repair (HDR) to introduce exhaustive nucleotide modifications at specific genomic sites in multiplex, enabling the functional analysis of genetic variants while preserving their native genomic context. Here, we present a protocol for SGE-based variant evaluation in HAP1-A5 cells. We describe the steps for designing variant libraries, single-guide RNAs (sgRNAs), and oligonucleotide primers for PCR. We also detail the sample preparation before the SGE screen, the cellular screening process, and subsequent next-generation sequencing (NGS) library preparation.

For complete details on the use and execution of this protocol, please refer to Radford *et al.*,<sup>1</sup> Waters *et al.*,<sup>2</sup> and Olvera-León *et al.*<sup>3</sup>

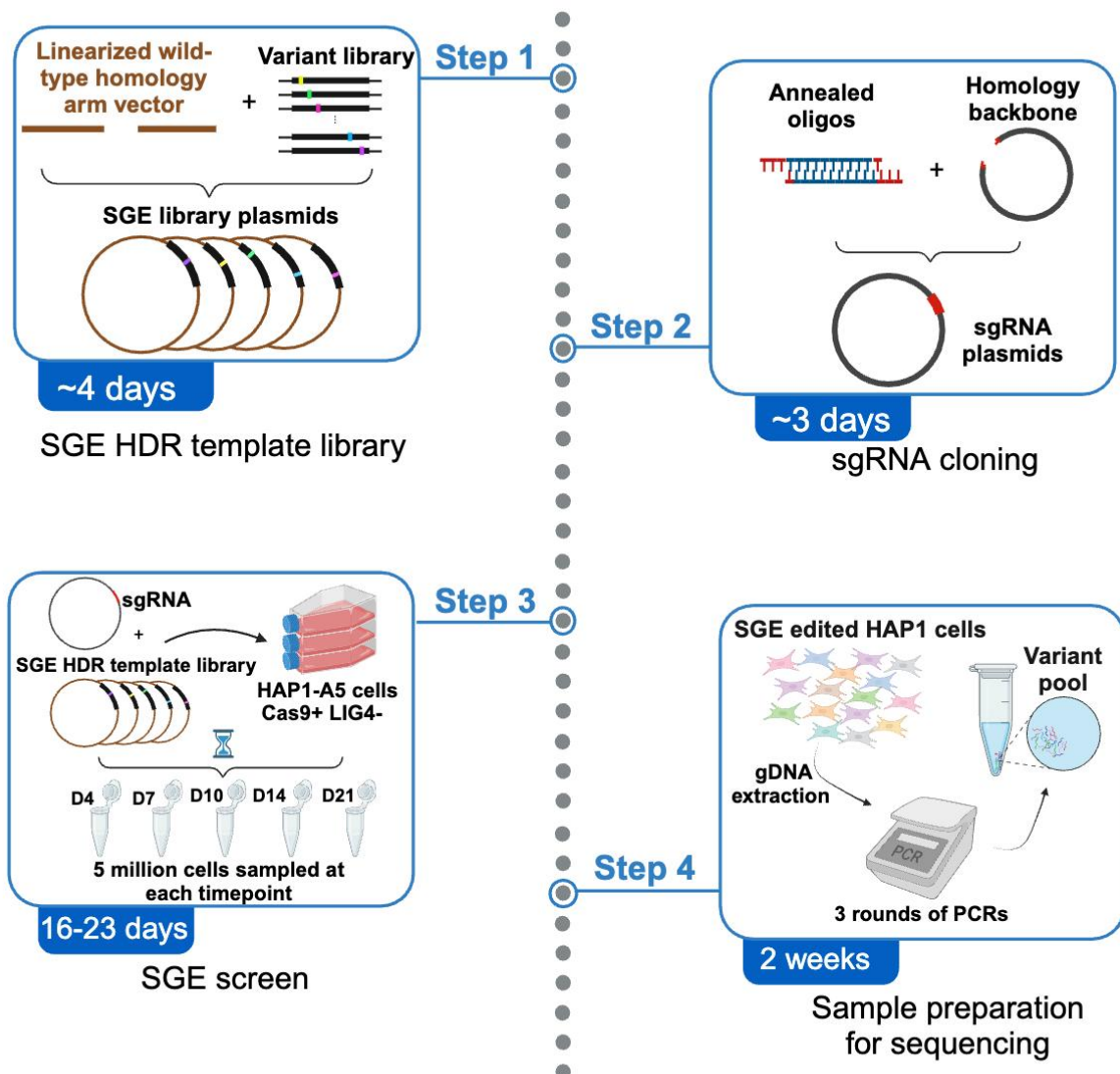
**Subject areas:** CRISPR, Genetics, Genomics, High-Throughput Screening

## HIGHLIGHTS

- SGE protocol in HAP1-A5 cells to assess variant effects on cell fitness over time
- Description of SGE library and sgRNA design, cloning, and HAP1-A5 cell culture
- Steps for SGE screen and SGE-edited gDNA library preparation for NGS
- SGE approach generates functional scores for all SNVs and key variants in CDS, intron, and UTRs

**Publisher's note:** Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

## GRAPHICAL ABSTRACT



## BEFORE YOU BEGIN

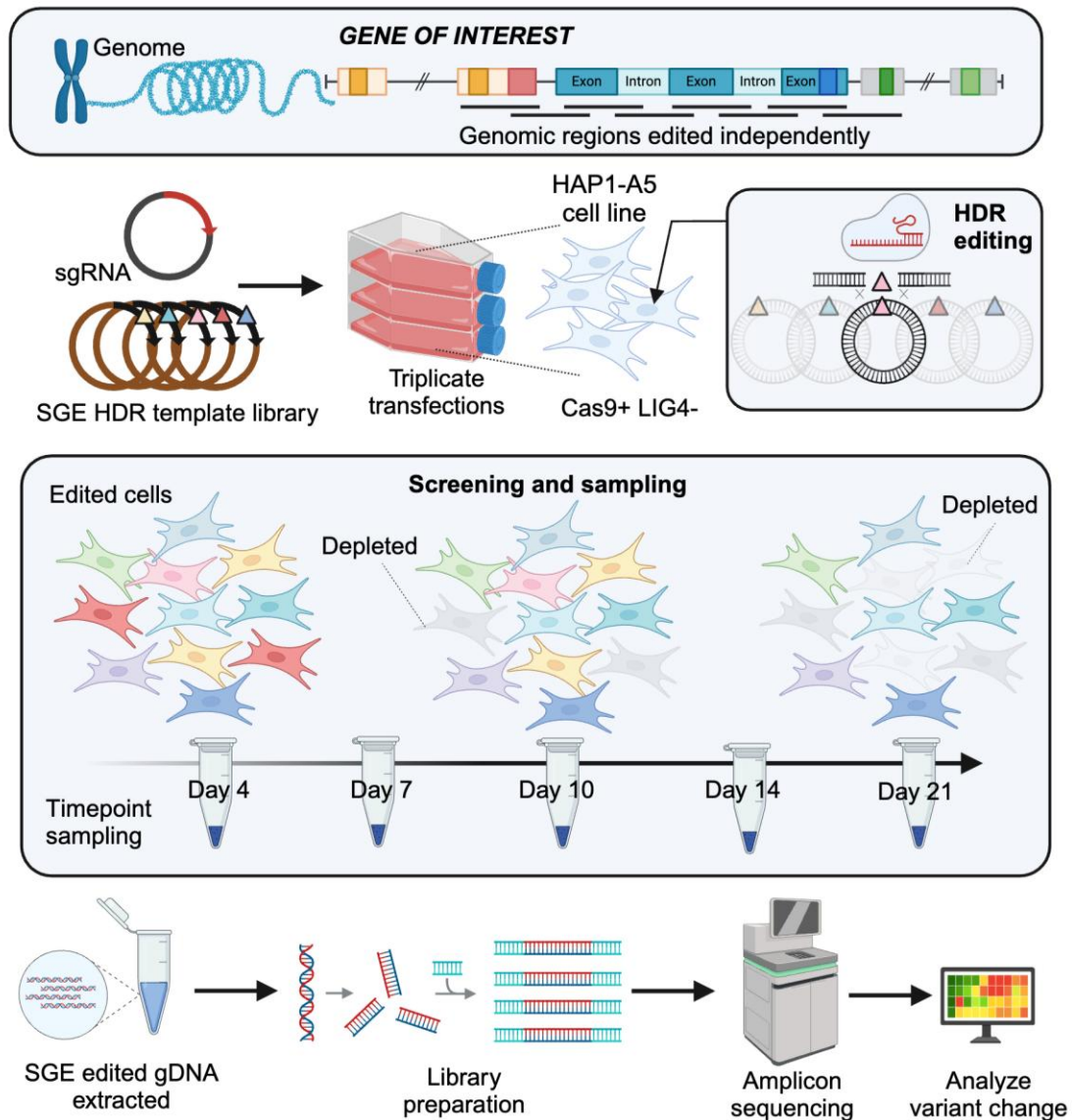
In the following protocol, we describe the laboratory experiments required to perform saturation genome editing (SGE) to study variant effects at scale on genes that are essential in HAP1-A5 cells (HZGHC-*LIG4*-Cas9). HAP1-A5 is an adherent, near-haploid cell line derived from the KBM-7 chronic myelogenous leukemia cell line<sup>4,5</sup>. It carries a DNA Ligase 4 (*LIG4*) gene knockout (KO) generated via a 10 base pair (bp) deletion within the gene, which biases DNA double-strand break repair away from non-homologous end joining (NHEJ) towards the desired homology-directed repair (HDR). HAP1-A5 cells also have a genomic Cas9 integration, resulting in high Cas9 activity and expression. This integration ensures high editing efficiency, as Cas9 activity and expression are not limiting<sup>2,6</sup>. A key advantage of HAP1-A5 cells is their stable maintenance of haploidy<sup>2</sup>, which is important for cell fitness readouts in functional genome editing screens. Haploidy at a locus allows recessive phenotypes to manifest with single-allele editing, reducing confounding wild-type allele interference. This is particularly appropriate for variants with loss-of-function (LoF) mechanisms, which include tumor suppressor genes<sup>7,8</sup>. However, HAP1 cells can increase in ploidy with prolonged cell culture<sup>1,9</sup>, leading to greater variability at later timepoints. HAP1-A5 cells sorted for high haploidy exhibit minimal haploidy loss (<3% between thawing and editing, <5% between editing and the final passage in a three-week SGE screen<sup>9</sup>), which is tolerable and unlikely to introduce appreciable noise into screens. Fluorescence-activated cell sorting (FACS) is a critical step to confirm that the cell stocks are haploid<sup>2,8</sup>. FACS analysis of thawed stocks is also sensible before conducting a large SGE screen.

**CRITICAL:** Prior to following the protocol described below, it is important to establish the essentiality of the target gene within HAP1-A5 cells. This can be achieved via CRISPR-Cas9-mediated KO of the target gene, followed by cell counting or colony assays, or by flow cytometry of cells stained with annexin-V and/or DAPI, to assess cellular apoptosis and/or necrosis. Additionally, a dataset<sup>10</sup> can be used to predict gene essentiality within HAP1 cells, the parental line of HAP1-A5 cells; a Bayes Factor of >6 is considered to have a probability of essentiality of 90%<sup>7,10</sup>.

An overview of the SGE experimental process is illustrated in **Figure 1**. As a first step in experimental design, an SGE variant oligonucleotide library is designed together with a corresponding single guide RNA (sgRNA). The SGE variant oligonucleotide library (when cloned into a target-region specific homology region extending plasmid, see below) acts as a repair template, incorporating saturating levels of variation or other desired allele types within target regions of  $\leq 245$  bp; each oligonucleotide making up the library contains a discrete

variant to be installed into the genome, in addition to synonymous changes at the sgRNA binding site to prevent cleavage of already-installed tracts (we call these fixed changes PAM/protospacer protection edits (PPEs)). SGE variant oligonucleotide library sequences for multiple target regions are then chemically synthesized *en masse* in an oligonucleotide pool, which is optionally amplified *en masse* (using **Primer Set 1**) to increase starting material and then specific target regions within the oligonucleotide pool are amplified using target region-specific **Primer Set 2 (key resources table)**, producing 'target specific SGE variant oligonucleotide libraries'. As a second step in this process, the wild-type homology regions of ~750-1000 bp flanking 5' and 3' of the variant containing the target region are generated using **Primer Sets 3 and 4 (key resources table)**. Subsequently, the wild-type homology region vectors are linearized using **Primer Sets 5 (key resources table)** and ligated with the target region-specific SGE variant oligonucleotide libraries, resulting in the SGE HDR template libraries.

In addition, appropriate sgRNAs are designed and cloned into a vector containing ampicillin and puromycin-resistant cassettes, for selection in *E. coli* and human cells, respectively. The genome for each target region is edited in triplicate by transfecting HAP1-A5 cells with an SGE HDR template library and the corresponding sgRNA, followed by puromycin selection to generate a population of cells, each containing a discrete variant. For genes that are essential in HAP1-A5, variants that compromise the gene function would be expected to be less represented within an edited cell population over time, as these variants result in a loss of fitness. Therefore, cells are cultured and sampled over a time course, with the cell fitness effect of each individual variant analyzed using deep amplicon next-generation sequencing (NGS) libraries. Depending on the kinetics of fitness/degree of essentiality in the target gene and the dynamics of variant effects, edited cells are cultured for a total of 14 or 21 days, on average. Day 4 is the baseline (first) timepoint collected, with additional timepoints collected between the baseline and terminal samples, which allows for variant kinetics to be calculated. To maintain a good representation of SGE variant installation complexity, 5–6 million cells are collected for each replicate timepoint. Genomic DNA (gDNA) is then extracted from timepoint-replicates, and SGE-edited gDNA is converted to NGS libraries using **Primer Sets 6 – 8 (key resources table)**. These libraries are then sequenced and analyzed to assess the cell fitness effect of the variants, followed by downstream analyses.



**Figure 1: Overview of the complete SGE process workflow.**

In a saturation genome editing (SGE) experiment, each region to be edited in a specific gene is targeted using a single guide RNA (sgRNA) in combination with an SGE homology-directed repair (HDR) template library in HAP1 cells deficient in DNA Ligase 4 (*LIG4*) and expressing endogenous Cas9 (HAP1-A5 cell line). Cas9 induces a double-strand break, which is repaired by the SGE HDR template libraries. Transfections are performed in triplicate, and genomic DNA (gDNA) samples are collected at varying timepoints depending on the screen length, typically at days 4, 7, 10, 14 and 21. This approach enables functional characterization through variant depletion kinetics, as deleterious alleles in essential loci are expected to decrease over time. The gDNA is then processed for edited gDNA library preparation and sequencing to generate functional scores. This figure is adapted from Olvera-León *et al.*<sup>3</sup>, with permission obtained.

## SOURCE OF *LIG4*- HAP1-A5 CELLS (ENDOGENOUSLY EXPRESSING CRISPR-CAS9)

The cell line used in the SGE experiment described below can be obtained from Horizon Discovery/Revvity (HAP1-A5 cell line (*LIG4*- Cas9+ clonal line) (HZGHC-*LIG4*-Cas9)).

## INSTITUTIONAL PERMISSIONS

All experiments were conducted with permission from the Wellcome Sanger Institute. Researchers will need to acquire permission from their institution prior to performing these experiments.

## SGE VARIANT OLIGONUCLEOTIDE LIBRARY DESIGN

**Timing: [1 week design, vendor-dependent time for ordering SGE variant oligonucleotide library pool]**

This section briefly outlines the SGE variant oligonucleotide library design and generation using the software VaLiAnT<sup>11</sup>.

### 1. Design the SGE variant oligonucleotide library using VaLiAnT<sup>11</sup>:

- a. When editing multiple regions in a gene with SGE, target each region separately.
- b. A target region typically includes a coding sequence exon and adjacent intronic or untranslated regions (UTRs).

**Note:** Due to the ~300 bp limit of high-quality oligonucleotide synthesis (with a maximum of ~245 bp reserved for variant edits), larger exons/SGE target regions should be divided into overlapping segments for full SGE coverage.

- c. Define variant-free regions at the target region ends for adding Illumina sequences and to enable targeted PCR amplification from the oligonucleotide pool (one of the reasons why a maximum of ~245 bp contains variants within a total of ~300 bp the defined target region in 1.b. above).

**Note:** Variant-free regions should ideally lack polymorphisms or HAP1-A5 specific variants.

- d. Select the desired VaLiAnT<sup>11</sup> 'mutator' functions for each target region.

**Note:** Libraries may include SNVs, in-frame codon deletions, alanine and stop-codon scans, all possible missense changes (based on a default or user-defined codon usage table), 1 bp deletions, and tandem deletions for splice-site scanning. Custom variants from databases like ClinVar<sup>12</sup> and gnomAD<sup>13</sup>, or genome studies, or user-defined indels can also be added as Variant Call Format (VCF) files.

- e. Provide input files, including the genomic coordinates of SGE target regions, gene transfer format (GTF) coding sequence (CDS) annotations, and reference genome sequence into the VaLiAnT software.
- f. Synthesize oligonucleotide pools via commercial vendors, such as Twist Bioscience.
- g. Resuspend the oligonucleotide pool to 10 ng/μL, following the manufacturer's instructions.
  - i. For Twist Bioscience, briefly centrifuge the tube and resuspend in nuclease-free Tris-EDTA (TE) Buffer, pH 8.0.

**Note:** Resuspended oligonucleotide pools can be stored at -20°C for long-term storage.

## SGRNA DESIGN

**Timing: [1–2 days, depending on number of target regions]**

This section outlines the sgRNA design process<sup>11</sup> for SGE experiments.

### 2. Design SGE sgRNAs

- a. For each SGE HDR template library, select a corresponding sgRNA that targets the specific genomic region for editing.

