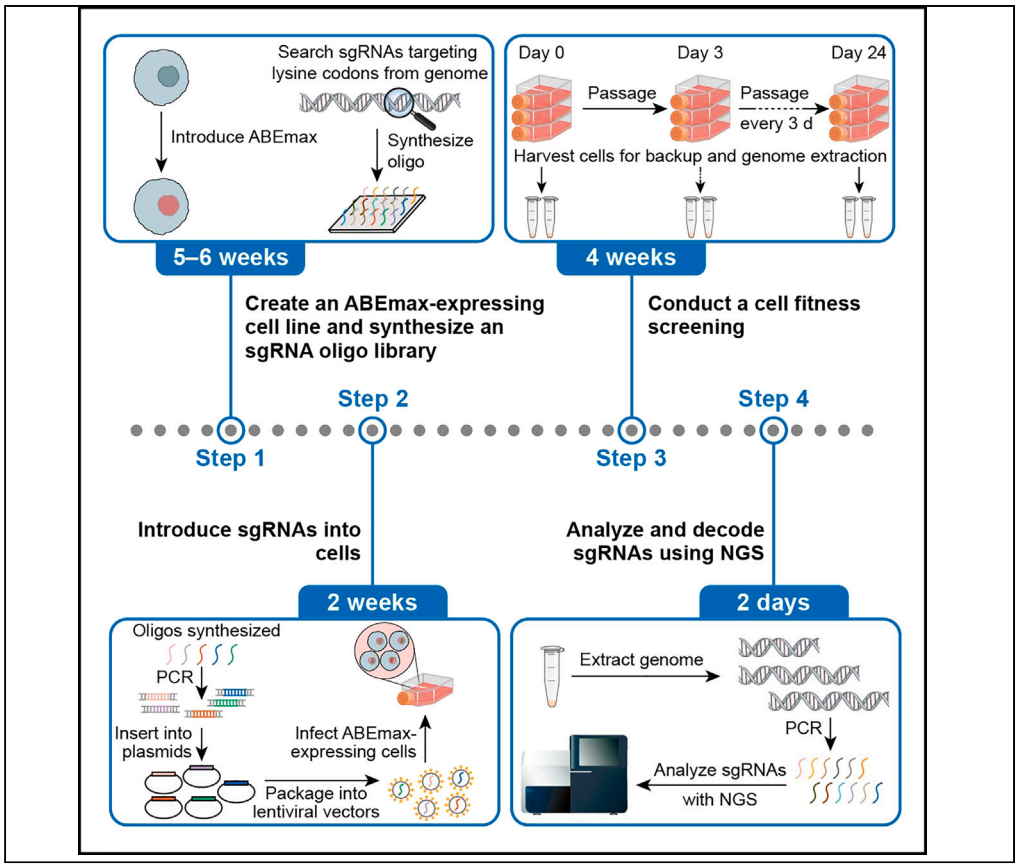
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for high-throughput screening of functional lysine residues in cell fitness



Amino acid residues are crucial to protein structure and function and have links to various human diseases. Here, we present a protocol for screening functional lysine residues across the human genome. We describe steps for designing lysine codon-targeting single-guide RNAs (sgRNAs), constructing an sgRNA library, conducting cell fitness screenings, and acquiring screening results. This approach leverages base editing and high-throughput screening techniques to systematically examine functional amino acid residues.

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

Ying Bao, Wensheng Wei

wswei@pku.edu.cn

Highlights

Steps for designing a lysine codon-targeting sgRNA library across the human genome

Guidance on constructing the library using iBAR technology at high MOI

Instructions for the execution of cell fitness screening in RPE1 cells

Bao & Wei, STAR Protocols 5, 103418
December 20, 2024
https://doi.org/10.1016/j.xpro.2024.103418

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for high-throughput screening of functional lysine residues in cell fitness

Ying Bao[1,3] and Wensheng Wei[1,2,4,*]

[1]Changping Laboratory, Beijing 102206, China

[2]Biomedical Pioneering Innovation Center, Peking-Tsinghua Center for Life Sciences, Peking University Genome Editing Research Center, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China

[3]Technical contact

[4]Lead contact

*Correspondence: wswei@pku.edu.cn
https://doi.org/10.1016/j.xpro.2024.103418

SUMMARY

Amino acid residues are crucial to protein structure and function and have links to various human diseases. Here, we present a protocol for screening functional lysine residues across the human genome. We describe steps for designing lysine codon-targeting single-guide RNAs (sgRNAs), constructing an sgRNA library, conducting cell fitness screenings, and acquiring screening results. This approach leverages base editing and high-throughput screening techniques to systematically examine functional amino acid residues.

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

BEFORE YOU BEGIN

This protocol outlines the detailed steps for assessing the functional lysine residues in hTERT-RPE1 (hereafter referred to as RPE1) cell fitness through high-throughput screening. It employs the adenine base editor (ABEmax) to induce mutations in lysine residues by converting adenines in lysine codons. This method can also be adapted for targeting other amino acid residues in various cell lines under different screening conditions.

Cell culture

Timing: 1 week

1. On day 1, thaw and revive cryopreserved RPE1 and HEK293T cells.
   a. Prepare the culture medium as follows:

Cell culture medium	
Reagent	Amount
Penicillin-streptomycin	5.6 mL (1% vol/vol)
Fetal bovine serum	56 mL (10% vol/vol)
DMEM or DMEM/F-12	500 mL

Note: Unless otherwise specified, DMEM or DMEM/F-12 hereafter refers to the prepared cell culture medium as indicated in the table.

Note: HEK293T cells are maintained in DMEM, while RPE1 cells are cultured in DMEM/F-12.

b. Immediately after retrieving the cells from the liquid nitrogen tank, thaw them in a 37°C water bath.
c. Centrifuge the thawed cells at 200 ×*g* for 5 min.
d. Remove the supernatant and resuspend the cells in the prepared medium.
e. Plate the appropriate number of cells into dishes containing the culture medium.
f. On day 2, examine the cells under an inverted microscope to assess their condition.

