ETHzürich

Protocol: Genome-Scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening

bioRxiv Preprint

Protocol: Genome-scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening

Julia Joung[1,2,3,4,*], Silvana Konermann[2,3,4,*], Jonathan S. Gootenberg[2,3,4,5],

Omar O. Abudayyeh[2,3,4,6], Randall J. Platt[2,3,4], Mark D. Brigham[2,3,4],

Neville E. Sanjana[2,3,4], and Feng Zhang[1,2,3,4,†]

[1] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[2] Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA

[3] McGovern Institute for Brain Research at MIT, Cambridge, Massachusetts 02139, USA

[4] Departments of Brain and Cognitive Science and Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[5] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

[6] Department of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

* These authors contributed equally to this work.

† Correspondence should be addressed to zhang@broadinstitute.org (F.Z.)

Forward genetic screens are powerful tools for the unbiased discovery and functional characterization of specific genetic elements associated with a phenotype of interest. Recently, the RNA-guided endonuclease Cas9 from the microbial immune system CRISPR (clustered regularly interspaced short palindromic repeats) has been adapted for genome-scale screening by combining Cas9 with guide RNA libraries. Here we describe a protocol for genome-scale knockout and transcriptional activation screening using the CRISPR-Cas9 system. Custom- or ready-made guide RNA libraries are constructed and packaged into lentivirus for delivery into cells for screening. As each screen is unique, we provide guidelines for determining screening parameters and maintaining sufficient coverage. To validate candidate genes identified from the screen, we further describe strategies for confirming the screening phenotype as well as genetic perturbation through analysis of indel rate and transcriptional activation. Beginning with library design, a genome-scale screen can be completed in 6-10 weeks followed by 3-4 weeks of validation.

<p style="text-align:center">*   *   *   *   *</p>

INTRODUCTION

Systematic and high-throughput genetic perturbation technologies within live model organisms are necessary for fully understanding gene function and epigenetic regulation[1-3]. Forward genetic screens allow for a "phenotype-to-genotype" approach to mapping specific genetic perturbations to a phenotype of interest. Generally, this involves perturbing many genes at once, selecting cells or organisms for a desired phenotype, and then sequencing to identify the genetic features involved in the phenotypic changes. Initial screening approaches relied on chemical DNA mutagens to induce genetic changes, but this process was inefficient and mutations were costly to identify. More recently, tools that utilize the RNA interference (RNAi) pathway, specifically through short hairpin RNAs (shRNAs)[4-7], to perturb transcript levels have revolutionized screening approaches[8-13]. ShRNAs exploit the endogenous RNAi machinery to knock down sequence-complementary mRNA (**Fig. 1**). Despite the contribution of RNAi

screens to many biological advances, this approach is hampered by incomplete knockdown of transcripts and high off-target activity, resulting in low signal-to-noise and limited interpretations[14-16].

Cas9 as a tool for precise genome editing

Programmable nucleases have emerged as a promising new genetic perturbation technology capable of precisely recognizing and cleaving target DNA[17-19]. Particularly, the RNA-guided endonuclease Cas9 from the microbial immune system CRISPR (clustered regularly interspaced short palindromic repeat) has proved powerful for precise DNA modifications[20-23]. Cas9 is guided to specific genomic targets by guide RNAs through Watson-Crick base pairing. As with RNAi, Cas9 is thus easily retargetable; in contrast, however, extensive characterization has shown that Cas9 is much more robust and specific than RNAi[24-28].

Cas9 generates precise double-strand breaks (DSBs) at target loci that are repaired through either homology-directed repair (HDR) or more often, non-homologous end-joining (NHEJ)[29]. HDR precisely repairs the DSB using a homologous DNA template, whereas NHEJ is error-prone and introduces indels. When Cas9 is targeted to a coding region, loss-of-function (LOF) mutations can occur as a result of frameshifting indels that produce a premature stop codon and subsequent nonsense-mediated decay of the transcript or generate a non-functional protein (**Fig. 1**)[22, 23]. These features make Cas9 ideal for genome editing applications.

High-throughput loss-of-function screening using Cas9

Together with large pooled single guide RNA (sgRNA) libraries, Cas9 can mediate high-throughput LOF dissection of many selectable phenotypes. Indeed, Cas9 LOF screens have provided insight into the molecular basis of gene essentiality, drug and toxin resistance, as well as the hypoxia response[27, 30-42]. Previously, we constructed a genome-scale CRISPR-Cas9 knockout (GeCKO) library to probe 18,080 genes in the human genome for roles in BRAF-inhibitor vemurafenib resistance in a melanoma cell line[27]. A comparison of GeCKO with shRNA screening indicated that GeCKO had higher levels of consistency amongst guides targeting the same gene and higher validation rates.

Similarly, Cas9 knockout screening has been shown to be more consistent and effective than shRNA screening for the identification of essential genes[28]. The Cas9 system is also effective for screening *in vivo*. For instance, Chen et al. identified key factors in cancer metastasis by infecting a non-metastatic mouse lung cancer cell line with a mouse GeCKO library, transplanting it into mice, and sequencing metastases in the lung[34].

Genome-scale transcriptional activation with Cas9

In addition to Cas9-based knockout screens, catalytically inactive Cas9 (dCas9) fused to transcriptional activation and repression domains can be used to modulate transcription without modifying the genomic sequence[41-48]. CRISPR activation (CRISPRa) and CRISPR inhibition (CRISPRi) can be achieved by direct fusion or recruitment of activation and repression domains, such as VP64 and KRAB respectively[44, 49]. Of these alternative CRISPR screening approaches, CRISPRa is perhaps the most robust and reliable. Up until CRISPRa, gain-of-function (GOF) screens were primarily limited to cDNA overexpression libraries, which suffered from incomplete representation, overexpression beyond physiological levels and endogenous regulation, lack of isoform diversity, and high cost of construction. CRISPRa overcomes these limitations because it activates gene transcription at the endogenous locus and simply requires the synthesis and cloning of small sgRNAs.

The first generation of CRISPRa fused dCas9 to a VP64 or p65 activation domain to produce modest transcriptional upregulation, but was not suitable for genome-scale screening[44-47, 49]. Second generation CRISPRa designs produced more robust upregulation by recruiting multiple activation domains to the dCas9 complex. For instance, SunTag recruits multiple VP64 activation domains via a repeating peptide array of epitopes paired with single-chain variable fragment antibodies[41]. Another activation method, VPR, uses a tandem fusion of three activation domains, VP64, p65, and Rta to dCas9 to enhance transcriptional activation[43]. We devised an alternative approach to CRISPRa that involved incorporating MS2 binding loops in the sgRNA backbone to recruit several different activation domains, p65 and HSF1, to a dCas9-VP64 fusion (**Fig. 1**)[42]. By recruiting three distinct transcriptional effectors, this synergistic activation mediator (SAM) complex could robustly and reliably drive transcriptional upregulation.

As a proof of principle to demonstrate SAM's utility for GOF screening, similar to GeCKO we performed a genome-scale GOF screen to identify genes that, upon activation, confer vemurafenib resistance in a melanoma cell line. A comparison of SunTag, VPR, and SAM across various cell types and species suggested that SAM is the most robust and effective transcriptional activator[50].

Screening strategies

In general, there are two formats for conducting a screen: arrayed and pooled. For arrayed screens, individual reagents are aliquoted into separate wells in multi-well plates. This format allows for a diverse range of measured phenotypes such as fluorescence, luminescence, or even direct imaging of cellular phenotypes[2, 51-53], but at the same time it is costly and time-consuming. An alternative format, and one that has been widely used for Cas9-based screens, is pooled screening in which pooled lentiviral libraries are transduced at a low multiplicity of infection (MOI) to ensure that most cells receive only one stably-integrated RNA guide. After the screen is complete, deep sequencing of the sgRNAs in the bulk genomic DNA identifies changes in the sgRNA distribution due to the applied screening selection pressure. While pooled screens are generally limited to growth phenotypes or to fluorescence-activated cell sorting (FACS)-selectable phenotypes, they require less cost and effort than arrayed screens. Additional design considerations for screening include the type of selection pressure applied. This can be categorized as positive (e.g. resistance to a drug, toxin, or pathogen), negative (e.g. essential genes, toxicity), or marker gene selection (e.g. reporter gene expression) (**Box 1**).

Here we explain in detail how to set up and perform a pooled genome-scale knockout and transcriptional activation screen using Cas9. We describe protocols for designing and cloning an sgRNA library, packaging lentivirus for transduction, analyzing screening results, and validating candidate genes identified from the screen (**Fig. 2**). Although we specifically focus on knockout and activation screening using the GeCKO and SAM systems, the protocol can be applied to other types of screens (e.g. other CRISPRa systems, Cas9 knockdown, and saturated mutagenesis). For reference, we have compiled a table of previously published screens (**Table 1**).

Experimental Design

Design and selection of the sgRNA library. Although each sgRNA library is computationally designed for a specific purpose, the basic design process is consistent across libraries. First, the genomic regions of interest for targeting the sgRNA library are identified based on known sgRNA targeting rules (e.g. 5' conserved exons for gene knockout, upstream or downstream of the transcriptional start site for transcriptional activation or repression respectively). Second, all possible sgRNA targets with the Cas9 ortholog-specific protospacer adjacent motif (PAM) are identified and selected based on four criteria: (i) minimization of off-target activity, (ii) maximization of on-target activity, (iii) avoidance of homopolymer stretches (e.g. AAAA, GGGG) and (iv) GC content. Recent work has begun to elucidate the features that govern sgRNA specificity and efficiency[33, 39]. Although specificity and efficiency will likely vary across experimental settings, false positive sgRNAs in screens can still be mitigated by including redundant sgRNAs in the library and requiring multiple distinct sgRNAs targeting the same gene to display the same phenotype when identifying screening hits. Once the targeting sgRNAs have been chosen, additional non-targeting guides that do not target the genome should be included as negative controls.

We provide several genome-scale libraries for knockout and activation screening through Addgene. For knockout screening, the GeCKO v2 libraries target the 5' conserved coding exons of 19,050 human or 20,611 mouse coding genes with 6 sgRNAs per gene (Fig. 3)[54]. In addition to targeting coding genes, the GeCKO v2 libraries also target 1,864 human miRNAs or 1,175 mouse miRNAs with 4 sgRNAs per miRNA. Each species-specific library contains 1,000 non-targeting control sgRNAs. The GeCKO library is available in a 1 vector (lentiCRISPR v2) or 2 vector (lentiCas9-Blast and lentiGuide-Puro) format. For activation screening, the SAM libraries target the 200bp region upstream of the transcriptional start site of 23,430 human or 23,439 mouse RefSeq coding isoforms with 3 sgRNAs per isoform (Fig. 3)[42]. The library has to be combined with additional SAM effectors in a 2 vector (lentiSAM v2 and MS2-P65-HSF1) or 3 vector (dCas9-VP64, sgRNA(MS2), and MS2-P65-HSF1) format. Both GeCKO v2 and SAM libraries prioritize sgRNAs with minimal off-target activity.

In cases when a subset of genes is known to be involved in the screening phenotype and/or when the cell number is limited, one can consider performing a targeted screen that captures a subset of the genes in the genome-scale screens provided. We have included a python script for isolating the sgRNA target sequences corresponding to the genes in the targeted screen and adding flanking sequences for cloning. Additionally, one can consider adapting the sgRNA library plasmid backbone to the needs of the screen. For instance, when screening *in vivo* in complex tissues, one can use a cell-type specific promoter to ensure that only the cell type of interest is perturbed. To select for successful transduction by FACS, one can replace the antibiotic selection marker with a fluorescent marker. For these situations, we provide a protocol for cloning a custom sgRNA library.

Approaches for sgRNA library construction and delivery. Throughout the sgRNA library cloning and amplification process, it is important to minimize any potential bias that may affect screening results. For example, the number of PCR cycles in the initial amplification of the pooled oligo library synthesis should be limited to prevent introducing bias during amplification. Scale each step of the cloning procedure provided according to the size of the library to reduce loss of sgRNA representation. After sgRNA library transformation, limit the growth time to avoid intercolony competition which can result in plasmid amplification bias. We provide a protocol and accompanying python script for assessing sgRNA library distribution by next-generation sequencing (NGS) prior to screening.

Depending on the desired application, the sgRNA library can be delivered with lentivirus, retrovirus, or adeno-associated virus (AAV). Lentivirus and retrovirus integrate into the genome, whereas AAV does not integrate and thus for screening, AAV delivery is limited to non-dividing cells. In contrast, retrovirus only transduces dividing cells. In addition, AAV has a smaller insert size capacity compared to lentivirus and retrovirus. As a result, to date most of the screens have relied on lentiviral delivery and we have provided two methods for lentivirus production and transduction.

Screening. Since the parameters for each screen differs according to the screening phenotype, in lieu of providing a protocol for screening we have outlined general

considerations for setting the relevant screening parameters as well as technical advice for carrying out a screen in **Box 2**. Additional *in vivo* screening considerations are described in **Box 3**. We also provide guidelines for saturated mutagenesis screening design and analysis in **Box 4**.

Analysis of screening results. For examples of anticipated results, we provide data from genome-scale knockout and transcriptional activation screening for genes that confer BRAF inhibitor vemurafenib (PLX) resistance in a BRAF^{V600E} (A375) cell line[27, 42]. As a result of the screening selection pressure, at the end of the screen the sgRNA library distribution in the experimental condition should be significantly skewed compared to the baseline and control conditions, with some sgRNAs enriched and others depleted (**Fig. 4a,b**). The sgRNA representation, which is measured by NGS, in the experimental relative to the control condition determines the enrichment or depletion of the sgRNA. Then, depending on the type of screen (positive, negative, or marker gene selection), the enrichment or depletion of sgRNAs will be used to identify candidate genes that confer the screening phenotype.

Screening analysis methods such as RNAi gene enrichment ranking (RIGER) and redundant siRNA activity (RSA) typically select candidate genes with multiple enriched or depleted sgRNAs to reduce the possibility that the observed change in sgRNA distribution was due to off-target activity of a single sgRNA[55, 56]. RIGER ranks sgRNAs according to their enrichment or depletion and for each gene, examines the positions of the sgRNAs targeting that gene in the ranked sgRNA list[55]. The algorithm then assesses whether the set of positions is biased towards the top of the list using a Kolmogorov-Smirnov statistic and calculates an enrichment score and gene ranking based on a permutation test. RSA is similar to RIGER, except that it assigns statistical significance based on an iterative hypergeometric distribution formula[56]. In this protocol we describe in detail how to identify candidate genes using RIGER.

