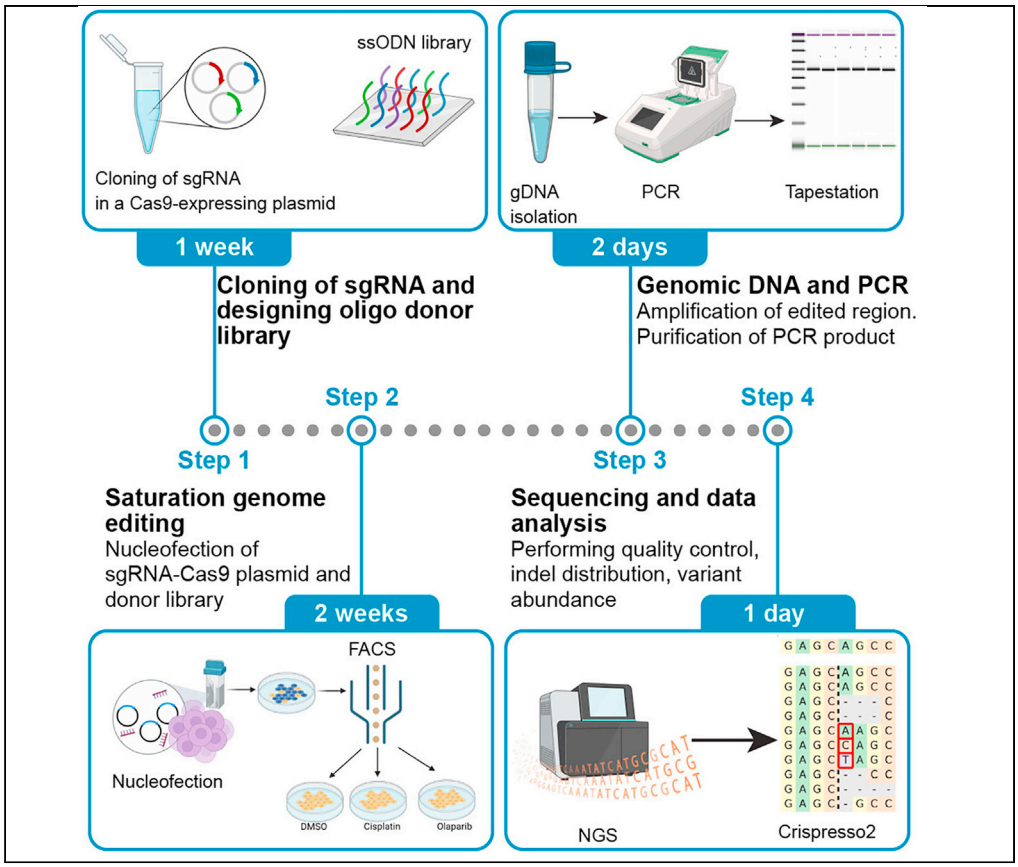
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for the saturation and multiplexing of genetic variants using CRISPR-Cas9



Here, we present a multiplexed assay for variant effect protocol to assess the functional impact of all possible genetic variations within a particular genomic region. We describe steps for saturation genome editing by designing and cloning of single-guide RNA (sgRNA). We then detail steps for nucleofection of sgRNA, testing drug response on variants, and amplification of genomic DNA for next-generation sequencing.

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

Sounak Sahu, Teresa Sullivan, Eileen Southon, Dylan Caylor, Josephine Geh, Shyam K. Sharan

sounak.sahu@nih.gov (S.S.)
sharans@mail.nih.gov (S.K.S.)

Highlights

Single-nucleotide variant generation for multiplexed assays for variant effects (MAVEs)

Designing single-guide RNA and ssDNA donor library for saturation genome editing (SGE)

Multiplexing of genetic variants for next-generation sequencing for variant analysis

Computational analysis to calculate indel formation and variant abundance

Sahu et al., STAR Protocols 4, 102702
December 15, 2023 © 2023
https://doi.org/10.1016/j.xpro.2023.102702

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for the saturation and multiplexing of genetic variants using CRISPR-Cas9

Sounak Sahu,[1,2,*] Teresa Sullivan,[1] Eileen Southon,[1] Dylan Caylor,[1] Josephine Geh,[1] and Shyam K. Sharan[1,3,*]

[1]Mouse Cancer Genetics Program, Centre for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

[2]Technical contact: sounak.sahu@nih.gov

[3]Lead contact

*Correspondence: sounak.sahu@nih.gov (S.S.), sharans@mail.nih.gov (S.K.S.)
https://doi.org/10.1016/j.xpro.2023.102702

SUMMARY

Here, we present a multiplexed assay for variant effect protocol to assess the functional impact of all possible genetic variations within a particular genomic region. We describe steps for saturation genome editing by designing and cloning of single-guide RNA (sgRNA). We then detail steps for nucleofection of sgRNA, testing drug response on variants, and amplification of genomic DNA for next-generation sequencing.

For complete details on the use and execution of this protocol, please refer to Sahu et al.[1]

BEFORE YOU BEGIN

The Multiplexed Assays for Variant Effect allows the study of multiple variants down to the single nucleotide level to examine how specific genetic variations affect gene function.[2–4] The protocol described here has been optimized and performed using a mouse embryonic stem cell line containing a single copy of human *BRCA2* integrated into the mouse genome. Please refer to Sahu et al. *Plos Genetics* 2023[1] for validation of the humanized mouse ES cell. However, this approach can be used for any other haploidized cell line like human haploid cell line (HAP1)[5] or generating a haploid copy of human cells like HEK293T[6] or pluripotent stem cells.[7]

sgRNA sequence design

Timing: 15 min

1. Import the fully annotated gene of interest (for example, BRCA2 (Ensembl: ENASG00000139618) from GRch38 human genome database) into a www.benchling.com worksheet.

   Note: Other cloud-based platforms can be used to import and store sequence information.

   a. Select a desired exon of BRCA2 and use the Benchling algorithm to generate a 20 bp sgRNA sequence. Benchling uses a guide RNA cutting efficiency algorithm and provides on-target and off-target scores for all sgRNA sequences.
   b. Select the sgRNA sequences which have high on-target and low off-target scores. Selected sgRNAs should be able to have a synonymous change that can disrupt the Protospacer Adjacent Motif (PAM) site. The synonymous PAM site acts as a fixed marker and is present in the donor library blocking Cas9 recutting after the editing event.

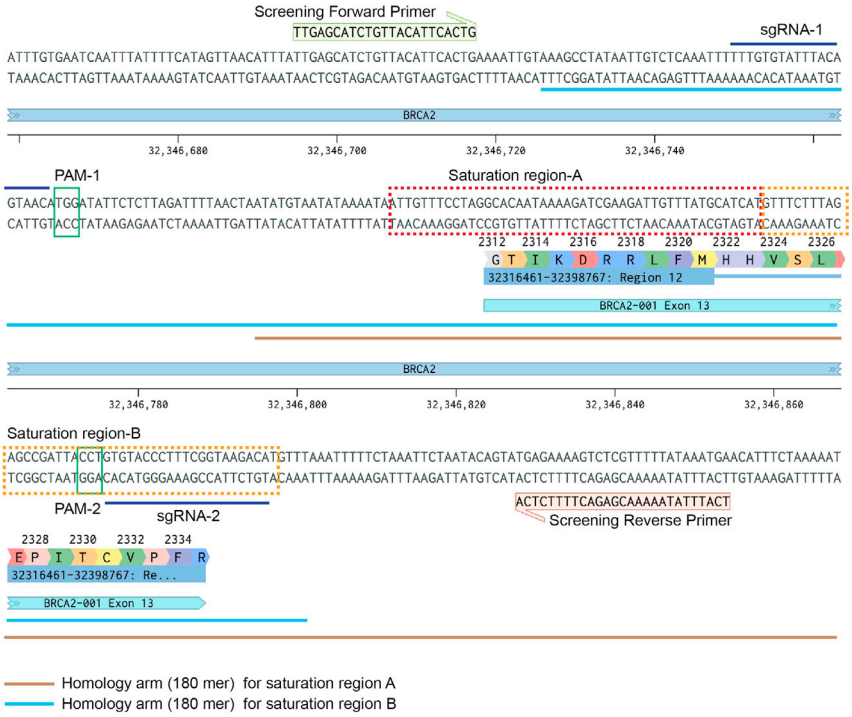


Figure 1. Designing sgRNA and oligo donor library

The nucleotide sequences exported from www.benchling.com illustrate the two sgRNAs used for saturating the exon 13 region, denoted by a dark blue line. PAM sites associated with the sgRNAs, enclosed within the green boxes, have synonymous mutations incorporated in order to disrupt the PAM sites. Region A targeted for saturation, enclosed in the red dotted box, employed both sgRNA1 and sgRNA2. The brown line denotes the homology arm used in designing the region A donor library.

For saturation of region B, enclosed in an orange dotted box, sgRNA2 was used. The blue line denotes the homology arm used in designing the region B donor library. To quantify the abundance of each variant generated following saturation genome editing (SGE), screening primers were designed to amplify this region.

c. From the selected sgRNAs, check for SpliceAI prediction to ensure that the PAM modification does not impact splicing. Refer to website https://spliceailookup.broadinstitute.org/ and put the genomic coordinate of the synonymous PAM substitution. If SpliceAI score is > 0.2 then the substitution is most likely to affect splicing. A threshold of 0.2 was used based on the recommendations of the ClinGen SVI Splicing Subgroup.[8]

d. Order the sgRNA sequence and forward and reverse primers from Integrated DNA technologies (IDT) or other oligonucleotide synthesis company. The forward primer must have 5′-CACCG followed by the 20 bp sgRNA sequence and the reverse primer must have 5′-AAAC and the reverse complement of the forward primer, followed by a "C" at the end. The PAM sequence is not included in the sgRNA sequence.