2. On day 3, passage the revived cells.
   a. Remove the supernatant and rinse the cells with PBS.
   b. Discard the PBS and add 0.25% trypsin to the dishes, then incubate at 37°C for 1–2 min.
   c. Once most cells have rounded up and detached, add culture medium (three times the volume of trypsin) to neutralize the trypsin.
   d. Transfer the cells to tubes and centrifuge at 200 ×*g* for 5 min.
   e. Remove the supernatant and resuspend the cells in fresh culture medium.
   f. Plate the appropriate number of cells into new dishes with culture medium. These cells are designated as passage 1 (P1).
3. Passage the cells every 2–3 days, keeping a record of the passage numbers.
4. Check for mycoplasma contamination in the cells using PlasmoTest-Mycoplasma Detection Kit following the manufacturer's instructions (https://www.invivogen.com/sites/default/files/invivogen/products/files/plasmotest-v2.pdf). If contamination is detected, treat the cells with Plasmocin prophylactic as per the manufacturer's guidelines (https://www.invivogen.com/sites/default/files/invivogen/products/files/plasmocin_prophylactic_tds.pdf).

⚠ CRITICAL: Ensure that all cells used in subsequent steps are confirmed to be free of mycoplasma.

Note: Cell used in subsequent steps should be at or below passage 10 (P10).

Prepare plasmids

Timing: 1 week

5. Obtain or clone the necessary plasmids.
   a. Prepare a base editor-expressing plasmid in a lentivirus vector with an appropriate expression marker (e.g., pLenti_ABEmax_EGFP, as used in this protocol).
   b. Prepare empty vectors for sgRNA expression in lentiviral vectors, including appropriate expression markers and internal barcodes (e.g., pCG_2.0_SV40_Puro_iBAR-1/2/3, as used in this protocol).
   c. Prepare envelope and packaging plasmids for lentivirus production (e.g., pVSV-G and pR8.74, as used in this protocol).
6. Amplify the plasmids by transforming them into *Trans*1-T1 and then extracting the plasmids.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CRISPR-Cas9, 1:1,000	Abcam	Cat#ab204448
Goat anti-rabbit IgG-HRP secondary antibody, 1:10,000	Jackson ImmunoResearch	Cat#111035003
Bacterial and virus strains		
E. coli HST08 premium electro-cells	Takara	Cat#9028

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
*Trans*1-T1 phage resistant chemically competent cell	TransGen	Cat#CD501
Chemicals, peptides, and recombinant proteins		
BASIC DMEM, high glucose, pyruvate	Gibco	Cat#C11995500BT
DMEM/F-12	Gibco	Cat#11320033
Opti-MEM reduced serum medium	Gibco	Cat#31985070
Polybrene	Sigma-Aldrich	Cat#107689
T4 DNA ligase	New England Biolabs	Cat#M0202
Tango buffer	Thermo Fisher Scientific	Cat#BY5
BsmBI	Thermo Fisher Scientific	Cat#ER0452
X-tremeGENE HP DNA transfection reagent	Roche	Cat#06366236001
Puromycin	Solaribio	Cat#P8230
Passive lysis buffer	Promega	Cat#E1941
Reducing SDS loading buffer	CWBIO	Cat#CW0027
Critical commercial assays		
PlasmoTest-mycoplasma detection kit	InvivoGen	Cat#rep-pt1
Plasmocin prophylactic	InvivoGen	Cat#ant-mpp
EndoFree plasmid maxi kit	QIAGEN	Cat#12362
KAPA HiFi HotStart ReadyMix PCR kit	Roche	Cat#KK2602
DNeasy blood and tissue kit	QIAGEN	Cat#69506
DNA Clean & Concentrator-5	Zymo Research Corporation	Cat#D4013
NEBNext Ultra DNA library prep kit for Illumina	NEB	Cat#E7805
NEBNext Ultra II Q5 master mix	New England Biolabs	Cat#M0544
Clarity western ECL substrate kit	Bio-Rad	Cat#1705060
Pierce BCA protein assay kit	Thermo Fisher Scientific	Cat#23225
TGX FastCast acrylamide starter kit	Bio-Rad	Cat#1610172
Experimental models: Cell lines		
HEK293T	C. Zhang's laboratory, Peking University	N/A
hTERT-RPE1	Y. Sun's laboratory, Peking University	N/A
hTERT-RPE1-ABEmax-SC	Bao et al., 2023[1]	N/A
Oligonucleotides		
Primers for PCR, see Table 1	RuiBiotech	N/A
Library oligos, Bao et al.[1]	Synbio Technologies	N/A
Recombinant DNA		
pCMV_ABEmax_P2A_GFP	Koblan et al.[2]	Addgene #112101
pLenti_ABEmax_EGFP	Bao et al.[1]	N/A
pCG_2.0_CMV_mCherry	Bao et al.[1]	N/A
pVSV-G	Addgene	N/A
pR8.74	Addgene	N/A
pCG_2.0_SV40_Puro_iBAR-1	Bao et al.[1]	N/A
pCG_2.0_SV40_Puro_iBAR-2	Bao et al.[1]	N/A
pCG_2.0_SV40_Puro_iBAR-3	Bao et al.[1]	N/A
Software and algorithms		
FlowJo	BD Biosciences	https://www.bdbiosciences.com/en-us/products/software/flowjo-v10-software
Library design	Bao et al.[1]	Zenodo: https://doi.org/105281/zenodo.8320606
Other		
C1000 Touch thermal cycler	Bio-Rad	Cat#1851197
Gene Pulser Xcell microbial system	Bio-Rad	Cat#1652662
Gene Pulser/MicroPulser electroporation cuvettes	Bio-Rad	Cat#1652089
Falcon 14 mL round bottom high clarity PP test tube	Corning	Cat#352059
Corning 225 cm^2 angled neck cell culture flask with vent cap	Corning	Cat#431082