Each candidate gene identified from the screening analysis should have multiple significantly enriched or depleted sgRNAs in the experimental condition relative to the control (**Fig. 4c,d**). The RIGER *P* values of the candidate genes should also be significantly lower than the rest of the genes (**Fig. 4e,f**).

Validation of candidate genes. Given that the screening process can be noisy and the analysis produces a ranked list of candidate genes, it is necessary to verify that perturbation of the identified candidate genes confers the phenotype of interest. For validation, each of the sgRNAs that target the candidate gene can be individually cloned into the plasmid backbone of the sgRNA library and validated for the screening phenotype. In addition, the perturbation induced by each sgRNA, indel rate and transcriptional activation for knockout and activation screening respectively, will be quantified to establish a phenotype-to-genotype relationship.

Indel rates can be detected either by the SURVEYOR nuclease assay or by NGS. Compared to SURVEYOR, which we have described previously[57], NGS is more suitable for sampling a large number of sgRNA target sites and therefore described here. For measuring indel rates, it is important to design primers situated at least 50 bp from the target cleavage site to allow for the detection of longer indels. Our protocol for targeted NGS outlines a two-step PCR in which the first step uses custom primers to amplify the genomic region of interest and the second step uses universal, barcoded primers for multiplex deep sequencing on the Illumina platform. Relative to the one-step PCR method recommended for preparing sgRNA libraries for NGS, the two-step PCR method is more versatile and less costly for assessing many different target sites because custom primers for each target site can be readily combined with different universal, barcoded primers.

After NGS, indel rates can be calculated by running the provided python script that implements two different algorithms. The first aligns reads using the Ratcliff-Obershelp algorithm and then finds regions of insertion or deletion from this alignment[58]. The second method, adapted from the Geneious aligner scans k-mers across the read and maps the alignment to detect indels[59]. In practice, Ratcliff-Obershelp alignment algorithm is more accurate, while k-mer based alignment algorithm is faster. These indel rates are then adjusted to account for background indel rates via a maximum likelihood estimation (MLE) correction[25]. The MLE correction models the observed indel rate as a combination of the true indel rate resulting from Cas9 cleavage and a separately measured background indel rate. The true indel rate is that which maximizes the probability of the observed

read counts under the assumption that they obey a binomial distribution with the background rate.

Measurement of transcriptional activation usually entails isolation of RNA, reverse transcription of the RNA to cDNA, and quantitative PCR (qPCR). Various different methods have been described for each step of the process. In this protocol, we provide a method for reverse transcription followed by qPCR that is rapid, high-throughput, and cost-effective and thus ideal for quantifying fold upregulation for validation. Our method involves direct lysis of cells grown on a 96-well plate followed by reverse transcription and TaqMan qPCR. TaqMan-based detection is more specific and reproducible than SYBR-based detection because it relies on a fluorogenic probe specific to the target gene, whereas SYBR depends on a dsDNA-binding dye. TaqMan also allows for multiplexing with control probes that measure housekeeping gene expression as a proxy for total RNA concentration.

After validation of the screening phenotype and perturbation, we recommend verifying down- or up-regulation of protein expression for knockout or transcriptional activation screening respectively. Immunohistochemistry and Western blot are two of the most common methods for verifying protein expression. Immunohistochemistry requires fixing the validation cell lines and detecting the target protein using a specific antibody, whereas Western blot involves harvesting protein and separating by electrophoresis before staining with the specific antibody. Although immunohistochemistry provides additional information on protein localization, it often requires a more specific antibody than Western blot because proteins are not separated by size. Thus, Western blot is preferable for verifying protein expression of candidate genes.

MATERIALS
REAGENTS
SgRNA libraries and backbones
- lentiCRISPR v2 (Addgene, cat. no. 52961)
- lentiCas9-Blast (Addgene, cat. no. 52962)
- lentiGuide-Puro (Addgene, cat. no. 52963)
- lenti dCAS-VP64_Blast (Addgene, cat. no. 61425)
- lenti MS2-P65-HSF1_Hygro (Addgene, cat. no. 61426)

- lenti sgRNA(MS2)_zeo backbone (human; Addgene, cat. no. 61427)
- lenti sgRNA(MS2)_puro backbone (human; Addgene, cat. no. 73795)
- lenti sgRNA(MS2)_puro optimized backbone (mouse; Addgene, cat. no. 73797)
- lentiSAMv2 backbone (human; Addgene, cat. no. 75112)
- Human GeCKO v2 Library, 1 plasmid system (Addgene, cat. no. 1000000048)
- Human GeCKO v2 Library, 2 plasmid system (Addgene, cat. no. 1000000049)
- Mouse GeCKO v2 Library, 1 plasmid system (Addgene, cat. no. 1000000052)
- Mouse GeCKO v2 Library, 2 plasmid system (Addgene, cat. no. 1000000053)
- Human SAM Library, Zeo, 3 plasmid system (Addgene, cat. no. 1000000057)
- Human SAM Library, Puro, 3 plasmid system (Addgene, cat. no. 1000000074)
- Mouse SAM Library, Puro optimized, 3 plasmid system (Addgene, cat. no. 1000000075)
- Human SAM library, lentiSAMv2, 2 plasmid system (Addgene, cat. no. 75117)

Custom sgRNA library cloning

- Pooled oligo library (Twist Bioscience or CustomArray)
- PCR primers for amplifying oligo library for cloning are listed in **Table 2**. Primers longer than 60bp can be ordered as 4-nmol ultramers (Integrated DNA technologies)
- NEBNext High Fidelity PCR Master Mix, 2× (New England BioLabs, cat. no. M0541L)
- UltraPure DNase/RNase-free distilled water (Thermo Fisher, cat. no. 10977023)
- QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
- QIAquick gel extraction kit (Qiagen, cat. no. 28704)
- UltraPure TBE buffer, 10x (Thermo Fisher, cat. no. 15581028)
- SeaKem LE agarose (Lonza, cat. no. 50004)
- SYBR Safe DNA stain, 10,000x (Thermo Fisher, cat. no. S33102)
- 1-kb Plus DNA ladder (Thermo Fisher, cat. no. 10787018)
- 50bp DNA ladder (Thermo Fisher, cat. no. 10416014)
- TrackIt Cyan/Orange Loading Buffer (Thermo Fisher, cat. no. 10482028)
- FastDigest Esp3I (BsmBI; Thermo Fisher, cat. no. FD0454)
- FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher, cat. no. EF0651)
- DTT, Molecular Grade (Promega, cat. no. P1171)
- Gibson Assembly Master Mix (New England BioLabs, cat. no. E2611L)
- GlycoBlue Coprecipitant (Thermo Fisher, cat. no. AM9515)
- 2-Propanol (Sigma-Aldrich, cat. no. I9516-25ML)
- Sodium chloride solution (Sigma-Aldrich, cat. no. 71386-1L)
- Ethyl alcohol, Pure (Sigma-Aldrich, cat. no. 459844-500ML)
- Tris-EDTA buffer solution (Sigma-Aldrich, cat. no. 93283-100ML)

SgRNA plasmid amplification

- LB Agar, Ready-Made Powder (Affymetrix, cat. no. 75851)

- LB Broth, Ready-Made Powder (Affymetrix, cat. no. 75852)
- Ampicillin, 100 mg ml^{-1}, sterile filtered (Sigma-Aldrich, cat. no. A5354)
- Endura ElectroCompetent Cells (Lucigen, cat. no. 60242-1)
- NucleoBond Xtra Maxi EF (Macherey-Nagel, cat. no. 740424.10)
- One Shot Stbl3 chemically competent *E. coli* (Thermo Fisher, cat. no. C737303)
- SOC outgrowth medium (New England BioLabs, cat. no. B9020S)
- NucleoBond Xtra Midi EF (Macherey-Nagel, cat. no. 740420.50)

Next-generation sequencing

- Primers for amplifying the library for NGS are listed in **Table 3**. Primers for amplifying indels for NGS are listed in **Table 5**. Primers longer than 60bp can be ordered as 4 nmol ultramers (Integrated DNA technologies)
- NEBNext High Fidelity PCR Master Mix, 2× (New England BioLabs, cat. no. M0541L)
- KAPA HiFi HotStart ReadyMix, 2× (Kapa Biosystems, cat. no. KK2602)
- UltraPure DNase/RNase-free distilled water (Thermo Fisher, cat. no. 10977023)
- QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
- QIAquick gel extraction kit (Qiagen, cat. no. 28704)
- UltraPure TBE buffer, 10× (Thermo Fisher, cat. no. 15581028)
- SeaKem LE agarose (Lonza, cat. no. 50004)
- SYBR Safe DNA stain, 10,000x (Thermo Fisher, cat. no. S33102)
- 1-kb Plus DNA ladder (Thermo Fisher, cat. no. 10787018)
- TrackIt CyanOrange loading buffer (Thermo Fisher, cat. no. 10482028)
- Qubit dsDNA HS Assay Kit (Thermo Fisher, cat. no. Q32851)
- NextSeq 500/550 High Output Kit v2 (150 cycle; Illumina, cat. no. FC-404-2002)
- MiSeq Reagent Kit v3 (150 cycle; Illumina, cat. no. MS-102-3001)
- MiSeq Reagent Kit v2 (300 cycle; Illumina, cat. no. MS-102-2002)
- PhiX Control Kit v3 (Illumina, cat. no. FC-110-3001)
- Sodium hydroxide solution, 10 N (Sigma-Aldrich, cat. no. 72068-100ML)
- Tris, pH 7.0 (Thermo Fisher, cat. no. AM9850G)

Mammalian cell culture

- HEK293FT (Thermo Fisher, cat. no. R70007)
- DMEM, high glucose, GlutaMAX Supplement, pyruvate (Thermo Fisher, cat. no. 10569010)
- Penicillin-streptomycin, 100× (Thermo Fisher, cat. no. 15140122)
- Fetal bovine serum, premium grade (VWR, cat. no. 97068-085)
- TrypLE Express, no phenol red (Thermo Fisher, cat. no. 12604021)
- HUES 66 cell line (Harvard Stem Cell Science)
- Geltrex LDEV-free reduced growth factor basement membrane matrix (Thermo Fisher, cat. no. A1413202)

- mTeSR1 medium (Stemcell Technologies, cat. no. 05850)
- Normocin (InvivoGen, cat. no. ant-nr-1)
- Rho-associated protein kinase (ROCK) inhibitor (Y-27632; Millipore, cat. no. SCM075)
- Accutase (Stemcell Technologies, cat. no. 07920)
- Dulbecco's PBS (DPBS; Thermo Fisher, cat. no. 14190250)

Lentivirus production and titer

- Opti-MEM I reduced serum medium (Thermo Fisher, cat. no. 31985062)
- pMD2.G (Addgene, cat. no. 12259)
- psPAX2 (Addgene, cat. no. 12260)
- pcDNA3-EGFP transfection control plasmid (Addgene, cat. no. 13031)
- Lipofectamine 2000 transfection reagent (Thermo Fisher, cat. no. 11668019)
- PLUS Reagent (Thermo Fisher, cat. no. 11514015)
- Polyethylenimine HCl MAX, Linear, Mw 40,000 (PEI Max; Polysciences, cat. no. 24765-1)
- Sodium hydroxide solution, 10 N (Sigma-Aldrich, cat. no. 72068-100ML)
- Polybrene (Hexadimethrine bromide; Sigma-Aldrich, cat. no. 107689-10G)
- Blasticidin S HCl (Thermo Fisher, cat. no. A1113903)
- Puromycin dihydrochloride (Thermo Fisher, cat. no. A1113803)
- Hygromycin B (Thermo Fisher, cat. no. 10687010)
- Zeocin (Thermo Fisher, cat. no. R25001)
- CellTiter-Glo Luminescent Cell Viability Assay (Promega, cat. no. G7571)

Screening and Validation

- Quick-gDNA MidiPrep (Zymo Research, cat. no. D3100)
- DNA Binding Buffer (Zymo Research, cat. no. D4004-1-L)
- DNA Wash Buffer (Zymo Research, cat. no. D4003-2-24)
- DNA Elution Buffer (Zymo Research, cat. no. D3004-4-4)
- Ethyl alcohol, Pure (Sigma-Aldrich, cat. no. 459844-500ML)
- Primers for cloning the validation sgRNAs are listed in **Table 4**. (Integrated DNA technologies)
- T4 polynucleotide kinase (New England BioLabs, cat. no. M0201S)
- T4 DNA ligase reaction buffer, 10× (New England BioLabs, cat. no. B0202S)
- UltraPure DNase/RNase-free distilled water (Thermo Fisher, cat. no. 10977023)
- T7 DNA ligase with 2× rapid ligation buffer (Enzymatics, cat. no. L6020L)
- BSA, Molecular Biology Grade (New England BioLabs, cat. no. B9000S)
- FastDigest Esp3I (BsmBI; Thermo Fisher, cat. no. FD0454)
- DTT, Molecular Grade (Promega, cat. no. P1171)
- QuickExtract DNA extraction solution (Epicentre, cat. no. QE09050)
- RNase AWAY (VWR, cat. no. 53225-514)
- Proteinase K (Sigma-Aldrich, cat. no. P2308-25MG)

- Tris, 1 M, pH 8.0 (Thermo Fisher, cat. no. AM9855G)
- Deoxyribonuclease I bovine (Sigma-Aldrich, cat. no. D2821-50KU)
- UltraPure 1 M Tris-HCI Buffer, pH 7.5 (Thermo Fisher, cat. no. 15567027)
- Calcium chloride solution (Sigma-Aldrich, cat. no. 21115-1ML)
- Glycerol (Sigma-Aldrich, cat. no. G5516-100ML)
- $MgCl_2$, 1 M (Thermo Fisher, cat. no. AM9530G)
- Triton X-114 (Sigma-Aldrich, cat. no. X114-100ML)
- DTT, Molecular Grade (Promega, cat. no. P1171)
- Proteinase K Inhibitor (EMD Millipore, cat. no. 539470-10MG)
- Dimethyl sulfoxide (Sigma-Aldrich, cat. no. D8418-50ML)
- Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (Sigma-Aldrich, cat. no. E3889-10G)
- Sodium hydroxide solution, 10 N (Sigma-Aldrich, cat. no. 72068-100ML)
- RevertAid RT Reverse Transcription kit (Thermo Fisher, cat. no. K1691)
- Oligo dT (TTTTTTTTTTTTTTTTTTTTNN; Integrated DNA Technologies)
- TaqMan target probes, FAM dye (Thermo Fisher)
- Taqman endogenous control probe, VIC dye (e.g. Human GAPD, GAPDH, Endogenous Control VIC®/MGB probe, primer limited; Thermo Fisher, cat. no. 4326317E)
- TaqMan Fast Advanced Master Mix, 2x (Thermo Fisher, cat. no. 4444557)