For example: (refer to Figure 1)

sgRNA sequence: 5′- TTTGTGTATTTACAGTAACATGG-3′ (PAM sequence is underlined) (refer to sgRNA-1 in Figure 1). Order the following two primers to clone the sgRNA.

Forward primer: 5′- CACCGTTTGTGTATTTACAGTAACA-3′.

Reverse primer: 5′ AAACTGTTACTGTAAATACACAAAC-3′.

Note: If the target sequence is on the bottom strand, then orient the 20 bp sgRNA sequence accordingly.

For example, sgRNA-2 in Figure 1 is 5′-GTCTTACCGAAAGGGTACACAGG-3′ (PAM sequence is underlined). Order the following two primers to clone the sgRNA.

Forward primer: 5′- CACCGGTCTTACCGAAAGGGTACAC-3′.

Reverse primer: 5- AAACGTGTACCCTTTCGGTAAGAC-3′.

Oligo donor library design

Timing: 30 min

This protocol is designed for a cloning-free single stranded oligodeoxynucleotide (ssODN) library. These 180 bp ultramers can be directly used for nucleofection.

> ***Note:*** The oligo library can also be cloned into a plasmid as a donor library containing all possible nucleotide combinations. However, deep sequencing is required to confirm equal distribution of these variants.

2. Generate an antisense oligo of 180 bp long covering the Cas9 cut site and the region of saturation.
   a. Design homology arms such that the guide RNA cut site is in the middle. If a pair of sgRNAs was used, a 40–50 bp homology arm was kept ensuring efficient knock-in (see Figure 1).
   b. The PAM modification which acts as a fixed HDR marker is inserted into all the oligos (see key resources table).
   c. Each nucleotide position across the region of saturation is changed to all three non-wild-type bases (except the PAM site). Use the standard mixed bases nomenclature code designed by IUB (International Union of Biochemistry).

Mixed bases required	Mixed-base code for DNA
A, C, T	H
G, C, T	B
A, C, G	V
A, G, T	D

For example, if the 180 bp long antisense oligo is:

AACATGTCTTACCGAAAGGGTACACAGGTAATCGGCTCTAAAGAAACATGATGCATAAACAAT
CTTCGATCTTTTATTGTGCCTAGGAAACAATTATTTTATATTACATATTAGTTAAAATCTAAGAGAATA
TCCATGTTACTGTAAATACACAAAAAATTTGAGACAATTATAGGCTTT.

The PAM site is underlined. This antisense will be changed to the following incorporating the fixed synonymous PAM modification and the respective mixed base as following:

AACATGTCTTACCGAAAGGGTACACAtGTAATCGGCTCTAAAGAAACBTGATGCATAAACAATC
TTCGATCTTTTATTGTGCCTAGGAAACAATTATTTTATATTACATATTAGTTAAAATCTAAGAGAATAT
gCATGTTACTGTAAATACACAAAAAATTTGAGACAATTATAGGCTTT.

The fixed PAM modifications are underlined and the first nucleotide of the region of saturation, which is "A" is changed to "B" mixed base giving rise to G, C, T in that position.

3. The 180 bp oligo library containing all possible nucleotide combinations was ordered from IDT at a concentration of 50 pmol per oligo.

Note: Please take caution to insert the fixed PAM modification and the mixed bases, as these oligos generally cost $300–500.

Mouse embryonic stem cells (mESCs) culture

Timing: 1–2 weeks

Note: Sterile cell culture technique is followed throughout the experiment and cells were handled in a sterile hood and incubated in a 37°C incubator with 5% CO_2. mESCs must be routinely checked for mycoplasma or other contamination.

Day 1

Timing: 1.5 h

4. Prepare SNLP feeder dishes for mESC culture:
   The mouse SNLP feeder (STO cells that are neomycin and puromycin resistant and express Leukemia inhibitory factor or LIF) have previously been treated with Mitomycin C, cell number counted, aliquoted at 2.1 × 10^7 cells in 1 mL of freezing media and stored in liquid nitrogen tank. The volume of 0.2 mL will provide a feeder layer for three 6 cm dishes or one 10 cm dish or one 6-well dish or two 96-well dishes.
   a. Gelatinize 6 cm culture dishes with 0.1% gelatin for 1 h at RT.
   b. Prepare the mESC maintenance media in the 500 mL knockout DMEM bottle.
   c. Add 4 mL of fresh mESC maintenance media to each dish.
   d. Thaw a 1 mL vial of SNLP mouse feeders (frozen at 2.1 × 10^7 cells in one mL) in a 37°C water bath until a small piece of ice is left in the vial. Collect the cells in a 15 mL falcon tube with 5 mL of fresh mESC maintenance media.
   e. Centrifuge the tube at 250 × *g* for 5 min and remove the residual media.
   f. Resuspend the cell pellet with 15 mL of fresh mESC media.
   g. Distribute equally to each 6 cm dish (1.4 × 10^6 feeder cells/6 cm dish).
   h. Incubate dishes in a 37°C incubator with 5% CO_2.

Day 2

Timing: 1 h

5. Thaw mouse ESCs and culture them in maintenance media.
   a. Thaw a vial of F7/F7 mESCs (containing 1.5 × 10^6 cells per 1 mL) in a 37°C water bath and collect the cells in a 15 mL falcon tube with 8 mL of fresh mESC maintenance media.
   b. Centrifuge the tube for 250 × *g* for 5 min and remove the residual media.
   c. Re-feed each of the feeder dishes with 4 mL of fresh mESC maintenance media.
   d. Resuspend the ES cell pellet with 1 mL of fresh media, add to one 6 cm dish and incubate in a 37°C incubator with 5% CO_2.
   e. The ESCs must be routinely re-fed, checked for confluency status and trypsinized when the cultures reach 70%–80% confluency.

Day 4

Timing: 1 h

6. Trypsinize and passage mESCs.
   a. Obtain one 6 cm dish of mouse ES cells at 80% confluency that will be passaged to one 10 cm feeder-containing dish.

b. Aspirate the media on the dish, then wash the cells with 5 mL of DPBS. Gently overlay the DPBS onto the cells making sure to not allow detachment of the cells from the dish. Aspirate the DPBS and repeat once more.
c. Add 2 mL of 0.05% trypsin/EDTA such that it covers the entire cell surface area. Incubate the dish in a 37°C incubator with 5% CO_2 for 10 min. Gently tap the sides of the cell culture dish and observe under the microscope to determine if the cells are detached from the dish.
d. Add 5 mL of mESC maintenance media to stop the action of trypsin. Gently pipette to break up the ES colonies making a single cell suspension.
e. Collect the cells in a 15 mL tube and centrifuge at 200 × *g* for 5 min to remove the residual trypsin. Add 3 mL of mES cell maintenance media and pipette gently to resuspend the cell pellet.
f. Add the entire cell suspension into 10 cm feeder-coated dishes and 8 mL each of mESC maintenance media. Incubate the dishes in a 37°C, incubator with 5% CO_2 overnight.

Day 6

Timing: 1 h

7. Repeat step-6 to Trypsinize one 10 cm dish containing mESCs and divide it into two 10 cm gelatin-coated dishes. This ensures that the cells are in the dividing phase allowing a higher recombination efficiency.
8. Prepare a 10 cm dish containing SNLP feeders that will be used to culture cells after nucleofection.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
One Shot Top10 competent cells (species *E. coli*)	Invitrogen	Cat# C404003
Chemicals, peptides, and recombinant proteins		
Gelatin (0.1%)	STEMCELL Technologies	Cat# 07903
Knockout Dulbecco's modified Eagle's medium (DMEM)	Gibco	Cat# 10829-018
Fetal bovine serum (FBS)	Gibco	Cat# 16000-044
β-Mercaptoethanol	Sigma	Cat# M3148
Glutamine-Penicillin-Streptomycin	Thermo Fisher Scientific	Cat# 10378016
Trypsin EDTA (0.5%), no phenol red	Thermo Fisher Scientific	Cat# 15400054
DPBS (1×)	Gibco	Cat# 14190-144
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	Cat# D2650
Cisplatin	Sigma-Aldrich	Cat# P4394
Olaparib	Selleck Chemicals	Cat# AZD2281
T4 polynucleotide kinase enzyme	New England Biolabs	Cat# M0201S
BbsI (high fidelity) enzyme	New England Biolabs	Cat# R0539S
T7 DNA ligase enzyme	New England Biolabs	Cat# M0318S
Recombinant DNA		
PX458-Cas9-2A-GFP	Addgene # 62988 (kind gift from Dr. Feng Zhang)	RRID-48138
Critical commercial assays		
QIAprep spin miniprep kit	QIAGEN	Cat# 27104
QIAGEN plasmid maxi kit	QIAGEN	Cat# 12165
Mouse embryonic stem cell nucleofector kit	Lonza	Cat# VPH-1001
Zymo DNA isolation kit	Zymo Research	Cat# D3020
Platinum Taq Hifi DNA polymerase	Invitrogen	Cat# 11304011
QIAquick spin columns	QIAGEN	Cat# 28115
QIAquick PCR purification kit	QIAGEN	Cat# 28106
Deposited data		
Raw and analyzed data for Exon 13 SGE	This paper	GEO: GSE238143