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BD LSRFortessa cell analyzer	BD Biosciences	N/A
BD FACSAria III cell sorter	BD Biosciences	N/A
1300 Series A2 biological safety cabinet	Thermo Fisher Scientific	Cat#1332
Forma Series 2 water jacketed CO_2 incubator	Thermo Fisher Scientific	Cat#4131

STEP-BY-STEP METHOD DETAILS

Generation of ABEmax-expressing RPE1 cells

Timing: 5–6 weeks

1. Package the ABEmax lentivirus.
   a. On day 0, seed 4×10^6 to 5×10^6 HEK293T cells in a 100 mm dish with 10 mL DMEM, aiming for 60%–70% confluency by the time of transfection.
   b. On day 1, co-transfect the ABEmax, packaging, and envelope plasmids into the HEK293T cells.
      i. Prepare the transfection mixture as follows:

Transfection mixture	
Reagent	Amount
plenti-ABEmax-EGFP	4 μg
pR8.74	4 μg
pVSV-G	0.4 μg
Opti-MEM	1 mL

      ii. Add 24 μL of X-tremeGENE HP DNA transfection reagent (pre-equilibrated to 25°C) to the transfection mixture and mix gently by pipetting.
      iii. Allow the mixture to incubate at 25°C for 15 min.
      iv. Add the mixture to the HEK293T cell medium, then gently shake the dish to ensure even distribution.
   c. Within 8 h, replace the medium with fresh DMEM containing 30% fetal bovine serum (FBS).
   d. On day 4, collect the medium containing the ABEmax-packaged lentivirus.
   e. Centrifuge the viral supernatant at 500 ×*g* for 10 min to remove any residual HEK293T cells.
   f. Collect the viral solution, aliquot it, and store at −80°C.
2. Titrate the ABEmax lentivirus.
   a. On day 0, seed 1×10^5 RPE1 cells in 6-well plates with 2 mL of DMEM/F-12 cell culture medium, aiming for 60%–70% confluency at the time of infection.
   b. On day 1, add polybrene to the medium to a final concentration of 8 μg/mL, then introduce a gradient of ABEmax lentivirus volumes (e.g., 2 μL, 4 μL, 8 μL, etc.) to the RPE1 cells.
   c. On day 2, replace the medium with fresh DMEM/F-12.
   d. On day 5, assess the transduction percentage and calculate the multiplicity of infection (MOI) for each viral volume using the following formula:

$$m = -\ln(1 - P)$$

   where m is the MOI, and P is the probability of cells being infected, determined by the percentage of cells expressing marker proteins for ABEmax expression (e.g., drug resistance or fluorescence).

   Note: Ensure P remains below 90% to minimize variation when calculating m, as m increases rapidly when P approaches 90%.

   e. Perform a linear regression analysis between MOI and viral volume. Troubleshooting 1.

3. Transduce RPE1 cells with ABEmax lentivirus and isolate single-cell clones.
   a. Transduce RPE1 cells with ABEmax lentivirus at an MOI of 1–3. Troubleshooting 2.
   b. Sort the ABEmax-expressing cells and isolate single-cell clones into 96-well plates.

   ⚠ CRITICAL: Refresh the medium every 2–3 days to prevent evaporation.

   c. Expand the RPE1-ABEmax clonal populations for further characterization.
4. Verify ABEmax expression by Western blot and select the single clone with the highest editing efficiency for library construction and screening (Figure 1).

Design and synthesize the lysine codon-targeting sgRNA library

Timing: 1 week

Note: The following section outlines the design rules. For specific code implementations, please refer to Bao et al.[1]

5. Identify all lysine residue positions in protein sequences using the UniProt reference Fasta file.
6. Locate lysine codons (5′-AAR, where R is A or G) in the reference genome based on the positions of lysine residues in the protein sequences.
7. Search for PAM sequences (5′-NGG) that ABEmax can use to mutate lysine codons.
   a. Identify lysine codons by their proximity to splicing sites.

   Note: Codons within exons are classified as unspliced, while those interrupted by introns after the first or second A in the codon are classified as "A|AR" or "AA|R" spliced types, respectively.

   b. For unspliced and "AA|R" spliced types, scan for "NGG" PAM sequences located 12–18 nucleotides downstream from the first adenosine.

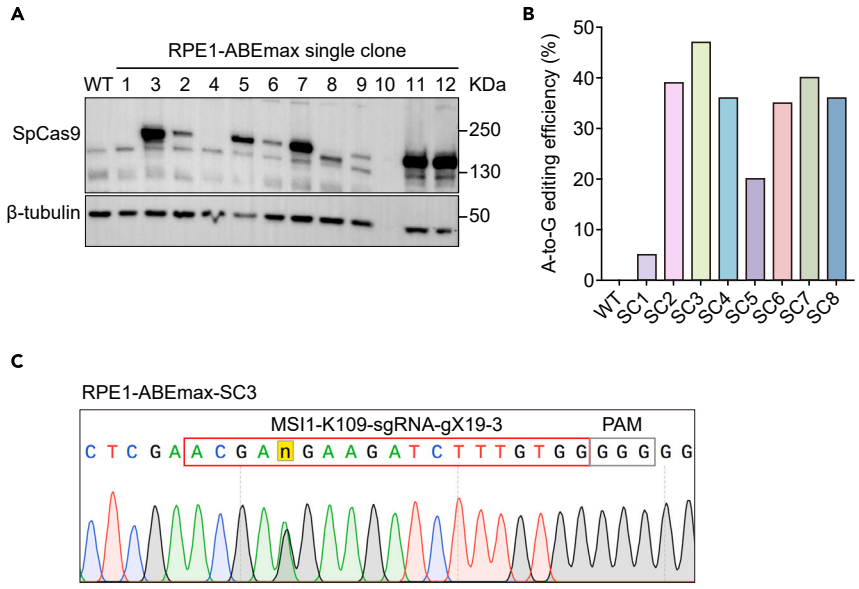


Figure 1. Selection of RPE1-ABEmax cells based on Western blot and editing efficiency

(A) Western blot analysis of SpCas9 from the ABEmax construct.