EQUIPMENT

- Axygen 8-Strip PCR tubes (Fisher Scientific, cat. no. 14-222-250)
- Axygen PCR plates, 96 well (VWR, cat. no. PCR-96M2-HS-C)
- 384-well optical plate (e.g. Roche, LightCycler 480 Multiwell plate 384, cat. no. 5102430001)
- Axygen 1.5 ml Boil-Proof Microcentrifuge Tubes, (VWR, cat. no. 10011-702)
- Falcon tubes, polypropylene, 15 ml (Corning cat. no. 352097)
- Falcon tubes, polypropylene, 50 ml (Corning, cat. no. 352070)
- Filtered sterile pipette tips (e.g. Rainin)
- 100 mm × 15 mm Not TC-Treated Bacteriological Petri Dish (Corning, cat. no. 351029)
- 245mm Square BioAssay Dish without Handles, not TC-Treated Culture (Corning, cat. no. 431111)
- VWR Bacti Cell Spreaders (VWR, cat. no. 60828-688)
- AirPore Tape Sheets (Qiagen, cat. no. 19571)
- Nunc EasYFlask $25cm^2$, Filter Cap, 7 ml working volume (T25 flask; Thermo Scientific, cat. no. 156367)
- Nunc EasYFlask $75cm^2$, Filter Cap, 25 ml working volume, (T75 flask; Thermo Scientific, cat. no. 156499)
- Nunc EasYFlask 225 cm^2, filter cap, 70 ml working volume (T225 flask; Thermo Scientific, cat. no. 159934)

- Stericup filter unit, 0.45 µM (Millipore, cat. no. SCHVU02RE)
- Syringe filter unit, 0.45 µM (Millipore, cat. no. SLHV013SL)
- Disposable Syringes with Luer-Lok Tip (Fisher Scientific, cat. no. 14-829-45)
- Falcon tissue culture plate, 6 wells (Corning, cat. no. 353224)
- Falcon tissue culture plate, 12 wells (Corning, cat. no. 353043)
- Falcon tissue culture dish, 100mm (Corning, cat. no. 353003)
- 96 well flat clear bottom black polystyrene TC-treated microplates (Corning, cat. no. 3904)
- BD BioCoat clear Poly-D-Lysine 96 well clear flat bottom TC-treated microplate (Corning, cat. no. 356461)
- Cellometer SD100 Counting Chambers (Nexcelom Bioscience, cat. no. CHT4-SD100-002)
- Zymo-Spin V with Reservoir (Zymo Research, cat. no. C1016-25)
- Collection Tubes, 2 ml (Zymo Research, cat. no. C1001-25)
- Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-100 membrane (Millipore, cat. no. UFC910008)
- Thermocycler with programmable temperature stepping functionality, 96 well (e.g. Applied Biosystems Veriti, cat. no. 4375786)
- Real-time PCR system, 384 well (e.g. Roche Lightcycler 480 cat. no. 05015243001)
- Desktop microcentrifuges (e.g. Eppendorf, cat. nos. 5424 and 5804)
- Eppendorf ThermoStat C (Eppendorf, cat. no. 5383000019)
- Gene Pulser Xcell Microbial System (Bio-Rad, cat. no. 1652662)
- Digital gel imaging system (GelDoc EZ, Bio-Rad, cat. no. 170-8270), and blue sample tray (Bio-Rad, cat. no. 170-8273)
- Blue-light transilluminator and orange filter goggles (SafeImager 2.0; Invitrogen, cat. no. G6600)
- Gel quantification software (Bio-Rad, ImageLab or open-source ImageJ frmo the National Institutes of Health (NIH), USA, available at http://rsbweb.nih.gov/ij/)
- UV spectrophotometer (e.g. NanoDrop 2000c, Thermo Scientific)
- Plate spectrophotometer (e.g. Synergy H4 Hybrid Multi-Mode Microplate Reader, BioTek)
- Qubit Assay Tubes (Thermo Fisher, cat. no. Q32856)
- Qubit Fluorometer (Thermo Fisher, cat. no. Q33216)
- MiSeq System (Illumina, cat. no. SY-410-1003)
- NextSeq 500/550 System (Illumina, cat. nos. SY-415-1001 and SY-415-1002)
- Cell counter (e.g. Cellometer Image Cytometer, Nexcelom Bioscience)
- Sorvall Legend XTR Centrifuge (Thermo Fisher, cat. no. 75004520)

REAGENT SETUP

TBE electrophoresis solution Dilute TBE buffer in distilled water to a $1\times$ working condition, and store it at room temperature (18-22 ºC) for up to 6 months.

Ethanol, 80% (vol/vol) Prepare 80% (vol/vol) ethanol in UltraPure water right before use.

D10 medium For culture of HEK 293FT cells, prepare D10 medium by supplementing DMEM with GlutaMAX and 10% (vol/vol) FBS. For routine cell line culture and maintenance, D10 can be further supplemented with $1\times$ penicillin-streptomycin. Store the medium at 4 ºC for up to 1 month.

mTeSR1 medium For culture of human embryonic stem cells (hESCs), prepare mTeSR1 medium by supplementing it with the supplement supplied with the medium and 100 µg ml^{-1} Normocin. Prepared medium can be stored at 4 ºC for up to 2 months.

Proteinase K, 300 U ml^{-1} Resuspend 25 mg of Proteinase K in 2.5 ml of 10 mM Tris, pH 8.0 for 10 mg ml^{-1} (300 U ml^{-1}) of Proteinase K. Store at 4 ºC for up to 1 year.

Deoxyribonuclease I, 50 KU ml^{-1} Resuspend 50 KU of Deoxyribonuclease I in a solution containing 50% (vol/vol) Glycerol, 10 mM CaCl$_2$, and 50 mM Tris-HCl (pH 7.5) for 50 KU ml^{-1} of Deoxyribonuclease I. See **Table S1** for a detailed setup of the Deoxyribonuclease I storage solution. Store at -20 ºC for up to 1 year.

RNA lysis buffer Prepare an RNAse-free solution of 9.6 mM Tris-HCl (pH 7.8), 0.5 mM MgCl$_2$, 0.44 mM CaCl$_2$, 10 µM DTT, 0.1% (wt/vol) Triton X-114, and 3 U ml^{-1} Proteinase K in UltraPure water. The final pH of the solution should be approximately 7.8. Store at 4 ºC for up to 1 year. See **Table S2** for a detailed setup.

RNA lysis stop solution Prepare solution under RNAse-free conditions. Resuspend 10 mg of Proteinase K Inhibitor in 150 µl of DMSO for a final concentration of 100 mM. Prepare 0.5 M EGTA (pH 8.3) in UltraPure water. **Critical** EGTA is light sensitive, and can be stored at 4 ºC protected from light for up to 2 years. Combine for a final solution with 1 mM Proteinase K inhibitor, 90 mM EGTA, and 113 µM DTT in UltraPure water. Aliquot into 8-strip PCR tubes to avoid freeze-thaw and facilitate sample processing with multichannel pipettes. Store at -20 ºC for up to 1 year. See **Table S3** and **Table S4** for a detailed setup.

Oligo dT, 100 µM Resuspend oligo dT to 100 µM in UltraPure water. Aliquot and store at -20 ºC for up to 2 years.

EQUIPMENT SETUP

Large LB agar plates (245 mm square bioassay dish, ampicillin) Reconstitute the LB Broth with agar at a concentration of 35 g L^{-1} in deionized water and swirl to mix. Autoclave to sterilize. Allow the LB agar to cool to 55 °C before adding ampicillin to a final concentration of 100 µg ml^{-1} and swirl to mix. On a sterile bench area, pour ~300 ml of LB agar per 245 mm square bioassay dish. Place the lids on the plates and allow them to cool for 30-60 min until solidified. Invert the plates and let sit for several more hours or overnight. Agar plates can be stored in plastic bags or sealed with parafilm at 4 °C for up to 3 months.

Standard LB agar plates (100 mm Petri dish, ampicillin) Preparation of standard LB agar plates is similar to large LB agar plates, except pour ~20 ml of LB agar per 100 mm Petri dish. Store at 4 °C for up to 3 months.

PROCEDURE

Prior to performing the screen, construct a pooled sgRNA library by designing and cloning a custom sgRNA library (Steps 1-20) or amplifying a ready-made library from Addgene (Skip to Step 21).

Designing a targeted sgRNA library o TIMING 1 d

1.   *Input target genes for library design.* We provide a python script design_oligos.py that extracts the sgRNA spacers that target an input set of genes from the genome-scale screen. Once a set of genes for the targeted screen has been identified, prepare a csv file containing the names of the target genes with each line corresponding to one gene. Prepare another csv file for the annotated genome-scale library with the names of each gene in the first column and respective spacer sequences in the second column. Each line contains a different spacer sequence. The gene names in the target genes file should be in the same format as the names of the annotated library file.

2.   Isolate the subset of spacers from the genome-scale library that correspond to the target genes by running python design_oligos.py with the following optional parameters:

Flag	Description	Default
-o	Output csv file with names for target genes, corresponding spacer sequences, and oligo library sequences in columns from left to right	oligos.csv
-l	Annotated library csv file with names in the first column and corresponding spacer sequences in the second column	annotated_library.csv
-g	Target genes csv file with names of target genes	target_genes.csv
-gecko or -sam	Specify the type of library and add the respective flanking sequences to the spacers for the oligo library synthesis	Neither

3.   After running design_oligos.py, the subset of spacers for the target genes will be written to an output csv file. If -gecko or -sam is specified, the full oligo library sequence

containing the spacers and respective flanking sequences for synthesis will be in the last column.

4.   Synthesize the oligo library as a pool on an array through a DNA synthesis platform such as Twist Bioscience or CustomArray. Parafilm and store pooled oligos at - 20 ºC.

Cloning a custom sgRNA library o TIMING 2 d

5.   Throughout the sgRNA library cloning process, refer to the table below for the number of reactions recommended at each cloning step for a library size of 100,000 sgRNAs and scale the number of reactions according to the size of the targeted sgRNA library.

Steps	Cloning process	Number of reactions
6-10	PCR amplification of pooled oligo library	12
11-13	Restriction digest of plasmid backbone	16
14-15	Gibson assembly	10 with sgRNA insert, 5 control
16-20	Isopropanol precipitation	10 with sgRNA insert, 5 control
21-32	Electroporation of library	20

6.   *PCR amplification of pooled oligo library.* Amplify the pooled oligo library using the Oligo-Fwd and Oligo-Rev primers (**Table 2**). Prepare a master mix using the reaction ratios outlined below:

Component	Amount per reaction (µl)	Final concentration
NEBNext High Fidelity PCR Master Mix, 2×	12.5	1×
Pooled oligo library template	1	0.04 ng µl^{-1}

Oligo-Fwd primer (Universal)	1.25	0.5 µM
Oligo-Knockout-Rev or Oligo-Activation-Rev primer	1.25	0.5 µM
UltraPure water	9	
Total	25	

Critical Step To minimize error in amplifying oligos, it is important to use a high-fidelity polymerase. Other high-fidelity polymerases, such as PfuUltra II (Agilent) or Kapa HiFi (Kapa Biosystems), may be used as a substitute.

7. Aliquot the PCR master mix into 25 µl reactions and perform a PCR by using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	98 ºC, 30 s		
2-21	98 ºC, 10 s	63 ºC, 10 s	72 ºC, 15 s
22			72 ºC, 2 min

Critical Step Limit the number of PCR cycles to 20 cycles during amplification to reduce potential biases introduced during amplification.

8. After the reaction is complete, pool the PCR reactions and purify the PCR product by using the QIAquick PCR purification kit according to the manufacturer's directions.

9. Run PCR purified oligo library on a 2% (wt/vol) agarose gel along with a 50bp ladder.

Critical Step Run on a 2% (wt/vol) agarose gel for long enough to separate the target library (140bp) from a possible primer dimer of ~120bp. Under the optimized PCR conditions suggested above the presence of primer dimers should be minimal.

10. Gel extract the purified PCR product using the QIAquick gel extraction kit according to the manufacturer's directions and quantify.

11. *Restriction digest of plasmid backbone.* Digest the desired library plasmid backbone with the restriction enzyme Esp3I (BsmBI) that cuts around the sgRNA target region. Refer to the master mix set up below for the reaction ratios:

Component	Amount per reaction (µl)	Final concentration
FastDigest Buffer, 10×	2	1×
Library Plasmid Backbone	1	50 ng µl^{-1}
FastDigest Esp3I (BsmBI)	1	
FastAP Thermosensitive Alkaline Phosphatase	1	
DTT, 100 mM	0.2	1 mM
UltraPure water	14.8	
Total	20	

12.   Aliquot 20 µl reactions from the master mix and incubate the restriction digest reaction at 37 ºC for 1 h.

13.   After the reaction has completed, run the restriction digest on a 2% (wt/vol) agarose gel and gel extract the library plasmid backbone using the QIAquick gel extraction kit according to the manufacturer's protocol and quantify. Note that the Gecko library backbones contain a 1880bp filler sequence which should be visible as a dropout. The SAM library backbones do not contain a filler sequence and the expected dropout of 20bp is usually not readily visible.

14.   *Gibson Assembly*. Set up a master mix for the Gibson reactions on ice according to the reaction ratios below. Be sure to include reactions without the sgRNA library insert as a control.

Component	Amount per reaction	Final concentration
Gibson Assembly Master Mix, 2×	10 µl	1×
Digested library Plasmid Backbone	330 ng	16.5 ng µl^{-1}
SgRNA library insert or UltraPure water control	50 ng	2.5 ng µl^{-1}

UltraPure water	To 20 µl
Total	20 µl

15.   Aliquot 20 µl reactions from the master mix and incubate the Gibson reaction at 50 ºC for 1h.

Pause Point Completed Gibson reactions can be stored at -20 ºC for at least 1 week.

16.   *Isopropanol precipitation.* Pool cloning and control reactions separately. Purify and concentrate the sgRNA library by mixing the following:

Component	Amount per reaction (µl)	Final concentration
Gibson Assembly Reaction	20	
Isopropanol	20	
GlycoBlue Coprecipitant	0.2	0.375 µg µl^{-1}
NaCl solution	0.4	50 mM
Total	~40	

Critical Step In addition to concentrating the library, purification by isopropanol precipitation removes salts from the Gibson reaction that can interfere with electroporation.