**Note:** Utilize previously reported<sup>11</sup> selection criteria.

- b. Incorporate synonymous PPEs into each SGE HDR template library target region within the sgRNA binding sites to enhance HDR efficiency; specify PPE coordinates and changes in VCF format for VaLiAnT usage.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE                                            | SOURCE                   | IDENTIFIER                            |
|----------------------------------------------------------------|--------------------------|---------------------------------------|
| <b>Bacterial and virus strains</b>                             |                          |                                       |
| One Shot TOP10 chemically competent <i>E. coli</i>             | Thermo Fisher Scientific | Cat#C404010                           |
| Endura competent cells                                         | Biosearch Technologies   | Item ID: 60242-2                      |
| <b>Chemicals, peptides, and recombinant proteins</b>           |                          |                                       |
| IDTE pH 8.0 (1X TE solution)                                   | IDT                      | Cat#11-01-02-05                       |
| Exonuclease I (20 U/μL)                                        | Thermo Fisher Scientific | Cat#EN0582                            |
| DpnI (20,000 U/mL)                                             | NEB                      | Cat#R0176S                            |
| rCutSmart buffer (10X)                                         | NEB                      | Cat#B6004S                            |
| NEBuilder HiFi DNA assembly master mix (2X)                    | NEB                      | Cat#E2621S;<br>Comp#M5520AVIAL        |
| BbsI-HF (20,000 U/mL)                                          | NEB                      | Cat#R3539S                            |
| T4 polynucleotide kinase (10,000 U/mL)                         | NEB                      | Cat#M0201L                            |
| T4 polynucleotide kinase reaction buffer (10X)                 | NEB                      | Cat#B0201S                            |
| Quick ligase                                                   | NEB                      | Cat#M2200S                            |
| Quick Ligation reaction buffer (2X)                            | NEB                      | Cat#B2200S                            |
| KAPA HiFi HotStart ReadyMix (2X)                               | Roche                    | Kit code: KK2601;<br>Mat.#07958927001 |
| ROX reference dye (%)                                          | Invitrogen               | Cat#12223012                          |
| EvaGreen dye, 20X in water                                     | Biotium                  | Cat#31000-T                           |
| UltraPure agarose                                              | Thermo Fisher Scientific | Cat#16500500                          |
| SYBR safe DNA gel stain                                        | Thermo Fisher Scientific | Cat#S33102                            |
| DNA gel loading dye (6X)                                       | Thermo Fisher Scientific | Cat#R0611                             |
| HyperLadder 1 kb                                               | Meridian Bioscience      | Cat#BIO-33053                         |
| S.O.C. (super optimal broth with catabolite repression) medium | Invitrogen               | Cat#15544034                          |
| Ampicillin (100 mg/mL)                                         | Merck                    | A5354; Cas#69-52-3                    |
| Recovery medium, 8 × 12 mL                                     | Biosearch Technologies   | Item ID: 80026-1                      |
| repliQa HiFi ToughMix                                          | Quantabio                | Mat.#95200-025                        |

Table continued.

| REAGENT or RESOURCE                             | SOURCE                                 | IDENTIFIER                                                                                |
|-------------------------------------------------|----------------------------------------|-------------------------------------------------------------------------------------------|
| Xfect transfection reagent                      | Takara                                 | Cat#631318                                                                                |
| Trypan blue solution, 0.4%                      | Gibco                                  | Cat#15250061                                                                              |
| Ampure XP                                       | Beckman Coulter                        | Cat#A63882                                                                                |
| Iscove's modified Dulbecco's medium (IMDM) (1X) | Thermo Fisher Scientific               | Cat#12440-053                                                                             |
| Fetal bovine serum (FBS)                        | Thermo Fisher Scientific               | Cat#10437-028                                                                             |
| Dulbecco's phosphate-buffered saline (DPBS)     | Sigma-Aldrich                          | 14190144; Cas#38249996                                                                    |
| Penicillin-streptomycin (5,000 U/mL)            | Gibco                                  | Cat#15070-063                                                                             |
| Blasticidin (10 mg/mL)                          | InvivoGen                              | Cat#ant-bl-05                                                                             |
| Puromycin (10 mg/mL)                            | InvivoGen                              | Cat#ant-pr-1                                                                              |
| TrypLE express enzyme (1X), phenol red          | Gibco                                  | Cat#12605010                                                                              |
| 0.05% trypsin-EDTA (1x)                         | Gibco                                  | Cat#25300-054                                                                             |
| pmaxGFP vector                                  | NovoPro                                | Cat#V014332                                                                               |
| Sodium acetate buffer solution                  | Thermo Fisher Scientific               | Cat#S7899-100ML                                                                           |
| Ampicillin (100 µg/mL) agar plates              | N/A                                    | N/A                                                                                       |
| Lysogenic broth (LB) media                      | N/A                                    | N/A                                                                                       |
| 50% glycerol                                    | N/A                                    | N/A                                                                                       |
| Ethanol absolute (100%)                         | N/A                                    | N/A                                                                                       |
| <b>Critical commercial assays</b>               |                                        |                                                                                           |
| QIAquick PCR purification Kit                   | QIAGEN                                 | Cat#28106                                                                                 |
| MinElute PCR purification kit                   | QIAGEN                                 | Cat#28006                                                                                 |
| QIAprep spin miniprep kit                       | QIAGEN                                 | Cat#27106                                                                                 |
| QIAquick gel extraction kit                     | QIAGEN                                 | Cat#28704                                                                                 |
| Plasmid plus maxi kit                           | QIAGEN                                 | Cat#12963                                                                                 |
| DNeasy blood & tissue kit                       | QIAGEN                                 | Cat#69506                                                                                 |
| <b>Deposited data</b>                           |                                        |                                                                                           |
| ClinVar                                         | Landrum <i>et al.</i> <sup>12</sup>    | <a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a> |
| gnomAD                                          | Karczewski <i>et al.</i> <sup>13</sup> | <a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a>       |

Table continued.

| REAGENT or RESOURCE                                                                                                                     | SOURCE                              | IDENTIFIER                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Experimental models: Cell lines</b>                                                                                                  |                                     |                                                                                                                                                                                         |
| HAP1 cell line ( <i>LIG4</i> -)                                                                                                         | Horizon Discovery                   | HZGHC- <i>LIG4</i>                                                                                                                                                                      |
| HAP1-A5 cell line ( <i>LIG4</i> - Cas9+ clonal line)                                                                                    | Horizon Discovery                   | HZGHC- <i>LIG4</i> -Cas9                                                                                                                                                                |
| <b>Oligonucleotides</b>                                                                                                                 |                                     |                                                                                                                                                                                         |
| sgRNA oligos                                                                                                                            | Sigma-Aldrich (100 µM in TE buffer) | N/A                                                                                                                                                                                     |
| SGE variant oligonucleotide library pool                                                                                                | Twist Bioscience                    | Lyophilized                                                                                                                                                                             |
| Primer Set 1<br>Generic Illumina adaptor sequences used for SGE variant oligonucleotide library pool amplification (300–350 bp product) | Sigma-Aldrich                       | Sequence:<br>Illumina_P5:<br>AATGATACGGCGACCACCG<br>A<br>Illumina_P7:<br>CAAGCAGAAGACGGCATAC<br>GA                                                                                      |
| Primer Set 2<br>Target region-specific primers (~245–300 bp product)                                                                    | Sigma-Aldrich                       | Target-region dependent                                                                                                                                                                 |
| Primer Set 3<br>Primers used to generate the SGE-pMin backbone (1831 bp product)                                                        | Sigma-Aldrich                       | Sequence:<br>SGE_pMin_backbone_F:<br>GCGGCCGCCGTCAGGTGG<br>SGE_pMin_backbone_R:<br>CCTGCAGGACATGTGAGCA<br>AAAGGCCA                                                                      |
| Primer Set 4<br>Primers used for wild-type region amplification (~1,200–2,000 bp product)                                               | Sigma-Aldrich                       | Sequence: homology_region_F:<br>TTTTGCTCACATGTCCTGCA<br>GG+ [REGION SPECIFIC<br>SEQUENCE]<br>homology_region_R:<br>GAAAAGTGCCACCTGACG<br>GCGGCCGC+ [REGION<br>SPECIFIC SEQUENCE]        |
| Primer Set 5<br>Primers used to linearize wild-type homology region plasmid (~245–300 bp product).                                      | Sigma-Aldrich                       | Sequence:<br>Linearization_F: Reverse<br>complement sequence<br>of <b>Primer Set 2</b> forward<br>Linearization_R: Reverse<br>complement sequence<br>of <b>Primer Set 2</b> reverse     |
| Primer Set 6<br>Primers for sampling/amplification of edited gDNA regions (~1,700–2,000 bp product)                                     | Sigma-Aldrich                       | Target-region dependent                                                                                                                                                                 |
| Primer Set 7<br>Primers for sequencing adapter addition/amplification of edited gDNA regions (~245–300 bp product)                      | Sigma-Aldrich                       | Sequence:<br>Sequencing_F: Indexing<br>barcode dependent<br>adaptors + ' <b>Primer Set 2_F</b> '<br>Sequencing_R: Indexing<br>barcode dependent<br>adaptors + ' <b>Primer Set 2_R</b> ' |
| Primer Set 8<br>Indexing primers (~350 bp product).                                                                                     | Integrated DNA Technologies         | N/A                                                                                                                                                                                     |

Table continued.

| REAGENT or RESOURCE                                                    | SOURCE                             | IDENTIFIER                                                                            |
|------------------------------------------------------------------------|------------------------------------|---------------------------------------------------------------------------------------|
| cDNA sequences for colony formation assay                              | GeneArt                            | N/A                                                                                   |
| <b>Recombinant DNA</b>                                                 |                                    |                                                                                       |
| pMin-U6-ccdb-hPGK-puro (pMin)                                          | Radford <i>et al.</i> <sup>9</sup> | N/A                                                                                   |
| <b>Software and algorithms</b>                                         |                                    |                                                                                       |
| ValiAnT                                                                | Barbon <i>et al.</i> <sup>11</sup> | <a href="https://github.com/cancerit/ValiAnT">https://github.com/cancerit/ValiAnT</a> |
| <b>Other</b>                                                           |                                    |                                                                                       |
| Thermocycler                                                           | N/A                                | N/A                                                                                   |
| Optional: quantitative PCR instrument                                  | N/A                                | N/A                                                                                   |
| 1.5 mL Eppendorf tubes                                                 | N/A                                | N/A                                                                                   |
| 15 mL conical Falcon tubes                                             | N/A                                | N/A                                                                                   |
| 50 mL conical Falcon tubes                                             | N/A                                | N/A                                                                                   |
| 50 mL conical Erlenmeyer flasks                                        | N/A                                | N/A                                                                                   |
| 500 mL storage bottle (sterile)                                        | N/A                                | N/A                                                                                   |
| T150 cell culture flasks                                               | N/A                                | N/A                                                                                   |
| T75 cell culture flasks                                                | N/A                                | N/A                                                                                   |
| 5, 10, 25, and 50 serological pipettes                                 | N/A                                | N/A                                                                                   |
| Waterbath (for bacterial work & sterile for cell culture work)         | N/A                                | N/A                                                                                   |
| Centrifuge                                                             | N/A                                | N/A                                                                                   |
| Vortex                                                                 | N/A                                | N/A                                                                                   |
| Optional: Countess 3 automated cell counter                            | Thermo Fisher Scientific           | N/A                                                                                   |
| Optional: Countess cell counting chamber slides and holder, disposable | Thermo Fisher Scientific           | Cat#C10312                                                                            |
| L-shaped spreaders                                                     | Medline Scientific                 | Cat#MLSP25                                                                            |
| Inoculating loops for microbiology                                     | N/A                                | N/A                                                                                   |
| 5% CO <sub>2</sub> incubator                                           |                                    |                                                                                       |
| 50 mL reagent reservoir                                                | N/A                                | N/A                                                                                   |
| 96-well cell culture cluster (flat bottom)                             | N/A                                | N/A                                                                                   |
| DynaMag-96 bottom magnet                                               | Thermo Fisher Scientific           | Cat#12332D                                                                            |

Table continued.

| REAGENT or RESOURCE                                                    | SOURCE  | IDENTIFIER      |
|------------------------------------------------------------------------|---------|-----------------|
| Optional: Gene Pulser Xcell electroporation systems                    | Bio-Rad | Cat#1652660     |
| Optional: Gene Pulser/MicroPulser electroporation cuvettes, 0.1 cm gap | Bio-Rad | Cat#1652083     |
| Optional: GelDoc Go imaging system                                     | Bio-Rad | Cat#12009077EDU |
| Optional: NanoDrop                                                     | N/A     | N/A             |
| (Shaking) incubator                                                    | N/A     | N/A             |

## MATERIALS AND EQUIPMENT

### Reaction setup for ExoI enzymatic cleanup of the PCR product of Primer Set 1

| Reagent                     | Final concentration | Amount |
|-----------------------------|---------------------|--------|
| PCR product of Primer Set 1 | N/A                 | 50 µL  |
| Exonuclease I (20 U/µL)     | 1 U/µL              | 5 µL   |
| 10X Reaction Buffer         | 1 X                 | 10 µL  |
| ddH <sub>2</sub> O          | N/A                 | 35 µL  |
| Total                       | N/A                 | 100 µL |

Keep on ice (0°C–4°C) during use, prepare fresh, and use immediately; do not store.

### Reaction setup for DpnI digestion of the PCR product of Primer Set 3

| Reagent                              | Final concentration | Amount |
|--------------------------------------|---------------------|--------|
| PCR product of Primer Set 3 template | N/A                 | 25 µL  |
| DpnI (20,000 U/mL)                   | 0.4 U/µL            | 1 µL   |
| rCutSmart Buffer (10 X)              | 1 X                 | 5 µL   |
| ddH <sub>2</sub> O                   | N/A                 | 19 µl  |
| Total                                | N/A                 | 50 µL  |

Keep on ice (0°C–4°C) during use, prepare fresh, and use immediately; do not store.

**Reaction setup for ligation of wild-type homology region plasmids**

| Reagent                                     | Final concentration | Amount                 |
|---------------------------------------------|---------------------|------------------------|
| PCR product of Primer Set 3                 | 50 ng               | X $\mu$ L              |
| PCR product of Primer Set 4                 | 50 ng               | X $\mu$ L              |
| NEBuilder HiFi DNA Assembly Master Mix (2X) | 1 X                 | 10 $\mu$ L             |
| ddH <sub>2</sub> O                          | N/A                 | Complete to 20 $\mu$ L |
| Total                                       | N/A                 | 20 $\mu$ L             |

Store at  $-20^{\circ}\text{C}$  for up to 1 year.

**Reaction setup for DpnI digestion of the PCR product of Primer Set 5**

| Reagent                     | Final concentration | Amount     |
|-----------------------------|---------------------|------------|
| PCR product of Primer Set 5 | N/A                 | 25 $\mu$ L |
| DpnI (20,000 U/mL)          | 0.4 U/ $\mu$ L      | 1 $\mu$ L  |
| rCutSmart Buffer (10 X)     | 1 X                 | 5 $\mu$ L  |
| ddH <sub>2</sub> O          | N/A                 | 19 $\mu$ L |
| Total                       | N/A                 | 50 $\mu$ L |

Keep on ice ( $0^{\circ}\text{C}$ – $4^{\circ}\text{C}$ ) during use, prepare fresh, and use immediately; do not store.

**Reaction setup for ligation of linearized wild-type homology regions with target-specific SGE variant oligonucleotide library**

| Reagent                                                                                  | Final concentration | Amount                 |
|------------------------------------------------------------------------------------------|---------------------|------------------------|
| Target region-specific SGE variant oligonucleotide library (PCR product of Primer Set 2) | 50 ng               | X $\mu$ L              |
| Linearized wild-type homology regions (PCR product of Primer Set 5)                      | 50 ng               | X $\mu$ L              |
| NEBuilder HiFi DNA Assembly Master Mix (2X)                                              | 1 X                 | 10 $\mu$ L             |
| ddH <sub>2</sub> O                                                                       | N/A                 | Complete to 20 $\mu$ L |
| Total                                                                                    | N/A                 | 20 $\mu$ L             |

Store at  $-20^{\circ}\text{C}$  for up to 1 year.

**Optional negative control: Reaction setup for ligation of the linearized wild-type homology regions (PCR product of Primer Set 5)**

| Reagent                                                             | Final concentration | Amount                 |
|---------------------------------------------------------------------|---------------------|------------------------|
| linearized wild-type homology regions (PCR product of Primer Set 5) | 50 ng               | X $\mu$ L              |
| NEBuilder HiFi DNA Assembly Master Mix (2X)                         | 1 X                 | 10 $\mu$ L             |
| ddH <sub>2</sub> O                                                  | N/A                 | Complete to 20 $\mu$ L |
| Total                                                               | N/A                 | 20 $\mu$ L             |

Keep on ice (0°C–4°C) during use, prepare fresh, and use immediately; do not store.

**Digestion setup for DpnI-digested SGE-pMin backbone**

| Reagent                       | Final concentration | Amount       |
|-------------------------------|---------------------|--------------|
| pMin-U6-ccdb-hPGK-puro (pMin) | 1 $\mu$ g           | 2.5 $\mu$ L  |
| BbsI-HF (20,000 U/ml)         | 3.33 U/ $\mu$ L     | 5 $\mu$ L    |
| rCutSmart Buffer (10 X)       | 1 X                 | 3 $\mu$ L    |
| ddH <sub>2</sub> O            | N/A                 | 19.5 $\mu$ L |
| Total Volume                  | N/A                 | 30 $\mu$ L   |

Store at –20°C for up to 3 years.

**Reaction setup for phosphorylation and annealing of complementary sgRNA oligos**

| Reagent                                         | Final concentration | Amount      |
|-------------------------------------------------|---------------------|-------------|
| Oligo 1 (100 $\mu$ M)                           | 10 $\mu$ M          | 1 $\mu$ L   |
| Oligo 2 (100 $\mu$ M)                           | 10 $\mu$ M          | 1 $\mu$ L   |
| T4 Polynucleotide Kinase Reaction Buffer (10 X) | 1 X                 | 1 $\mu$ L   |
| T4 Polynucleotide Kinase (10,000 U/ml)          | 0.5 U/ $\mu$ L      | 0.5 $\mu$ L |
| ddH <sub>2</sub> O                              | N/A                 | 6.5 $\mu$ L |
| Total Volume                                    | N/A                 | 10 $\mu$ L  |

Store at –20°C for up to 1 year.