(B) Editing efficiency of RPE1-ABEmax single clones #1–8 at an endogenous site (MSI-K109).

(C) Sanger sequencing showing the editing efficiency at the MSI-K109 site in RPE1-ABEmax single clone #3 (SC3), with adenosine at position 15 being edited at approximately 50% efficiency.

Note: Based on previous studies and our tests, the ABEmax editing window is defined as positions 13–18 (with position 1 being the nucleotide immediately upstream of the PAM sequence).[1,2]

Note: To effectively substitute targeted lysine residues, the first or second adenosine in the lysine codons should fall within positions 13–18. Therefore, the distance between the first adenosine of lysine codons and the PAM sequences should range from 12–18 nucleotides.

c. For the "A|AR" spliced type, scan for "NGG" PAM sequences located 12–17 nucleotides downstream from both the first and second adenosine.

8. Extract sgRNA sequences (Figure 2).

Note: Previous reports and our tests suggest that extending sgRNA length can enhance editing efficiency at positions beyond the central optimal sites. Specifically, 19-nt sgRNAs are more effective at positions 15 and 16, 20-nt sgRNAs are more effective at positions 12–14 and 17–18, and 21-nt sgRNAs can further improve efficiency at positions 18 and 19.[1,2]

a. Extract sgRNAs for unspliced and "AA|R" spliced lysine codons.
   i. If the first adenosine is at positions 15–17, extract a 19-nt sequence upstream of the PAM.
   ii. If the first adenosine is at positions 13, 14, or 18, extract a 20-nt sgRNA sequence upstream of the PAM.
   iii. If the first adenosine is at position 19, extract a 21-nt sgRNA sequence upstream of the PAM.

b. Extract sgRNAs for "A|AR" spliced lysine codons.
   i. If the targeted adenosine is at positions 14–16, extract a 19-nt sequence upstream of the PAM.
   ii. If the targeted adenosine is at positions 13 or 17, extract a 20-nt sequence upstream of the PAM.
   iii. If the targeted adenosine is at position 18, extract a 21-nt sequence upstream of the PAM.

9. Analyze on-target and 1-bp mismatch off-target sites using Bowtie software with the parameter "bowtie -a -v 1".
10. Filter the sgRNAs according to the following quality control criteria: Troubleshooting 3.
    a. GC content between 0.2 and 0.8.
    b. No thymidine homopolymers with a length of ≥ 4.
11. Set negative and positive controls to evaluate the false positive and false negative rates of the screening results.

Note: Negative controls are sgRNAs that do not affect cell fitness upon introduction, such as those targeting adenosines in the *AAVS1* locus or non-targeting sgRNAs with no binding sites in the human genome. Positive controls are sgRNAs that impact cell fitness, either enhancing

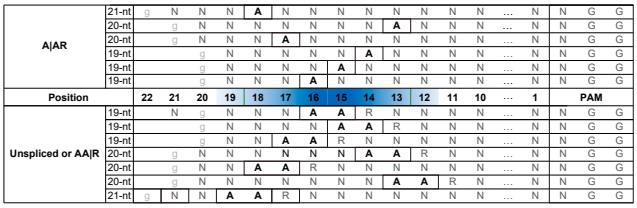


Figure 2. Design rules for sgRNAs to target lysine codons

The editing efficiency of ABEmax at different positions within the editing window is represented by color intensity, with darker shades indicating higher efficiency. The upper panel corresponds to the "A|AR" spliced type, while the lower panel pertains to both unspliced and "AA|R" spliced types. The length of the sgRNA varies based on the positions of the targeted lysine codons.

or inhibiting it, such as lysine residues known to regulate cell fitness. If such sites are limited, sgRNAs targeting start codons or splice sites of essential or tumor suppressor genes can be used, which is a common approach in ABE-based screens.

Note: According to our validation results and other BE-based screens, we recommend including 0.1%–1% negative and 0.1%–1% positive controls to ensure screening quality.[3,4]

12. Synthesize sgRNA oligos flanked by PCR fragments in the following format:

5′-CGCTCCGTGAACAGTAATTAGGTGCGTCTCAACCG[sgRNA]GTTTTGAGACGCAACGCATGG TATCACGATTCCTG-3′.

The library contains 283,169 sgRNAs in total.

Clone sgRNAs into plasmids

Timing: 3–4 days

Note: This protocol employs the iBAR strategy we previously developed to enhance screening quality and reduce the number of starting cells by using a high MOI.[5,6] IBAR refers to a 6-nt barcode embedded in the sgRNA scaffold, with three different iBARs assigned to each sgRNA, serving as internal replicates during the screen. Each sgRNA is evaluated according to the significance and consistency of abundance changes across the three sgRNAs[iBAR].

Note: Each iBAR vector is used for parallel library cloning, with subsequent bacterial transformation conducted separately.