17.   Vortex and incubate at room temperature for 15 min and centrifuge at >15,000 × g for 15 min to precipitate the plasmid DNA. The precipitated plasmid DNA should appear as a small light blue pellet at the bottom of the microcentrifuge tube.

18.   Aspirate the supernatant and gently wash the pellet twice without disturbing it using 1 ml of ice-cold (-20 ºC) 80% (vol/vol) ethanol in UltraPure water.

19.   Carefully remove any residual ethanol and air dry for for 1 min.

20.   Resuspend the plasmid DNA pellet in 5 µl of TE per reaction by incubating at 55 ºC for 10 min and quantify the targeted sgRNA library by nanodrop. Isopropanol-purified sgRNA libraries can be stored at -20 ºC for several months.

Amplification of pooled sgRNA library o TIMING 2 d

21. *Pooled sgRNA library transformation.* Electroporate the library at 50-100 ng μl^{-1} using Endura ElectroCompetent cells according to the manufacturer's directions. If amplifying a ready-made genome-scale library from Addgene, repeat for a total of 8 electroporations. If amplifying a targeted sgRNA library refer to Step 5 to scale the amplification steps appropriately and include an additional electroporation for the control Gibson reaction without sgRNA library insert.

22. Pre-warm 1 standard LB agar plate (100 mm Petri dish, ampicillin) and large LB agar plates (245 mm square bioassay dish, ampicillin). Each large LB agar plate can be substituted with 10 standard LB agar plates. For amplification of a targeted sgRNA library, include an additional standard LB agar plate for the control Gibson reaction.

23. After the recovery period, pool electroporated cells and mix well by inverting.

24. Plate a dilution for calculating transformation efficiency. To prepare the dilution mix, add 10 µl of the pooled electroporated cells to 990 µl of LB medium for a 100-fold dilution and mix well. Then add 100 µl of the 100-fold dilution to 900 µl of LB medium for a 1,000-fold dilution and mix well.

25. Plate 100 µl of the 1,000-fold dilution onto a pre-warmed standard LB agar plate (10 cm Petri dish, ampicillin). This is a 10,000-fold dilution of the full transformation that will be used to estimate the transformation efficiency. If amplifying a targeted sgRNA library, repeat Steps 24-25 for the control Gibson reaction.

26. Plate pooled electroporated cells. Add 1 volume of LB medium to the pooled electroporated cells, mix well, and plate on large LB agar plates (option A) or standard LB agar plates (option B).

    a. Plate 2 ml of electroporated cells on each of the pre-warmed large LB agar plates using a cell spreader. Spread the liquid culture until it is largely absorbed into the agar and does not drip when the plate is inverted. At the same time, make sure the liquid culture does not completely dry out as this will lead to poor survival.

    b. Alternatively, plate 200 µl of electroporated cells on each of the pre-warmed standard LB agar plates using the same technique as described in Step 26a.

Critical Step Plating the electroporated cells evenly is important for preventing intercolony competition that may skew the sgRNA library distribution.

27. Incubate all LB agar plates overnight at 37 ºC for 12-14 h.

Critical Step Limiting the bacterial growth time to 12-14 h ensures that there is sufficient growth for sgRNA library amplification without potentially biasing the sgRNA library distribution through intercolony competition or differences in colony growth rates.

28. Calculate electroporation efficiency.

　　a. Count the number of colonies on the dilution plate.

　　b. Multiply the number of colonies by 10,000 and the number of electroporations to obtain the total number of colonies on all plates. If amplifying a ready-made sgRNA library from Addgene, proceed if the total number of colonies is greater than 100 colonies per sgRNA in the library. If amplifying a custom targeted sgRNA library, proceed if there are more than 500 colonies per sgRNA in the library.

　　Critical Step Obtaining a sufficient number of colonies per sgRNA is crucial for ensuring that the full library representation is preserved and that sgRNAs do not drop out during amplification.

　　c. In addition, for amplification of a targeted sgRNA library, calculate the electroporation efficiency for the control Gibson reaction and proceed if there are at least 20 times more colonies per electroporation in the sgRNA library condition compared to the control Gibson reaction.

Troubleshooting

29. *Harvest colonies from the LB agar plates.* Pipette 10 ml of LB medium onto each large LB agar plate or 1 ml of LB medium onto each standard LB agar plate. Gently scrape the colonies off with a cell spreader and transfer the liquid with scraped colonies into a 50-ml Falcon tube.

30. For each LB agar plate, repeat Step 29 for a total of 2 LB medium washes to capture any remaining bacteria.

31. Maxiprep the amplified sgRNA library by using the Macherey-Nagel NucleoBond Xtra Maxi EF according to the manufacturer's directions. Calculate the number of maxipreps needed by measuring the OD600 of the harvested bacterial

suspension as follows: Number of maxipreps = OD600*(total volume of suspension)/1200.

Critical Step Using an endotoxin-free plasmid purification kit is important for avoiding endotoxicity in virus production and mammalian cell culture.

32.   Pool the resulting plasmid DNA and quantify. Maxiprepped sgRNA library can be aliquoted and stored at -20 ºC.

Next-generation sequencing of the amplified sgRNA library to determine sgRNA distribution o TIMING 2-3 d

33.   *Library PCR for NGS*. We have provided NGS primers that amplify the sgRNA target region with Illumina adapter sequences (**Table 3**). To prepare the sgRNA library for NGS, set up a reaction for each of the 10 NGS-Lib-Fwd primers and 1 NGS-Lib-KO-Rev or NGS-Lib-SAM-Rev barcode primer as follows:

Component	Amount per reaction (µl)	Final concentration
NEBNext High Fidelity PCR Master Mix, 2×	25	1×
Pooled sgRNA library template	1	0.4 ng µl^{-1}
NGS-Lib-Fwd primer (unique)	1.25	0.25 µM
NGS-Lib-KO-Rev or NGS-Lib-SAM-Rev primer (barcode)	1.25	0.25 µM
UltraPure water	21.5	
Total	50	

Critical step Using a different reverse primer with a unique barcode for each library allows for pooling and sequencing of different libraries in a single NextSeq or HiSeq run. This is more efficient and cost-effective than running the same number of libraries on multiple Miseq runs.

Critical Step To minimize error in amplifying sgRNAs, it is important to use a high-fidelity polymerase. Other high-fidelity polymerases, such as PfuUltra II (Agilent) or Kapa HiFi (Kapa Biosystems), may be used as a substitute.

34.   Perform a PCR by using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	98 ºC, 3 min		
2-23	98 ºC, 10 s	63 ºC, 10 s	72 ºC, 25 s
24			72 ºC, 2 min

35.   After the reaction is complete, pool the PCR reactions and purify the PCR product by using the QIAquick PCR purification kit according to the manufacturer's directions.

36.   Quantify the purified PCR product and run 2 µg of the product on a 2% (wt/vol) agarose gel. Successful reactions should yield a ~260-270bp product for the knockout library and a ~270-280bp product for the activation library. Gel extract using the QIAquick gel extraction kit according to the manufacturer's directions.

Pause Point Gel-extracted samples can be stored at -20 ºC for several months.

37.   Quantify the gel-extracted samples using the Qubit dsDNA HS Assay Kit according to the manufacturer's instructions.

38.   Sequence the samples on the Illumina MiSeq or NextSeq according to the Illumina user manual with 80 cycles of read 1 (forward) and 8 cycles of index 1. We recommend sequencing with 5% PhiX control on the MiSeq or 20% PhiX on the NextSeq to improve library diversity and aiming for a coverage of >100 reads per sgRNA in the library.

39.   *Analyze sequencing data with count_spacers.py*. Install biopython (http://biopython.org/DIST/docs/install/Installation.html). Prepare a csv file containing the guide spacer sequences with each line corresponding to one sequence.

40.   To determine the spacer distribution, run python count_spacers.py with the following optional parameters:

Flag	Description	Default

-f	Fastq file containing NGS data for analysis	NGS.fastq
-o	Output csv file with guide spacer sequences in the first column and respective read counts in the second column	library_count.csv
-i	Input csv file with guide spacer sequences	library_sequences.csv
-no-g	Indicate absence of a guanine before the guide spacer sequence	guanine is present

Place all relevant files are in the same folder before running python count_spacers.py. The human SAM libraries do not have a guanine before the guide spacer sequence, so make sure to run the script with the parameter -no-g when analyzing those libraries.

41. After running count_spacers.py, spacer read counts will be written to an output csv file. Relevant statistics including the number of perfect guide matches, non-perfect guide matches, sequencing reads without key, the number of reads processed, percentage of perfectly matching guides, percentage of undetected guides, and skew ratio will be written to statistics.txt. An ideal sgRNA library should have more than 70% perfectly matching guides, less than 0.5% undetected guides, and a skew ratio of less than 10.

Troubleshooting

Lentivirus production and titer o TIMING 8-10 d

42. *HEK 293FT maintenance*. Cells are cultured in D10 medium at 37 ºC with 5% CO_2 and maintained according to the manufacturer's recommendation.

43. To passage, aspirate the medium and rinse the cells by gently adding 5 ml of TrypLE to the side of the T225 flask, so as not to dislodge the cells. Remove the TrypLE and incubate the flask for 4-5 min at 37 ºC until the cells have begun to detach. Add 10 ml of warm D10 to the flask and dissociate the cells by pipetting them up and down gently, and transfer the cells to a 50-ml Falcon tube.

Critical Step We typically passage cells every 1-2 d at a split ratio of 1:2 or 1:4, never allowing cells to reach more than 70% confluency. For lentivirus production, we recommend using HEK 293FT cells with a passage number less than 10.

44.   *Preparation of cells for transfection.* Seed the well-dissociated cells into T225 flasks 20-24 h before transfection at a density of 1.8×10^7 cells per flask in a total volume of 45 ml D10 medium. Plate 5 T225 flasks for knockout screening and 9 T225 flasks for activation screening.

Critical Step Do not plate more cells than the recommended density, as doing so may reduce transfection efficiency.

45.   *Lentivirus plasmid transfection.* Cells are optimal for transfection at 80-90% confluency. Transfect 4 T225 flasks with the sgRNA library and 1 T225 flask with a constitutive GFP expression plasmid as a transfection control. If transfecting for activation screening, transfect an additional 4 T225 flasks with the Cas9 activator components that are not in the sgRNA library backbone, i.e. MS2-p65-HSF1. We outline below a transfection method using Lipofectamine 2000 and PLUS reagent. Alternatively, we describe a cost-effective method for lentivirus transfection with Polyethylenimine (PEI) in **Box 5**.

Critical Step Transfecting at the recommended cell density is crucial for maximizing transfection efficiency. Lower densities can result in Lipofectamine 2000 toxicity for cells, while higher densities can reduce transfection efficiency.

    a.   For each lentiviral target, combine the following lentiviral target mix in a 15-ml or 50-ml Falcon tube and scale up accordingly:

Component	Amount per T225 flask
Opti-MEM	2250 µl
pMD2.G (lentiviral helper plasmid)	15.3 µg
psPAX (lentiviral helper plasmid)	23.4 µg
Lentiviral target plasmid	30.6 µg

    b.   Prepare the PLUS reagent mix as follows and invert to mix:

Component	Amount per T225 flask
Opti-MEM	2250 µl

| PLUS reagent | 297 µl |

    c.   Add the PLUS reagent mix to the lentiviral target mix, invert, and incubate at room temperature for 5 min.

    d.   Prepare the Lipofectamine reagent mix as follows and invert to mix:

Component	Amount per T225 flask
Opti-MEM	4500 µl
Lipofectamine 2000	270 µl

    e.   Add the lentiviral target and PLUS reagent mix to the Lipofectamine reagent mix, invert, and incubate at room temperature for 5 min.

    f.   Pipette 9 ml of the lentiviral transfection mix into each T225 flask and shake gently to mix. Return the T225 flasks to the incubator.

    g.   After 4 h, replace the medium with 45 ml of pre-warmed D10 medium. The constitutive GFP expression plasmid transfection control indicates transfection efficiency.

Troubleshooting

46.   *Harvest and store lentivirus.* 2 d after the start of lentiviral transfection, pool the lentivirus supernatant from the T225 flasks with the same lentivirus and filter out cellular debris using Millipore's 0.45 µm Stericup filter unit.

Pause Point The filtered lentivirus supernatant can be aliquoted and stored at -80 ºC. Avoid freeze-thawing lentivirus supernatant.

47.   *Determine the lentiviral titer through transduction.* The CRISPR-Cas9 system has been used in a number of mammalian cell lines. Conditions may vary for each cell line. Lentiviral titer should be determined using the relevant cell line for the screen. Below we detail transduction conditions and calculation of viral titer for HEK 293FT cells (option A) and hESC HUES66 cells (option B).

Critical Step We recommend performing a kill curve for the antibiotic used to select the sgRNA library on your cells prior to determination of lentiviral titer. It is important to use the lowest concentration of antibiotic sufficient to to kill the negative control within 4-7

days to avoid excessively stringent selection that biases selection for cells transduced with multiple sgRNAs.

    a.   Lentiviral transduction and titer for HEK 293FT cells by spinfection

        i.   *HEK 293FT maintenance and passaging*. Refer to steps 42-43.

        ii.   *Preparation of cells for spinfection*. For each lentivirus, seed 6 wells of a 12-well plate at a density of 3×10^6 cells in 2 ml D10 medium per well with 8 µg ml^{-1} of polybrene. In each well, add 400 µl, 200 µl, 100 µl, 50 µl, 25 µl, or 0 µl of lentivirus supernatant. Mix each well thoroughly by pipetting up and down.

        iii.   Spinfect the cells by centrifuging the plates at $1000 \times g$ for 2 h at 33 ºC. Return the plates to the incubator after spinfection.

        iv.   *Replating spinfection for calculation of viral titer*. 24 h after the end of spinfection, remove the medium, gently wash with 400 µl TrypLE per well, add 100 µl of TrypLE, and incubate at 37 ºC for 5 min to dissociate the cells. Add 2 ml of D10 medium per well and resuspend the cells by pipetting up and down.

        v.   Determine the cell concentration for the 0 µl lentivirus supernatant condition.

        vi.   For each virus condition, seed 4 wells of a 96-well clear bottom black tissue culture plate at a density of 4×10^3 cells based on the cell count determined in the previous step in 100 µl of D10 medium. Add an additional 100 µl of D10 medium with the corresponding selection antibiotic for the virus at an appropriate final concentration to 2 wells and 100 µl of regular D10 medium to the other 2 wells for a total of 2 bioreps per virus condition.

        vii.   72-96 h after replating, when the no virus conditions contain no viable cells and the no antibiotic selection conditions are at 80-90% confluency, quantify the cell viability for each condition using CellTiter Glo according to the manufacturer's protocol. We have found that Cell Titer Glo can be diluted 1:4 in PBS to reduce cost while still achieving optimal results.

viii.   For each virus condition, the multiplicity of infection (MOI) is calculated as the luminescence, or viability, of the condition with antibiotic selection divided by the condition without antibiotic selection. A linear relationship between lentivirus supernatant volume and MOI is expected at lower volumes, with saturation achieved at higher volumes.

b.   Lentiviral transduction and titer for hESC HUES66 cells by mixing

i.   *HUES66 maintenance*. We routinely maintain HUES66 cells (a hESC cell line) in feeder-free conditions with mTeSR1 medium on GelTrex-coated tissue culture plates. To coat a 100-mm tissue culture dish, dilute cold GelTrex 1:100 in 5 ml of cold DMEM, cover the entire surface of the culture dish, and place the dish in the incubator for at least 30 min at 37 ºC. Aspirate the GelTrex mix prior to plating. During passaging and plating, mTeSR1 medium is supplemented further with 10 µM ROCK inhibitor. The mTeSR1 medium is refreshed daily.

ii.   *Passaging HUES66*. Aspirate the medium and rinse the cells once by gently adding 10 ml of DPBS to the side of the 100-mm tissue culture dish, so as not to dislodge the cells. Dissociate the cells by adding 2 ml of Accutase and incubate at 37 ºC for 3-5 min until the cells have detached. Add 10 ml of DMEM, resuspend the dissociated cells, and pellet the cells at $200 \times$ g for 5 min. Remove the supernatant, resuspend the cells in mTeSR1 medium with 10 µM ROCK inhibitor and replate the cells onto GelTrex-coated plates. Replace with normal mTeSR1 medium 24 h after plating.