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human reference genome NCBI build 37, GRCh38	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Experimental models: Cell lines		
Humanized mouse embryonic stem cells (*Brca2*$^{-/-}$, *Tg* [*BRCA2*]) (clone: F7/F7)	Sharan Laboratory	Sahu et al.[1]
SNLP mouse feeders	Sharan Laboratory	Kuznetsov et al., 2008
Oligonucleotides		
sgRNA1_Forward: CACCGGTCTTACCGAAAGGGTACAC	Integrated DNA Technologies	Sahu et al.[1]
sgRNA1_Reverse: aaacGTGTACCCTTTCGGTAAGAC	Integrated DNA Technologies	Sahu et al.[1]
sgRNA2_Forward: CACCGTTTGTGTATTTACAGTAACA	Integrated DNA Technologies	Sahu et al.[1]
sgRNA2_Reverse: AAACTGTTACTGTAAATACACAAAc	Integrated DNA Technologies	Sahu et al.[1]
Exon 13_Screening_forward: TTGAGCATCTGTTACATTCACTG	Integrated DNA Technologies	Sahu et al.[1]
Exon 13_Screening_reverse: TGAGAAAAGTCTCGTTTTTATAAATGA	Integrated DNA Technologies	Sahu et al.[1]
Saturation-A (180 mer Ultramer mixed base oligo library pool): AACATGTCTTACCGAAAGGGTACACAtGTAATCGGCTCTAA AGAAACBTGATGCATAAACAATCTTCGATCTTTTATTGTGCC TAGGAAACAATTATTTTATATTACATATTAGTTAAAATCTAAG AGAATATgCATGTTACTGTAAATACACAAAAAATTTGAGAC AATTATAGGCTTT	Integrated DNA Technologies	Sahu et al.[1]
Saturation-B (180 mer Ultramer mixed base oligo library pool): ATTTTTAGAAATGTTCATTTATAAAAACGAGACTTTTCTCATA CTGTATTAGAATTTAGAAAAATTTAAACATHTCTTACCGAAA GGGTACACAtGTAATCGGCTCTAAAGAAACATGATGCATAA ACAATCTTCGATCTTTTATTGTGCCTAGGAAACAATTATTTT ATATTACATATT	Integrated DNA Technologies	Sahu et al.[1]
Software and algorithms		
Genomic data analysis	RStudio	https://posit.co/download/rstudio-desktop/
Sanger sequencing analysis	4Peaks	https://nucleobytes.com/4peaks/
Gene retrieval, sgRNA design, ssODN design	Benchling	https://www.benchling.com/
Genome editing analysis	CRISPResso2	http://crispresso2.pinellolab.org/submission
Other		
Vac-Man vacuum manifold	Promega	Cat# A7231
Nucleofector transfection 2B device	Lonza	Cat# AAB-1001

MATERIALS AND EQUIPMENT

Lonza nucleofector 2B device

We have used Nucleofector 2B for our SGE experiments using the program code: A030

> *Alternatives:* Lipofectamine-based transfection or other electroporation platforms like Lonza-4D Nucleofector and Thermo Fisher Scientific Neon can be used. The voltage and pulse settings must be optimized depending on the cell line and transfection efficiency.

To make 1 L of Superbroth, add the following reagents to a 2-L flask; swirl until the solutes have dissolved.

Reagent	Final concentration	Amount
Deionized water		950 mL
Yeast extract	2%	20 g
NaCl	0.5%	5 g
NaOH (1 N)	1 N	5 mL
Tryptone	3.2%	32 g

Add a stir bar to the flask and mix on a stir plate until the reagents have dissolved. Adjust the volume of the solution to 1 L with deionized water, if needed. Sterilize by autoclaving for 20 min at 15 psi on liquid cycle. Allow the Superbroth to cool until it is at room temperature and use immediately or store in 1-L sterile bottle.

mESC maintenance media		
Reagent	Final concentration	Amount
Knockout DMEM		500 mL
Fetal Bovine Serum	15%	90 mL
Glutamine-Penicillin-streptomycin (100×)	1×	6 mL
β-Mercaptoethanol (100×)	1×	6 mL

Note: The media must be stored at 4°C and can be used for 2–3 weeks. The media must be warmed up prior to adding to ES cells.

STEP-BY-STEP METHOD DETAILS

Cloning of sgRNA

The protocol below describes steps for cloning 20 bp oligos into a sgRNA cassette and into the PX458-Cas9-2A-GFP plasmid. However, we have also used this protocol to clone into PX330-Cas9 and PX459-Cas9-2A-Puromycin.

Day 1 step by step method to clone sgRNA

Timing: 5 h

1. Anneal oligonucleotides and clone in PX458-Cas9-2A-GFP vector.

Note: Thaw all the reagents on ice and keep the enzyme in a StrataCooler benchtop cooler during sample preparation. If multiple sgRNAs are to be cloned, prepare the master mix as shown below and add each forward and reverse oligo pair into the appropriate PCR tube.

a. Briefly spin down the lyophilized oligo tubes prior to opening them.
b. Add non-DEPC treated water to lyophilized primers to equal a 100 μM concentration, vortex, and briefly centrifuge the tube. For example, if the oligo is 29.2 nmol, add 292 μL of distilled water to equal 100 μM.
c. Label a PCR tube for each sgRNA oligo pair and add reagents below:

Reagent	Amount
10× T4 Ligation Buffer *	1 μL
T4 Polynucleotide kinase enzyme (10 U/μL)	1 μL
sgRNA_Forward oligo (100 μM)	1 μL
sgRNA_Reverse oligo (100 μM)	1 μL
Water	6 μL
Total	**10 μL**

*Do not use T4 PNK buffer as this buffer does not contain ATP; T4 ligation buffer is essentially the same composition but contains ATP.

d. Briefly centrifuge PCR tube & run on thermocycler using the following program:

Temperature	Duration	Cycle
37°C	30 s	1
95°C	5 min	1
25°C	Forever (Ramp 0.1°C/s)	

e. At the end of run, briefly centrifuge tubes. The annealed phosphorylated oligos can be stored at -20°C.

f. Add 1 μL of annealed product to 199 μL of distilled water in a 0.2 mL PCR-strip tube. Vortex properly and briefly centrifuge the strip-tube before use.

2. Set up digestion/ligation reaction
   a. Label a PCR tube for each oligo duplex and add reagents below:

Reagent	Amount
Plasmid DNA (100 ng) (PX458-Cas9-2A-GFP)	1 μL
Diluted annealed oligo (from step-1f) (10 pmol)	2 μL
10× TANGO buffer	2 μL
10 mM DTT	1 μL
10 mM ATP	1 μL
BbsI Enzyme (5 U/μL)	1 μL
T7 DNA ligase enzyme (400 U/μL)	0.5 μL
Water	11.5 μL
Total	**20 μL**

   b. Include a negative control that is 2 μL of water instead of annealed oligo.
   c. Briefly centrifuge tube & run on thermocycler using the following program.

Temperature	Time	
37°C	5 min	6 cycles
21°C	5 min	
12°C	Forever	Hold

3. Set up exonuclease Reaction.
   a. Label a PCR tube for each ligation reaction and add reagents below:

10× plasmid safe buffer	1.5 μL
10 mM ATP	1.5 μL
Plasmid-Safe Exonuclease	1 μL
Ligation reaction (from Step-2c)	11 μL
Total	**15 μL**

   b. Briefly centrifuge tube & run on thermocycler using the following program:

Temperature	Duration	Cycle
37°C	30 min	1
70°C	30 min	1
12°C	Forever	

4. Use One Shot Top10 competent cells to transform the ligation mix.

Note: Ampicillin-containing LB plates must be prepared and stored at 4°C before transformation. One Shot Top10 competent cells must be taken out of −80°C and immediately put on ice.

   a. Add 1 μL of product from step-3b above and 10 μL of One Shot Top10 competent cells into a 1.5 mL Eppendorf tube.
   b. Incubate the tube on ice for 10 min.
   c. Perform heat shock step by adding the tube to a 42°C water bath for 30 s.
   d. Immediately incubate the tube on ice for 2 min.
   e. Add 100 μL of SOC (without antibiotic) to the tube and incubate for 1 h in a 37°C bacterial incubator.

f. Label Ampicillin-containing LB plates and pre-warm in a 37°C bacterial incubator.
g. After incubation, transfer all of the mixture from the 1.5 mL Eppendorf tube to the Ampicillin LB plate and using a bacterial spreader, evenly spread the mixture on top of the LB agar.
h. Incubate the plates overnight in a 37°C bacterial incubator.

Day 2: Bacterial colony screening

Timing: 30 min

5. Observe the colony numbers in the bacterial plates. Ideally the sgRNA cloning plates should have 5–10 times more colonies than the control plates. (See troubleshooting 1).
6. Using a sterile 10 μL pipet tip, pick up a colony, add to a bacterial culture tube containing 3 mL of LB and 3 μL of (1000×) Ampicillin and incubate overnight in a 37°C bacterial shaking incubator. Repeat picking 5 colonies total for each bacterial plate. In addition, make a replicate plate by adding a small volume of each picked colony to a second Ampicillin-containing LB plate. Label the plate at the spot where each clone was deposited and incubate the plate overnight in a 37°C bacterial incubator.