**Reaction setup for ligation of sgRNA oligos with pMin\_gRNA\_backbone**

| Reagent                                                   | Final concentration | Amount                 |
|-----------------------------------------------------------|---------------------|------------------------|
| gRNA_pMin_backbone (BbsI-digested)                        | 50 ng               | X $\mu$ L              |
| Phosphorylated and annealed oligo duplex (1:200 dilution) | N/A                 | 1 $\mu$ L              |
| Quick Ligation Reaction Buffer (2 X)                      | 0.5 X               | 5 $\mu$ L              |
| Quick Ligase                                              | N/A                 | 0.8 $\mu$ L            |
| ddH <sub>2</sub> O                                        | N/A                 | Complete to 10 $\mu$ L |
| Total Volume                                              | N/A                 | 10 $\mu$ L             |

Store at  $-20^{\circ}\text{C}$  for up to 1 year.

**Cas9-selection media for HAP1-A5 cell line ('B+ media')**

| Reagent                                         | Final concentration | Amount      |
|-------------------------------------------------|---------------------|-------------|
| Iscove's Modified Dulbecco's Medium (IMDM) (1X) | N/A                 | 500 mL      |
| Fetal Bovine Serum (FBS)                        | 10%                 | 50 mL       |
| Penicillin-Streptomycin (5,000 U/mL)            | 1%                  | 5.5 mL      |
| Blasticidin (10 mg/mL)                          | 10 $\mu$ g/mL       | 550 $\mu$ L |
| Total                                           | N/A                 | 556.05 mL   |

Store at  $4^{\circ}\text{C}$  for up to 1 month. Before use, warm the medium to  $37^{\circ}\text{C}$  in a water bath.

**Growth media for HAP1-A5 cell line ('noAB media')**

| Reagent                                         | Final concentration | Amount   |
|-------------------------------------------------|---------------------|----------|
| Iscove's Modified Dulbecco's Medium (IMDM) (1X) | N/A                 | 500 mL   |
| Fetal Bovine Serum (FBS)                        | 10%                 | 50 mL    |
| Penicillin-Streptomycin (5,000 U/mL)            | 1%                  | 5.5 mL   |
| Total                                           | N/A                 | 555.5 mL |

Store at  $4^{\circ}\text{C}$  for up to 1 month. Before use, warm the medium to  $37^{\circ}\text{C}$  in a water bath.

**sgRNA selection media for HAP1-A5 cell line ('B+P+ media')**

| Reagent                                         | Final concentration | Amount     |
|-------------------------------------------------|---------------------|------------|
| Iscove's Modified Dulbecco's Medium (IMDM) (1X) | N/A                 | 500 mL     |
| Fetal Bovine Serum (FBS)                        | 10%                 | 50 mL      |
| Penicillin-Streptomycin (5,000 U/mL)            | 1%                  | 5.5 mL     |
| Blasticidin (10 mg/mL)                          | 10 µg/mL            | 550 µL     |
| Puromycin (10 mg/mL)                            | 3 µg/mL             | 165 µL     |
| Total                                           | N/A                 | 556.215 mL |

Store at 4°C for up to 1 month. Before use, warm the medium to 37°C in a water bath.

**Reaction setup for ExoI enzymatic cleanup of the PCR product of Primer Set 6**

| Reagent                     | Final concentration | Amount |
|-----------------------------|---------------------|--------|
| PCR product of Primer Set 6 | N/A                 | 40 µL  |
| Exonuclease I (20 U/µL)     | 1 U/µL              | 5 µL   |
| 10X Reaction Buffer         | 1 X                 | 10 µL  |
| ddH <sub>2</sub> O          | N/A                 | 45 µL  |
| Total                       | N/A                 | 100 µL |

Keep on ice (0°C–4°C) during use, prepare fresh, and use immediately; do not store.

**Reaction setup for ExoI enzymatic cleanup of the PCR product of Primer Set 7**

| Reagent                     | Final concentration | Amount |
|-----------------------------|---------------------|--------|
| PCR product of Primer Set 7 | N/A                 | 30 µL  |
| Exonuclease I (20 U/µL)     | 0.61 U/µL           | 1 µL   |
| 10X Reaction Buffer         | 1 X                 | 3 µL   |
| Total                       | N/A                 | 33 µL  |

Keep on ice (0°C–4°C) during use, prepare fresh, and use immediately; do not store.

## STEP-BY-STEP METHOD DETAILS

### PCR OLIGONUCLEOTIDE POOL AMPLIFICATION USING PRIMER SET 1

**Timing: [1 day]**

In this step, a PCR is performed with the aim of amplifying the oligonucleotide pool in bulk (**Figure 1A**). For this PCR, Illumina adapter sequences are used as Primer Set 1 and are listed in the **key resources table** of this protocol. The Illumina adapter sequences are appended to all oligonucleotides to perform a generic amplification of the oligonucleotide pool *en masse*, to increase the starting oligonucleotide pool material for downstream cloning processes. Note, these Illumina adapter sequences are not used for sequencing in this step; they are used because they perform efficient, unbiased amplifications.

1. Amplify and increase the starting DNA concentration of the oligonucleotide pool via PCR. 3-4 reactions should be independently run and then pooled.
  - a. Prepare a PCR reaction master mix as indicated below.

#### PCR reaction master mix for amplification of the SGE variant oligonucleotide library—Using Primer Set 1

| Reagent                        | Amount                 |
|--------------------------------|------------------------|
| Oligonucleotide pool template  | X $\mu$ L (20 ng)      |
| Illumina_P5 (10 $\mu$ M)       | 1.5 $\mu$ L            |
| Illumina_P7 (10 $\mu$ M)       | 1.5 $\mu$ L            |
| KAPA Hifi HotStart ReadyMix 2x | 25 $\mu$ L             |
| ddH <sub>2</sub> O             | Complete to 50 $\mu$ L |

- b. Run a PCR following the conditions below.

**PCR cycling conditions for amplification of the oligonucleotide pool—Using Primer Set 1**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 12 cycles        |
| Annealing            | 60°C        | 15 s    |                  |
| Extension            | 72°C        | 30 s    |                  |
| Final extension      | 72°C        | 30 s    | 1                |
| Hold                 | 4°C         | forever |                  |

2. Perform an enzymatic PCR cleanup with Exonuclease (Exo) I (to remove ssDNA & primers).
  - a. Prepare a master mix, as described in the materials and equipment section under “Reaction setup for ExoI enzymatic cleanup of the PCR product of Primer Set 1”.
  - b. Vortex and mix the reaction.
  - c. In a thermocycler, incubate the reaction at 37°C for 20 min followed by 80°C for 20 min.
3. Purify the ExoI-digested PCR product(s) of Primer Set 1 via MinElute columns (QIAGEN), following the manufacturer’s instructions.
  - a. Elute the DNA in 10 µL EB Buffer. Pool independent reactions together.
  - b. Measure the DNA concentration of the purified PCR product(s) of Primer Set 1.

**Pause point:** The PCR product can be stored at –20°C for long-term storage.

**PRIMER DESIGN FOR SGE TARGET REGION AMPLIFICATION - PRIMER SET 2**

**Timing: [30 min, depending on number of target regions]**

This step describes the primer design for the second PCR of the protocol, which aims to amplify the target region-specific sequences from the SGE variant oligonucleotide library pool (product of Primer Set 1) (**Figure 2A**) using Primer Set 2 as shown in the **key resources table** of this protocol.

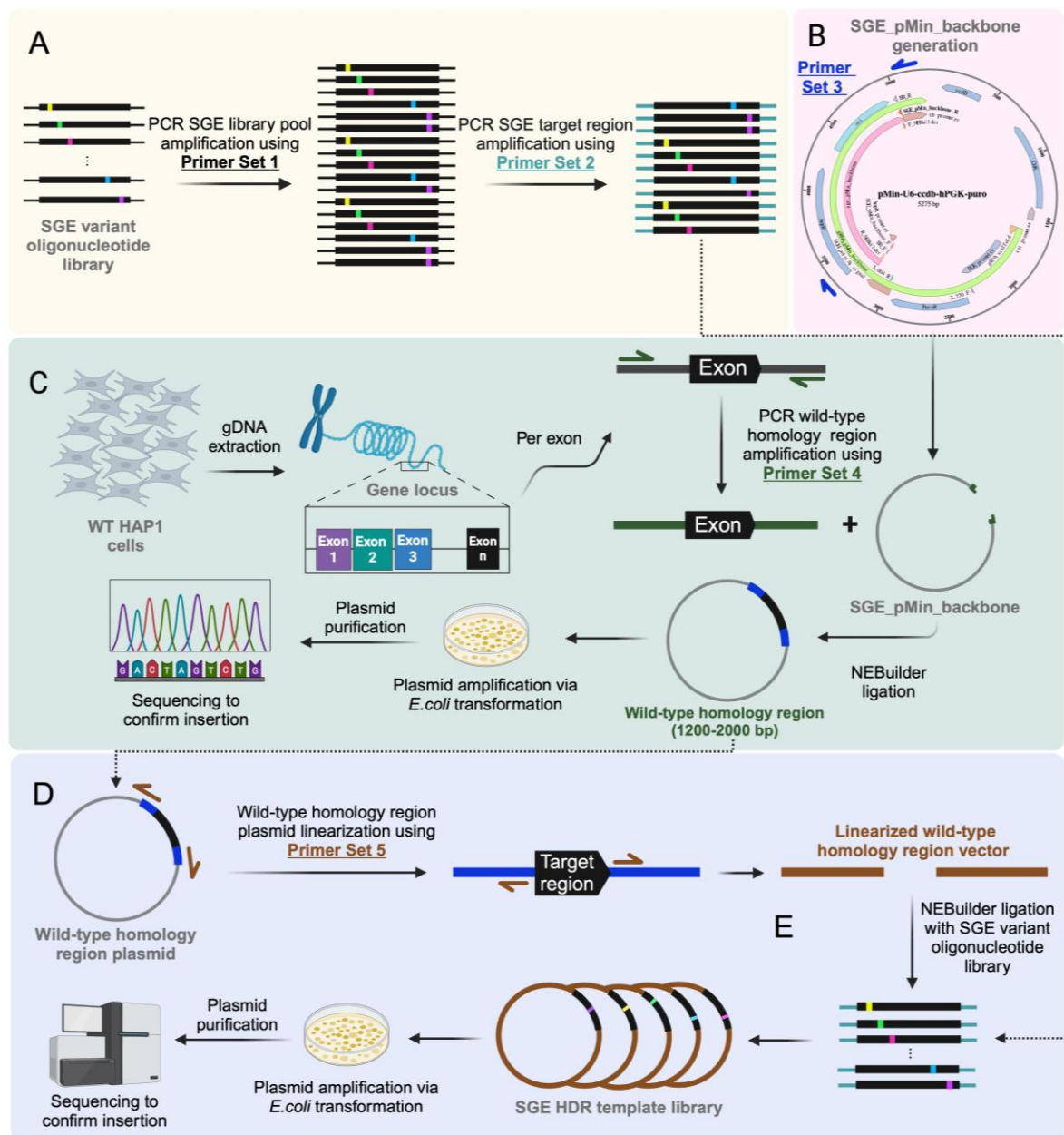
4. Design Primer Set 2 using any preferred primer design software.

**Note:** In this protocol, we used the 'Primer Design' function within Geneious Prime, a modified version of Primer 3.

- a. The primers are target region-specific and should define the boundaries of each target region.
- b. The expected product size should range from 245-300 bp.

**Note:** The total length should not exceed 300 bp, when including the Illumina P5/P7 adaptors, due to synthesis limitations ( $\leq 300$  bp).

- c. When designing primers for multiple target regions, maintain similar product sizes across all targets to ensure consistency and optimal amplification conditions



**Figure 2: Overview of SGE pre-screen process: Cloning SGE variant oligonucleotide libraries into linearized wild-type homology regions.**

**A)** The saturation genome editing (SGE) variant oligonucleotide library pool is PCR amplified using Primer Set 1 to increase the starting material of the SGE variant oligonucleotide library for subsequent cloning. Target region-specific SGE variant oligonucleotide libraries are generated by amplifying the SGE target region using 'Primer Set 2'. **B)** SGE\_pMin\_backbone is generated using Primer Set 3, which amplifies the promoter and coding sequence for ampicillin resistance as well as the origin of replication (ori) sequence from the pMin-U6-ccdb-hPGK-puro (pMin) plasmid. **C)** Wild-type homology arms are amplified from HAP1-A5 genomic DNA (gDNA) using Primer Set 4, with one set designed for each exon. The resulting PCR amplicons are then ligated into the SGE\_pMin\_backbone, which contains homology regions. The wild-type homology region plasmids are then amplified via *E. coli* transformation and verified by Sanger sequencing. **D)** Wild-type homology region plasmids are linearized using Primer Set 5, with specific primers for each target region. **E)** The resulting linearized vectors are ligated with the individual target region-specific SGE library oligos to create SGE homology-directed repair (HDR) template libraries. These libraries are amplified via *E. coli* transformation and verified by next-generation sequencing.

## PCR SGE TARGET REGION AMPLIFICATION USING PRIMER SET 2

**Timing:** [1–2 days, depending on number of target regions]

This step outlines the PCR protocol using Primer Set 2, as detailed in the **key resources table** of this protocol, to amplify specific target regions from the oligonucleotide pool (product of Primer Set 1) (**Figure 2A**). Prior to the PCR, we recommend performing an optional gradient PCR (e.g., at 58°C, 60°C, 63°C, and 65°C) to determine the optimal annealing temperature for each target region (steps not described) and an optional qPCR (outlined below) to determine the optimal number of amplification cycles, ensuring library complexity is maintained and avoiding over- or under-amplification. If the optional qPCR steps are skipped, proceed directly to step 8 of this protocol.

5. **Optional step:** Prepare a qPCR reaction master mix of the PCR product of Primer Set 1 using Primer Set 2 as indicated below.

**Note:** ROX dye is used as an internal reference; however, alternative dyes as well as varying dye concentrations can be used depending on the qPCR instrument employed. Individual reactions should be set up for each primer pair.

| qPCR reaction master mix for amplification of the PCR product of Primer Set 1—Using Primer Set 2 |                   |
|--------------------------------------------------------------------------------------------------|-------------------|
| Reagent                                                                                          | Amount            |
| PCR product of Primer Set 1 template                                                             | X µL (2 ng)       |
| target_region_F (10 µM)                                                                          | 0.45 µL           |
| target_region_R (10 µM)                                                                          | 0.45 µL           |
| KAPA Hifi Hotstart ReadyMix 2x                                                                   | 7.5 µL            |
| EvaGreen (5%)                                                                                    | 0.75 µL           |
| ROX (1%)                                                                                         | 0.3 µL            |
| ddH <sub>2</sub> O                                                                               | Complete to 15 µL |

6. **Optional step:** Run a qPCR following the conditions below.

qPCR cycling conditions for amplification of the PCR product of Primer Set 1—Using Primer Set 2

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 35 cycles        |
| Annealing            | 58°C–65°C   | 15 s    |                  |
| Extension            | 72°C        | 30 s    |                  |
| Final extension      | 72°C        | 30 s    | 1                |
| Hold                 | 4°C         | forever |                  |

7. **Optional step:** Analyze qPCR results.

- Identify the most optimal amplification cycle (to maintain library complexity and avoid over/under-amplification<sup>6,14–16</sup>).

**Note:** An example of a qPCR amplification of 4 target regions is indicated in **Figures 3A–3D**.

- To assess the purity of the qPCR products a melting curve analysis can be performed.
  - Run 5 µL of the qPCR product of Primer Set 2 and 2 µL of loading dye on a 2% agarose-TAE gel (30 min at 120 V) to check for correct fragment size (~245–300 bp), band intensity and potential non-specific bands that may have been co-amplified during the PCR (**Figure 3E**).
8. Extract and amplify 'target specific SGE variant oligonucleotide libraries' from the oligonucleotide pool via PCR.
- Prepare a PCR reaction master mix as indicated below.

**CRITICAL:** The PCR steps for each exon/target region are performed in separate experiments.

**Note:** Apply the identified amplification cycle number (via qPCR, optional) and annealing temperature (via gradient PCR, optional).