Critical Step We typically passage cells every 4-5 d at a split ratio of 1:5 or 1:10, never allowing cells to reach more than 70% confluency.

iii.   *Preparation of cells for lentiviral transduction*. For each lentivirus, plate 6 wells of a Geltrex-coated 6-well plate at a density of 5×10^5 cells in 2 ml mTeSR1 medium per well. In each well, add 400 µl, 200 µl, 100 µl, 50 µl, 25 µl, or 0 µl of lentivirus supernatant, fill to a total volume to 3 ml with DPBS, and supplement with 10 µM ROCK inhibitor. Plate an

additional no antibiotic selection control well at the same seeding density without virus. Mix each well thoroughly by pipetting up and down.

iv.   24 h after lentiviral transduction, replace the medium with mTeSR1 containing the relevant antibiotic selection. For the no antibiotic selection control well, replace the medium with normal mTeSR1 medium. Refresh the mTeSR1 medium with and without antibiotic selection every day until the plate is ready for the next step.

v.   *Calculation of viral titer*. 72-96 h after starting the antibiotic selection, when the no virus condition contains no viable cells and the no antibiotic selection control condition is 80-90% confluent, rinse the cells with 2 ml of DPBS, add 500 µl of Accutase, and incubate at 37 ºC for 3-5 min to dissociate the cells. Add 2 ml of DMEM and mix well.

vi.   Count and record the number of cells in each well.

vii.   For each virus condition, the MOI is calculated as the number of cells in the antibiotic selection condition divided by the number of cells in the no antibiotic selection control. A linear relationship between lentivirus supernatant volume and MOI is expected at lower volumes, with saturation achieved at higher volumes.

Troubleshooting

Lentiviral transduction and screening o TIMING 3-6 w

Skip to Step 50 if performing a knockout screen.

48.   *Generation of a cell line with stably expressed Cas9 activation components*. Prior to performing an activation screen, transduce the relevant cell line with the additional Cas9 activation components that are not present in the sgRNA library backbone at an MOI < 0.7. If two additional Cas9 activation components are required, both components can be transduced at the same time. Scale up as necessary to generate sufficient cells for maintaining sgRNA representation after sgRNA library transduction and selection. We have found that generation of a clonal line with Cas9 or SAM components is not necessary for successful screening. The cells can therefore be transduced and selected as a bulk population at the desired scale. Below we describe lentiviral transduction and cell

line generation methods for HEK 293FT cells (option A) and hESC HUES66 cells (option B).

    a.   Generation of HEK 293FT cell lines

        i.   Seed cells in 12-well plates at a density of 3×10^6 cells in 2 ml D10 medium per well with 8 µg ml^{-1} of polybrene. Add the appropriate volume of lentivirus supernatant to each well, and make sure to include a no virus control. Mix each well thoroughly by pipetting up and down.

        ii.   Spinfect the cells by spinning the plates at $1000 \times$ g for 2 h at 33 ºC. Return the plates to the incubator after spinfection.

        iii.   24 h after the end of spinfection, remove the medium, gently wash with 400 µl TrypLE per well, add 100 µl of TrypLE, and incubate at 37 ºC for 5 min to dissociate the cells. To each well, add 2 ml of D10 medium with the appropriate selection antibiotic for the lentivirus and resuspend the cells by pipetting up and down.

        iv.   Pool the resuspended cells from the wells with virus and seed the cells into T225 flasks at a density of 9×10^6 cells per flask in 45 ml of D10 medium with selection antibiotic.

        v.   Transfer the resuspended cells from the no virus control into a T75 and add 13 ml of D10 medium with selection antibiotic.

        vi.   Refresh the selection antibiotic every 3 d and passage as necessary for 4-7 d and until there are no viable cells in the no virus control.

    b.   Generation of HUES66 cell lines

        i.   Seed cells in Geltrex-coated 6-well plates at a density of 5×10^5 cells in of 2 ml mTeSR1 medium per well. Add the appropriate volume of lentivirus supernatant to each well, and make sure to include a no virus control. Fill up the total volume to 3 ml with DPBS, and supplement with 10 µM ROCK inhibitor. Mix each well thoroughly by pipetting up and down.

        ii.   24 h after lentiviral transduction, replace the medium with mTeSR1 containing the relevant selection antibiotic. Refresh the mTeSR1

medium with selection antibiotic every day and passage as necessary for 4-7 d and until there are no viable cells in the no virus control.

Critical Step The lentiviral transduction method for generating a cell line for screening should be consistent with the method for titering the virus to ensure that cells are transduced at the appropriate MOI.

49.   After selecting for successfully transduced cells, allow the cells to recover from selection by culturing in normal medium for 2-7 d before transducing with the sgRNA library. If culturing for more than 7 d after selection or after freezing cells, re-select the Cas9 activation cell line with the appropriate selection antibiotic to ensure expression of the Cas9 activation components and allow the cells to recover before sgRNA library transduction.

Pause Point Cells can be frozen down according to the manufacturer's protocol.

50.   *Transduction of cells with the sgRNA library.* Refer to Steps 48-49 for lentiviral transduction at the appropriate MOI and selection of transduced cell lines. To ensure that most cells receive only one genetic perturbation, transduce the sgRNA library at an MOI < 0.3. Scale up the transduction such that the sgRNA library has a coverage of >500 cells expressing each sgRNA. For example, for a library size of 100,000 unique sgRNAs, transduce 1.67×10^8 cells at an MOI of 0.3. After the appropriate selection for 4-7 days, the cells are ready for screening. For knockout screening, we have found that maximal knockout efficiency is achieved 7 days after sgRNA transduction and therefore recommend selecting for 7 days before starting the screen selection. In contrast, maximal SAM activation is achieved as early as 4 days after sgRNA transduction. If selection is complete based on the no virus control, gain-of-function screening can be started 5 days after transduction. We generally recommend to perform 4 independent screening replicates (i.e. 4 separate sgRNA library infections followed by separate screening selection). Multiple bioreps are critical for determining screening hits with a high rate of validation.

Critical Step It is important to aim for a coverage of >500 cells per sgRNA to guarantee that each perturbation will be sufficiently represented in the final screening readout. Increase the coverage as necessary if the screening selection pressure is not very strong or if performing a negative selection screen. Transducing the sgRNA library at an MOI <

0.3 ensures that most cells receive at most one genetic perturbation. Transducing at higher MOI's may confound screening results.

51.   Since the parameters of each screen depends on the screening phenotype of interest, we provide guidelines and technical considerations for the screen (**Box 2-4**).

Harvest genomic DNA for screening analysis o TIMING 3-4 d

52.   *Harvest genomic DNA.* At the end of the screen, harvest genomic DNA (gDNA) from a sufficient number of cells to maintain a coverage of >500. For a library size of 100,000 unique sgRNAs, harvest gDNA from at least 5×10^7 cells for downstream sgRNA analysis using the Zymo Research Quick-gDNA MidiPrep according to the manufacturer's protocol. Make sure to tighten the connection between the reservoir and the column and centrifuge at a sufficient speed and time to remove any residual buffer. Addition of a final dry spin is recommended to remove residual wash buffer. Elution should be performed twice with 150-200 µl each for maximum recovery of gDNA. **Pause Point** Frozen cell pellets or isolated gDNA can be stored at -20 ºC for several months.

53.   *Preparation of the gDNA for NGS analysis.* Refer to Steps 33-34 for how to amplify the sgRNA for NGS. Scale up the number of reactions such that all of the gDNA harvested from the screen is amplified. Each 50 µl reaction can hold up to 2.5 µg of gDNA. Barcoded NGS-Lib-Rev primers enable sequencing of different screening conditions and bioreps on the same sequencing run.

Troubleshooting

54.   *Purification of amplified screening NGS library.* For large-scale PCR purification, we recommend using the Zymo-Spin V with Reservoir. Add 5 volumes of DNA Binding Buffer to the PCR reaction, mix well, and transfer to Zymo-Spin V with Reservoir in a 50 ml Falcon tube. Each Zymo-Spin V column can hold up to 12 ml. Make sure to tighten the connection between the reservoir and the Zymo-Spin V column.

55.   Centrifuge at $500 \times$ g for 5 min. Discard the flow-through.

56.   Add 2 ml of DNA Wash Buffer and centrifuge at $500 \times$ g for 5 min. Discard the flow-through and repeat for an additional wash.

57.   Remove the reservoir from the Zymo-Spin V column and transfer the column to a 2 ml collection tube. In a microcentrifuge, spin at the maximum speed (>12,000 × g) for 1 min to remove residual wash buffer.

58.   Transfer the Zymo-Spin V column to a 1.5 ml microcentrifuge tube. Add 150 µl of elution buffer, wait for 1 min, and spin at the maximum speed (>12,000 × g) for 1.5 min to elute the purified PCR reaction.

59.   Pool the purified PCR reactions and quantify. Refer to Steps 36-41 for NGS analysis of the sgRNA distribution.

60.   *Analysis of screening results with RNAi gene enrichment ranking (RIGER)*. Prior to RIGER analysis, determine the sgRNA fold change due to screening selection. For each biorep of screening experimental or control condition, add a pseudocount of 1 to the NGS read count of each sgRNA and normalize by the total number of NGS read counts for that condition. To obtain the sgRNA fold change, divide the experimental normalized sgRNA count by the control and take the base 2 logarithm.

61.   Prepare a RIGER input csv file with the column headers WELL_ID, GENE_ID, and biorep 1, biorep 2, etc from left to right. WELL_ID is a list of sgRNA identification numbers, GENE_ID the genes that the sgRNAs target, and biorep columns the sgRNA fold change. Each row contains a different sgRNA in the library.

62.   RIGER is launched through GENE-E from the Broad Institute (http://www.broadinstitute.org/cancer/software/GENE-E/download.html). Start GENE-E and import the input csv file by navigating to File > Import > Ranked Lists. Click the table cell containing the first data row and column as instructed. Launch RIGER by going to Tools > RIGER. Adjust the RIGER settings to the following recommended values:

- Number of permutations: 1,000,000
- Method to convert hairpins to genes: Kolmogorov-Smirnov
- Gene rank order: Positive to negative for positive selection screens; negative to positive for negative selection screens
- Select adjust gene scores to accommodate variation in hairpin set size
- Select hairpins are pre-scored
- Hairpin Id: WELL_ID
- Convert hairpins to: GENE_ID

63.   Once the RIGER analysis has completed, export the gene rank dataset. Determine the top candidates genes based on either the overlap or the average ranking between the screening bioreps.

Validation of candidate genes for screening phenotype o TIMING 3-4 w

64.   *Cloning validation sgRNAs into the plasmid backbone of the sgRNA library.* Design top and bottom strand primers for cloning the top 3 sgRNAs for each candidate gene individually into the plasmid backbone of the sgRNA library used for screening according to **Table 4** as we have previously described[57]. Primers for cloning 2 non-targeting sgRNAs (NT1 and NT2) for control are also provided in **Table 4**.

65.   Resuspend the top and bottom strand primers to a final concentration of 100 μM. Prepare the following mixture for phosphorylating and annealing the top and bottom strand primers for each validation sgRNA:

Component	Amount (μl)	Final concentration
sgRNA-top, 100 μM	1	10 μM
sgRNA-bottom, 100 μM	1	10 μM
T4 ligation buffer, 10×	1	1×
T4 PNK	0.5	
UltraPure water	6.5	
Total	10	

66.   Phosphorylate and anneal the primers in a thermocycler by using the following conditions: 37 ºC for 30 min; 95 ºC for 5 min; ramp down to 25 ºC at 5 ºC min^{-1}.

67.   After the annealing reaction is complete, dilute the phosphorylated and annealed oligos 1:10 by adding 90 μl of UltraPure water.

Pause Point Annealed oligos can be stored at -20 ºC for at least 1 week.

68.   Clone the annealed sgRNA inserts into the sgRNA library backbone by setting up a Golden Gate assembly reaction for each sgRNA. We have found that when cloning

many sgRNAs, Golden Gate assembly is efficient and offers a high cloning success rate. Mix the following for each sgRNA:

Component	Amount (μl)	Final concentration
Rapid Ligase Buffer, 2×	12.5	1×
FastDigest Esp3I (BsmBI)	1	
DTT	0.25	1 mM
BSA, 20 mg ml^{-1}	0.125	0.1 mg ml^{-1}
T7 ligase	0.125	
Diluted oligo duplex from Step 67	1	0.04 μM
sgRNA library backbone	1	1 ng μl^{-1}
UltraPure water	9	
Total	25	

Critical Step We recommend using FastDigest BsmBI (Fermentas) as we have had reports of BsmBI from other vendors not working as efficiently in the Golden Gate assembly reaction setup described. It is not necessary to perform a negative control (no insert) Golden Gate assembly reaction as it will always contain colonies and therefore is not a good indicator of cloning success.