Day 3: Plasmid isolation and Sanger sequencing

Timing: 6 h

7. Isolate plasmid DNA and confirm presence of correct sgRNA insert by Sanger sequencing.
   a. After overnight culture, centrifuge the tubes at 4500 × *g* for 10 min to pellet the bacterial cells.

   Pause point: The tubes can be stored at -20°C for a few weeks after removing the supernatant.

   b. Isolate the plasmid using the QIAprep Spin Miniprep Kit using the manufacturer's instruction https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/plasmid-dna/qiaprep-spin-miniprep-kit. Elute the plasmid DNA in 50 μL of distilled water.
   c. Quantify the concentration of the plasmid using a Nanodrop and perform a Sanger sequencing using the U6 forward primer. Human U6 promoter forward primer: GACTATCA TATGCTTACCGT.
   d. Using 4Peaks (or other sequencing analysis software like SnapGene or ApE-A plasmid Editor), check for the insertion of 20 bp sgRNA sequence following CACCG-sequence.

Day 4: MaxiPrep to produce high yield of plasmid DNA for nucleofection

Timing: 6 h

8. After identifying the clone with correctly inserted sgRNA, pick the same clone from the replicate dish and add to 50 mL of Superbroth containing 50 μL of 1000× Ampicillin in a 200 mL flask. Incubate overnight in a shaking 37°C bacterial incubator.

   Note: LB medium can be used as a substitute if Superbroth is not accessible, however, it is worth noting that using Superbroth has been observed to result in higher plasmid concentrations.

   a. Collect the 50 mL bacterial culture in a falcon tube.
   b. Make a glycerol stock of each sgRNA clone by adding 500 μL of bacterial culture containing sgRNA from the 50 mL tube with 500 μL of 80% glycerol in a labeled 1.5 mL Eppendorf tube. Vortex the tube and store at −80°C freezer for future use.

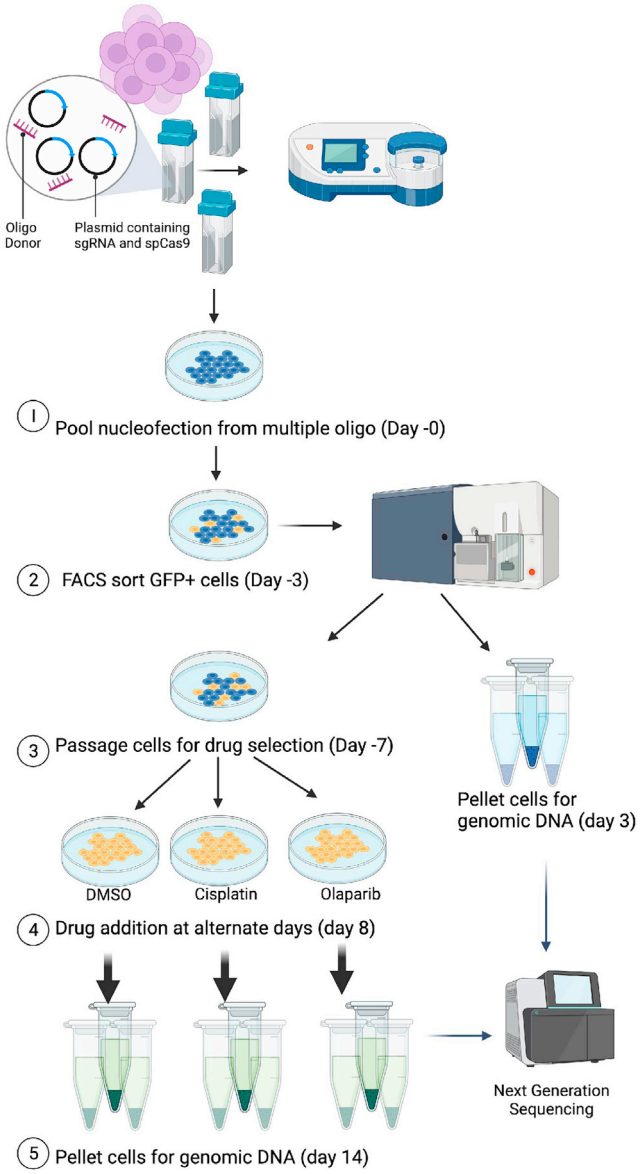


Figure 2. Experimental schematic showing different steps of saturation genome editing (SGE)

In this experimental workflow, mouse embryonic stem cells (mESCs) were initially nucleofected with a combination of sgRNA and a plasmid carrying GFP-tagged-Cas9 along with single-stranded DNA oligo donors. At day 3 post-nucleofection, GFP-expressing mESCs were FACS-sorted and half of this population was utilized for genomic DNA isolation. The edited locus was amplified with screening primers and subsequently deep sequencing was performed via Next-Generation Sequencing (NGS) to quantify the abundance of variants generated after CRISPR-SGE.The remaining GFP+ cells were cultured and treated with cisplatin and olaparib treatment from day 8 to day 14, aiming to induce DNA damage and evaluate the response of edited cells to these compounds. After the treatment period, genomic DNA was once again isolated from these cells at day 14 and the edited locus was amplified with the screening primers. Deep sequencing via Next-Generation Sequencing (NGS) was performed to calculate the dropout frequency of the genetic variants, providing comprehensive insights into the outcomes of genome editing and the cellular response to DNA-damaging agents.

c. Centrifuge the 50 mL falcon tube for 20 min at 15000 g in an ultracentrifuge.
d. After centrifugation, discard the supernatant and either proceed with plasmid isolation or store the pellet at −20°C.

e. Isolate the plasmid using the QIAGEN Plasmid Maxi kit using the manufacturer's instruction. https://www.qiagen.com/us/resources/resourcedetail?id=0bc0a0fe-480e-4e5e-b6e2-e75a654dd791&lang=en. Ideally 300 μg of endotoxin-free plasmid DNA should be obtained from 50 mL of overnight culture.
f. Check the concentration and quality of the plasmid DNA using NanoDrop. The A260/280 ratio should be 1.9.
g. Store the isolated plasmid DNA in −20°C in 50 μL aliquots to reduce multiple freeze thaw cycles.

Saturation genome editing of *BRCA2*

Saturation genome editing (SGE) is a powerful high-throughput genome editing technique designed to systematically introduce every possible nucleotide variant at a specific genomic location. It leverages the precision of the CRISPR-Cas9 technology and employs synthetic donor molecules to facilitate the targeted mutation process.

In the SGE workflow, cells are transfected with a sgRNA-Cas9 plasmid and successful transfection is identified by the presence of GFP fluorescence at day 3. The cell population is then divided into two groups: one for immediate genomic DNA isolation to determine the initial variant composition and another cultured until day 14 and in the presence of DNA damaging agents like cisplatin and olaparib. After day14 genomic DNA was isolated to assess the abundance of variants that may have been lost or maintained over time (Figure 2).

The key principle of SGE imparts that variants exhibiting a loss in abundance over time are considered non-functional, indicating that they negatively impact cell fitness. Conversely, variants that do not show a fitness defect are deemed functional. This systematic approach allows for the classification of genetic variants based on their effect on cell fitness. Furthermore, multiplexing enables the study of numerous variants simultaneously within a single experiment.

Nucleofection of sgRNA and donor library

Note: To reduce variability of knock-in efficiency, the same set of sgRNA and oligo library is used for three nucleofections each containing 3.3 million mESCs. A total of 10 million mESCs from these three replicates were pooled into one 10 cm dish after nucleofection which represents replicate 1. The same experimental strategy is repeated to yield a replicate 2.

mESCs preparation and nucleofection

Day 0

Timing: 1 h

Note: Before starting nucleofection, add 10 mL of fresh media to the appropriate number of 10 cm feeder plates and keep 5 mL of warm media in a 37°C water bath. All the reagents for nucleofection are kept on ice and the cell culture medium is warmed up before adding to the cells.

9. Prepare mESC and the sgRNA for nucleofection.
   a. Prepare sgRNA mix for nucleofection by adding the following reagents in a 1.5 mL Eppendorf tube. Store all tubes at RT.

Reagents	Amount for 1 nucleofection (~3.3 million mESCs)	Amount for 3 nucleofection (~10 million mESCs)
sgRNA-Cas9 Plasmid (Px458)	3 μg	9 μg
Oligo Pool library	6.5 μg	19.5 μg

Note: Keep the plasmid concentration and the oligo donor concentration at 2 μg/μL such that the total volume to be added in step-9d is less than 15 μL.

b. Trypsinize mESCs and collect the cells in a 15ml falcon tube. Do not centrifuge the cells. Quantify the cell number in the trypsin/media/cell mixture using a cell counter apparatus. Calculate the volume needed to equal 10 million cells. Add the cells to a new 15 mL falcon tube and centrifuge at 100 × *g* for 8 min more than 1x.
c. Discard the media and add 280 μL of nucleofector solution to the cell pellet containing 10 million ESCs. Gently mix to resuspend the mESC pellet.
d. Add the sgRNA-Cas9 plasmid/oligo donor library mix from step-9a to the cell suspension and gently mix.
e. Transfer 100 μL of the cells/plasmid mix to the bottom of a nucleocuvette and give a gentle tap such that the entire solution settles to the bottom.
f. Repeat step-9e for the other two nucleocuvettes.
g. Set up the Lonza nucleofector 2B machine with program A030 and transfect each cuvette.
h. Immediately add 500 μL of warm mESC maintenance media (from the 15 mL falcon tube incubated in the 37°C water bath) to each nucleocuvette. Collect the transfected cells from the 3 nucleocuvettes and transfer them all into the 10 cm feeder dish. Incubate the dish in a cell-culture incubator.
i. Keep a small dish of mESCs as a non-transfected control for gating cells during FACS on day 3.