**qPCR reaction master mix for amplification of the PCR product of Primer Set 1—Using Primer Set 2**

| Reagent                              | Amount                 |
|--------------------------------------|------------------------|
| PCR product of Primer Set 1 template | X $\mu$ L (40 ng)      |
| target_region_F (10 $\mu$ M)         | 1.5 $\mu$ L            |
| target_region_R (10 $\mu$ M)         | 1.5 $\mu$ L            |
| KAPA Hifi Hotstart ReadyMix 2x       | 25 $\mu$ L             |
| ddH <sub>2</sub> O                   | Complete to 50 $\mu$ L |

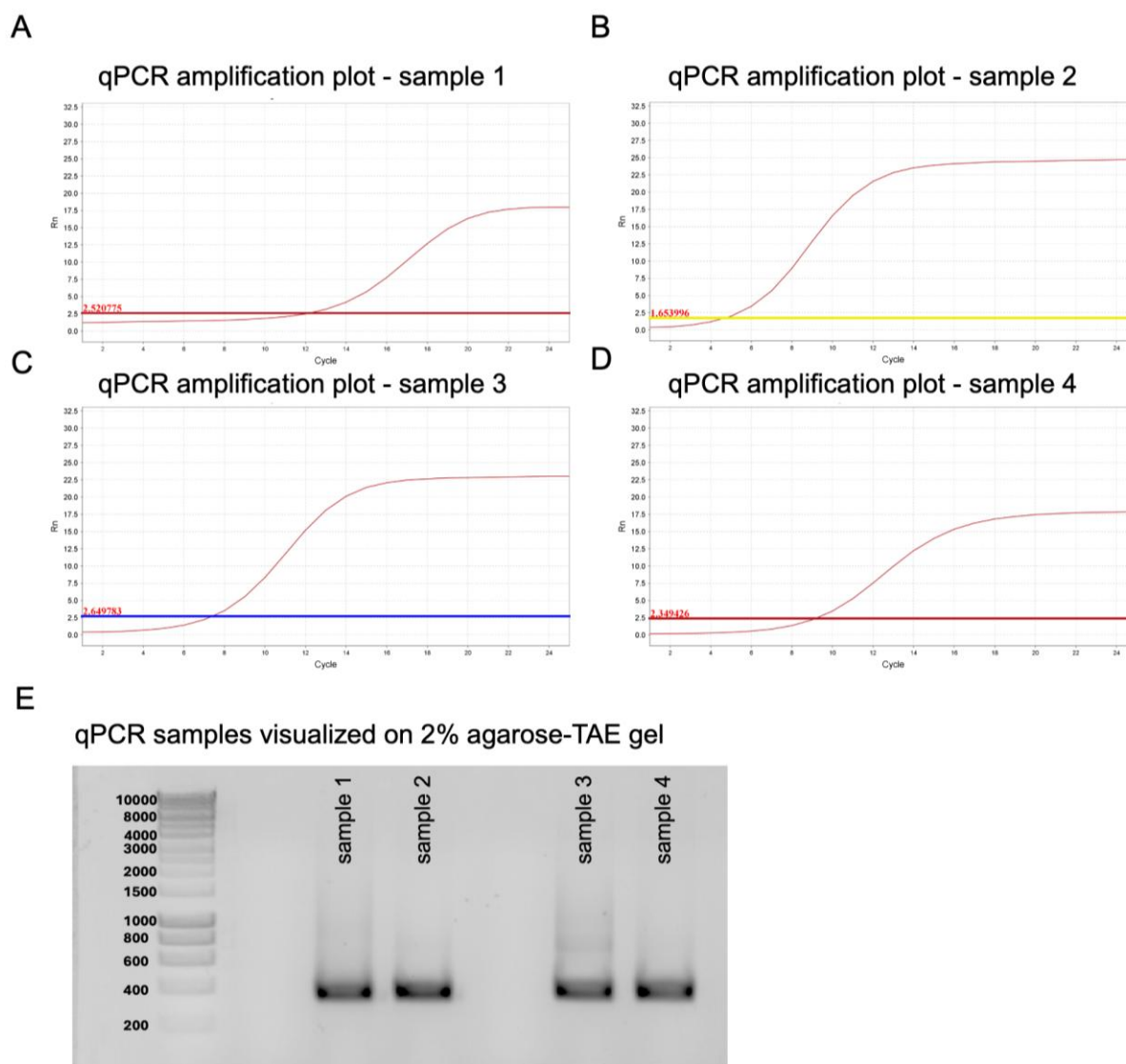
b. Run a PCR following the conditions below.

**qPCR cycling conditions for amplification of the PCR product of Primer Set 1—Using Primer Set 2**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 10-15 cycles     |
| Annealing            | 58°C–65°C   | 15 s    |                  |
| Extension            | 72°C        | 30 s    |                  |
| Final extension      | 72°C        | 30 s    | 1                |
| Hold                 | 4°C         | forever |                  |

9. Purify the PCR product(s) of Primer Set 2 via MinElute columns (QIAGEN), following the manufacturer's instructions.

a. Measure the DNA concentration of the purified PCR product(s) of Primer Set 2.



**Figure 3: Optimization of qPCR conditions to identify optimal amplification cycles for target-specific primers.**

**A–D)** qPCR amplification plots for samples 1–4, illustrating the exponential phase of amplification. The optimal cycle number is determined by identifying the point within the exponential phase prior to the onset of the non-exponential plateau phase. Based on the amplification curves, the optimal cycle numbers for maintaining saturation genome editing (SGE) variant oligonucleotide library complexity and avoiding over- or under-amplification are 19, 11, 13, and 15 cycles for samples 1, 2, 3, and 4, respectively, y-axis is normalized reporter (EvaGreen) fluorescence (Rn) and x-axis is cycle number, the threshold is also displayed by horizontal coloured line. **E)** Agarose gel electrophoresis (2% agarose in TAE buffer) of qPCR products after 35 cycles. The expected band size (~350 base pair (bp)) is observed with appropriate band intensity and no evidence of non-specific amplification.

## PCR FOR SGE\_PMIN\_BACKBONE GENERATION USING PRIMER SET 3

**Timing: [1 day]**

In the step below, we generate an SGE\_pMin\_backbone fragment, which can subsequently be used to create the wild-type homology region plasmids. Use Primer Set 3 here, as listed in

the **key resources table** of this protocol, which anneals to sequences within the pMin plasmid backbone (pMin-U6-ccdb-hPGK-puro (pMin))<sup>1</sup> and amplifies the backbone regions containing a promoter sequence for ampicillin resistance as well as an origin of replication (*ori*) sequence, 1831 bp in length (**Figures 2B and 4**).

10. Amplify the SGE\_pMin\_backbone region from the pMin plasmid backbone.

a. Prepare a PCR reaction master mix as indicated below.

**PCR reaction master mix for generation of SGE\_pMin backbone—Using Primer Set 3**

| Reagent                                | Amount            |
|----------------------------------------|-------------------|
| pMin plasmid (1:100 dilution of stock) | 1 µL              |
| SGE_pMin_backbone_F (10 µM)            | 0.75 µL           |
| SGE_pMin_backbone_R (10 µM)            | 0.75 µL           |
| KAPA HiFi Hotstart ReadyMix (2X)       | 12.5 µL           |
| ddH <sub>2</sub> O                     | Complete to 25 µL |

b. Run a PCR following the conditions below.

**PCR cycling conditions for generation of SGE\_pMin backbone—Using Primer Set 3**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 25 cycles        |
| Annealing            | 63°C        | 15 s    |                  |
| Extension            | 72°C        | 1 min   |                  |
| Final extension      | 72°C        | 1 min   | 1                |
| Hold                 | 4°C         | forever |                  |

11. Digest the PCR product of Primer Set 3 with DpnI restriction enzyme (to remove methylated bacterial DNA template, which prevents cloning contamination in subsequent steps).

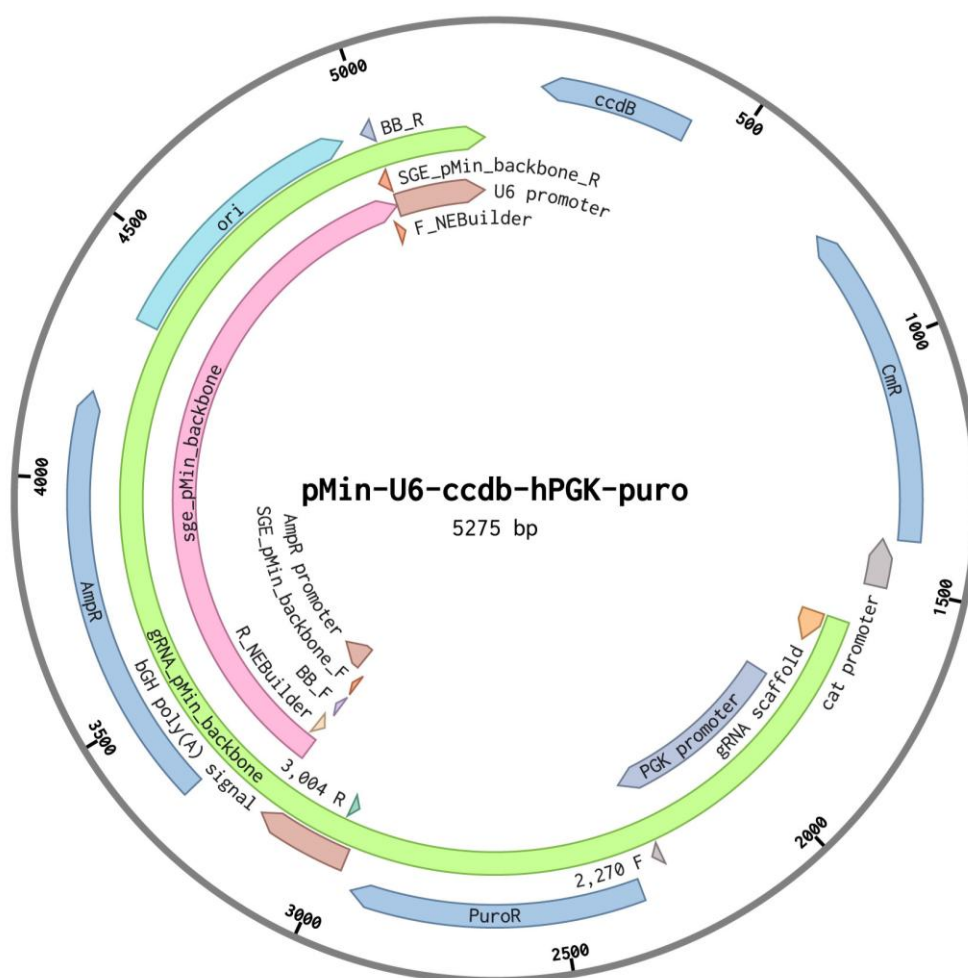
a. Prepare a restriction digestion master mix, as described in the materials and equipment section under “Reaction setup for DpnI digestion of the PCR product of Primer Set 3”.

b. Vortex and mix the reaction.

- c. In a thermocycler, incubate at 37°C for 1 h and subsequently inactivate at 80°C for 20 min.

**Note:** The 'sge\_pMin\_backbone' fragment is 1831 bp long (pink annotation 'sge\_pMin backbone' in **Figure 4**).

- Extract the 1831 bp band and purify using QIAquick Gel Extraction Kit (QIAGEN), following the manufacturer's instructions.
- Measure the DNA concentration of the purified PCR product of Primer Set 3.



**Figure 4: Schematic representation of the pMin-U6-ccdb-hPGK-puro plasmid.**

The pMin plasmid, with a total length of 5.275 kb, serves as the backbone for incorporating the wild-type homology arms (indicated by the pink annotation). Primer Set 3 is employed to amplify the ampicillin resistance (AmpR) gene and the origin of replication (ori) sequence. Additionally, the 'pMin' backbone is utilized to construct the 'gRNA\_pMin\_backbone' (green annotation). BbsI-HF restriction enzyme digestion facilitates the inclusion of the AmpR gene, ori sequence, and the puromycin resistance gene, and exclusion of the *ccdB* cassette in downstream clones.

## PRIMER DESIGN FOR WILD-TYPE HOMOLOGY REGION AMPLIFICATION—PRIMER SET 4

**Timing:** [30 min, depending on number of target regions]

This step describes the design of Primer Set 4, as listed in the **key resources table** of this protocol, which is used in a PCR reaction to amplify the wild-type homology sequence on either side of the target region (**Figure 2C**).

### 13. Design Primer Set 4 using any preferred primer design software

**Note:** In this protocol, we used the ‘Primer Design’ function within Geneious Prime, a modified version of Primer 3.

- a. Product size should be between 1200-2000 bp.

**Note:** When designing several target regions, it is recommended to keep the product sizes similar.

- b. The forward primer is 750-1000 bp 5' of the CDS exon/target region to be saturated with variants and the reverse primer is 750-1000 bp 3'.
- c. The primers define the limits of the target region, including homology arms.
- d. The target region to be saturated with variants should be roughly central.
- e. For the forward primer, the following adaptor sequence should be added 5':  
TTTTGCTCACATGTCCTGCAGG+[SEQUENCE].

**Note:** The adapter sequences are homologous to the amplified fragment of pMin-U6-ccdb-hPGK-puro (pMin)<sup>1</sup> (DpnI-digested PCR product of Primer Set 3) and are added to the 5' of the forward primer to allow for NEBuilder cloning.

- f. For the reverse primer, the following adaptor sequence should be added 5':  
GAAAAGTGCCACCTGACGGCGGCCGC+[SEQUENCE].

**Note:** The adapter sequences are homologous to the amplified fragment of pMin-U6-ccdb-hPGK-puro (pMin)<sup>1</sup> (DpnI-digested PCR product of Primer Set 3) and are added to the 5' of the reverse primer to allow for NEBuilder cloning.

## PCR FOR WILD-TYPE HOMOLGY REGION AMPLIFICATION USING PRIMER SET 4

### Timing: [1 day]

In this step, a PCR is performed that uses Primer Set 4, as listed in the **key resources table** of this protocol, to amplify wild-type homology regions surrounding the SGE target regions from HAP1-A5 gDNA. The PCR product of Primer Set 4 will subsequently be ligated into the plasmid backbone (PCR product of Primer Set 3) to generate the wild-type homology region plasmids.

**CRITICAL:** Extract gDNA from HAP1-A5 wild-type cells. Please refer to **steps 46 – 48** of this protocol.

14. Amplify the wild-type-sequence flanking and inclusive of the target region from HAP1-A5 gDNA.
  - a. Prepare a PCR reaction master mix as indicated below.

**CRITICAL:** The PCR steps for each exon/target region are performed in separate experiments.

| PCR reaction master mix for wild-type homology region amplification—Using Primer Set 4 |                        |
|----------------------------------------------------------------------------------------|------------------------|
| Reagent                                                                                | Amount                 |
| Wild-type HAP1-A5 gDNA                                                                 | X $\mu$ L (500 ng)     |
| homology_region_F (10 $\mu$ M)                                                         | 1.5 $\mu$ L            |
| homology_region_R (10 $\mu$ M)                                                         | 1.5 $\mu$ L            |
| KAPA HiFi HotStart ReadyMix (2X)                                                       | 25 $\mu$ L             |
| ddH <sub>2</sub> O                                                                     | Complete to 50 $\mu$ L |

- b. Run PCR following the conditions below.

**CRITICAL:** Optimal annealing temperature and cycle number need to be standardized for each primer pair per target region.

**PCR cycling conditions for wild-type homology region amplification—Using Primer Set 4**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 12–25 cycles     |
| Annealing            | 58°C–65°C   | 15 s    |                  |
| Extension            | 72°C        | 2 min   |                  |
| Final extension      | 72°C        | 5 min   | 1                |
| Hold                 | 4°C         | forever |                  |

15. Run PCR product(s) of Primer Set 4 on a 1% agarose-TAE gel at 100 V for 30 min.
  - a. Extract the corresponding band(s) (~1200-2000 bp) and purify using QIAquick Gel Extraction Kit (QIAGEN), following the manufacturer's instructions.
  - b. Measure the DNA concentration of the purified PCR product(s) of Primer Set 4.

## GENERATION OF WILD-TYPE HOMOLOGY REGION PLASMIDS USING PCR PRODUCTS OF PRIMER SET 3 AND PRIMER SET 4

### Timing: [2 days]

In the following steps, we generate the wild-type homology region plasmid for each SGE target region by ligating the SGE\_pMin\_backbone (PCR product of Primer Set 3) with wild-type homology regions (PCR product(s) of Primer Set 4).

16. Ligate the SGE\_pMin\_backbone (PCR product of Primer Set 3) with wild-type homology regions (PCR product(s) of Primer Set 4) via NEBuilder HiFi DNA assembly.

**CRITICAL:** The PCR steps for each exon / target region are performed in separate experiments

- a. Prepare a NEBuilder ligation master mix, as described in the materials and equipment section under “Reaction setup for ligation of wild-type homology region plasmids”.
  - b. In a thermocycler, incubate the reaction at 50°C for 1 h.
  - c. Dilute the ligated product 1:4 in ddH<sub>2</sub>O.
17. Clone the wild-type homology region plasmids in *E. coli* One Shot TOP10 Chemically Competent cells.

**CRITICAL:** Set the water bath to 42°C.

**CRITICAL:** The cloning steps for each exon / target region are performed in separate experiments.

- a. Place ampicillin (100 µg/mL) agar plates into an incubator set to 37°C.
- b. Thaw a vial of 50 µL TOP 10 cells on ice.
- c. Add 2 µL of the wild-type homology region plasmid 1:4 diluted NEBuilder ligation reaction from **step 16.c** to cells.
- d. Incubate cells on ice for 30 min, heat shock at 42°C for 30 s and incubate on ice for 2 min.
- e. Add 950 µL of Super Optimal broth with Catabolite repression (SOC) medium (stored at 20°C–25°C) and place cells at 37°C for 40 min in a shaking incubator at 210 RPM.
- f. Spread 100 µL of cells on pre-warmed ampicillin (100 µg/mL) agar plates and leave at 37°C for 12–16 h.
- g. Next day, inoculate five colonies into separate 15 mL conical Falcon tubes containing 5 mL of Lysogenic Broth (LB) media (supplemented with ampicillin (100 µg/mL)) and incubate in a shaking incubator at 210 RPM at 37°C for 12–16 h.
- h. Prepare glycerol stock from 12 – 16 h grown cultures. Combine 500 µL culture with 500 µL glycerol (50%), mix well and store at –80°C.
- i. Isolate plasmids from the remaining culture using the QIAprep Spin Miniprep Kit (QIAGEN), following the [manufacturer's instructions](#).
- j. Measure the DNA concentration of the miniprep DNA product.
- k. Confirm correct insertion of the 'wild-type\_homology\_region\_plasmid' for each target region via Sanger sequencing using commercial vendors.

## PRIMER DESIGN FOR THE LINEARIZATION OF WILD-TYPE HOMOLOGY REGION PLASMID—PRIMER SET 5

**Timing: [10 min]**

In the section below the design for Primer Set 5, as listed in the **key resources table** of this protocol, is described. These primers aid in the linearization of the 'wild-type\_homology\_region\_plasmid' (generated in step **17** of this protocol), so that the individual target-specific SGE variant oligonucleotide libraries (generated in step 6 of this protocol) can subsequently be cloned into the linearized homology backbone.

## 18. Design Primer Set 5

**Note:** For the design of this set of primers, no design tool is needed.

- a. Reverse-complement the sequences of Primer Set 2. These will be the new sequences of Primer Set 5 primers.

**Note:** In this way, the wild-type homology plasmid will be linearized exactly at the border of the target-specific regions (**Figure 2D**).

## LINEARIZATION OF THE WILD-TYPE HOMOLOGY REGION PLASMID USING PRIMER SET 5

**Timing:** [1–2 days, depending on number of target regions]

This step describes a PCR for the linearization of homology region plasmids, which will be used as a backbone to be ligated with the individual target-specific SGE variant oligonucleotide libraries, using Primer Set 5, as listed in the **key resources table** of this protocol, to produce SGE HDR template libraries.

### 19. Linearize the wild-type homology region plasmids.

- a. Prepare a PCR reaction master mix as indicated below.