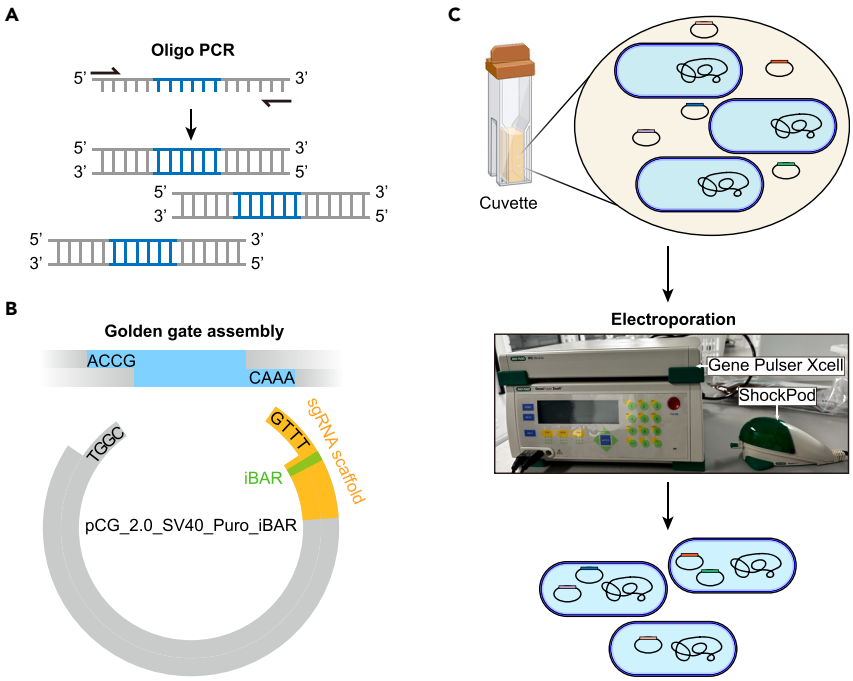


Figure 3. Cloning of sgRNAs into plasmids
(A) Amplification of the sgRNA oligos by PCR.
(B) Ligation of the amplified sgRNA oligos into plasmids.
(C) Amplification of the plasmids by transforming E. coli via electroporation.

13. Amplify sgRNA oligos by PCR (Figure 3A).
    a. Dissolve the synthesized sgRNA oligo powders in TE buffer, aliquot, and store at −80°C.
    b. Dilute the oligos to ensure 1000-fold library coverage in 2.5 μL, which will serve as the PCR template.

Note: For illustration, consider a library consisting of 300,000 sgRNAs, each synthesized as 90-nt. The oligo weight required for 1000-fold coverage is calculated as follows:

$$Oligo\ weight = ssDNA\ molecular\ weight \times \frac{sgRNA\ count \times coverage}{N_A}$$
$$= (90 \times 303.7 + 79.0)\ \frac{g}{mol} \times \frac{300,000 \times 1,000}{6.02 \times 10^{23}\ mol^{-1}} = 0.0137\ ng$$

For subsequent PCR, dilute the oligos to a final concentration of 0.0137 ng/2.5 μL.

    c. Amplify the sgRNA oligos under the following conditions:

PCR reaction master mix	
Reagent	Amount
DNA template	1,000-fold coverage of the library in 2.5 μL
NEBNext Ultra II Q5 Master Mix	25 μL
Forward primer	5 μL
Reverse primer	5 μL
ddH_2O	12.5 μL
Total volume	50 μL

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial Denaturation	98°C	30 s	1
Denaturation	98°C	10 s	20 cycles
Annealing	69°C	20 s	
Extension	72°C	10 s	
Final extension	72°C	2 min	1
Hold	4°C	forever	

    d. Purify the PCR products using the DNA Clean & Concentrator-5 according to the manufacturer's protocol (https://files.zymoresearch.com/protocols/_d4003t_d4003_d4004_d4013_d4014_dna_clean_concentrator_-5.pdf), and adjust the concentration to 10 ng/μL.

14. Ligate the purified fragments into pCG_2.0_SV40_Puro_iBAR vectors using Golden Gate assembly (Figure 3B).
    a. Prepare the Golden Gate reaction mix and perform the reaction as follows:

Golden Gate assembly reaction master mix		
Reagent	Amount	10 × reactions
Vector (100 ng/μL)	2 μL	20 μL
Purified sgRNA oligo PCR products (10 ng/μL)	4 μL	40 μL
DTT (50 mM)	0.4 μL	4 μL
ATP (10 mM)	2 μL	20 μL
BsmBI	1.5 μL	15 μL
T4 DNA ligase	0.5 μL	5 μL
Tango buffer	2 μL	20 μL
ddH_2O	7.6 μL	76 μL
Total volume	20 μL	200 μL

Golden Gate assembly conditions			
Steps	Temperature	Time	Cycles
Digestion	37°C	5 min	20 cycles
Ligation	16°C	5 min	
Final incubation	37°C	5 min	1
Hold	4°C	forever	

b. Purify the Golden Gate assembly products using the DNA Clean & Concentrator-5 and adjust the concentration to 100 ng/μL.

15. Transform *E. coli* with the purified Golden Gate products by electroporation (Figure 3C).

Note: Use the following formula to calculate the number of reactions required:

$$Number\ of\ reactions = \frac{sgRNA\ count \times Plasmid\ library\ coverage}{Transformation\ efficiency}$$

Where the transformation efficiency is the number of transformants produced per reaction and should be determined beforehand. For the *E. coli* HST08 premium electro-cells used in this protocol, the transformation efficiency is 2×10^6. Each reaction consists of 50 μL of electro-cells and 1 μL (100 ng) of purified Golden Gate products, recovered in 1 mL of SOC medium (included with the *E. coli* HST08 premium electro-cells kit). In this protocol, the plasmid library coverage is set to 300-fold, and the number of reactions is calculated to be approximately 43 (45–50 reactions are needed for potential loss during subesequent steps).

Note: Pre-chill the cuvettes and SOC medium on ice before starting.

a. Thaw the *E. coli* HST08 premium electro-cells on ice just before use.
b. For each reaction, add 1 μL of the purified Golden Gate products to 50 μL of the thawed cell suspension.
c. Transfer the mixture to a chilled electroporation cuvette and incubate on ice for 2 min.

Note: Avoid introducing bubbles during the transfer.

d. Set up the Gene Pulser Xcell Microbial System for electroporation with the following conditions:

Electroporation condition	
Parameter	Value
Voltage	1.5 kV
Capacitance	25 μF
Resistance	200 Ω
Cuvette	1 mm

e. Place the cuvette in the ShockPod, close the chamber lid, and apply the pulse.
f. Immediately add 1 mL of prechilled SOC medium to the pulsed cuvette and gently mix by pipetting.