Day 1

Timing: 30 min

Note: Before you start, warm up the fluorescence microscope for 15 min for the laser to strike off.

10. After 24 h, observe the mESCs under a fluorescence microscope to check for GFP^+ cells. Compare the transfected cells dish with the untransfected control. However, the number increases after 48 h. Cell death is also evident at 24 h post nucleofection. (see troubleshooting 2).

FACS sorting for GFP+ mESCs post nucleofection

Day 3

Timing: 5 h

Note: Six 6 cm feeder dishes with 5 mL of fresh mES cells maintenance media are required.

Before you start, warm up the fluorescence microscope for 15 min for the laser to strike off.

11. Re-feed the 6 cm dish with fresh mES cells maintenance media and incubate the dish in a cell-culture incubator.
    a. Observe the increase in GFP^+ mESCs in the 10 cm dish.
    b. Aspirate the media on the dish, then wash the cells with 2 mL of DPBS. Gently overlay the DPBS onto the cells making sure to not allow detachment of the cells from the dish. Aspirate the DPBS and repeat once more.
    c. Add 2 mL of 0.05% trypsin/EDTA such that it covers the entire cell surface area. Incubate the dish in a 37°C incubator with 5% CO_2 for 10 min. Gently tap the sides of the cell culture dish and observe under the microscope to determine if the cells are detached from the dish.
    d. Add 5 mL of mESC maintenance media to stop the action of trypsin. Gently pipette to break up the ES colonies making a single cell suspension.

e. Collect the cells in a 15 mL falcon tube and centrifuge at 200 × *g* for 5 min to remove the residual trypsin.
f. Discard the supernatant, add 1 mL of PBS and resuspend the pellet.
g. Gently pass the cell suspension through 40 μm cell filter to remove ESC clusters.
h. Using forward vs. side scatter and the intensity of GFP fluorescence, sort all the GFP^+ cells using BD FACS Flow cytometer (Figure 3) (see troubleshooting 3).
i. Count the total number of GFP+ mESCs and collect directly into mESC maintenance media. Do not centrifuge the cells after FACS sorting. Collect 10^6 mESCs for genomic DNA isolation, centrifuge to pellet the cells and store at −80°C until all time points have been completed.
j. In addition, re-plate 1–1.5 million mESCs into one new 6 cm feeder dish for subsequent culture and addition of drug at day 8.
k. Re-feed the 6 cm dish with mESC maintenance media on day 4 and day 6.

Testing drug sensitivity of variants

Day 7

Timing: 1 h

Note: Prepare six 10 cm feeder dishes with 10 mL of fresh mESC maintenance media per replicate.

12. Prepare variant-containing mESCs to test their sensitivity to cisplatin and olaparib.
    a. Aspirate the media on the dish, then wash the cells with 1 mL of DPBS. Gently overlay the DPBS onto the cells making sure to not allow detachment of the cells from the dish. Aspirate the DPBS and repeat once more.
    b. Add 1 mL of 0.05% trypsin/EDTA such that it covers the entire cell surface area. Incubate the dish in a 37°C incubator with 5% CO_2 for 10 min. Gently tap the sides of the cell culture dish and observe under the microscope to determine if the cells are detached from the dish.
    c. Add 4 mL of mESC maintenance media to stop the action of trypsin. Gently pipette to break up the ES colonies making a single cell suspension.
    d. Collect the cells in a 15 mL falcon tube and centrifuge at 200 × *g* for 5 min to remove the residual trypsin. Add 5 mL of mES cell maintenance media and pipette gently to resuspend the cell pellet.
    e. Count the total number of cells present in each sample using preferred cell counter.
13. Add 1 million cells into six 10 cm feeder dishes making sure to distribute the cells evenly over the dish. Two dishes are used each for DMSO (control), cisplatin and olaparib treatment. Incubate the dishes in a 37°C incubator with 5% CO_2.

Treatment with DNA damaging drugs

Day 8

Timing: 1 h

14. Make working concentrations for cisplatin and olaparib as follows.

Reagent	Amount	Final concentration
cisplatin (33 mM stock)	10 μL	33 μM
mESC maintenance media	10 mL	

Reagent	Amount	Final concentration
olaparib (10 mM stock)	10 μL	10 μM
mESC maintenance media	10 mL	

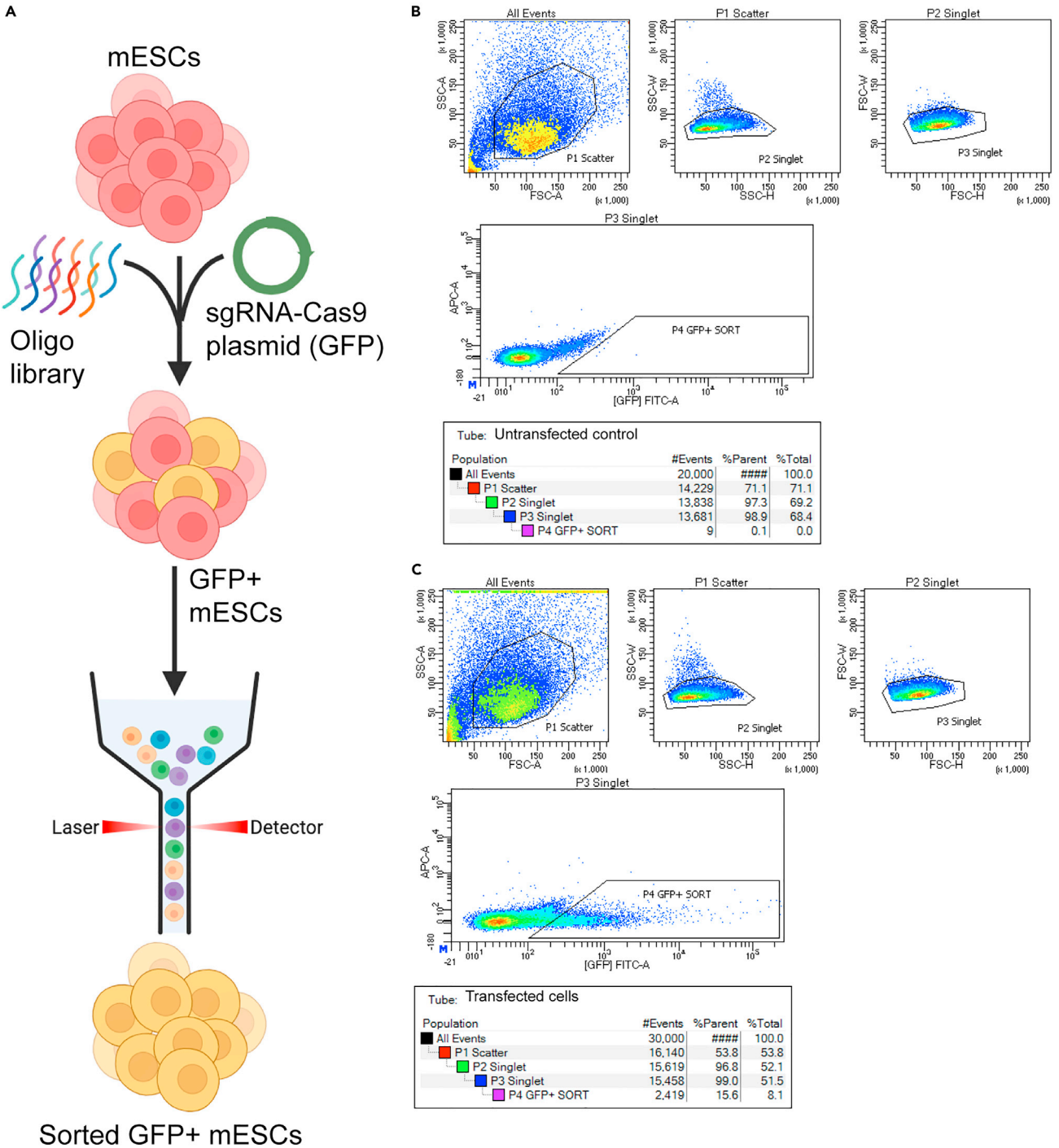


Figure 3. Gating strategy to collect GFP+ mESCs post nucleofection using FACS

(A) Mouse embryonic stem cells (mESCs) were initially nucleofected with sgRNAs, which were cloned into the Px458 plasmid. This plasmid carried Cas9 fused with 2A-GFP. At day 3 following nucleofection, transfected mESCs were selected based on the expression of GFP using Fluorescence-Activated Cell Sorting (FACS). Cells were gated initially using forward vs. side scatter parameters, and subsequent gating included considerations for cell size and exclusion of doublets.

(B) Untransfected cells that did not exhibit any GFP fluorescence were utilized as a reference to establish gating parameters for GFP fluorescence.

(C) Cells displaying high GFP expression were gated, and individual cells were sorted into separate tubes. It's noteworthy that the transfection efficiency for this procedure typically ranged from 8% to 15% of the total cell population.

Reagent	Amount	Final concentration
DMSO (final stock)	10 μL	10 μM
mESC maintenance media	10 mL	

Note: The working concentration tubes can be stored at 4°C for a week.