**Note:** The PCR steps for each exon/target region are performed in separate experiments.

**PCR reaction master mix for wild-type homology region plasmid linearization—Using Primer Set 5**

| Reagent                           | Amount                 |
|-----------------------------------|------------------------|
| wild-type homology region plasmid | X $\mu$ L (10 pg)      |
| Linearization_F (10 $\mu$ M)      | 1.5 $\mu$ L            |
| Linearization_R (10 $\mu$ M)      | 1.5 $\mu$ L            |
| KAPA HiFi HotStart ReadyMix (2X)  | 25 $\mu$ L             |
| ddH <sub>2</sub> O                | Complete to 50 $\mu$ L |

- b. Run a PCR following the conditions below.

**PCR cycling conditions for wild-type homology region plasmid linearization—Using Primer Set 5**

| Steps                | Temperature | Time       | Number of cycles |
|----------------------|-------------|------------|------------------|
| Initial Denaturation | 95°C        | 3 min      | 1                |
| Denaturation         | 98°C        | 20 s       | 35 cycles        |
| Annealing            | 55°C–65°C   | 15 s       |                  |
| Extension            | 72°C        | 2 min 30 s |                  |
| Final extension      | 72°C        | 1 min      | 1                |
| Hold                 | 4°C         | forever    |                  |

20. Digest the linearized PCR product(s) of Primer Set 5 with DpnI restriction enzyme (to remove methylated wild-type plasmid template).
  - a. Prepare a restriction digestion master mix, as described in the materials and equipment section under “Digestion of PCR product of Primer Set 5 with DpnI”.
  - b. In a thermocycler, incubate the reaction at 37°C for 1 h and subsequently inactivate at 80°C for 20 min.
21. Run the digested PCR product(s) of Primer Set 5 on a 1% agarose-TAE gel at 120 V for 30 min.
22. Extract the corresponding band and purify using QIAquick Gel Extraction Kit (QIAGEN), following the manufacturer’s instructions.
23. Measure the DNA concentration of the gel-extracted PCR product(s) of Primer Set 5.

**CLONING TARGET-SPECIFIC SGE VARIANT OLIGONUCLEOTIDE LIBRARIES INTO LINEARIZED WILD-TYPE HOMOLOGY REGIONS**
**Timing: [2 days]**

This step describes the cloning step of the amplified target-specific SGE variant oligonucleotide libraries (PCR product of Primer Set 2) into linearized wild-type homology regions (PCR product of Primer Set 5) to produce the final SGE HDR template libraries that can subsequently be transfected into HAP1-A5 cells (**Figure 2E**).

**Note:** The cloning steps for each exon/target region are performed in separate experiments.

24. Prepare a ligation reaction mix, as described in the materials and equipment section under ‘Reaction setup for ligation of linearized wild-type homology regions with target-specific SGE variant oligonucleotide libraries’.

- a. In a thermocycler, incubate the NEBuilder ligation reaction at 50°C for 1 h.

**Pause point:** The NEBuilder ligation can be used directly for bacterial cell transformations or can be stored at –20°C.

## 25. Endura cell transformation to clone SGE HDR template libraries.

**Note:** In this procedure, we use the Endura Electrocompetent Cells with some alterations to the manufacturer's instructions, as outlined below. For the electroporation of cells, the Gene Pulser Xcell Electroporation System is used. For further details on materials used, please refer to the **key resources table**.

**Note:** The transformation steps for each exon/target region are performed in separate experiments.

**Note:** Thaw the recovery media 1–2 h before the transformation.

- a. Add 2 µL of the NEBuilder ligation reaction (**from step 24.a**) to 25 µL of Endura cells on ice.

**CRITICAL:** Stir briefly with pipette tip. Do not pipette up and down to mix, which can introduce air bubbles and warm the cells.

- b. Add the cell/ligation mix carefully to the Gene Pulser Cuvette (0.1 cm gap).
- c. Electroporate cells at 1800 V (capacity: 10 µF, resistance: 600 Ω).
- d. Add 975 µL of recovery media to each sample. Pipette up and down. Transfer the transformation mixture to a 15 mL conical Falcon tube.
- e. Following electroporation of cells, transformed cells are placed in a shaking incubator at 210 RPM at 37°C for 1 h.
- f. Grow 990 µL of transformed cells in 125 mL of LB media (supplemented with ampicillin (100 µg/mL)) for 12–16 h.
- g. Plate 1% of the transformation reaction (10 µL transformed cells with 90 µL SOC medium) onto ampicillin-containing plates and leave at 37°C for 12–16 h. After 12–16 h incubation, evaluate the cultured plates for adequate colony growth/coverage (>500 colonies = >50x cloning coverage for an SGE variant oligonucleotide library containing 1000 variants) to determine if they are suitable for the next experiments.

- h. **Optional step:** Include a negative control to verify the effectiveness of the DpnI digestion in removing any remaining plasmid.
  - i. Plate 2  $\mu$ L of the gel-extracted and DpnI-digested PCR product of Primer Set 5 (**step 23**) onto ampicillin-containing plates and incubate at 37°C for 12–16 h.
  - ii. Since the product from **step 23** is linear and should not contain any plasmid, no colony growth is expected if the DpnI digestion was successful.
- i. **Optional step:** Include a negative control to verify the ligation in **step 24a** was performed correctly.
  - i. Prepare a ligation reaction mix, as described in the materials and equipment section under ‘Optional negative control: Reaction setup for ligation of the linearized wild-type homology regions (PCR product of Primer Set 5)’.
  - ii. In a thermocycler, incubate the NEBuilder ligation reaction at 50°C for 1 h.
  - iii. Plate 2  $\mu$ L of the NEBuilder ligation reaction onto ampicillin-containing plates and incubate at 37°C for 12–16 h.
  - iv. No colony growth should be expected, since the ligation is not expected to succeed due to the non-complementary nature of the linearized wild-type homology regions.
- j. Prepare glycerol stock from 12–16 h grown cultures of transformed cells. 500  $\mu$ L culture combined with 500  $\mu$ L glycerol (50%), mix well, and store at –80°C.
- k. Extract plasmid DNA from cultures using the QIAGEN Plasmid Plus Maxi Kit (QIAGEN), following the manufacturer’s instructions, to isolate the final ‘SGE HDR template library’.
  - i. Elute the plasmid DNA in 200  $\mu$ L EB Buffer.
- l. Confirm the correct insertion of the ‘SGE HDR template library’ for each target region via NGS.

## SGRNA CLONING

**Timing:** [2–3 days, depending on number of target regions]

This step describes the cloning of target region-specific sgRNAs into the gRNA\_pMin\_backbone pMin-U6-ccdb-hPGK-puro,<sup>1</sup> 3653 bp in length, as illustrated by the green annotation 'gRNA\_pMin\_backbone' in **Figure 4**. The sgRNA design is explained in Section 'before you begin' of this protocol. In this procedure, we use the Zhang Lab General Protocol, with some alterations as outlined below.

**CRITICAL:** Use the same pMin-U6-ccdb-hPGK-puro (pMin)<sup>1</sup> plasmid as used earlier in this protocol.

26. Digest the pMin-U6-ccdb-hPGK-puro (pMin) backbone with the BbsI-HF restriction enzyme, as described in the materials and equipment section under “Digestion setup for DpnI-digested SGE-pMin backbone”.
  - a. In a thermocycler, incubate the reaction at 37°C for 30 min and subsequently inactivate at 65°C for 20 min.
  - b. Run the product on a 1% agarose-TAE gel at 120 V for 30 min.

**Note:** The digest will result in two bands: one smaller at 1599 bp and one larger at 3653 bp. The gRNA\_pMin\_backbone fragment is the 3653 bp band (green annotation 'gRNA\_pMin\_backbone' in **Figure 4**).

- c. Extract the ~3.6 kb band and purify using QIAquick Gel Extraction Kit (QIAGEN), following the manufacturer's instructions.
  - d. Assess the concentration and purity of the extracted DNA.
27. Phosphorylate and anneal complementary oligos, as described in the materials and equipment section under “Reaction setup for phosphorylation and annealing of complementary sgRNA oligos”.
  - a. Incubate at 37°C for 30 min, then at 95°C for 5 min and finally ramp down to 25 °C at 5 °C/min.
  - b. Dilute the reaction 1:200 in ddH<sub>2</sub>O.
28. Ligate the annealed oligos into the BbsI-digested gRNA\_pMin\_backbone, as described in the materials and equipment section under “Reaction setup for ligation of sgRNA oligos with pMin\_gRNA\_backbone”.
  - a. In a thermocycler, incubate the reaction at 20°C–25°C for 30 min.
  - b. Dilute the ligated products 1:4 in ddH<sub>2</sub>O.

**Pause point:** The ligation can be used for subsequent cell transformations or can be stored at  $-20^{\circ}\text{C}$ .

29. Clone the sgRNA plasmids in *E. coli* One Shot TOP10 Chemically Competent cells.

**CRITICAL:** Set a water bath to  $42^{\circ}\text{C}$ .

- a. Place ampicillin ( $100\text{ }\mu\text{g/mL}$ ) agar plates into an incubator set to  $37^{\circ}\text{C}$ .
- b. Thaw  $25\text{ }\mu\text{L}$  of TOP10 cells on ice (5 min–10 min).
- c. Add  $1\text{ }\mu\text{L}$  of 1:4 diluted ligation to a vial containing  $25\text{ }\mu\text{L}$  of thawed TOP10 cells on ice.
- d. Incubate cells/DNA mixture on ice for 30 min, heat shock at  $42^{\circ}\text{C}$  for 30 s and incubate on ice for 2 min.
- e. Add  $125\text{ }\mu\text{L}$  of SOC medium (stored at  $20^{\circ}\text{C}$ – $25^{\circ}\text{C}$ ) and leave cells at  $37^{\circ}\text{C}$  for 45 min in a shaking incubator at 210 RPM.
- f. Spread  $100\text{ }\mu\text{L}$  of cells on pre-warmed ampicillin ( $100\text{ }\mu\text{g/mL}$ ) agar plates and leave at  $37^{\circ}\text{C}$  for 12–16 h.
- g. Next day, inoculate three colonies into separate 15 mL conical Falcon tubes containing 5 mL of LB media (supplemented with ampicillin ( $100\text{ }\mu\text{g/mL}$ )) and incubate tubes at  $37^{\circ}\text{C}$  in a shaking incubator for 12–16 h.
- h. After 12–16 h prepare glycerol stock from grown cultures.
  - i.  $500\text{ }\mu\text{L}$  culture combined with  $500\text{ }\mu\text{L}$  glycerol (50%), mix well and store at  $-80^{\circ}\text{C}$ .
- i. Isolate sgRNA plasmids from the remaining culture using the QIAprep Spin Miniprep Kit (QIAGEN), following the manufacturer's instructions.
- j. Sanger sequence them to confirm correct insertion.
- k. After sequencing, re-culture correct clones in 250 mL conical Erlenmeyer Flasks using  $5\text{ }\mu\text{L}$  of glycerol stock in 125 mL of LB media (supplemented with ampicillin ( $100\text{ }\mu\text{g/mL}$ )) and place flasks into a shaking incubator at 210 RPM at  $37^{\circ}\text{C}$  for 12–16 h.

**Note:** Culture every exon/SGE target region in an individual conical Erlenmeyer Flask.

- l. Purify sgRNA plasmids using the QIAGEN Plasmid Plus Maxi Kit (QIAGEN), following the manufacturer's instructions.
- m. Measure the DNA concentration, hereafter referred to as the 'sgRNA plasmid'.

## THAW AND EXPAND HAP1-A5 CELLS

### Timing: [6–9 days]

Below we outline the procedure for thawing and expanding HAP1-A5 cells before the SGE screen. HAP1-A5 cells need to be thawed and expanded for 2–3 passages, passaging upon reaching 70–80% confluency.

**Note:** As per Good Laboratory Practice (GLP) guidelines, it is recommended to test HAP1-A5 cells for mycoplasma contamination prior to use in SGE experiments.

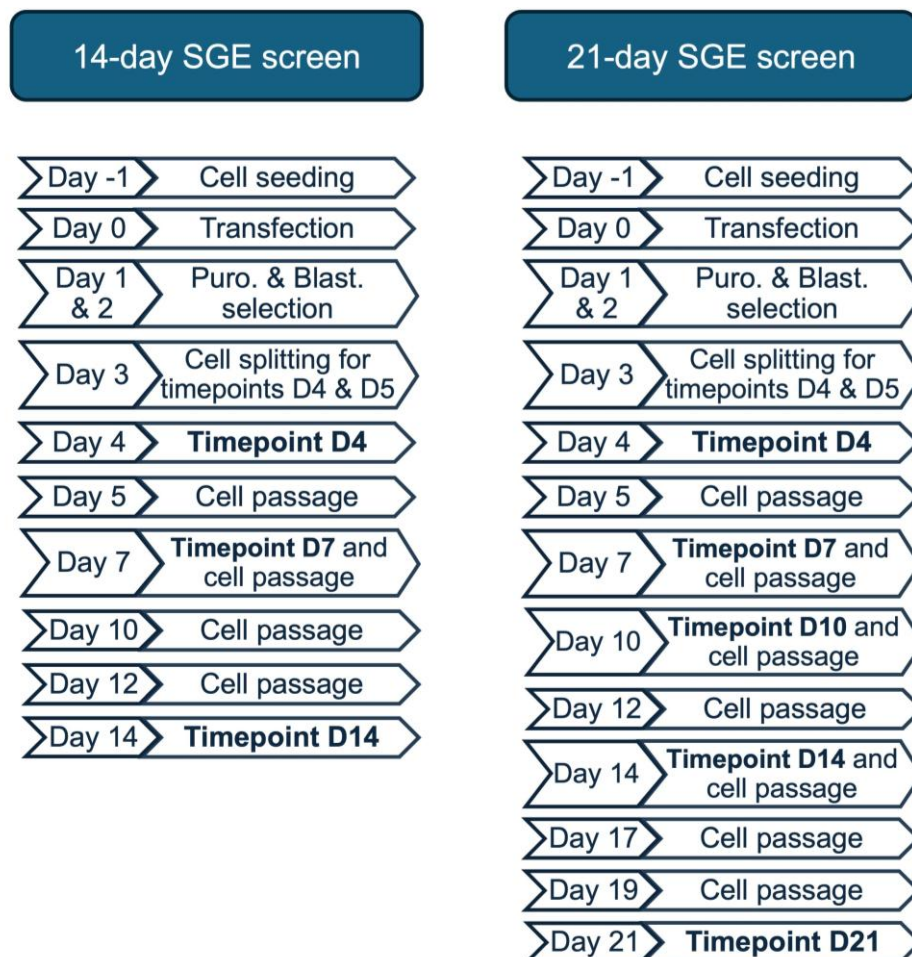
30. Thaw vial of haploid-sorted HAP1-A5 cells from liquid nitrogen 8–10 days prior to SGE cell culture experiments.
  - a. Prepare fresh 'B+ media', as described in the materials and equipment section under "Cas9-selection media for HAP1-A5 cell line ('B+ media')".
  - b. Thaw cells in a water bath set to 37°C for 2–3 min.
  - c. Add 5 mL of 'B+ media' into a 15 mL conical Falcon tube.
    - i. Add 1 mL of thawed cell suspension to the 15 mL conical Falcon tube.
    - ii. Wash the cryovial with 1 mL of 'B+ media' and add to the same 15 mL conical Falcon tube.
  - d. Spin the cells at 300 x g for 3 min at 20°C–25°C, aspirate liquid.
  - e. Resuspend cell pellet in 1 mL of 'B+ media' and add to a T150 flask filled with fresh 30 mL of 'B+ media'.
31. Expand cells for 7–9 days in T75/T150 flasks in 15/35 mL 'B+ media'.
  - a. Aspirate old media.
  - b. Wash cells 2 times with 10 mL of Dulbecco's phosphate-buffered saline (DPBS), aspirate and discard DPBS.
  - c. Detach cells from culture surface using TrypLE, following the manufacturer's instructions.
  - d. Add 2–4 mL of 'noAB media' to the detached cells.
  - e. Transfer the detached cells into a 50 mL conical Falcon tube.
  - f. Wash the flasks with 5 mL of 'B+ media' and add the wash to the same 50 mL conical Falcon tube.
  - g. Split the cell suspension across a suitable number of flasks (depending on the confluency of the flask/number of required transfections). Incubate in T75/T150 flasks in 15/35 mL of fresh 'B+ media' at 37°C in a 5% CO<sub>2</sub> incubator.
  - h. Passage the cells every ~2 days (upon reaching 70–80% confluency).

## SGE SCREEN

### Timing: [16 or 23 days]

This section outlines the tissue culture procedures performed during the SGE screen. For this experiment, standard tissue culture facilities (CL1) are required, including a ducted fume hood, microscope, water bath, and cell counting equipment.

**Note:** Before initiating the experiments, it is recommended to assess the essentiality of the target gene in HAP1-A5 cells to determine the appropriate screen duration. This assessment can be based on the Bayes Factor<sup>7,9</sup> or CRISPR KO screens. For instance, CRISPR KO screens in HAP1-A5 cells can provide insights into the depletion rate of the target gene. Some genes, such as BAP1<sup>2</sup>, may require up to 21 days, whereas others (e.g. RAD51C<sup>3</sup>) may be sufficiently analyzed within 14 days (**Figure 5**). The chosen duration will also dictate the timing of gDNA timepoint sampling.



**Figure 5: Schematic representation of two SGE tissue culture screening strategies.**

This figure illustrates the two potential saturation genome editing (SGE) screening approaches in tissue culture. For genes with strong essentiality in the HAP1-A5 cell line, a 14-day SGE screen is appropriate. In contrast, for genes with slower depletion dynamics, a 21-day SGE screen is recommended.

### 32. Day –1: Cell seeding.

- a. Prepare fresh 'noAB media', as described in the materials and equipment section under "Growth media for HAP1-A5 cell line ('noAB media')".
- b. Aspirate old media from all flasks you have in culture.
- c. Wash each flask 2 times with 5 mL of DPBS, aspirate, and discard DPBS.
- d. Detach cells from culture surfaces using TrypLE, following the manufacturer's instructions.
- e. Add 2–4 mL of 'noAB media' to the detached cells.

**Note:** Add at least the same volume of media as the volume of trypsin used.

- f. Transfer the detached cells into (several) 50 mL conical Falcon tube(s).