69.   Perform a Golden Assembly reaction using the following cycling conditions:

Cycle number	Condition
1-15	37 ºC for 5 min, 20 ºC for 5 min

Pause Point Completed Golden Gate assembly reactions can be stored at -20 ºC for at least 1 week.

70.   *Transformation and midiprep.* Transform the Golden Gate assembly reaction into a competent *E. coli* strain, according to the protocol supplied with the cells. We

recommend the Stbl3 strain for quick transformation. Thaw the chemically competent Stbl3 cells on ice, add 2 µl of the product from Step 69 into ice-cold Stbl3 cells, and incubate the mixture on ice for 5 min. Heat-shock the mixture at 42 ºC for 30 s and return to ice immediately for 2 min. Add 100 µl of SOC medium and plate it onto a standard LB agar plate (100 mm Petri dish, ampicillin). Incubate it overnight at 37 ºC.

71.   The next day, inspect the plates for colony growth. Typically, there should be tens to hundreds of colonies on each plate.

Troubleshooting

72.   From each plate, pick 1 or 2 colonies for midiprep to check for the correct insertion of sgRNA and for downstream lentivirus production. To prepare a starter culture for midiprep, use a sterile pipette tip to inoculate a single colony into a 3-ml culture of LB medium with 100 µg ml^{-1} ampicillin. Incubate the starter culture and shake it at 37 ºC for 4-6 h.

73.   Expand each starter culture by transferring the starter culture to 2 separate 25-ml cultures of LB medium with 100 µg ml^{-1} ampicillin in a 50-ml Falcon tube. Remove the cap and seal the top of the tube with AirPore Tape Sheets. Incubate the culture and shake it at 37 ºC overnight at >200 rpm.

74.   12-16 h after seeding the starter culture, isolate the plasmids using an endotoxin-free midiprep kit such as the Macherey-Nagel NucleoBond Xtra Midi EF kit according to the manufacturer's protocol.

Critical Step Using an endotoxin-free plasmid purification kit is important for avoiding endotoxicity in virus preparation and mammalian cell culture.

Pause Point Midiprepped validation sgRNA constructs can be stored at -20 ºC for at least 1 year.

75.   *Sequence validation of sgRNA cloning.* Verify the correct insertion of the validation sgRNAs by sequencing from the U6 promoter using the U6-fwd primer. Compare the sequencing results to the sgRNA library plasmid sequence to check that the 20-nt sgRNA target sequence is properly inserted between the U6 promoter and the remainder of the sgRNA scaffold.

Troubleshooting

76.   *Generation of validation cell lines.* Prepare lentivirus for validation by scaling down the lentivirus production in Steps 42-46 to T25 flasks or 2 wells of a 6-well plate. Filter the lentivirus supernatant using 5-ml syringes and 0.45 μm filters.

77.   Titer the lentivirus according to Step 47. If preparing multiple validation sgRNA lentivirus in the same plasmid backbone at the same time, titer lentivirus from 2-3 different sgRNAs and extend the average titer to the rest of the lentivirus.

78.   Similar to during screening, transduce either naive cells or MS2-p65-HSF1-expressing cells for knockout or activation screening with validation sgRNA lentivirus at an MOI < 0.5 according to Steps 48-50. For knockout validation, select for 7 d to allow for sufficient time for indel saturation.

79.   *Validation of candidate genes for screening phenotype.* Once the antibiotic selection for validation cell lines is complete, verify the screening phenotype. In addition, determine the indel rate for knockout screens (Steps 80-96) or fold activation for activation screens (Steps 97-108).

80.   *Indel rate analysis for validating a knockout screen.* We describe a two-step PCR for NGS in which the first step uses custom primers to amplify the genomic region of interest and the second step uses universal, barcoded primers for multiplexed sequencing of up to 96 different samples in the same NGS run. For each validation sgRNA, design custom round 1 NGS primers (NGS-indel-R1) that amplify the 100-300bp region centered around the sgRNA cut site according to **Table 5**. It is important to design primers situated at least 50 bp from the target cleavage site to allow for the detection of longer indels. Aim for an annealing temperature of 60 ℃ and check for potential off-target sites using Primer-BLAST. If necessary, include a 1-10bp staggered region to increase the diversity of the library.

81.   *Harvest gDNA from validation cell lines*. Seed the validation cells at a density of 60% confluency with 3 bioreps in a 96-well clear bottom black tissue culture plate.

82.   1 d after seeding, when the cells have reached confluency, aspirate the media and add 50 μl of QuickExtract DNA Extraction Solution. Incubate at room temperature for 2-3 min.

83.   Scrape the cells with a pipette tip, mix thoroughly by pipetting up and down, and transfer the mixture to a 96-well PCR plate.

84.   Extract the genomic DNA by running the following cycling conditions:

Cycle number	Condition
1	65 ºC, 15 min
2	68 ºC, 15 min
3	98 ºC, 10 min

Pause Point Extracted genomic DNA can be stored at -20 ºC for up to several months.

85.   *First round PCR for indel analysis by NGS.* Amplify the respective target regions for each validation and control cell line by using custom NGS-indel-R1 primers (**Table 5**) in the following reaction:

Component	Amount (µl)	Final concentration
KAPA HiFi HotStart ReadyMix, 2×	10	1×
QuickExtract	1	
NGS-indel-R1-Fwd	1	0.5 µM
NGS-indel-R1-Rev	1	0.5 µM
UltraPure water	7	
Total	20	

Critical Step To minimize error in amplifying sgRNAs, it is important to use a high-fidelity polymerase. Other high-fidelity polymerases, such as PfuUltra II (Agilent) or NEBNext (New England BioLabs), may be used as a substitute.

86.   Perform a PCR with the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 ºC, 5 min		

| 2-13 | 98 ºC, 20 s | 60 ºC, 15 s | 72 ºC, 15 s |
| 14 | | | 72 ºC, 1 min |

87. *Second round PCR for indel analysis by NGS.* Barcode the first round PCR for NGS by amplifying the product with different NGS-indel-R2 primers (**Table 5**) in the following reaction:

Component	Amount (µl)	Final concentration
KAPA HiFi HotStart ReadyMix, 2×	10	1×
First round PCR	1	
NGS-indel-R2-Fwd	1	0.5 µM
NGS-indel-R2-Rev	1	0.5 µM
UltraPure water	7	
Total	20	

88. Perform a PCR using the same cycling conditions as described in Step 76.

89. After the reaction is complete, run 1 µl of each amplified target on a 2% (wt/vol) agarose gel to verify successful amplification of a single product at the appropriate size.

Troubleshooting

90. Pool the PCR products and purify the pooled product using the QIAquick PCR purification kit.

91. Gel extract the appropriate sizes using the QIAquick gel extraction kit according to the manufacturer's directions.

Pause Point Gel-extracted product can be stored at -20 ºC for several months.

92. Sequence gel-extracted samples on the Illumina MiSeq according to the Illumina user manual with 260 cycles of read 1, 8 cycles of index 1, and 8 cycles of index 2. We recommend aiming for >10,000 reads per sgRNA.

93. *Indel analysis of validation sgRNAs with calculate_indel.py.* Install biopython (http://biopython.org/DIST/docs/install/Installation.html) and SciPy

(https://www.scipy.org/install.html). Construct a sample sheet with each line corresponding to a separate file. The structure of each line should be: <sample name>,<fasta/fastq file name>,<guide sequence>,<PCR target amplicon>,<Experimental or Control>. The last column is only required when performing the maximum likelihood estimate (MLE) correction. When MLE is performed, control samples which reflect the background indel rate should be labeled "Control" and experimental samples should be labeled "Experimental".

94.   If processing all files with a single command, run python calculate_indel.py, with following optional parameters:

Flag	Description	Default
-f	Indicates input file is fasta format	Fastq file format
-a	Uses alternative hashing algorithm for calculation[59]	Ratcliff-Obershelp based algorithm[58]
-o	Output file name	calc_indel_out.csv
-i	Input file name	sample_sheet.csv
-v or -q	Increase or decrease reporting as script runs	Standard reporting
-no-m	Does not perform MLE correction	MLE correction is performed

95.   Place all files in the same folder before running calculate_indel.py. For processing individual samples, such as in the case of parallelization, run python calculate_indel.py --sample <sample name>. This will produce a file <sample name>_out.csv, which can be combined by calling python calculate_indel.py --combine

96.   After running calculate_indel.py, calculated indels will be in the output file, which also contains counts of reads that matched perfectly, failed to align, or were rejected due to quality, or had miscalled bases/replacements. There will also be three columns corresponding to the MLE corrected indel rate, as well as the upper and lower bounds for the 95% confidence interval of indels

97.   *Determine the fold activation for validating an activation screen.* Prepare cells by seeding the validation cells at a density of 60% confluency with 4 bioreps for each validation cell line in a 96-well poly-d-lysine coated tissue culture plate.

98.   *Reverse transcription to cDNA.* When the cells are confluent approximately 1 d after seeding, prepare the following reagents for each well:

- Complete RNA lysis buffer:

Component	Amount (µl)	Final concentration
RNA lysis buffer	100	
Proteinase K, 300 U ml^{-1}	1	3 U ml^{-1}
DNAse I, 50 KU ml^{-1}	0.6	300 U ml^{-1}
Total	101.6	

- Reverse Transcription Mix:

Component	Amount (µl)	Final concentration
Reaction Buffer, 5×	5	1×
dNTP Mix, 10 mM	1.25	0.5 mM
Random Hexamer Primer, 100 µM	1.09	4.4 µM
Oligo dT, 100 µM	0.88	3.5 µM
RiboLock RNAse Inhibitor, 40 U µl^{-1}	0.125	0.2 U µl^{-1}
RevertAid Reverse Transcriptase, 200 U µl^{-1}	1.25	10 U µl^{-1}
UltraPure water	10.41	
Total	20	

Except for Oligo dT, all components can be found in the Thermo RevertAid RT Reverse Transcription kit.

Critical Step Make sure all reagents are RNAse free and take proper precautions when working with RNA.

99.   Aliquot 20 µl of the Reverse Transcription Mix into each well of a 96-well PCR plate. Thaw the RNA lysis stop solution, prepare cold DPBS, and keep all reagents except for the complete RNA lysis buffer on ice until needed.

100.   Aspirate the media from each well of the 96-well poly-d-lysine tissue culture plate, wash with 100 µl of cold DPBS, and add 100 µl of room temperature complete RNA lysis buffer. Incubate at room temperature while mixing thoroughly for 6-12 min to lyse the cells.

Critical Step It is important to limit the lysis time to less than 12 min to prevent RNA degradation.

101.   Transfer 30 µl of the cell lysate to a new 96-well PCR plate. Add 3 µl of RNA lysis stop solution to terminate lysis and mix thoroughly. The cell lysate with RNA lysis stop solution can be stored at -20 ºC for additional reverse transcription reactions.

102.   Then, add 5 µl of the cell lysate with RNA lysis stop solution to the Reverse Transcription Mix for a total volume of 25 µl and mix thoroughly.

103.   Reverse transcribe the harvested RNA to cDNA with the following cycling conditions:

Cycle number	Condition
1	25 ºC, 10 min
2	37 ºC, 60 min
3	95 ºC, 5 min

Pause Point The cDNA can be stably stored at -20 ºC.

104.   *Perform a TaqMan qPCR for fold activation analysis.* Thermo Fisher Scientific provides design ready TaqMan Gene Expression Assays for candidate genes as well as for endogenous control genes such as GAPDH or ACTB. Make sure that the

experimental and control gene expression assays have different probe dyes (i.e. VIC and
FAM dyes) that allow for running in the same reaction.

105.   Prepare the following qPCR master mix per reverse transcription reaction. We
recommend pre-mixing the TaqMan Fast Advanced Mastermix and gene expression
assays for all samples with the same target gene first.

Component	Amount (µl)	Final concentration
TaqMan Fast Advanced Master Mix, 2×	12	1×
TaqMan Gene Expression Assays for candidate gene, 20×	1.2	1×
TaqMan Gene Expression Assays for control gene, 20×	1.2	1×
cDNA	9.6	
Total	24	

106.   Aliquot 4 × 5 µl of the qPCR master mix into a 384-well optical plate for
technical replicates.

107.   Perform a qPCR with the following cycling conditions:

Cycle number	Hold	Denature	Anneal/Extend
1	50 ºC, 2 min		
2	95 ºC, 20 s		
3-42		95 ºC, 3 s	60 ºC, 30 s

108.   Once the qPCR is complete, calculate the candidate gene expression fold change
relative to control using the ddCt method according to the instrument manufacturer's
protocol.

109.   *Additional steps for combining candidate genes from knockout and activation
screens using dead sgRNAs (dRNAs)*. dRNAs are sgRNAs with 14- or 15-nt spacer

sequences, which are truncated versions of the standard sgRNAs with 20-nucleotide spacer sequences, that are still capable of binding DNA[60, 61]. dRNAs are considered catalytically 'dead' because they can guide wild-type Cas9 without inducing double-stranded breaks. Adding MS2 binding loops to the dRNA backbone allows for wild-type Cas9 to activate transcription without cleavage. To achieve simultaneous knockout and activation, refer to Steps 42-49 to generate a cell line that stably expresses wild-type Cas9 and MS2-p65-HSF1.

110. Transduce the Cas9- and MS2-p65-HSF1-expressing cell line with a standard sgRNA for knocking out a candidate gene and a 14-nt dRNA with MS2 binding loops for activating a second candidate gene according to Steps 64-78.

111. Verify the screening phenotype, indel percentage, and fold activation according to Steps 79-108.

Troubleshooting

Troubleshooting advice can be found in **Table 5**.

Timing

Steps 1-20, designing and cloning a targeted screen: 3 d

Steps 21-32, amplification of pooled sgRNA library: 2 d

Steps 33-41, next-generation sequencing of the amplified sgRNA library: 2-3 d

Steps 42-47, lentivirus production and titer: 8-10 d

Steps 48-51, lentiviral transduction and screening: 3-6 w

Steps 52-63, harvest genomic DNA for screening analysis: 3-4 d

Steps 64-111, validation of candidate genes: 3-4 w

ANTICIPATED RESULTS

As a reference for screening results, we provide data from genome-scale knockout and transcriptional activation screens for genes that confer BRAF inhibitor vemurafenib (PLX) resistance in a BRAF^{V600E} (A375) cell line. After applying vemurafenib selection, the sgRNA library distribution, which is measured by NGS, in the experimental condition is more skewed than the baseline and vehicle control conditions, with some sgRNAs

enriched and others depleted (**Fig. 4a,b**). SgRNAs targeting genes involved in vemurafenib resistance are enriched because they provide a proliferation advantage upon vemurafenib treatment. RIGER analysis of enriched sgRNAs in the vemurafenib condition relative to the control identified several candidate genes responsible for resistance. Each candidate gene has multiple significantly enriched sgRNAs (**Fig. 4c, d**) and *P* values that are significantly lower than the rest of the genes (**Fig. 4e, f**). To clarify potential points of confusion when performing genome-scale screens using CRISPR-Cas9, we have compiled a list of most-frequently asked questions from our web-based CRISPR forum (discuss.genome-engineering.org) (**Box 6**). (**Fig. 4**)

Author contributions J.J., S.K., J.S.G., O.O.A., R.J.P., M.D.B., N.S., and F.Z. designed and performed the experiments. J.J., S.K., and F.Z. wrote the manuscript with help from all authors.