15. Aspirate media from each of the dishes and add 0.4 μM final concentration of cisplatin or 50 nM final concentration of olaparib in the appropriate dishes. For the DMSO control dishes, add the same volume as the cisplatin drug.
    For example, make 25 mL of media containing the drugs and add 10 mL of media to the appropriate dish per drug condition.
    a. Add 302.5 μL of 33 μM cisplatin solution to 25 mL of mESC maintenance media.
    b. Add 125 μL of 10 μM olaparib solution to 25 mL of mESC maintenance media.
    c. Add 302.5 μL of DMSO solution to 25 mL of mESC maintenance media.
    d. Add 10 mL of drug-containing media to the appropriate dish.
    e. Incubate in 37°C incubator with 5% CO_2.
    f. Replace media with fresh cisplatin-containing, olaparib-containing or DMSO-containing media on day 10 and day 12.

Note: Add the same volume of media containing the drugs each day the media is replaced.

Collecting cells after drug treatment

Day 14

Timing: 3 h

16. On day 14, collect the mESCs for genomic DNA isolation.
    a. Aspirate the media on the dish, then wash the cells with 2 mL of DPBS. Gently overlay the DPBS onto the cells making sure to not allow detachment of the cells from the dish. Aspirate the DPBS and repeat once more.
    b. Add 2 mL of 0.05% trypsin/EDTA such that it covers the entire cell surface area. Incubate the dish in a 37°C incubator with 5% CO_2 for 10 min. Gently tap the sides of the cell culture dish and observe under the microscope to determine if the cells are detached from the dish.
    c. Add 5 mL of mESC maintenance media to stop the action of trypsin. Gently pipette to break up the ES colonies making a single cell suspension.
    d. Collect the cells in a 15 mL falcon tube and centrifuge at 200 × *g* for 5 min to remove the residual trypsin.
    e. Discard the media and wash once with 10 mL of DPBS. Centrifuge the tubes at 200 × *g* for 5 min, discard the supernatant and store the cell pellet in −80°C for genomic DNA isolation.

Note: The cell pellets can be stored for 4–5 months if kept at −80°C.

Isolation of genomic DNA

Timing: 3 h

In this protocol, the genomic DNA is isolated using a Zymo Quick DNA miniprep kit. Other genomic DNA isolation kit or in-house prepared reagents can be used for isolation, but the quality of genomic DNA plays a major role in PCR and for NGS library preparation.

Buffer preparation: Add β-mercaptoethanol (BME) to the genomic lysis buffer (Zymo) to a final dilution of 0.5% (v/v). Add 250 μL per 50 mL of BME.

17. Isolate genomic DNA from cell pellets collected at day 3 and at day 14 from DMSO control, cisplatin and olaparib treatment.
    a. Thaw the falcon tubes containing cell pellets at room temperature after removing from −80°C freezer.
    b. Add 10 mL of Genomic lysis buffer to the cell pellet. Mix completely by vortexing for 10–20 s and mix the entire suspension using a 10 mL serological pipette, then let stand for 5 min at room temperature. Repeat vortexing for 10 s and then let stand for another 5 min at room temperature.

    Note: The solution becomes viscous during lysis step so caution must be taken to reduce loss of sample.

    c. Transfer 500 μL of the mixture to twenty Zymo-spin IICR columns in collection tubes. Centrifuge at 10000 × *g* for 4 min and discard the flow through.
    d. Add 300 μL of DNA pre-wash buffer to the spin column. Centrifuge at 10000 × *g* for 2 min and discard the flow through.
    e. Add 500 μL of g-DNA wash buffer to the spin column. Centrifuge at 10000 × *g* for 4 min and discard the flow through.
    f. Centrifuge the empty column at 10000 × *g* for 1 min and discard the residual flow through.
    g. Transfer the spin column to a new 1.5 mL microcentrifuge tube. Add 40 μL of nuclease-free water and incubate at room temperature for 5 min to completely saturate the column.
    h. Centrifuge at 16000 × *g* for 1 min and collect the genomic DNA from all twenty columns.
    i. Using a NanoDrop UV-Vis spectrophotometer quantify the DNA concentration and the quality using A260/280 ratio. For example, using 10^7 mESCs you would expect a concentration of 350–500 μg/μL in a total volume of 800 μL. (see troubleshooting 4).
    j. Store the genomic DNA at 4°C if used for PCR in a week or store in a −80°C freezer for long-term storage.

PCR to amplify edited region of interest

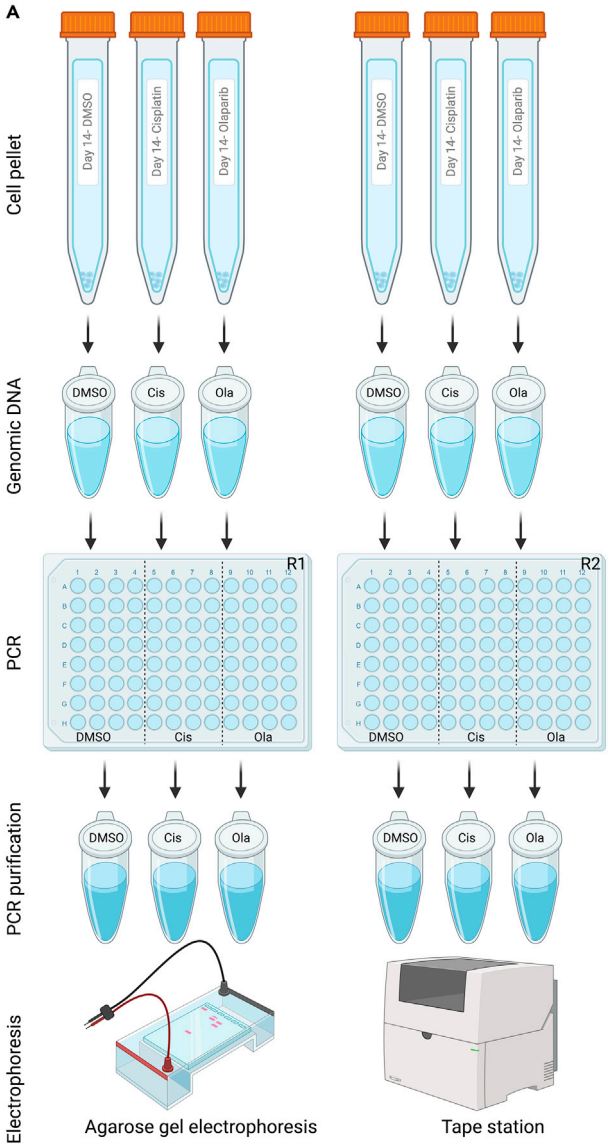
In this protocol, we have used Platinum Taq Hifi DNA polymerase. Other polymerases can be used but caution must be taken to ensure the Taq polymerase does not introduce nucleotide errors. Screening primers were designed using Primer 3 Plus software such that the amplicon size ranges between 250-450 bp. For screening exon 13 we have used the following primers:

Exon 13_Screening_forward: TTGAGCATCTGTTACATTCACTG.

Exon 13_Screening_reverse: TGAGAAAAGTCTCGTTTTTATAAATGA.

The size of the amplicon will be 265 bp.

18. Prepare PCR master mix
    a. Label 6 tubes corresponding to D1, C1, O1 and D2, C2 and O2 corresponding to DMSO-replicate-1, Cisplatin-replicate-1, Olaparib-replicate-1, DMSO-replicate-2, Cisplatin-replicate-2, Olaparib-replicate-2. Mix the following reagents for each samples using the following reagents-


A
Cell pellet
Genomic DNA
PCR
PCR purification
Electrophoresis
Day 14- DMSO
Day 14- Cisplatin
Day 14- Olaparib
DMSO
Cis
Ola
R1
R2
Agarose gel electrophoresis
Tape station

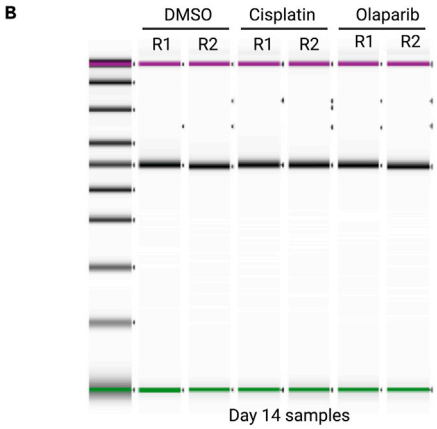

B
DMSO
Cisplatin
Olaparib
R1
R2
R1
R2
R1
R2
Day 14 samples

Figure 4. Experimental set up for isolating genomic DNA and purification of PCR amplicons

(A) On day 14 post-nucleofection, cell pellets from both replicates (R1 and R2) were collected for each of the following conditions: DMSO treatment, cisplatin treatment, and olaparib treatment. Genomic DNA was isolated from these cell pellets. For each condition, PCR amplification was performed and distributed across 32 wells. A 96-well plate was employed for the PCR, partitioned into three sections, each designated for DMSO, cisplatin, or olaparib. Subsequently, the PCR products were purified, and their separation was achieved through agarose gel electrophoresis and a TapeStation.

(B) An illustrative electropherogram was provided, demonstrating the successful separation of purified PCR products obtained from the day 14 samples.