**Note:** Depending on the number of flasks, use the appropriate number of conical Falcon tubes. Since the cells are wild-type at this stage, cells from multiple T75 flasks can be pooled into a single 50 mL conical Falcon tube.

- g. Wash each flask with 5 mL of 'noAB media' and add the wash to the same 50 mL conical Falcon tube(s).
- h. Count the cells.

**Note:** We use Countess 3 Automated Cell Counter.

- i. Add 10  $\mu$ L of a 1:1 mixture of trypan blue and the cell suspension.
- i. Make a suspension that contains 8 million cells in 15 mL of 'noAB media'.
- j. Example calculation (for 3 exons in replicate and 3 control flasks: sgRNA-only transfection, sgRNA + SGE HDR template library transfection, HAP1-A5 wild-type untransfected):

Calculate the total number of cells needed for all flasks:

$$11(\text{total of flasks}) \times 8 \times 10^6 / \text{concentration of cells per ml measured} = X$$

Calculate the total volume of media required:

$$11(\text{flasks}) \times 15 \text{ ml (final volume per flask)} = 165 \text{ ml}$$

Subtract the calculated volume of the cell suspension (X) from the total volume:

$165\text{ ml} - X = \text{media volume to add to } X \text{ to make stock of cell suspension}$

The remainder is 'noAB media'.

- k. Seed 15 mL of cell suspension into 12 T75 flasks.
- l. Maintain the cells at 37°C in a 5% CO<sub>2</sub> incubator.
- m. Thaw Xfect Buffer at +4°C for 16–20 h.

### 33. Day 0: Transfection.

- a. Leave Xfect Buffer at 20°C–25°C for 1–2 h before starting tissue culture work.
- b. Aspirate old 'noAB media' and add 10 mL of fresh 'noAB' media. Allow the cells to incubate with fresh media for at least 1 h before transfection.
- c. Prepare the transfection mixture toward the end of the 1 h incubation.
  - i. Combine 7.5 µg sgRNA and 15 µg SGE HDR template library per exon/target region in a 1.5 mL Eppendorf tube.

**Note:** Perform the transformation mixtures in biological triplicates per exon / SGE target region sample.

- ii. Prepare the sgRNA-only transfection control sample by adding 7.5 µg sgRNA of a selected sgRNA together with 15 µg of a GFP plasmid to a 1.5 mL Eppendorf tube.
- iii. Add Xfect Buffer to adjust the final volume to 750 µL.
- iv. Vortex the samples briefly.
- v. Thaw the Xfect polymer briefly before adding 13.5 µL to each reaction (calculated as 0.6 µL of Xfect polymer per 1 µg of DNA, with 22.5 µg of DNA requiring 13.5 µL Xfect polymer).
- vi. Vortex each sample and incubate at 20°C–25°C for 10 min.

**CRITICAL:** Ensure the incubation time does not exceed 30 min.

- vii. Very briefly spin samples at <300 x g for 2–3 s at 20°C–25°C.
- d. After the 1 h incubation, add the prepared transfection mixture to the cells.
  - i. Position the T75 flask vertically and add the transfection mixture dropwise to the media. Then tilt the flask horizontally and gently swirl the media to ensure even distribution over the cells.

**CRITICAL:** Maintain the flask horizontally while processing other flasks.

**Note:** Perform the transformation in biological triplicates per exon/SGE target region sample.

**Note:** For the sgRNA-only transfection control, apply the corresponding transfection mixture. Do not add any transfection mixture to the HAP1-A5 wild-type untransfected T75 control flask.

- e. Incubate the flasks at 37°C for 4 h.
- f. Aspirate media and replace with 15 mL of fresh 'noAB media'.
- g. Leave cells to recover at 37°C in a 5% CO<sub>2</sub> incubator for 16–20 h.

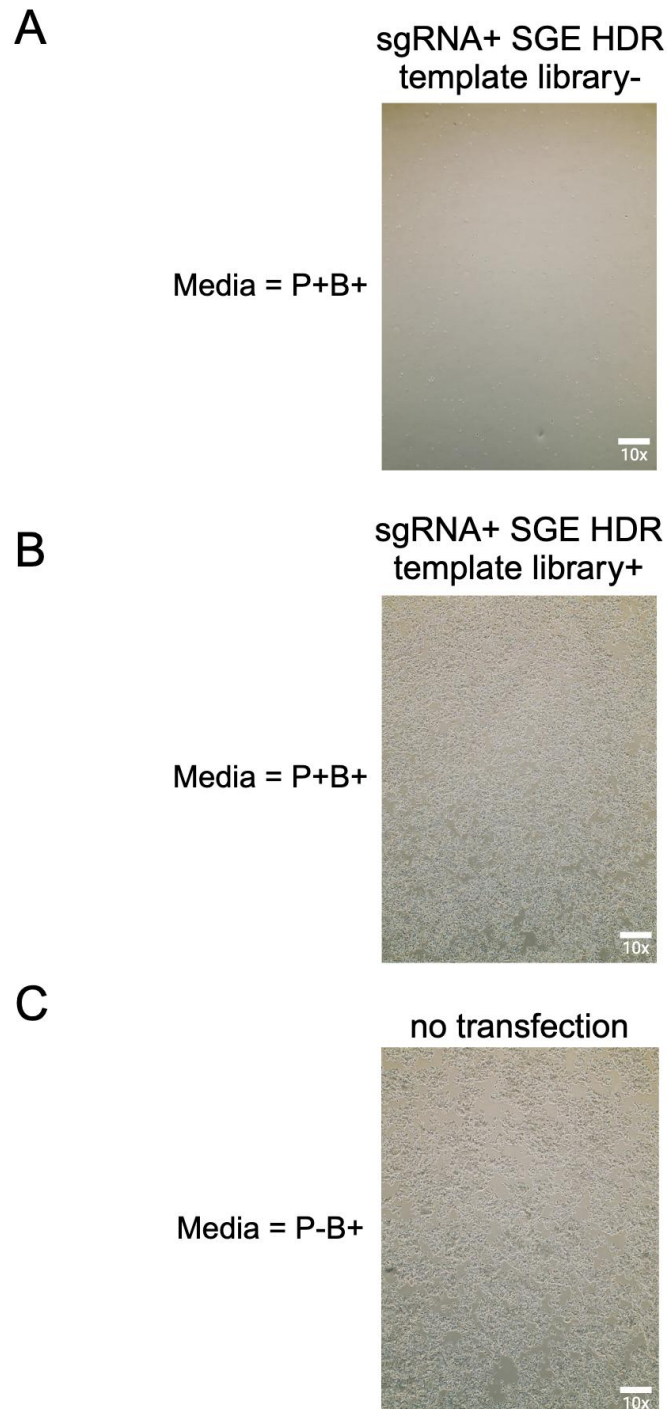
#### 34. Day 1: Selection.

- a. Aspirate old media and replace with 15 mL of fresh 'B+P+ media', as described in the materials and equipment section under "sgRNA selection media for HAP1-A5 cell line ('B+P+ media')".
- b. For the sgRNA-only transfection control: add 15 mL of 'B+P+ media'.
- c. For the HAP1-A5 wild-type untransfected control: add 15 mL of 'B+ P+media', if using as a control for puromycin selection (i.e., to confirm complete cell death is achieved in this control), which is recommended to ensure efficient sample selection. If using as a control for wild-type cell growth use 'B+ media'.
- d. Leave cells at 37°C in a 5% CO<sub>2</sub> incubator for 16–20 h.

#### 35. Day 2: Selection.

- a. Aspirate old media from all T75 flasks and replace with 15 mL of fresh 'B+P+ media'.
- b. Examine cells under a microscope to assess cell viability.

**Note:** See **Figure 6** for representative images of cell populations under 'B+P+ media' selection, with and without transfection. The sgRNA-induced effects on cell fitness/death are evident in sgRNA-only transfection control samples, often showing reduced cell density (**Figure 6A**). In contrast, sgRNA + SGE HDR template library transfection generally enhances survival, with most variants rescuing the fitness phenotype, and resulting in confluency similar to HAP1-A5 wild-type untransfected cells (**Figures 6B** and **6C**). While controls are optional, differences in cell fitness can confirm transfection efficiency, especially in multiple larger experiments (as often some guides will show this effect); in addition, replicates should all show similar effects, which can give a rough indication that experimental accuracy is acceptable.



**Figure 6: Imaging of HAP1-A5 cells under different tissue culture conditions for the SGE screen.** **A)** HAP1-A5 cells transfected with single guide RNA (sgRNA) alone and selected using puromycin (P+) and blasticidin (B+) media. The efficacy of the sgRNA in inducing a cell fitness or cell death phenotype may result in fewer cells in control samples, where only sgRNAs are transfected. **B)** HAP1-A5 cells transfected with both the saturation genome editing (SGE) homology-directed repair (HDR) template library and the corresponding sgRNA, and selected with puromycin (P+) and blasticidin (B+) media, exhibit 70–80% confluency. This is because the SGE HDR template library serves as an HDR template, and most variants in the library rescue the cell fitness phenotype. **C)** Untransfected wild-type HAP1-A5 cells also display 70–80% confluency. All images were captured using the Evos XL microscope at 10x magnification (scale bar indicated).

- c. Leave cells at 37°C in a 5% CO<sub>2</sub> incubator for 16–20 h.

36. Day 3: Split cells for subsequent time points.

- a. Aspirate old media from all T75 flasks.
- b. Wash each flask 2 times with 5 mL of DPBS, aspirate and discard DPBS.
- c. Detach cells from culture surface using TrypLE, following the manufacturer's instructions.
- d. Add 2–4 mL of 'B+ media' to the detached cells.

**Note:** Add at least the same volume of media as the volume of trypsin used.

- e. Transfer the detached cells into a 15 mL conical Falcon tube.

**Note:** Use a separate 15 mL conical Falcon tube for each flask.

- f. Wash the flasks with 5 mL of 'B+ media' and add the wash to the same 15 mL conical Falcon tube.
  - i. Resuspend the cell mixture gently by pipetting up and down.
- g. Divide the cell suspension of each 15 mL Falcon tube into two equal portions, transferring 50% of the cells to one T75 flask and 50% to another.

**Note:** Label one flask with the corresponding sample name followed by 'Day 4', and label the other flask with the sample name followed by 'Day 5' to indicate the respective time points.

- h. Incubate all T75 flasks in 15 mL of fresh 'B+ media', as described in the materials and equipment section under "Cas9-selection media for HAP1-A5 cell line ('B+ media')", and leave at 37°C in a 5% CO<sub>2</sub> incubator for 16–20 h.

37. Day 4: Time point 'Day 4' – gDNA extraction

- a. Aspirate old media from T75 flasks labelled 'Day 4'.
- b. Wash the T75 flasks with 5 mL of DPBS, aspirate and discard.
- c. Detach cells from culture surface using TrypLE, following the manufacturer's instructions.
- d. Add 2–4 mL of 'B+ media' to the detached cells.

**Note:** Add at least the same volume of media as the volume of trypsin used.

- e. Transfer the detached cells into a 15 mL conical Falcon tube.

**Note:** Use a separate 15 mL Falcon tube for each flask.

- f. Wash each flask with 5 mL of 'B+ media' and add the wash to the same 15 mL conical Falcon tube.
- g. Count the cells.
  - i. Add 10  $\mu$ L of a 1:1 mixture of trypan blue and the cell suspension.
- h. Centrifuge the 15 mL Falcon tubes at 300 x g for 3 min.
- i. Carefully aspirate the supernatant and resuspend the cell pellets in DPBS to achieve a concentration of 5–6 million cells per mL.
- j. Aliquot 5–6 million cells per mL into a 1.5 mL Eppendorf tube.

**Note:** It is recommended to aliquot several, at least 3, technical replicates per sample.

- k. Spin the 1.5 mL Eppendorf tubes at 300 x g for 3 min at 20°C–25°C, carefully remove the DPBS, and store pellets at –80°C.
- l. Discard any excess cells.

#### 38. Day 5: Passage cells.

- a. Aspirate old media from T75 flasks labelled 'Day 5'.
- b. Wash the T75 flasks with 5 mL of DPBS, aspirate and discard.
- c. Detach cells from culture surface using TrypLE, following the manufacturer's instructions.
- d. Add 2–4 mL of 'B+ media' to the detached cells.

**Note:** Add at least the same volume of media as the volume of trypsin used.

- e. Transfer the detached cells into a 15 mL conical Falcon tube.

**Note:** Use a separate 15 mL Falcon tube for each flask.

- f. Wash each flask with 5 mL of 'B+ media' and add the wash to the same 15 mL conical Falcon tube.
- g. Count the cells.
  - i. Add 10  $\mu$ L of a 1:1 mixture of trypan blue and the cell suspension.
- h. Centrifuge the 15 mL Falcon tubes at 300 x g for 3 min.
- i. Carefully aspirate the supernatant and resuspend the cell pellets in fresh 'noAB media' to achieve a concentration of 5–6 million cells per mL.

- j. Seed 5–6 million cells into one T75 flask in 15 mL of fresh 'noAB media'.

**Note:** Label each flask with the corresponding sample name followed by 'Day 7'.

- k. Leave cells at 37°C in a 5% CO<sub>2</sub> incubator until 'Day 7'.

39. Day 7: Time point 'Day 7' – extraction of gDNA and cell passage.

- a. Aspirate old media from T75 flasks labelled 'Day 7'.
- b. Wash the T75 flasks with 5 mL of DPBS, aspirate and discard.
- c. Detach cells from culture surface using TrypLE, following the manufacturer's instructions.
- d. Add 2–4 mL of 'B+ media' to the detached cells.

**Note:** Add at least the same volume of media as the volume of trypsin used.

- e. Transfer the detached cells into a 15 mL conical Falcon tube.

**Note:** Use a separate 15 mL Falcon tube for each flask.

- f. Wash each flask with 5 mL of 'B+ media' and add the wash to the same 15 mL conical Falcon tube.
- g. Count the cells.
  - i. Add 10 µL of a 1:1 mixture of trypan blue and the cell suspension.
- h. Centrifuge the 15 mL Falcon tubes at 300 x g for 3 min.
- i. Carefully aspirate the supernatant and resuspend the cell pellets in fresh 'noAB media' to achieve a concentration of 5–6 million cells per mL.
- j. Seed 5–6 million cells into one T150 flask in 35 mL of fresh 'noAB media'.

**Note:** Label each flask with the corresponding sample name followed by 'Day 10'.

- k. Leave cells at 37°C in a 5% CO<sub>2</sub> incubator until 'Day 10'.
- l. Spin the remaining cell suspension, 300 x g for 3 min at 20°C–25°C.
- m. Add 5 mL of DPBS, spin at 300 x g for 3 min at 20°C–25°C.
- n. Carefully aspirate the supernatant and resuspend the cell pellets in DPBS to achieve a concentration of 5–6 million cells per mL.
- o. Aliquot 5–6 million cells per mL into a 1.5 mL Eppendorf tube.

**Note:** It is recommended to aliquot several, at least 3, technical replicates per sample.

- p. Spin the 1.5 mL Eppendorf tubes at 300 x g for 3 min at 20°C–25°C, carefully remove the DPBS, and store pellets at –80°C.
- 40. Day 10: Time point ‘Day 10’ – extraction of gDNA and cell passage.
  - a. Repeat the same steps as for ‘Day 7’ with one modification: seed 5–6 million cells into a T75 flask rather than a T150 flask to provide optimal surface area for 2 days of growth rather than 3 days, as the cells will be passaged on ‘Day 12’.
- 41. Day 12: Passage cells.
  - a. Repeat the same steps as for ‘Day 5’.
- 42. Day 14: Time point ‘Day 14’ – extraction of gDNA and cell passage.
  - a. Repeat the same steps as for ‘Day 7’.
- 43. Day 17: Passage cells.
  - a. Repeat the same steps as for ‘Day 5’.
- 44. Day 19: Passage cells.
  - a. Repeat the same steps as for ‘Day 5’.
- 45. Day 21: Time point ‘Day 21’ – extraction gDNA.
  - a. Repeat the same steps as for ‘Day 4’.

## GDNA EXTRACTION FROM CELL PELLETS

**Timing: [1–2 days, depending on number of samples]**

After completing the SGE tissue culture experiments, gDNA is extracted and prepared for NGS submission. The following section outlines the steps for extracting gDNA from HAP1-A5 cell pellets obtained during the previous SGE screen experiment.

**CRITICAL:** Set a thermocycler to 56°C.

- 46. Thaw cell pellets for each time point-replicate.
- 47. Resuspend in 200 µL DPBS.
- 48. Extract gDNA from each pellet using the DNeasy Blood & Tissue Kit (QIAGEN), following the manufacturer’s instructions, with a few adaptations to the protocol as outlined below.

**Note:** All collection tubes used for the washing steps are provided, except for the final 1.5 mL Eppendorf tubes where the DNA product is eluted.

- a. Add 4 µL of RNase A (100 mg/mL). Incubate at 37°C for 15 min.
- b. Add 20 µL of Proteinase K.

- c. Add 200  $\mu\text{L}$  of Buffer AL. Mix thoroughly (vortex and a quick spin at 300 x g for 30 s at 20°C–25°C).
- d. Incubate at 56°C for 15 min.
- e. Add 200  $\mu\text{L}$  of Ethanol (96–100%). Mix thoroughly. (vortex and a quick spin at 300 x g for 30 s at 20°C–25°C)
- f. Mix the mixture by pipetting up and down for 5 times. Pipette 700  $\mu\text{L}$  (all) of the mixture into the DNAEasy Mini spin column placed in a 2 mL collection tube (provided within the kit). Centrifuge at 9500 RPM for 1 min.
- g. Place the spin column in a new collection tube. Add 500  $\mu\text{L}$  of Buffer AW1. Centrifuge at 9500 RPM for 1 min. Discard the flow-through and the collection tube.
- h. Place the spin column in a new collection tube. Add 500  $\mu\text{L}$  of Buffer AW2. Centrifuge at 13,200 RPM for 1 min. Discard the flow-through. Centrifuge for another 2 min at 13,200 RPM. Discard the flow-through and the collection tube.
- i. Place the spin column into a new 1.5 mL Eppendorf tube (not provided within the kit).
- j. Elute gDNA in 100  $\mu\text{L}$  of Buffer AE.