⚠ CRITICAL: The short time between pulsing and adding SOC medium is crucial for optimal cell recovery.

g. Transfer the cell suspension to a Falcon 14 mL Round Bottom High Clarity PP Test Tube and incubate at 37°C with shaking at 225 rpm for 1 h.
h. Mix all cell suspensions and incubate in liquid LB medium at 37°C with shaking at 225 rpm for 15–16 h.

Note: Before mixing, take several tubes and spread a diluted cell suspension (e.g., 1 μL in 200 μL of liquid LB) onto an LB agar plate to confirm the transformation efficiency.

16. Extract plasmids using the EndoFree Plasmid Maxi Kit following the manufacturer's protocol (https://www.qiagen.com/us/resources/resourcedetail?id=a48e64ab-27cf-4576-bb93-98bbd0e1229e&lang=en).
17. Verify the quality of the plasmid library using NGS. SgRNA coverage should reach 98%, and the distribution should follow a normal distribution.

Transduce sgRNAs into cells and conduct screen

Timing: 5–6 weeks

This step facilitates the introduction of the sgRNA library into mammalian cell lines for subsequent screening. In this protocol, RPE1-ABEmax cells are used, with regular passaging for cell fitness screening.

18. Package the sgRNA library into lentivirus (Figure 4A).

Note: Combine the three iBAR library plasmids in equal proportions for transfection.

19. Titrate the lentivirus.
    a. On day 0, seed 1×10^5 RPE1-ABEmax cells in 6-well plates.
    b. On day 1, add polybrene to the medium to a final concentration of 8 μg/mL, then introduce varying volumes of the sgRNA library lentivirus (e.g., 2 μL, 4 μL, 8 μL, etc.) to the RPE1 cells.
    c. On day 2, replace the medium with fresh DMEM/F-12.
    d. On day 3, re-seed the infected RPE1-ABEmax cells into 6-well plates at a density of 1×10^5 cells per well.
    e. On day 4, select sgRNA-expressing cells by adding 15 μg/mL puromycin.
    f. On day 6, count the viable cells and calculate the MOI for each viral volume. Perform a linear regression analysis between MOI and viral volume.

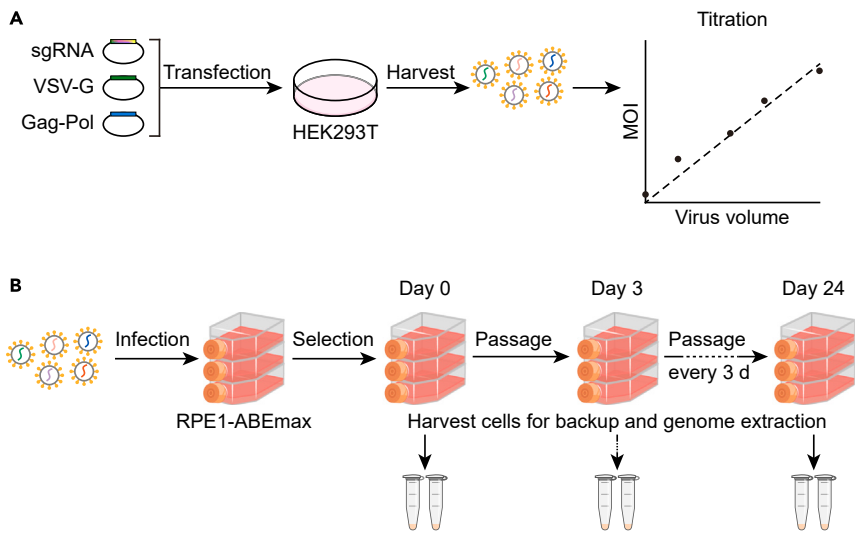


Figure 4. Construction of cell library and cell fitness screen

(A) Packaging of the sgRNA-expressing plasmids into lentivirus and titration of the virus.

(B) Infection of cells with the sgRNA-packaged lentivirus and cell fitness screening.

⚠ CRITICAL: Set up two control wells: one without virus treatment (and no puromycin) to calculate infection percentage, and another with puromycin treatment to confirm its effectiveness in killing uninfected cells.

20. Infect RPE1 cells with the sgRNA library lentivirus and conduct a cell fitness screen (Figure 4B).
    a. On day 0, seed 2.4×10^6 RPE1-ABEmax cells into T-225 flasks for infection. Calculate the total cell number for infection using the following formulas:

$$\textit{Cell number for infection} = \frac{\textit{sgRNA count} \times \textit{Cell library coverage}}{\textit{MOI}}$$

Note: The seeding density should match the titration conditions and be adjusted according to the culturing area.

    b. On day 1, infect the RPE1-ABEmax cells with the sgRNA library lentivirus at an MOI of 3.

Note: The library consists of 283 million sgRNAs (283,169 sgRNA at 1,000-fold coverage), with 944 million cells used for library infection at an MOI of 0.3. In this protocol, iBAR technology is employed to reduce the starting cell number to 94 million with an MOI of 3, while ensuring high screening quality.

    c. On day 2, replace the medium with fresh DMEM/F-12.
    d. On day 3, re-seed the sgRNA library-transduced RPE1-ABEmax cells into T-225 flasks at a density of 2.4×10^6 cells per flask, maintaining the same total cell number as on day 0.
    e. On day 4, add 15 μg/mL of puromycin.
    f. On day 6, re-seed the puromycin-selected cells into T-225 flasks at a density of 2.4×10^6 cells per flask to achieve the desired cell library size for screening. Harvest cells for sequencing and cryopreserve several backups.

$$\textit{Cell library size} = \textit{Cell number for infection} \times \textit{infection efficiency}$$

Note: The infection efficiency is determined by the MOI. For an MOI of 3, the infection efficiency is 95.0%. To account for potential loss during screening, the cell library size should exceed the calculated number.

Note: This day is designated as Day 0, the reference group for analysis.