Reagent	Amount
10× Hifi buffer	100 μL
50 mM $MgSO_4$	40 μL
10 mM dNTP mix	20 μL
Forward primer (100 μM)	3 μL
Reverse primer (100 μM)	3 μL
Platinum Taq Hifi DNA polymerase	4 μL
Water	Volume make up to 1000 μL

i. Add 25 μg of genomic DNA per condition to each tube, vortex the tubes followed by quick spin.
ii. Using a stepper pipette add 32 μL of PCR mix into 32 wells of a 96 well plate (as described in Figure 4A).

b. PCR and purification of amplicon.
i. Spin the 96 well plate & run on thermocycler using the following program.

Steps	Temperature	Time	Cycles
Initial Denaturation	94°C	2 min	1
Denaturation	94°C	30 s	24 cycles
Annealing	55°C	30 s	
Extension	68°C	5 min	
Final extension	68°C	5 min	1
Hold	4°C	Forever	

Pause point: At the end of run, quick spin the PCR plates and the plates can be stored at 4°C overnight.

Note: Hifi Taq polymerase is optimized for 55°C annealing and this temperature must be optimized if different polymerase like Q5 (NEB) is used. The number of cycles must not be increased more than 24 as they may increase the chance of preferential amplification.

19. Collect the PCR product from individual wells of a 96 well plate into a 15 mL falcon tube corresponding to individual samples. PCR purify the product using QIAGEN PCR purification kit as per manufacturer's protocol. https://www.qiagen.com/us/resources/resourcedetail?id=95f10677-aa29-453d-a222-0e19f01ebe17&lang=en.

Note: PCR primers must be optimized such that a single band of the amplicon is obtained. If multiple bands are observed, then design a new pair of primers or use gel extraction to purify the band of interest.

20. Elute the PCR product with 30 μL of nuclease-free water and run 2 μL of product on 1% agarose gel to confirm correct size of the amplicon.

21. The amplicon products are sent for Next Generation Sequencing and DNA libraries are prepared with Illumina adapters.

EXPECTED OUTCOMES

Figure 4B shows a representative image from TapeStation showing purified PCR amplicons. Our lab uses Azenta for NGS, but other NGS providers can be used. Azenta performs quality control of the amplicon product: TapeStation, Qubit concentration. This protocol produces 70 μL of purified sample with 350–500 ng/μL of PCR product. Azenta recommends 20 μL of purified sample at 25 ng/μL to run a quality control test, prepare libraries and to load onto the flow cell. We deep sequence the libraries for 3 million reads per sample necessary for accurate variant calling. Custom scripts can be written for variant calling or CRISPResso[9] can be used to study the abundance of variants using the web interface or command line.

QUANTIFICATION AND STATISTICAL ANALYSIS

Timing: 2 h (for step 1)

Timing: 5 h (for step 2)

In this section, we describe a step-by-step approach for data analysis including quality control metrics. We have used CRISPResso2, a versatile genome editing analysis tool that automates various tasks, including filtering low-quality reads, trimming adapters, aligning reads to reference sequences, quantifying editing outcomes.[9] Additionally, it offers visualizations of indel distributions and positions for cleaving nucleases and substitution distributions, aiding in the comprehensive assessment of genome editing outcomes and allele frequencies.

1. Installation of CRISPResso2.
   CRISPResso2 can be used as a web interface http://crispresso2.pinellolab.org/submission or set up through the Bioconda package manager using conda, or it can be executed within the Docker container system.

   Note: The web interface is time consuming if the sequencing files are large and it is recommended to use the conda or docker system.

   In this protocol, we have used CRISPResso2 through the Docker containerization system, which simplifies its installation and configuration. To get started, follow these steps.
   a. Begin by downloading and installing Docker from https://docs.docker.com/engine/installation/.
   b. After Docker is installed, configure it to access the hard drive and allocate sufficient memory (at least 4 GB).
   c. To run CRISPResso2, make sure that Docker is running, then open a command prompt (for Mac) or Powershell (for Windows). Change current directory to the location where the NGS sequencing files are stored.
   d. Execute the following command:

```
>docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso -h
```

This will allow access to CRISPResso2 seamlessly through Docker without the need for additional package configuration or installation.

2. NGS data quality control and sequencing analysis.

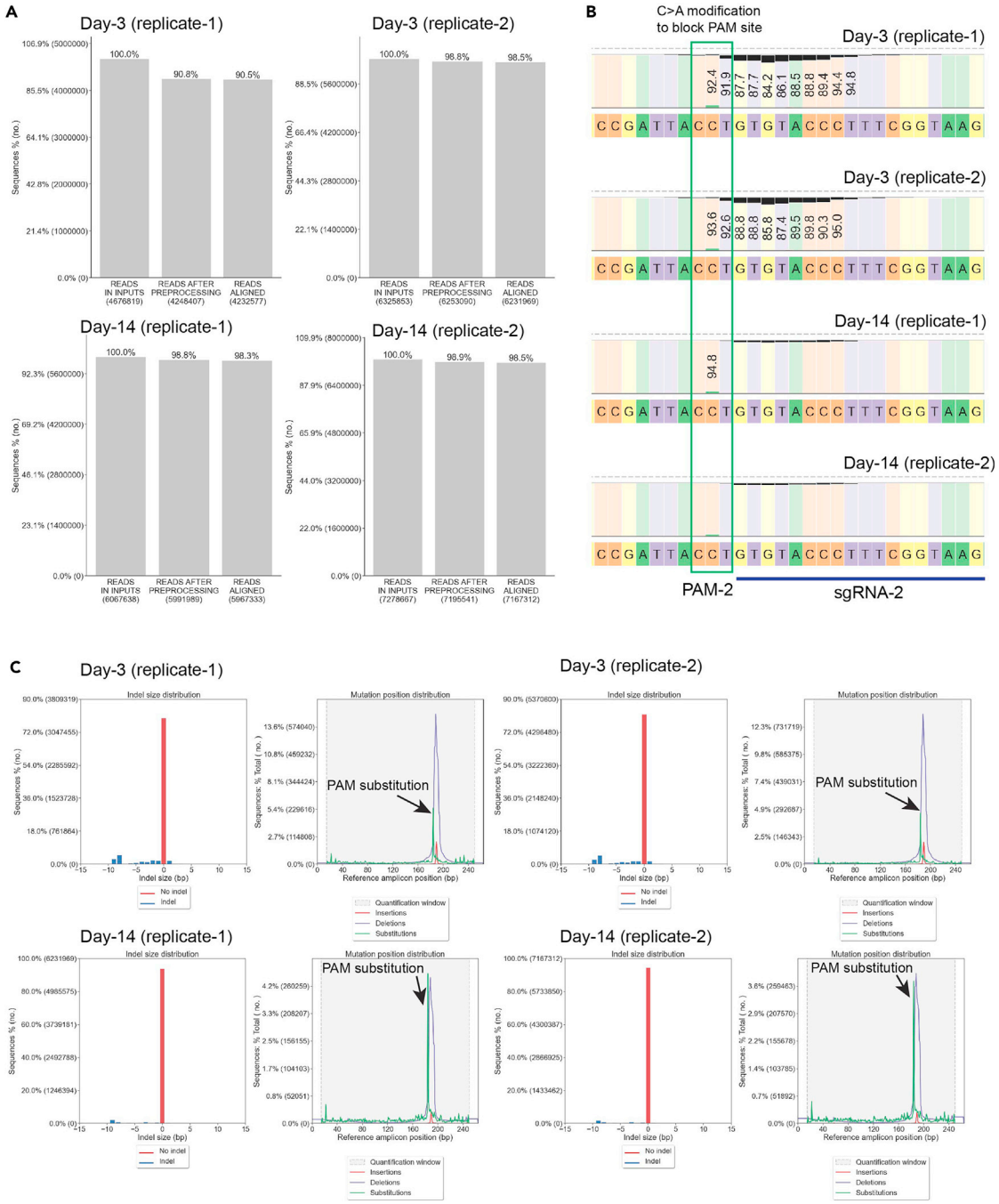


Figure 5. Sequencing data analysis to quantify genome editing outcome

(A) Barplots representing the total number of reads sequenced per sample and the reads aligned to the reference amplicon after adapter trimming and quality control.

(B) Nucleotide abundance plot showing the presence of deletions around the sgRNA cut site. The PAM site is represented by a box where the TCC PAM is changed to TAC to block PAM site recognition.

(C) Plot representing the percentage of sequences with different sizes of indel and the presence of fixed PAM-synonymous substitution (arrowhead). The number of indels leading to loss-of-function is reduced from day 3 to day 14.

a. NGS can be performed using an Illumina MiSeq instrument. The resulting sequencing data (base call files, or .bcl files) should be demultiplexed and converted into zipped fastq files using bcl2fastq v2.0. At 1000× coverage, more than 3 million paired end reads per sample can be obtained (Figure 5A).
b. Perform the initial sequencing quality control analysis using CRISPResso2 which generates a plot describing the total number of reads aligned to the reference amplicon after adapter trimming and reads filtering. CRISPResso2 uses Needleman-Wunsch alignment with gap-open -20, gap-extend-2 and gap incentive 1. Generally, 90%–98% of the reads are aligned (Figure 5A).
c. Example data for exon 13 dataset is available in the NCBI GEO database (Accession number: GSE238143). Download the sequence fastq.gz files from exon 13b folder change the current directory using the command prompt to the location where the NGS sequencing files are stored.
d. Execute the following command:

```
>docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso --fastq_r1 13bDMSOR2_R1_001.fastq.gz --fastq_r2
13bDMSOR2_R2_001.fastq.gz --amplicon_seq TTGAGCATCTGTTACATTCACTGAAAATTGTAAAGCCTATAATTGTCTCAAATTTTTTGTGTATTTACAGTAACATG
GATATTCTCTTAGATTTTAACTAATATGTAATATAAAATAATTGTTTCCTAGGCACAATAAAAGATCGAAGATTGTTTATGCATCATGTTTCTTTAGAGCCGATTACCTGTGTAC
CCTTTCGGTAAGACATGTTTAAATTTTTCTAAATTCTAATACAGTATGAGAAAAGTCTCGTTTTTATAAATGA --plot_window_size 50 --quantification_window_
size 0 --max_rows_alleles_around_cut_to_plot 1000 --min_frequency_alleles_around_cut_to_plot 0.001 --place_report_in_
output_folder --write_detailed_allele_table
```

e. For different replicates (Replicate-1 and Replicate-2) and time points (day 3 and day 14), change the fastq.gz file such that the name matches exactly with the sequencing files.