**Note:** For samples that have less than 5–6 million cells, elute in 50  $\mu\text{L}$  Buffer AE.

- k. Incubate at 20°C–25°C for 5 min.
- l. Centrifuge at 13,200 RPM for 1 min.

**Note:** The solution should look viscous.

- m. Measure the DNA concentration.

**Note:** From a 5–6 million-cell pellet, gDNA concentrations of  $\geq 10 \mu\text{g}$  in 100  $\mu\text{L}$  are to be expected.

## PRIMER DESIGN FOR PCR FOR SAMPLING/AMPLIFICATION OF EDITED GDNA REGIONS-PRIMER SET 6

**Timing: [30 min, depending on number of samples]**

The following step outlines the primer design for Primer Set 6, as shown in the **key resources table** of this protocol. This primer set is used in a PCR to sample/amplify SGE targeted regions from the HAP1-A5 extracted, edited gDNA.

49. Design Primer Set 6 using any preferred primer design software.

**Note:** In this protocol, we used the 'Primer Design' function within Geneious Prime, a modified version of Primer 3.

- a. Product size should be between 1700-2000 bp.

**Note:** When designing several target regions, it is recommended to keep the product sizes similar.

- b. The Forward primer should be positioned outside the sequence range of Primer Set 4, as shown in the **key resources table** of this protocol. This ensures that amplification does not occur from any residual SGE HDR template library plasmid present in the gDNA extractions.

3

## PCR FOR SAMPLING/AMPLIFICATION OF EDITED GDNA REGIONS USING PRIMER SET 6

**Timing: [2 days]**

In this step, a PCR is performed using Primer Set 6, as shown in the **key resources table** of this protocol, to amplify the edited SGE genomic region from the collected gDNA pool. Please refer to **Figure 7A** for an illustration of this process.

**Note:** Before performing the PCR below, optimal primer annealing temperature and PCR cycle number need to be standardized for each primer pair per target region. Please refer to **steps 5 – 7** of this protocol.

50. Amplify the area surrounding the specific SGE target region from the collected gDNA pool.

- a. Prepare a PCR reaction master mix as indicated below.

**Note:** Set up three 50  $\mu$ L PCR reactions for each time point-replicate. Each 50  $\mu$ L reaction should contain 1  $\mu$ g of gDNA, resulting in 3  $\mu$ g of gDNA per time point-replicate (which equates to roughly 1 million haploid genomes, with a mass of  $\sim$ 3 pg).

**PCR reaction master mix to amplify edited gDNA enrichment—Using Primer Set 6**

| Reagent                          | Amount                 |
|----------------------------------|------------------------|
| Time point-replicate gDNA        | X $\mu$ L (1 $\mu$ g)  |
| gdamp_F (10 $\mu$ M)             | 1.5 $\mu$ L            |
| gdamp_R (10 $\mu$ M)             | 1.5 $\mu$ L            |
| KAPA HiFi HotStart ReadyMix (2X) | 25 $\mu$ L             |
| ddH <sub>2</sub> O               | Complete to 50 $\mu$ L |

- a. Run a PCR following the conditions below.

**PCR cycling conditions for amplifying edited gDNA enrichment—Using Primer Set 6**

| Steps                | Temperature | Time       | Number of cycles |
|----------------------|-------------|------------|------------------|
| Initial Denaturation | 95°C        | 3 min      | 1                |
| Denaturation         | 98°C        | 20 s       | 20 - 25 cycles   |
| Annealing            | 58°C–65°C   | 15 s       |                  |
| Extension            | 72°C        | 1 min 30 s |                  |
| Final extension      | 72°C        | 1 min 30 s | 1                |
| Hold                 | 4°C         | forever    |                  |

51. Pool the 3 x 50  $\mu$ L reactions per time point-replicate into a 1.5 mL Eppendorf tube.

52. Purify each sample using the QIAquick PCR purification Kit (QIAGEN), following the manufacturer's instructions.

**Note:** Total volume per sample: 150  $\mu$ L.

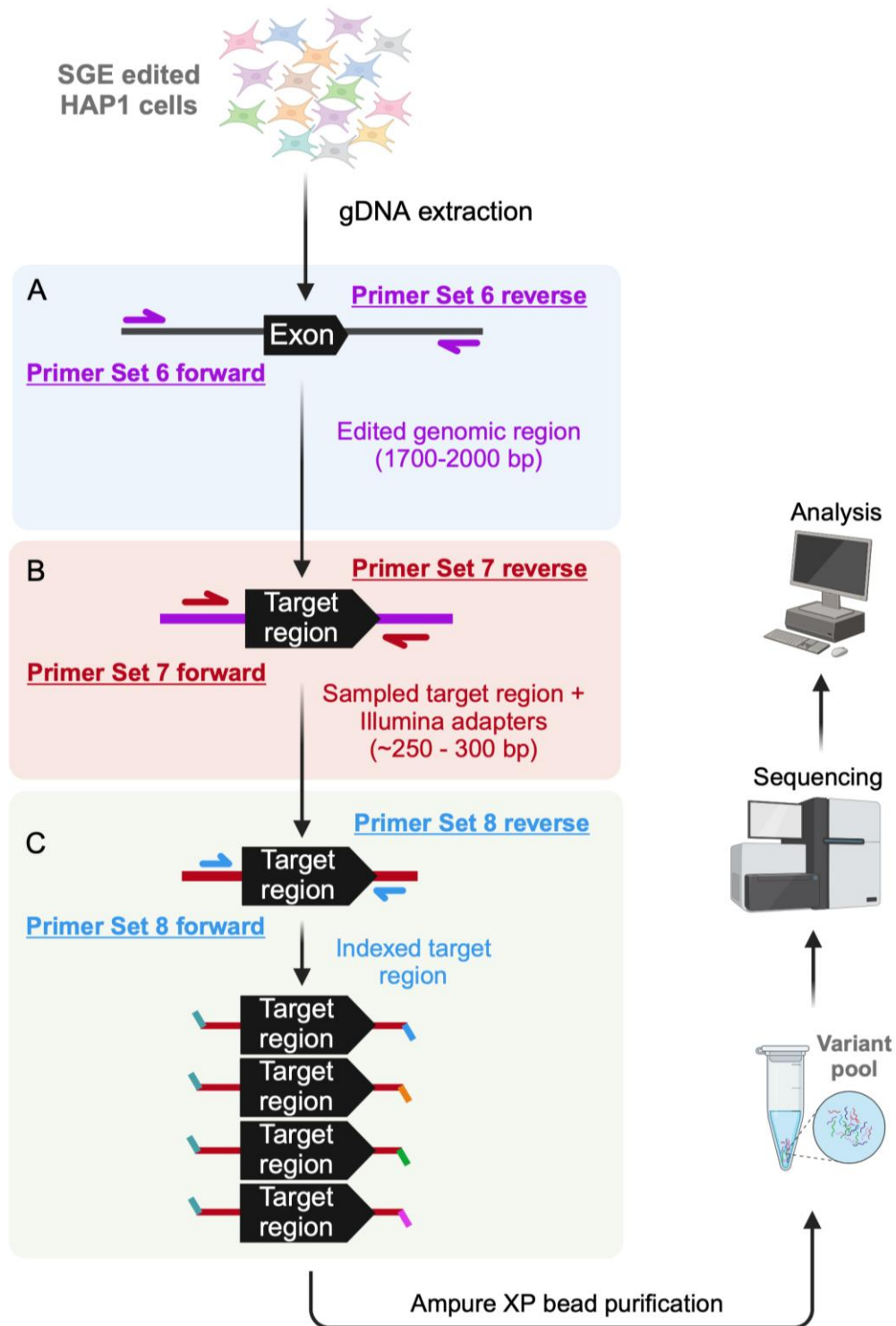
- a. Elute the DNA in 40  $\mu$ L EB Buffer.

53. Perform an enzymatic PCR cleanup with ExoI (to remove ssDNA & primers).

- Prepare a master mix, as described in the materials and equipment section under "Reaction setup for ExoI enzymatic cleanup of the PCR product of Primer Set 6".
- Vortex and mix the reaction.
- In a thermocycler, incubate the reaction at 37°C for 20 min followed by 80°C for 20 min.

54. Purify the ExoI-digested PCR product(s) of Primer Set 6 via MinElute columns (QIAGEN), following the manufacturer's instructions.
- Elute the DNA in 15  $\mu$ L of EB Buffer.
  - Measure the DNA concentration of the purified PCR product(s) of Primer Set 6.
  - Dilute samples to 10 ng/ $\mu$ L in EB Buffer.

***Pause point:*** The PCR product can be stored at  $-20^{\circ}\text{C}$  for long-term storage.



**Figure 7: Overview of SGE-edited genomic DNA library samples preparation for sequencing.**

Once the saturation genome editing (SGE) tissue culture experiments are completed, cell pellets are collected, and genomic DNA is extracted to prepare the SGE-edited genomic DNA (gDNA) library for next-generation sequencing (NGS). Library preparation includes three rounds of PCR. **A)** Primary PCR: Amplification of the genomic region edited by SGE is performed using Primer Set 6, with a specific set of primers designed for each exon. **B)** Secondary PCR: To amplify the specific target region from the sample and add Illumina adapters, Primer Set 7 is used. **C)** Indexing PCR: To add unique barcodes to all samples for NGS (including replicates and time points), Primer Set 8 is used for each sample. After indexing, all samples are bead-purified, pooled, and sequenced for subsequent analysis.

## PRIMER DESIGN FOR PCR FOR SEQUENCE ADAPTER ADDITION/AMPLIFICATION OF EDITED GDNA REGIONS—PRIMER SET 7

**Timing: [20 min, depending on number of samples]**

The following steps outline the primer design for Primer Set 7, as shown in the **key resources table** of this protocol. This primer set further amplifies the specific target region from the Primer Set 6 PCR product and incorporates Illumina adapters (for sequencing), allowing for the addition of individual sequencing barcodes in subsequent steps.

55. Design Primer Set 7 using any preferred primer design software.

**Note:** In this protocol, we used the 'Primer Design' function within Geneious Prime, a modified version of Primer 3.

- a. Product size should be between ~245-300 bp, the same as for the individual SGE target regions.
- b. The primers flank the target region, both facing inwards, orientated 5'>3'.
- c. The sequencing forward primer should be designed as: Indexing barcode dependent adaptors + 'Primer Set 2\_F', 5' and 3', respectively.
- d. The sequencing reverse primer should be designed as: Indexing barcode dependent adaptors + 'Primer Set 2\_R', 5' and 3', respectively.

## PCR FOR SEQUENCE ADAPTER ADDITION/AMPLIFICATION OF EDITED GDNA REGIONS USING PRIMER SET 7

**Timing: [2 days, depending on number of target regions]**

In this step, a PCR is performed that uses Primer Set 7, as shown in the **key resources table** of this protocol, to amplify the edited target-specific gDNA regions. Please refer to **Figure 7B** for an illustration of this process.

**Note:** Before performing the PCR below, optimal primer annealing temperature and PCR cycle number need to be standardized for each primer pair per target region. Please refer to **steps 5 – 7** of this protocol.

56. Amplify the target-specific gDNA region from 'Primers Set 7' PCR product.

- a. Prepare a PCR reaction master mix as indicated below.

**Note:** Set up three 50  $\mu$ L PCR reactions for each time point-replicate. Each 50  $\mu$ L reaction should contain 25 ng of DNA, resulting in 75 ng of DNA per time point-replicate.

**PCR reaction master mix for edited target-specific region amplification and Illumina adapter addition—Using Primer Set 7**

| Reagent                                            | Amount                 |
|----------------------------------------------------|------------------------|
| PCR product of Primer Set 6 (10 ng/ $\mu$ L stock) | 2.5 $\mu$ L (25 ng)    |
| sequencing_F (10 $\mu$ M)                          | 1.5 $\mu$ L            |
| sequencing_R (10 $\mu$ M)                          | 1.5 $\mu$ L            |
| KAPA HiFi HotStart ReadyMix (2X)                   | 25 $\mu$ L             |
| ddH <sub>2</sub> O                                 | Complete to 50 $\mu$ L |

b. Run a PCR following the conditions below.

**PCR cycling conditions for edited target-specific region amplification and Illumina adapter addition—Using Primer Set 7**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 6-15 cycles      |
| Annealing            | 58°C–65°C   | 15 s    |                  |
| Extension            | 72°C        | 30 s    |                  |
| Final extension      | 72°C        | 30 s    | 1                |
| Hold                 | 4°C         | forever |                  |

57. Pool the 3 x 50  $\mu$ L reactions per time point-replicate into a 1.5 mL Eppendorf tube.

58. Purify each sample using the QIAquick PCR purification Kit (QIAGEN), following the manufacturer's instructions.

**Note:** Total volume per sample: 150  $\mu$ L.

a. Elute the DNA in 30  $\mu$ L EB Buffer.

59. Perform an enzymatic PCR cleanup with ExoI (to remove ssDNA & primers).

- a. Prepare a master mix, as described in the materials and equipment section under “Reaction setup for Exol enzymatic cleanup of the PCR product of Primer Set 7”.
  - b. Vortex and mix the reaction.
  - c. In a thermocycler, incubate the reaction at 37°C for 20 min followed by 80°C for 20 min.
60. Purify the Exol-digested PCR product(s) of Primer Set 7 via MinElute columns (QIAGEN), following the manufacturer’s instructions.

**Note:** After adding PB Buffer, add 3 µL of 3M NaOAC pH 5.2 (to adjust the pH) to each sample.

- a. Elute the DNA in 12 µL of EB Buffer.
- b. Measure the DNA concentration of the purified PCR product(s) of Primer Set 7.
- c. Dilute samples to 5 ng/µL in EB Buffer.

**Pause point:** The PCR product can be stored at –20°C for long-term storage.

## INDEXING PCR TO ADD INDIVIDUAL BARCODES TO SAMPLES FOR NGS USING PRIMER SET 8

**Timing:** [1 day, depending on number of samples]

In this step a PCR is performed that uses Primer Set 8, as shown in the **key resources table** of this protocol, which allows for the addition of individual sequencing barcodes to each time point-replicate. Please refer to **Figure 7C** for an illustration of this process.

61. Add an indexing barcode to each time point-replicate sample.
  - a. Prepare a PCR reaction master mix as indicated below.

**CRITICAL:** Use individual indexing primers for each sample. No design or optimizations are needed for these primers. The primers are ordered either combined, pre-mixed, or can be added to reactions individually.

**Indexing PCR reaction master mix—Using Primer Set 8**

| Reagent                                     | Amount            |
|---------------------------------------------|-------------------|
| PCR product of Primer Set 7 (5 ng/μL stock) | 5 μL (25 ng)      |
| IDT_indexing_F (10 μM)                      | 1.5 μL            |
| IDT_indexing_R (10 μM)                      | 1.5 μL            |
| KAPA HiFi HotStart ReadyMix (2X)            | 25 μL             |
| ddH <sub>2</sub> O                          | Complete to 50 μL |

- b. Run PCR following the conditions below.

**Indexing PCR cycling conditions—Using Primer Set 8**

| Steps                | Temperature | Time    | Number of cycles |
|----------------------|-------------|---------|------------------|
| Initial Denaturation | 95°C        | 3 min   | 1                |
| Denaturation         | 98°C        | 20 s    | 7 cycles         |
| Annealing            | 59°C        | 15 s    |                  |
| Extension            | 72°C        | 30 s    |                  |
| Final extension      | 72°C        | 30 s    | 1                |
| Hold                 | 4°C         | forever |                  |

62. Purify the PCR products of Primer Set 8 using Ampure XP beads.

**Note:** This process can be done either in a Corning 96-well plate (round bottom) or in individual 1.5 mL Eppendorf tubes.

**CRITICAL:** Leave Ampure XP beads at 20°C–25°C for ~30 min before start.

- Vortex Ampure XP beads to resuspend.
- Add 45 μL (0.9X) of resuspended beads to 50 μL indexed PCR reaction. Mix well by pipetting up and down at least 10 times.

**Note:** Be careful to expel all the liquid out of the tip during the last mix.

- Incubate samples at 20°C–25°C for 15 min.
- During the incubation, prepare fresh 80% ethanol, ensuring enough volume to wash each sample with 400 μL.

- e. After the incubation, place the plate on a magnetic stand to separate the beads from the supernatant. After 5 min, carefully remove 30  $\mu$ L of the supernatant from the top.

**Note:** The removed supernatant should look clear.

- f. Seal and spin the plate at 1000 RPM for 20 s.

**Note:** Most of the beads should be at the bottom and the solution should look clear.

- g. Remove the seal and place it on the magnetic rack. After 5 min, remove all the solution.

**Note:** Repeat **step 62f**, if necessary.

- h. Add 200  $\mu$ L of the 80% freshly prepared ethanol to the plate while in the magnetic stand.

**CRITICAL:** Add to the side opposite the pellet so as not to disturb the pellet.

- i. Incubate at 20°C–25°C for 30 s, and then carefully remove and discard the supernatant.

**CRITICAL:** Be careful not to disturb the beads that contain the DNA.

- j. Repeat **Step 62h** and **i** for a total of two washes.
- k. Seal and spin the plate at 1000 RPM for 20 s.
- l. Remove the seal and place it on the magnetic rack. Remove traces of ethanol with a long-reach p10 pipette tip.

**Note:** Repeat **step 62k**, if necessary.

- m. Air dry the beads for up to 5 min while the plate is on the magnetic stand.

**CRITICAL:** Do not over-dry the beads. This may result in lower recovery of DNA. Elute the samples when the beads are still dark brown and glossy looking, but when all visible liquid has evaporated. When the beads turn lighter brown and start to crack, they are too dry.

- n. Remove the plate from the magnetic stand. Elute the DNA by adding 33  $\mu\text{L}$  of EB Buffer.
- o. Mix well by pipetting up and down 10 times. Incubate at 20°C–25°C for 10 min.
- p. Place the plate on the magnetic stand. After 5 min (or when the solution is clear) transfer 30  $\mu\text{L}$  to a new plate or a new 1.5 mL Eppendorf tube per time point-replicate.