Note: Each aliquot of harvested cells and cryopreserved cells should exceed the cell library size to account for potential loss during subsequent steps.

    g. Passage the cell library every 3 days for a total of 24 days, maintaining one cell library size at each round of cell seeding.
    h. On day 30, harvest the cells and cryopreserve several backups for potential additional rounds of screening.

Note: This day is designated as Day 24, the experimental group for analysis.

Note: Harvest and cryopreserve cells at various stages during the screen to create backups.

Capture and sequence the sgRNAs

Timing: 2 days

This step is crucial for acquiring the screen results by extracting, amplifying, and sequencing sgRNAs from the samples collected on Day 0 and Day 24, with each group processed in parallel.

Table 1. PCR primers

Usage	Forward	Reward
Amplify sgRNA oligos	CGCTCCGTGAACAGTAATTAGGTG	CAGGAATCGTGATACCATGCGTTG
Decode sgRNAs from genomic DNA	TACACGACGCTCTTCCGATCT TAAGTAGAG TATCTTGTGGAAAGGACGAAACACC	AGACGTGTGCTCTTCCGATCT TAAGTAGAG TCGACCTGCTGGAATCTCGTG
	TACACGACGCTCTTCCGATCT ATCATGCTTA TATCTTGTGGAAAGGACGAAACACC	AGACGTGTGCTCTTCCGATCT ATCATGCTTA TCGACCTGCTGGAATCTCGTG
	TACACGACGCTCTTCCGATCT GATGCACATCT TATCTTGTGGAAAGGACGAAACACC	AGACGTGTGCTCTTCCGATCT GATGCACATCT TCGACCTGCTGGAATCTCGTG
	TACACGACGCTCTTCCGATCT CGATTGCTCGAC TATCTTGTGGAAAGGACGAAACACC	AGACGTGTGCTCTTCCGATCT CGATTGCTCGAC TCGACCTGCTGGAATCTCGTG
	TACACGACGCTCTTCCGATCT TCGATAGCAATTC TATCTTGTGGAAAGGACGAAACACC	AGACGTGTGCTCTTCCGATCT TCGATAGCAATTC TCGACCTGCTGGAATCTCGTG

21. Extract genomic DNA from cell pellets harvested on Day 0 and Day 24 using DNeasy Blood and Tissue Kit according to the manufacturer's protocol (https://www.qiagen.com/us/resources/resourcedetail?id=aa250d94-fc4b-4e27-bb74-d32391ff8a48&lang=en).
22. Amplify sgRNA sequences from the extracted genomic DNA using the KAPA HiFi HotStart ReadyMix PCR Kit under the following conditions:

PCR reaction master mix

Reagent	Amount
DNA template	1–2 μg
2 × KAPA HiFi HotStart ReadyMix	25 μL
Forward primer	2.5 μL
Reverse primer	2.5 μL
ddH_2O	As needed
Total volume	50 μL

Note: To ensure high-quality NGS, add a 9- to 13-nt variant sequence to the primers to increase sequence diversity. In this protocol, the total genomic DNA for each group is divided into five equal parts, with each part amplified using a different pair of primers containing distinct additional sequences (Table 1).

PCR cycling conditions

Steps	Temperature	Time	Cycles
Initial Denaturation	95°C	3 min	1
Denaturation	98°C	20 s	26 cycles
Annealing	60°C	15 s	
Extension	72°C	15 s	
Final extension	72°C	1 min	1
Hold	4°C	forever	

23. Mix the PCR products from the five parts for each group and purify them using the DNA Clean & Concentrator-5 according to the manufacturer's protocol (https://files.zymoresearch.com/protocols/_d4003t_d4003_d4004_d4013_d4014_dna_clean_concentrator_-5.pdf).
24. Prepare the library using the NEBNext Ultra™ II FS DNA Library Prep Kit for Illumina according to the manufacturer's protocol (https://www.neb.com/en-us/protocols/2017/10/25/protocol-for-fs-dna-library-prep-kit-e7805-e6177-with-inputs-less-than-or-equal-to-100-ng), and perform NGS with 200- to 500-fold coverage using an Illumina HiSeq2500.

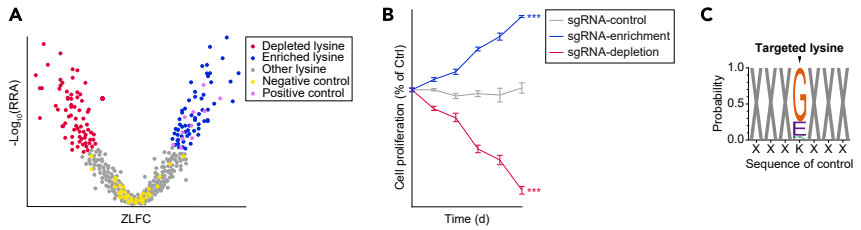


Figure 5. Expected screening outcomes

(A) Expected distribution of screening results, with red dots representing negatively selected sgRNAs and blue dots representing positively selected sgRNAs.

(B) Expected CPA outcomes for the screen hits.

(C) Expected NGS results of sgRNA editing, depicted as sequence logos. The x-axis represents the reference sequence from cells expressing sg*AAVS1*, with the targeted lysine residue centered in the sequence. "X" denotes surrounding residues, and gray indicates residues that were not edited.

25. Analyze the screen data using the ZFC^{iBAR} algorithm.[5] Rank sgRNAs based on their Fitness Score (FS):

$$FS = sign(zLFC) \times (-log_{10}(RRA) + |zLFC|))$$

where zLFC is the z score of the log fold-change, and RRA is the robust rank aggregation for the three iBARs per sgRNA.

EXPECTED OUTCOMES

Expected screen results distribution

Present the screen results using a volcano plot, where the x-axis represents zLFC and the y-axis represents -log$_{10}$(RRA). A well-conducted screen should display clear enrichment of hits in both directions of the volcano, with negative controls positioned at the bottom and positive controls appearing towards the top of the volcano (Figure 5A).