References

1.   Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* **9**, 554-566 (2008).
2.   Echeverri, C.J. & Perrimon, N. High-throughput RNAi screening in cultured cells: a user's guide. *Nat Rev Genet* **7**, 373-384 (2006).

3.   Shalem, O., Sanjana, N.E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* **16**, 299-311 (2015).
4.   Elbashir, S.M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498 (2001).
5.   Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811 (1998).
6.   McManus, M.T. & Sharp, P.A. Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* **3**, 737-747 (2002).
7.   Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343-349 (2004).
8.   Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431-437 (2004).
9.   Boutros, M. et al. Genome-wide RNAi analysis of growth and viability in Drosophila cells. *Science* **303**, 832-835 (2004).
10.   Moffat, J. & Sabatini, D.M. Building mammalian signalling pathways with RNAi screens. *Nat Rev Mol Cell Biol* **7**, 177-187 (2006).
11.   Paddison, P.J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427-431 (2004).
12.   Root, D.E., Hacohen, N., Hahn, W.C., Lander, E.S. & Sabatini, D.M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods* **3**, 715-719 (2006).
13.   Silva, J.M. et al. Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet* **37**, 1281-1288 (2005).
14.   Birmingham, A. et al. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods* **3**, 199-204 (2006).
15.   Jackson, A.L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**, 635-637 (2003).
16.   Jackson, A.L. & Linsley, P.S. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov* **9**, 57-67 (2010).
17.   Wright, A.V., Nunez, J.K. & Doudna, J.A. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **164**, 29-44 (2016).
18.   Marraffini, L.A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55-61 (2015).
19.   Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278 (2014).
20.   Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
21.   Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-2586 (2012).
22.   Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* (2013).
23.   Mali, P. et al. RNA-Guided Human Genome Engineering via Cas9. *Science* (2013).

24. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826 (2013).

25. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827-832 (2013).

26. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843 (2013).

27. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).

28. Evers, B. et al. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* **34**, 631-633 (2016).

29. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* **14**, 8096-8106 (1994).

30. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84 (2014).

31. Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera Mdel, C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267-273 (2014).

32. Zhou, Y. et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487-491 (2014).

33. Doench, J.G. et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262-1267 (2014).

34. Chen, S. et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246-1260 (2015).

35. Shi, J. et al. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* **33**, 661-667 (2015).

36. Parnas, O. et al. A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell* **162**, 675-686 (2015).

37. Wong, A.S. et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc Natl Acad Sci U S A* **113**, 2544-2549 (2016).

38. Wang, T. et al. Identification and characterization of essential genes in the human genome. *Science* **350**, 1096-1101 (2015).

39. Doench, J.G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191 (2016).

40. Jain, I.H. et al. Hypoxia as a therapy for mitochondrial disease. *Science* **352**, 54-61 (2016).

41. Gilbert, L.A. et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647-661 (2014).

42. Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).

43. Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326-328 (2015).

44. Gilbert, L.A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451 (2013).

45.	Maeder, M.L. et al. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* **10**, 977-979 (2013).
46.	Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833-838 (2013).
47.	Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**, 973-976 (2013).
48.	Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
49.	Konermann, S. et al. Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472-476 (2013).
50.	Chavez, A. et al. Comparison of Cas9 activators in multiple species. *Nat Methods* (2016).
51.	Hasson, S.A. et al. High-content genome-wide RNAi screens identify regulators of parkin upstream of mitophagy. *Nature* **504**, 291-295 (2013).
52.	Moffat, J. et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283-1298 (2006).
53.	Neumann, B. et al. High-throughput RNAi screening by time-lapse imaging of live human cells. *Nat Methods* **3**, 385-390 (2006).
54.	Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784 (2014).
55.	Konig, R. et al. A probability-based approach for the analysis of large-scale RNAi screens. *Nat Methods* **4**, 847-849 (2007).
56.	Luo, B. et al. Highly parallel identification of essential genes in cancer cells. *Proc Natl Acad Sci U S A* **105**, 20380-20385 (2008).
57.	Ran, F.A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).
58.	Ran, F.A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).
59.	Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771 (2015).
60.	Dahlman, J.E. et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat Biotechnol* **33**, 1159-1161 (2015).
61.	Kiani, S. et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat Methods* **12**, 1051-1054 (2015).
62.	Canver, M.C. et al. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192-197 (2015).
63.	Korkmaz, G. et al. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol* **34**, 192-198 (2016).

Boxes

Box 1: Different types of screening selection

Before setting up the screen, it is important to determine the type of screening selection based on the phenotype of interest and available selection pressures for the screen. Positive selection screens rely on enrichment of sgRNAs for genetic perturbations that produce the screening phenotype as a result of cell proliferation. These typically have the highest signal-to-noise ratio compared to the other types of screens because the number of phenotypically relevant sgRNAs increases relative to the rest of the sgRNAs. On the other hand, negative selection screens involve depletion of sgRNAs that correspond to the phenotype due to cell death. However, for a large number of screens the phenotype of interest will not result in cell proliferation or cell death and thus the phenotypically relevant sgRNAs are not enriched or depleted. For these phenotypes, the screen may be read out by changes in protein expression using either endogenous-tagged fluorescent proteins or a highly specific antibody and FACS. Regardless of the type of screening selection, NGS is used to compare the number of reads for each sgRNA in the perturbed experimental condition relative to a control to identify candidate genes for validation.

Box 2: Considerations for setting screening parameters

Optimal screening parameters should maximize the difference in sgRNA distribution between the experimental and control conditions. Selection conditions such as drug dosage or FACS bin cutoff should be predetermined, if possible, using positive and negative controls from the literature and set to the level at which the greatest difference is observed. As for determining the duration of the screen, collection of time points throughout the screen helps identify the best time point for harvesting and analyzing the screen. These time points are also informative for assessing whether it is necessary to increase the duration to enhance the difference between experimental and control conditions.

Throughout the screen, it is imperative to maintain sufficient coverage to avoid losing sgRNA representation or bias the screening results. Try to maintain sufficient coverage at >500 cells per sgRNA in the library during library transduction, screening

selection, and screening harvest. In addition, we recommend 2-4 infection replicates per screen to account for stochastic noise. Increase the coverage and number of infection replicates if the screening selection is noisy. Finally, consistency of screening conditions such as sgRNA representation and passaging reduces the variability between infection replicates.

Box 3: Additional considerations for *ex vivo* and *in vivo* pooled screening

Ex vivo screening involves removing a primary cell type of interest from a living animal, culturing them in vitro and then performing the screen. For example, Parnas et al. demonstrated this strategy by deriving immune dendritic cells from Cas9 mice, transducing them with a CRISPR knockout library, triggering an immune response with lipopolysaccharide (LPS), and then FACS sorting different populations of cells based on immune response (e.g. TNF expression)[36]. This *ex vivo* screen identified many known as well as novel regulators of LPS response. When performing an *ex vivo* screen, it is necessary to be able to obtain enough cells to maintain library representation, deliver appropriate reagents to the cells, and culture the cells for long enough to perform the screen. In cases where these conditions cannot be met, adapt the screening strategy by, for instance, reducing the library size to capture a subset of genes.

In vivo screening is performed with either a) transduction of cells *in vitro* followed by *in vivo* cell transplantation, or b) direct transduction of tissues *in vivo* . The first strategy was demonstrated by Chen et al., whereby a cancer cell line was transduced with a CRISPR knockout library and injected subcutaneously in immunocompromised mice[34]. NGS analysis of harvested tumors identified known and novel tumor suppressors associated with tumor growth and metastasis. The main challenge of this approach is engrafting cells *in vivo*. Special care must be taken to ensure that the library is not only maintained upon infection of cells *in vitro* but also after engraftment of cells *in vivo*. While it is not required to maintain library representation on a per animal basis, a sufficient number of animals should be used such that library representation is maintained for each experimental cohort. Because the engraftment efficiency and time of engraftment can change for each application it is necessary to sequence the library at several time points after injection of cells in vivo. The optimal time point is one where

engraftment is complete and selection (i.e. proliferation, death, or migration) has not yet occurred. Identifying this time point is critical as it is used as a reference to identify enriched and/or depleted perturbations.

For the second method of *in vivo* screening, special considerations will vary widely depending on the specific animal model, tissue, cell type, developmental time point, or biological question. Thus, each screen should be uniquely designed. In addition to the screening considerations outlined previously, the additional challenge for this strategy is the delivery of reagents in a complex environment while maintaining library representation and also infecting cells at a low MOI. Beyond specific circumstances, it may not be feasible to achieve appropriate cell numbers suitable for a genome-scale library. In these cases it is recommended to design smaller, targeted libraries with a specific hypothesis in mind. The complexity of the in vivo environment makes it difficult to meet the critical requirements for performing an informative screen. In assessing whether a direct in vivo screening strategy is feasible for any particular application, consider these guiding questions: 1) Is there a delivery strategy for infecting the target cells at low MOI? 2) Can enough of the target population be infected and purified to maintain library representation? 3) Can a reference population be identified before the guide RNA abundance changes?

Box 4: Designing and analyzing a saturated mutagenesis screen

Although most pooled CRISPR screens to date have focused on knockout or activation of protein-coding genes, CRISPR screens can also be used to identify functional elements in noncoding regions of the genome such as enhancers or repressors. These functional elements are often inferred using biochemical hallmarks associated with function (e.g. chromatin accessibility, transcription factor binding sites, or post-translational histone modifications). In contrast, CRISPR screens enable direct testing of how mutagenesis at a specific noncoding site impacts phenotype.

Several strategies can be used to design libraries to target noncoding regions. For understanding regulation of a particular gene, tiling mutagenesis libraries were designed to include many or all possible target sites within a noncoding region near a gene[62]. This allows unbiased identification of all regulatory elements in regions near a gene that has

already been established to be important for the screening phenotype (see illustration). Another approach is to design the library to target all instances across the genome of a specific biochemical hallmark, such as all binding sites of a transcription factor like p53[63]. With this kind of library, it is possible to identify specific binding sites or regulatory elements associated with a phenotype of interest.

As with screening the coding genome, a key factor in assessing the performance of the screen is to find multiple sgRNAs targeting the same element that are enriched or depleted together. In coding regions, this is straightforward, as the library is designed to have multiple sgRNAs that target the same gene. In noncoding regions, the same principle of consistent enrichment or depletion can be applied to multiple sgRNAs that target neighboring regions, as the indels are of variable length. Once a functional element is validated using multiple sgRNAs with adjacent target sites, expression of nearby genes and potential mechanisms such as alterations in transcription factor binding at the site can be used to gain further insight into biological mechanisms.

Box 5: Lentivirus production with Polyethylenimine

1.   Prepare polyethylenimine (PEI) transfection reagent:

    a.   Dissolve 50 mg PEI Max in 45 ml UltraPure Water.

    b.   Adjust pH to 7.1. First add 10 M NaOH dropwise until the pH approaches 6 and then add 1 M NaOH dropwise until final pH reaches 7.1.

    c.   Adjust final volume to 50 ml with UltraPure Water.

    d.   Sterilize using Millipore's 0.45 µm Steriflip filter.

    e.   Prepare 50×1 ml aliquots and store at -20 ºC until use. Note: PEI is stable for up to one year and can undergo 5 freeze-thaw cycles without a drop in transfection efficiency.

2.   Prepare HEK293FT cells for lentivirus transfection as described in Steps 41-44.

    a.   For each lentiviral target, combine the following lentiviral target mix in a 50-ml Falcon tube and scale up accordingly:

Component	Amount per T225 flask
DMEM (serum free)	651 µl

pMD2.G (lentiviral helper plasmid)	3.4 µg
psPAX2 (lentiviral helper plasmid)	6.8 µg
Lentiviral target plasmid	13.6 µg

    b.   Add 195 µl PEI transfection reagent, vortex, and incubate at room temperature for 10 min.

    c.   Add 25 ml D10 media to the transfection reagent mixture.

    d.   Aspirate old media from cells, gently add new media containing the transfection reagent mixture and shake gently to mix. Return the T225 flask to the incubator.

3.   2 d after transfection, harvest and store lentivirus as described in Step 46.

Box 6: Frequently asked screening questions from the CRISPR Forum

The following questions are selected from the CRISPR Discussion Forum (discuss.genome-engineering.org).

Q1: Can I use liquid culture amplification of the library rather than solid plates?

We recommend using plates because liquid culture can generate more bias in the plasmid library. Beta-lactamase, the enzyme responsible for ampicillin resistance, is secreted and eventually in liquid culture the selective pressure on the plasmid is decreased causing bias. Additionally, it is more difficult for certain clones to predominate on solid plates because they are spatially limited in growth and each clone is spatially separated to prevent potential intercolony competition. However, it is important to note that some studies have had success with liquid culture amplification[38].

Q2: Is there a difference between using HEK293FT vs HEK293T cells for lentivirus production?

Yes, HEK293FT cells are generally more ideal for lentivirus production. HEK293T cells are a cell line stably expressing the SV40 large T antigen, which helps to boost protein production off expression constructs containing the SV40 enhancer element. HEK293FT

cells are a fast-growing, highly transfectable clonal derivation of HEK293T cells that yield higher lentivirus titer than the HEK293T line.

Q3: For activation, how do I design guides relative to the TSS of the transcript? Additionally, can I expect these guides to work with transient transfection of dCas9-VP64 and MS2-VP65-HSF1 plasmids?

The TSS is the first base of the transcript, i.e. beginning of the 5' UTR. The UCSC table browser is a good resource for TSS annotations. We have observed the most robust transcriptional upregulation when sgRNAs are designed to target the 200bp region upstream of the TSS. We have created a web tool using these parameters to simplify activation sgRNA design for human and mouse genes (http://sam.genome-engineering.org/database/). SAM is highly robust and should yield significant activation levels even in the case of transient transfection[50].