**Note:** Ensure that you did not take out any beads, if possible.

- q. Measure the DNA concentration of the bead-purified DNA.

**Pause point:** The PCR product can be stored at  $-20^{\circ}\text{C}$  for long-term storage.

### 63. Control gel of bead-purified DNA samples.

- a. For each SGE edited gDNA library, prepare 15  $\mu\text{L}$  of 10 ng/ $\mu\text{L}$  SGE edited gDNA library with EB Buffer.
- b. Run 8  $\mu\text{L}$  of the 10 ng/ $\mu\text{L}$  SGE edited gDNA library on a 2% agarose-TAE gel at 120 V for 30 min.
  - i. As a control, run 80 ng of the purified PCR product of Primer Set 7 on the gel (~245-300 bp) to confirm the expected band size increase and assess relative concentration.
- c. Confirm the fragment size, clarity and intensity on the gel.

**Note:** You should see a single band at ~350-400 bp (**Figure 8**). Each SGE-edited gDNA library should exhibit similar band intensity if DNA quantification was performed correctly.

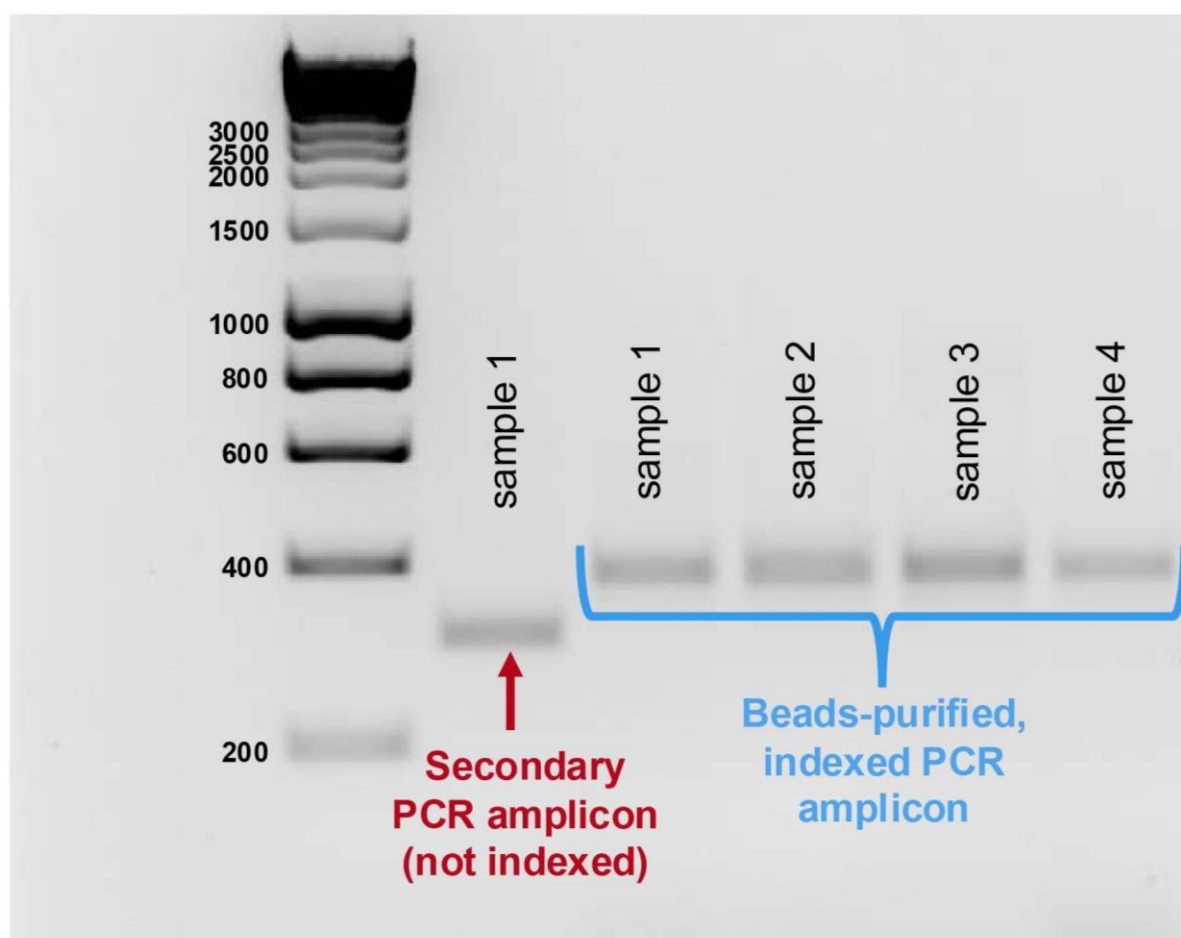
**Note:** If some bands on the gel exhibit significant variation in intensity compared to other bands, the DNA concentration of these samples should be re-evaluated to ensure that uniform concentrations (10 ng/ $\mu\text{L}$ ) across all samples are used for subsequent steps. This helps to normalize sequencing depth in the pooled NGS library.

- d. **Optional step:** Samples can also be analyzed using the Agilent TapeStation system to assess the integrity, size, and concentration of samples.
- e. Pool 5  $\mu\text{L}$  of each 10 ng/ $\mu\text{L}$  SGE edited gDNA library into one 1.5 mL Eppendorf tube to create a pooled NGS library.

- f. Sequence the samples using NGS at sufficient depth to achieve roughly 500x variant coverage. Read allocation will depend on the specific platform and number of NGS libraries in the pool. Please refer to Waters *et al.*<sup>2</sup> for an example.

64. Analyze data to calculate the depletion kinetics of the variants within target regions.

**Note:** SGE variant quantification, functional scoring, classifications, and data analysis have been previously conducted for the DDX3X,<sup>1</sup>BAP1,<sup>2</sup> and RAD51C<sup>3</sup> genes, providing established examples for the analysis of SGE datasets.



**Figure 8: Control gel electrophoresis of bead-purified samples before sequencing submission.** 80 ng of DNA sample and an appropriate volume of loading dye were used for gel electrophoresis. Lane 1: 1 kb HyperLadder. Lane 2: Sample 1 following secondary PCR (non-indexed), with an expected fragment size between ~245-300 base pair (bp). Lane 3: Sample 1 after indexing in the PCR round, showing an expected larger fragment size due to the addition of a unique barcode. Lanes 4–6: Three additional indexed sample examples, which show larger fragment sizes compared to the secondary PCR product. All band intensities are within an acceptable range, and no non-specific bands are observed.

## EXPECTED OUTCOMES

This protocol facilitates the generation of high-throughput mutational data, which can be analyzed using various approaches. Comprehensive methodologies for SGE data analysis, including functional score derivation, for the *DDX3X*<sup>1</sup>, *BAP1*<sup>2</sup>, and *RAD51C*<sup>3</sup> genes, are available for future reference. Briefly, SGE high-throughput sequencing data is processed into read counts, with SGE edited gDNA library-specific counts generated for each indexed time point and replicate. These counts are subsequently aggregated, quality-checked, and integrated with SGE HDR template library counts into a data frame. The resulting counts are utilized as input for DESeq2<sup>18</sup> to compute log<sub>2</sub>-fold changes between time points, enabling the assessment of variant dynamics over time. Variants are then correlated with the mutational consequences incorporated into the SGE oligo library design, followed by additional quality control steps. Functional scores for variants can be derived using analytical frameworks described in prior studies<sup>1–3</sup>.

The overarching goal of SGE data is to produce clinically relevant variant functional scores, thereby advancing efforts to resolve variants of uncertain significance (VUS). This contributes to improved clinical management, personalized medicine, and genetics research.

## LIMITATIONS

### CELL ESSENTIALITY IN HAP1-A5 CELLS

An analysis of the Cancer Gene Census identified that 59% of genes are predicted to exhibit a strong growth phenotype in HAP1 cells<sup>17</sup>. However, not all genes are essential. Notably, 46% of non-essential genes in HAP1 cells, including clinically relevant oncogenes, require alternative phenotyping approaches to the here presented protocol, for accurate functional characterization. Although we have no direct evidence that cell context influences editing readout at the gene level, factors such as cell type, the patient background from which HAP1 cells were derived, and the conditions in which the cells are grown could all impact the functional readout.

### POTENTIAL PCR OPTIMIZATIONS PER TARGET REGION

Despite the troubleshooting strategies and optimization suggestions provided in this protocol, researchers may still encounter challenges with specific primer sets. These difficulties may arise due to issues such as primer design, template quality or quantity, reaction conditions,

non-specific binding, or polymerase selection. If primer-related issues persist, it is recommended to redesign the primers by selecting a new primer pair with a more suitable annealing temperature, balanced GC content, and reduced likelihood of template mismatches.

## TROUBLESHOOTING

### PROBLEM 1

SGE primers (Primer Set 2, Primer Set 4, or Primer Set 7) produce non-specific results, such as multiple bands or smeared gel patterns.

### POTENTIAL SOLUTION

First, experiment with different annealing temperatures and consider using alternative polymerases, such as the Q5 High-Fidelity DNA Polymerase, with and without DMSO. If these adjustments do not improve results, redesign the primers for the target exon, shifting the binding sites while maintaining the same insert size. Additionally, ensure the target region remains relatively centralized to optimize amplification efficiency.

### PROBLEM 2

KAPA HiFi HotStart ReadyMix 2X is not performing well during SGE post-screen sample preparation for NGS.

### POTENTIAL SOLUTION

Consider using an alternative polymerase, such as repliQa HiFi ToughMix, which may offer better performance under the given conditions. When switching polymerases, it is essential to optimize the reaction conditions and conduct comparative testing to select the polymerase that provides the best results.

### PROBLEM 3

Target regions are difficult to amplify during SGE post-screen sample preparation for NGS.

### POTENTIAL SOLUTION

A first step might be to redesign Primer Set 6 as there is flexibility with these primers as they do not inform the experimental design beyond the need for one of the primers to be outside of

cloned homology arm regions. An alternative might be to implement a PCR touchdown protocol, as shown below, where the initial annealing temperature is set high and gradually reduced over successive cycles. This approach improves primer specificity and enhances DNA amplification efficiency, as supported by previous studies. However, this solution may lead to less representation of editing as amplification potentially proceeds more from amplicons than from gDNA. This should only be used as a last resort and carefully checked for quality<sup>19,20</sup>.

| PCR cycling conditions for Touchdown PCR |                  |            |
|------------------------------------------|------------------|------------|
| Step                                     | Temp             | Time       |
| 1                                        | 95°C             | 3 min      |
| 2                                        | 98°C             | 20 s       |
| 3                                        | 67 (–1/cycle) °C | 15 s       |
| 4                                        | 72°C             | 1 min 30 s |
| Back to step 2 for 11 total cycles       |                  |            |
| 5                                        | 98°C             | 20 s       |
| 6                                        | 57°C             | 15 s       |
| 7                                        | 72°C             | 1 min 30 s |
| Back to step 5 for 13 total cycles       |                  |            |
| 5                                        | 72°C             | 1 min 30 s |

#### PROBLEM 4

Endura cell transformation for cloning SGE HDR template libraries does not produce sufficient colony growth/coverage (i.e., 1% of transformation contains <500 colonies = <50x cloning coverage for SGE variant libraries containing 1000 variants).

#### POTENTIAL SOLUTION

First, repeat the transformation, performing 2–3 reactions per sample and pooling them. If unsuccessful, run a gel for the NEBuilder ligation reaction from **step 24** to verify band size. If the size is incorrect, repeat the ligation. If the problem persists after confirming a successful ligation, try one or more of the following: use a higher concentration of the amplified 'target specific SGE variant oligonucleotide library' and 'linearized wild-type homology region' (e.g., 75ng and 150ng, respectively) for the ligation or use a different strain of competent cells

(i.e., Stellar chemically competent cells). If necessary, re-prepare individual fragments and follow the protocol closely.

## RESOURCE AVAILABILITY

### LEAD CONTACT

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Andrew J. Waters (aw28@sanger.ac.uk).

### TECHNICAL CONTACT

Questions about the technical specifics of performing the protocol should be directed to the technical contacts, Dr. Andrew J. Waters (aw28@sanger.ac.uk) and Sofia Obolenski (so16@sanger.ac.uk).

### MATERIALS AVAILABILITY

This study did not generate new unique reagents.

### DATA AND CODE AVAILABILITY

This study did not generate/analyze datasets/code.

## ACKNOWLEDGMENTS

This research was funded by the Cancer Research UK CG-MAVE programme and the Wellcome Trust (grant no. 220540/Z/20/A, “Wellcome Sanger Institute Quinquennial Review 2021–2026”).

## AUTHOR CONTRIBUTIONS

S.O. and R.O.-L. optimized the protocol, performed the experiments, and prepared the manuscript. D.S. performed the experiments and optimized the Endura cell transformations. D.J.A. designed the experiments and prepared the manuscript. A.J.W. designed the experiments, optimized the protocol, performed the experiments, and prepared the manuscript. All authors reviewed, approved, and commented on the manuscript.

## REFERENCES

1. Radford, E.J., Tan, H.-K., Andersson, M.H.L., Stephenson, J.D., Gardner, E.J., Ironfield, H., Waters, A.J., Gitterman, D., Lindsay, S., Abascal, F., et al. (2023). Saturation genome editing of DDX3X clarifies pathogenicity of germline and somatic variation. *Nat. Commun.* 14, 7702. <https://doi.org/10.1038/s41467-023-43041-4>.
2. Waters, A.J., Brendler-Spaeth, T., Smith, D., Offord, V., Tan, H.K., Zhao, Y., Obolenski, S., Nielsen, M., Van Doorn, R., Murphy, J.-E., et al. (2024). Saturation genome editing of BAP1 functionally classifies somatic and germline variants. *Nat. Genet.* 56, 1434–1445. <https://doi.org/10.1038/s41588-024-01799-3>.
3. Olvera-León, R., Zhang, F., Offord, V., Zhao, Y., Tan, H.K., Gupta, P., Pal, T., Robles-Espinoza, C.D., Arriaga-González, F.G., Matsuyama, L.S.A.S., et al. (2024). High-resolution functional mapping of RAD51C by saturation genome editing. *Cell* 187, 5719–5734.e19. <https://doi.org/10.1016/j.cell.2024.08.039>.
4. Andersson, B.S., Collins, V.P., Kurzrock, R., Larkin, D.W., Childs, C., Ost, A., Cork, A., Trujillo, J.M., Freireich, E.J., and Siciliano, M.J. (1995). KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts. *Leukemia* 9, 2100–2108.
5. Kotecki, M., Reddy, P.S., and Cochran, B.H. (1999). Isolation and characterization of a near-haploid human cell line. *Exp. Cell Res.* 252, 273–280. <https://doi.org/10.1006/excr.1999.4656>.
6. Head, S.R., Komori, H.K., LaMere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., and Ordoukhanian, P. (2014). Library construction for next-generation sequencing: Overviews and challenges. *BioTechniques* 56, 61–77. <https://doi.org/10.2144/000114133>.
7. Cooper, S., Obolenski, S., Waters, A.J., Bassett, A.R., and Coelho, M.A. (2024). Analyzing the functional effects of DNA variants with gene editing. *Cell Rep. Methods* 4, 100776. <https://doi.org/10.1016/j.crmeth.2024.100776>.
8. Findlay, G.M., Daza, R.M., Martin, B., Zhang, M.D., Leith, A.P., Gasperini, M., Janizek, J.D., Huang, X., Starita, L.M., and Shendure, J. (2018). Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 562, 217–222. <https://doi.org/10.1038/s41586-018-0461-z>.
9. Olbrich, T., Vega-Sendino, M., Murga, M., de Carcer, G., Malumbres, M., Ortega, S., Ruiz, S., and Fernandez-Capetillo, O. (2019). A Chemical Screen Identifies Compounds Capable of Selecting for Haploidy in Mammalian Cells. *Cell Rep.* 28, 597–604.e4. <https://doi.org/10.1016/j.celrep.2019.06.060>.
10. Hart, T., Tong, A.H.Y., Chan, K., Van Leeuwen, J., Seetharaman, A., Aregger, M., Chandrashekar, M., Hustedt, N., Seth, S., Noonan, A., et al. (2017). Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 GenesGenomesGenetics* 7, 2719–2727. <https://doi.org/10.1534/g3.117.041277>.
11. Barbon, L., Offord, V., Radford, E.J., Butler, A.P., Gerety, S.S., Adams, D.J., Tan, H.K., and Waters, A.J. (2022). Variant Library Annotation Tool (VaLiAnT): an oligonucleotide library design and annotation tool for saturation genome editing and other deep mutational scanning experiments. *Bioinforma. Oxf. Engl.* 38, 892–899. <https://doi.org/10.1093/bioinformatics/btab776>.
12. Landrum, M.J., Chitipiralla, S., Brown, G.R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., et al. (2020). ClinVar: improvements to accessing data. *Nucleic Acids Res.* 48, D835–D844. <https://doi.org/10.1093/nar/gkz972>.
13. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. <https://doi.org/10.1038/s41586-020-2308-7>.
14. Aird, D., Ross, M.G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C., and Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* 12, R18. <https://doi.org/10.1186/gb-2011-12-2-r18>.
15. Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J. Biosci. Bioeng.* 96, 317–323. [https://doi.org/10.1016/S1389-1723\(03\)90130-7](https://doi.org/10.1016/S1389-1723(03)90130-7).
16. Thompson, J.R. (2002). Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by “reconditioning PCR.” *Nucleic Acids Res.* 30, 2083–2088. <https://doi.org/10.1093/nar/30.9.2083>.
17. Cooper, S.E., Schwartzentruber, J., Bello, E., Coomber, E.L., and Bassett, A.R. (2020). Screening for functional transcriptional and splicing regulatory variants with GenIE. *Nucleic Acids Res.* 48, e131. <https://doi.org/10.1093/nar/gkaa960>.

18. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
19. Green, M.R., and Sambrook, J. (2018). Touchdown Polymerase Chain Reaction (PCR). *Cold Spring Harb. Protoc.* 2018, pdb.prot095133. <https://doi.org/10.1101/pdb.prot095133>.
20. Korbie, D.J., and Mattick, J.S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* 3, 1452–1456. <https://doi.org/10.1038/nprot.2008.133>.