Expected validation results for screen hits

Select several high-ranking and low-ranking sgRNAs for validation. Validation involves assessing both phenotype and editing. Transduce selected sgRNAs individually into RPE1-ABEmax cells to perform a Cell Fitness Assay (CPA) and NGS of the targeting sites.[1] True hits from the screen should exhibit significantly higher or lower fitness compared to cells transduced with negative controls (Figure 5B), and the targeted lysine residues should be edited more efficiently than the surrounding sites (Figure 5C). Troubleshooting 4.

LIMITATIONS

Since this protocol employs ABEmax to construct the lysine-mutating library, the editing properties of ABEmax may limit its application.

First, due to the PAM restriction of SpCas9, not all lysine codons in the human genome can be mutated with this library. To achieve a more comprehensive screen, base editors with more flexible PAM requirements should be considered.

Second, ABEmax's editing window spans several nucleotides, so sgRNAs targeting lysine codons might also edit adenosines in nearby codons. As a result, high-ranking sgRNAs might influence cell fitness by mutating residues other than the targeted lysines.

TROUBLESHOOTING

Problem 1

The infection percentage remains low (<10%) despite increasing viral volume (related to Step 2e).

Potential solution

This may be due to low virus titer resulting from the large size and low quality of cargo plasmids.

- Avoid endotoxin contamination during plasmid extraction by using endotoxin-free kits.
- Concentrate the virus using a centrifugal filter and titrate with the concentrated virus.

Problem 2

Infected cells are contaminated by residual HEK293T cells from the lentivirus preparation (related to Step 3a).

Potential solution

HEK293T cells, commonly used for packaging lentivirus, can easily detach and remain suspended in the medium. It is important to remove these residual HEK293T cells before infecting the target cells.

- Filter the lentivirus using a 0.45 μm microfiltration membrane.
- Freeze the lentivirus at −80°C for 12 h to help kill any remaining cells.

Problem 3

The number of sgRNA per sites is inconsistent. Due to the constraints of PAM sequence availability and the editing window of ABEmax, most lysine residues can only be targeted by a single sgRNA, while a few sites can be targeted by multiple sgRNAs (related to Step 10).

Potential solution

- Include all sgRNAs targeting the same site that pass quality control and analyze them individually. As different sgRNAs targeting the same lysine codon may induce different mutation types, they should be evaluated separately.
- The iBAR strategy provides a tool for accurately evaluating sgRNAs at each site. Each sgRNA is attached with three independent iBARs, allowing for the independent tracking of the sgRNA behavior in different cells. With the specifically devised ZFC^{iBAR} algorithm, the effect of each sgRNA can be evaluated based on the significance and consistency of abundance changes across the three sgRNAs^{iBAR}.
- For sites targeted by a single sgRNA with lower editing efficiency, the edited cellular population will be enriched during screening. According to our validation results, some edited mutation populations can increase from approximately 10% on Day 1 to 40% by Day 28.

Problem 4

Both targeted lysine and surrounding codons are edited with a single sgRNA (related to expected outcomes, expected validation results for screen hits).

Potential solution

Further verification is required to determine whether the targeted lysine residues contribute to cell fitness.

- Track sgRNA editing patterns over time and analyze changes in the proportion of lysine mutants. If a specific mutation type exhibits an advantage over time, the targeted lysine residue is likely functional in cell fitness.
- Construct plasmids expressing each mutant and verify their function through overexpression in cells.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Wensheng Wei (wswei@pku.edu.cn).

Technical contact

For technical questions related to this protocol, contact Ying Bao (baoying2017@pku.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

This study did not generate datasets or code.

ACKNOWLEDGMENTS

This project was funded by the National Natural Science Foundation of China (NSFC31930016), the Peking-Tsinghua Center for Life Sciences, and the Changping Laboratory. The authors offer special thanks to Y. Sun and C. Zhang (PKU) for providing cell lines, the BIOPIC High-throughput Sequencing Center for NGS analysis, and the National Center for Protein Sciences (PKU) for flow cytometry support (particularly H. Yang and H. Lyu).

AUTHOR CONTRIBUTIONS

Writing, Y.B.; revision and supervision, W.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Bao, Y., Pan, Q., Xu, P., Liu, Z., Zhang, Z., Liu, Y., Xu, Y., Yu, Y., Zhou, Z., and Wei, W. (2023). Unbiased interrogation of functional lysine residues in human proteome. Mol. Cell *83*, 4614–4632.e6. https://doi.org/10.1016/j.molcel.2023.10.033.

2. Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. *36*, 843–846. https://doi.org/10.1038/nbt.4172.

3. Sangree, A.K., Griffith, A.L., Szegletes, Z.M., Roy, P., DeWeirdt, P.C., Hegde, M., McGee, A.V., Hanna, R.E., and Doench, J.G. (2022). Benchmarking of SpCas9 variants enables deeper base editor screens of BRCA1 and BCL2. Nat. Commun. *13*, 1318. https://doi.org/10.1038/s41467-022-28884-7.

4. Kim, Y., Lee, S., Cho, S., Park, J., Chae, D., Park, T., Minna, J.D., and Kim, H.H. (2022). High-throughput functional evaluation of human cancer-associated mutations using base editors. Nat. Biotechnol. *40*, 874–884. https://doi.org/10.1038/s41587-022-01276-4.

5. Xu, P., Liu, Z., Liu, Y., Ma, H., Xu, Y., Bao, Y., Zhu, S., Cao, Z., Wu, Z., Zhou, Z., and Wei, W. (2021). Genome-wide interrogation of gene functions through base editor screens empowered by barcoded sgRNAs. Nat. Biotechnol. *39*, 1403–1413. https://doi.org/10.1038/s41587-021-00944-1.

6. Zhu, S., Cao, Z., Liu, Z., He, Y., Wang, Y., Yuan, P., Li, W., Tian, F., Bao, Y., and Wei, W. (2019). Guide RNAs with embedded barcodes boost CRISPR-pooled screens. Genome Biol. *20*, 20. https://doi.org/10.1186/s13059-019-1628-0.