Note: Each sample takes 20 min depending on the speed of the internet connection and memory storage.

f. **Nucleotide abundance around sgRNA site:** The PAM site 3-CCT-5′ is changed to 3-CAT-5′ to disrupt the PAM (Figure 5B). The different abundance and deletions are also observed around the sgRNA cut site region. (see troubleshooting 5).
g. **Indel distribution and mutation position:** Samples for day 3 are expected to have higher number of indels as compared to samples at day 14 (Figure 5C), because indels leading to loss-of-function will be eventually lost from the pool. The corresponding plot showing insertion, deletion and substitution clearly shows around 3%–4% of substitution that corresponds to the fixed PAM synonymous substitution.
h. **Variant analysis:** Use the allele frequency table to identify the number of substitutions. Each sequencing read will have the fixed PAM mutation and the presence of a single nucleotide change along the region of saturation. Sort the excel file based on deletion = 0, insertion = 0 and substitution = 2.
i. **Variant annotation:** Execute the codes written in github page and follow the AnnotateCRISPResso.R script to identify the exact point mutation present in individual reads.
j. **Drop out frequency:** To identify the relative drop out frequency of individual variants follow the following steps.
    i. Add a pseudocount of 1 to the reads of all the variants and for all conditions at day 3 and at day 14. This read count is normalized to the total number of reads present in the library. For example, if no. of reads for variant-A in day 3 is M and total reads in the sample library is X, then normalized read count for variant at day 3 (E) = (M+1)/X. Similarly, if no. of reads for variant-A in day 14 is N and total reads in the sample library is Y, then normalized read count for variant at day 14 (F) = (N+1)/Y.
k. Calculate the drop out frequency (function score) by taking ratio of day 14 to day 3 = F/E.
l. Express this value in the log scale (base 2). Based on this score, variants that are present in the pool will have a positive function score and will be considered to be functional. In contrary, variants that are lost from the pool will have a negative score and are considered to be non-functional.

LIMITATIONS

This protocol can be adapted to saturate other genes in different cell lines. However, the efficiency of HDR-based knock in of variants must be optimized. Specifically, number of cells and the concentration of plasmid and donor library must be optimized.

TROUBLESHOOTING

Problem 1: Reduced cloning efficiency

A significantly higher number of bacterial colonies must be present in the cloning plates as compared to control plates.

Potential solution

- If a lot of colonies are visible in the control plates, then check the efficiency of *Bbs*I digestion or use a new vial of the enzyme.
- If a lawn of colonies is seen then use fresh LB plates containing ampicillin (100 μg/mL).
- If similar number of colonies are present between experimental and control plates, repeat the cloning step.

Problem 2: Reduced recovery of cells after nucleofection

Cells must be closely monitored for 2–3 days post nucleofection. Excessive cell death could be due to variable voltage/pulse setting.

Potential solution

- Check cells for mycoplasma contamination and use a fresh vial of cells if contamination is detected.
- Check the voltage and pulse settings. For F7/F7 mESCs, we recommend using program A030 in the Lonza (Amaxa) Nucleofector 2B machine.

Problem 3: Reduced GFP+ cells after nucleofection

Very few GFP+ cells were recovered after FACS sorting.

Potential solution

- Check the correct voltage and pulse settings used for nucleofection.
- Check if the sgRNA is cloned into PX458-GFP-tagged to Cas9 plasmid.

Problem 4: Reduced concentration of genomic DNA

If DNA concentration is reduced, it is an indication that cells were lost during the process of nucleofection and subsequent culturing. This will also affect the frequency of editing outcome.

Potential solution

- Check if enough cells were present after nucleofection and FACS sorting.
- Check if the lysis solution is clogging the genomic DNA isolation column.

Problem 5: Reduced editing after CRISPR-Cas9 editing

The rate of variant knock-in via HDR is rate limiting and depends on the efficiency of sgRNA and the distance of the mutation from the Cas9 cut site.

Potential solution

- Check the efficiency of Indel formation to estimate the efficiency of the sgRNA.

- Confirm if the PAM modification is present in the DNA donor oligo.
- Increase the homology arm to enhance the efficiency of HDR.
- If the distance of the variant to the Cas9 cut site is > 50 bp, then use a pair of sgRNAs to cover the region of saturation.

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. Shyam K. Sharan (sharans@mail.nih.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data generated or analyzed during this study are included in the manuscript and in the supplementary files. The accession number for the sequencing data reported in this paper is GSE238143, available in the GEO database.The code developed for data analysis is available at https://zenodo.org/doi/10.5281/zenodo.10018770.

ACKNOWLEDGMENTS

We thank members of our laboratory for helpful discussions and suggestions. We acknowledge the support from Jeff Carrel and Megan Karwan from the CCR Flow Cytometry Core, Dr. Elizabeth Conner from the CCR Genomics Core, Dr. Raj Chari from the Genome Modification Core, NCI, Frederick for suggestions on CRISPR editing, Dr. Kajal Biswas for generating the humanized mouse embryonic stem cells, and Melissa Galloux for sequencing analysis. Next-generation sequencing was performed using a paid service to Azenta (New Jersey). Illustrations were made using a paid subscription to BioRender.com. The project is funded by the Intramural Research Program, Center for Cancer Research, National Cancer Institute, U.S. National Institutes of Health. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

AUTHOR CONTRIBUTIONS

S.S. and S.K.S. conceived the project; S.S. designed the protocol; T.S., E.S., D.C., and J.G. validated and adapted the protocol; J.G. performed the cloning of sgRNA; E.S., T.S., and D.C. performed nucleofection and genomic DNA isolation; S.S. performed PCR for sequencing; and S.S., T.S., and S.K.S. wrote the manuscript. All authors commented and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Sahu, S., Sullivan, T.L., Mitrophanov, A.Y., Galloux, M., Nousome, D., Southon, E., Caylor, D., Mishra, A.P., Evans, C.N., Clapp, M.E., et al. (2023). Saturation genome editing of 11 codons and exon 13 of BRCA2 coupled with chemotherapeutic drug response accurately determines pathogenicity of variants. PLoS Genet. *19*, e1010940. https://doi.org/10.1371/journal.pgen.1010940.

2. Starita, L.M., Ahituv, N., Dunham, M.J., Kitzman, J.O., Roth, F.P., Seelig, G., Shendure, J., and Fowler, D.M. (2017). Variant Interpretation: Functional Assays to the Rescue. Am. J. Hum. Genet. *101*, 315–325. https://doi.org/10.1016/j.ajhg.2017.07.014.

3. Tabet, D., Parikh, V., Mali, P., Roth, F.P., and Claussnitzer, M. (2022). Scalable Functional Assays for the Interpretation of Human Genetic Variation. Annu. Rev. Genet. *56*, 441–465. https://doi.org/10.1146/annurev-genet-072920-032107.

4. Weile, J., and Roth, F.P. (2018). Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. Hum. Genet. *137*, 665–678. https://doi.org/10.1007/s00439-018-1916-x.

5. 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.

6. Erwood, S., Bily, T.M.I., Lequyer, J., Yan, J., Gulati, N., Brewer, R.A., Zhou, L., Pelletier, L.,

Ivakine, E.A., and Cohn, R.D. (2022). Saturation variant interpretation using CRISPR prime editing. Nat. Biotechnol. *40*, 885–895. https://doi.org/10.1038/s41587-021-01201-1.

7. Li, H., Bartke, R., Zhao, L., Verma, Y., Horacek, A., Rechav Ben-Natan, A., Pangilinan, G.R., Krishnappa, N., Nielsen, R., and Hockemeyer, D. (2023). Functional annotation of variants of the BRCA2 gene via locally haploid human pluripotent stem cells. Nat. Biomed. Eng. https://doi.org/10.1038/s41551-023-01065-7.

8. Walker, L.C., Hoya, M.d.l., Wiggins, G.A.R., Lindy, A., Vincent, L.M., Parsons, M.T., Canson, D.M., Bis-Brewer, D., Cass, A., Tchourbanov, A., et al. (2023). Using the ACMG/AMP framework to capture evidence related to predicted and observed impact on splicing: Recommendations from the ClinGen SVI Splicing Subgroup. Am. J. Hum. Genet. *110*, 1046–1067. https://doi.org/10.1016/j.ajhg.2023.06.002.

9. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., and Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224–226. https://doi.org/10.1038/s41587-019-0032-3.