Q4: What are important considerations for NGS PCR amplification?

When designing primers, it is important to include stagger between the primer binding site and the Illumina adapter sequence such that the sequencing regions of different amplicons are offset, improving the sequence diversity and quality. For genomic DNA amplification, it can be helpful to optimize the DNA input for the sequencing readout PCR step. Generally, it is recommended for any given instance of the screen to titrate the DNA input and use the highest possible input without a decrease in the target band intensity. It is critical to minimize amplification bias. The optimal cycle number should always be determined by doing a series of different cycle numbers (e.g. 5, 10, and 15) and identifying the lowest cycle number that generates a visible band by gel electrophoresis. Avoid conditions that yield additional bands at higher cycle numbers.

Q5: My screening design requires too many cells. Can I reduce the coverage?

We recommend screening at a coverage of >500 cells per sgRNA. Because there is always variability in the copy number of each sgRNA in a given library, it is important to have high coverage to overcome any bias. If it is impossible to screen at this coverage

(e.g. insufficient primary cells or cells are difficult to transduce), consider screening with a smaller, targeted library.

Q6: How do you measure the quality of a cloned plasmid library?

While there are many methods for determining the quality of a library, we typically use the following measures for a sequencing depth of >100 reads per sgRNA:

1.   Overall representation: <0.5% of sgRNAs have dropped out with no reads.

2.   Library uniformity: <10-fold difference between the 90th percentile and the 10th percentile.

Figure Legends

Figure 1. Approaches to genetic perturbation: shRNA knockdown, Cas9 knockout, and Cas9 transcriptional activation. Schematic of the mechanisms behind shRNA knockdown, Cas9 knockout, and Cas9 transcriptional activation. ShRNA knockdown begins with processing of the shRNA by Drosha/Dicer machinery and results in degradation of an RNA transcript with a complementary target site by the RISC complex. Cas9 knockout is accomplished by targeted indel formation at a genomic site complementary to the sgRNA. An indel can result in a frameshift, causing early termination, and either production of non-functional protein or non-sense mediated decay of the mRNA transcript. Programmable transcriptional activation can be achieved using dCas9 and activation domains (e.g. VP64/p65/HSF1) to recruit transcriptional machinery to the transcriptional start site of the desired gene target, resulting in upregulation of the target transcript.

Figure 2. Timeline and overview of experiments. Genome-scale Cas9 knockout and transcriptional activation screens begin with the construction of a plasmid library encoding the effector protein and sgRNAs. These plasmid libraries are packaged into lentivirus and then transduced into the cell type of interest to generate stably expressing lines for the screen, along with an accessory transcriptional activator complex (MS2-p65-HSF1) lentivirus for the case of activation screening. A selection pressure is applied depending on the nature of the screen and at given timepoints, genomic DNA is harvested. The sgRNA regions are amplified from genomic DNA and then analyzed by next generation sequencing followed by statistical analyses (e.g. RIGER) to identify candidate genes. Candidate genes are then validated by various forms of analysis, including testing individual sgRNAs for the screening phenotype, indel formation by targeted sequencing, or transcript upregulation by qPCR.

Figure 3. SgRNA library design for genome-scale knockout or activation screens. For knockout screening, the GeCKO v2 libraries target the 5' conserved coding exons of 19,050 human or 20,611 mouse coding genes with 6 sgRNAs per gene. For activation

screening, the SAM libraries target the 200bp region upstream of the transcriptional start site of 23,430 human or 23,439 mouse RefSeq coding isoforms with 3 sgRNAs per isoform. Both libraries select sgRNAs with minimal off-target activity.

Figure 4. Anticipated results for genome-scale knockout and activation screens. Genome wide knockout and activation screens are performed to identify drivers of resistance to the BRAF inhibitor vemurafenib (PLX) in a BRAF^{V600E} (A375) melanoma cell line. A significant number of guides are seen enriched and depleted in the PLX day 14 condition, revealing depletion of guides essential for cell growth and enrichment of guides that promote resistance to BRAF inhibitor (**a**,**b**). RIGER identification of candidate enriched genes from the screens are highlighted. Each gene has multiple sgRNAs that are enriched. Many of these genes are known tumor suppressors or oncogenes that play a role in PLX4720 resistance (**c**,**d**). The top hits of the screen are seen as distributed across the genome, revealing the necessity of genome-scale screens for identifying drivers of resistance. RIGER p-values for candidate enriched genes are significantly lower (**e**,**f**).

Tables

Table 1. Previously published screens using Cas9.

Type of screen	Selection	Organism	Cas9 variant	Reference
Knockout	Vemurafenib resistance (positive)	*In vitro*; A375 (human melanoma cell line)	Wild-type Cas9	27
Knockout	6-thioguanine/etoposide resistance (positive); essential gene targeting (negative)	*In vitro*; HL60, KBM70 (human leukemic cell line)	Wild-type Cas9	30
Knockout	6-thioguanine/Clostridium septicum alpha-toxin resistance (positive)	*In vitro*; mouse embryonic stem cells	Wild-type Cas9	31
Knockout	Anthrax/diphtheria toxin resistance (positive)	*In vitro*; HeLa (human adenocarcinoma cell line)	Wild-type Cas9	32
Knockout	Surface receptor expression (negative)	*In vitro*; EL4 (mouse thymic cell line) and MOLM13/NB4/TF1 (human acute myeloid leukemia cell lines)	Wild-type Cas9	33
Knockout	Metastasis (positive)	*In vivo*; mouse	Wild-type Cas9	34
Knockout	Chromatin regulatory domain dependence	*In vitro*; RN2 (murine acute myeoloid	Wild-type Cas9	35

	(negative)	leukemia cell line)		
Knockout	Response to LPS stimulation (both)	*Ex vivo*; bone-marrow derived dendritic cells (mouse)	Wild-type Cas9	[36]
Knockout	Combinatorial gene dependency (negative)	*In vitro*; OVCAR8-ADR (human ovarian cancer cell line)	Wild-type Cas9	[37]
Knockout	Gene dependency for essential genes (negative)	*In vitro*; KBM7 (human chronic myelogenous leukemia cell line)	Wild-type Cas9	[38]
Knockout	Vemurafenib resistance (positive)/essential genes (negative)/6-thioguanine resistance (positive)/interferon survival (positive)	*In vitro*; A375 (human melanoma cell line)/HT29/HEK293T cells; BV2 (mouse cell line)	Wild-type Cas9	[39]
Knockout	Survival under oxidative stress (positive)	*In vitro*; K562 (human leukemic cell line)	Wild-type Cas9	[40]
Activation	Ricin sensitivity (both)/cell growth (both)	*In vitro*; K562 (human leukemic cell line)	sunCas9-VP64	[41]
Activation	Vemurafenib resistance (positive)	*In vitro*; A375 (human melanoma cell line)	dCas9-VP64/P65/HSF1	[42]
Knockdown	Ricin resistance (positive)/essential genes (negative)/cholera	*In vitro*; K562 (human leukemic cell line)	dCas9 or dCas9-KRAB fusion	[41]

sensitivity (both)

Table 2. Primers for amplifying the sgRNA oligo library.

Primer	Sequence (5'-3')	Purpose
Oligo-Fwd	GTAACTTGAAAGTATTTCGATTTCTT GGCTTTATATATCTTGTGGAAAGGAC GAAACACC	Targeted knockout or activation library cloning
Oligo-Knockout-Rev	ACTTTTTCAAGTTGATAACGGACTAG CCTTATTTTAACTTGCTATTTCTAGCT CTAAAAC	Targeted knockout library cloning
Oligo-Activation-Rev	ATTTTAACTTGCTAGGCCCTGCAGAC ATGGGTGATCCTCATGTTGGCCTAGC TCTAAAAC	Targeted activation library cloning

Table 3. Primer sequences for amplifying sgRNA library and NGS. The NGS-Lib-Fwd primers contain 1-10bp staggered nucleotides designed to increase the diversity of the NGS library, and the NGS-Lib-Rev primers provide unique barcodes for distinguishing different sgRNA libraries (i.e. from different screening conditions) in a pooled sequencing run. Since the sgRNA backbone is different between the GeCKO and SAM libraries, separate NGS-Lib-KO-Rev and NGS-Lib-SAM-Rev primers have been provided for each library.

Primer	Sequence (5'-3')	Purpose
NGS-Lib-Fwd-1	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTTAAGTAGAGGCTTTATATATCT TGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-2	AATGATACGGCGACCACCGAGATCTA	GeCKO or SAM

	CACTCTTTCCCTACACGACGCTCTTCC GATCTATCATGCTTAGCTTTATATATC TTGTGGAAAGGACGAAACACC	sgRNA library NGS
NGS-Lib-Fwd-3	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTGATGCACATCTGCTTTATATAT CTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-4	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTCGATTGCTCGACGCTTTATATA TCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-5	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTTCGATAGCAATTCGCTTTATAT ATCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-6	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTATCGATAGTTGCTTGCTTTATA TATCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-7	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTGATCGATCCAGTTAGGCTTTAT ATATCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-8	AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTCGATCGATTTGAGCCTGCTTTA TATATCTTGTGGAAAGGACGAAACAC	GeCKO or SAM sgRNA library NGS

C

NGS-Lib-Fwd-9	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACGATCGATACACGATCGCTTTATATATCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-Fwd-10	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGATCGATGGTCCAGAGCTTTATATATCTTGTGGAAAGGACGAAACACC	GeCKO or SAM sgRNA library NGS
NGS-Lib-KO-Rev-1	CAAGCAGAAGACGGCATACGAGAT**TCGCCTTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCGACTCGGTGCCACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-2	CAAGCAGAAGACGGCATACGAGAT**ATAGCGTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCGACTCGGTGCCACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-3	CAAGCAGAAGACGGCATACGAGAT**GAAGAAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCGACTCGGTGCCACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-4	CAAGCAGAAGACGGCATACGAGAT**ATTCTAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCGACTCGGTGCCACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)

NGS-Lib-KO-Rev-5	CAAGCAGAAGACGGCATACGAGAT**C GTTACCA**GTGACTGGAGTTCAGACGT GTGCTCTTCCGATCTCCGACTCGGTGC CACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-6	CAAGCAGAAGACGGCATACGAGAT**GT CTGATG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-7	CAAGCAGAAGACGGCATACGAGAT**TT ACGCAC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-KO-Rev-8	CAAGCAGAAGACGGCATACGAGAT**TT GAATAG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA	GeCKO sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-1	CAAGCAGAAGACGGCATACGAGAT**TC GCCTTG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-2	CAAGCAGAAGACGGCATACGAGAT**AT AGCGTC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-3	CAAGCAGAAGACGGCATACGAGAT**G AAGAAGT**GTGACTGGAGTTCAGACGT GTGCTCTTCCGATCTGCCAAGTTGATA	SAM sgRNA library NGS and barcode (bold)

ACGGACTAGCCTT

NGS-Lib-SAM-Rev-4	CAAGCAGAAGACGGCATACGAGATA**T TCTAGG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-5	CAAGCAGAAGACGGCATACGAGAT**C GTTACCA**GTGACTGGAGTTCAGACGT GTGCTCTTCCGATCTGCCAAGTTGATA ACGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-6	CAAGCAGAAGACGGCATACGAGAT**GT CTGATG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-7	CAAGCAGAAGACGGCATACGAGAT**TT ACGCAC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)
NGS-Lib-SAM-Rev-8	CAAGCAGAAGACGGCATACGAGAT**TT GAATAG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT	SAM sgRNA library NGS and barcode (bold)

Table 4. Primers for sgRNA cloning and validation.

Primer	Sequence (5'-3')	Purpose
sgRNA-top	CACCgNNNNNNNNNNNNNNNNNN NNNN	Top strand primer for cloning sgRNA into sgRNA library backbone; appended guanine in lowercase

sgRNA-bottom	AAACNNNNNNNNNNNNNNNNN NNNNc	Bottom strand primer for cloning sgRNA into sgRNA library backbone; appended cytosine in lowercase
NT1-top	CACCGCTGAAAAAGGAAGGAG TTGA	Top strand primer for cloning non-targetting sgRNA into sgRNA library backbone
NT1-bottom	AAACTCAACTCCTTCCTTTTTC AGC	Bottom strand primer for cloning non-targetting sgRNA into sgRNA library backbone
NT2-top	CACCGAAGATGAAAGGAAAGG CGTT	Top strand primer for cloning non-targetting sgRNA into sgRNA library backbone
NT2-bottom	AAACAACGCCTTTCCTTTCATC TTC	Bottom strand primer for cloning non-targetting sgRNA into sgRNA library backbone
U6-Fwd	CGTGACGTAGAAAGTAATAAT TTCTTGGG	Sanger sequencing of sgRNA cloned into sgRNA library backbone

Table 5. Primers for amplifying target sites to determine percentage of indels by NGS. Custom first round primers amplify the target region and second round universal, barcoded primers amplify the first round products for multiplexed NGS.

Primer	Sequence (5'-3')	Purpose
NGS-indel-R1-Fwd	CTTTCCCTACACGACGCTCTTCCGAT CT(stagger)[priming_site]	Custom first round amplification for NGS
NGS-indel-	GACTGGAGTTCAGACGTGTGCTCTTC	Custom first round

R1-Rev	CGATCT[priming_site]	amplification for NGS
NGS-indel-R2-Fwd-1	AATGATACGGCGACCACCGAGATCT ACAC**TATAGCCT**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-2	AATGATACGGCGACCACCGAGATCT ACAC**ATAGAGGC**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-3	AATGATACGGCGACCACCGAGATCT ACAC**CCTATCCT**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-4	AATGATACGGCGACCACCGAGATCT ACAC**GGCTCTGA**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-5	AATGATACGGCGACCACCGAGATCT ACAC**AGGCGAAG**ACACTCTTTCCCT ACACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-6	AATGATACGGCGACCACCGAGATCT ACAC**TAATCTTA**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-7	AATGATACGGCGACCACCGAGATCT ACAC**CAGGACGT**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-R2-Fwd-8	AATGATACGGCGACCACCGAGATCT ACAC**GTACTGAC**ACACTCTTTCCCTA CACGACGCTCTTCC	Universal second round amplification for NGS
NGS-indel-	CAAGCAGAAGACGGCATACGAGAT**C**	Universal second round

R2-Rev-1	**GAGTAAT**GTGACTGGAGTTCAGACG TGTGCTCTTC	amplification for NGS
NGS-indel-R2-Rev-2	CAAGCAGAAGACGGCATACGAGAT**T CTCCGGA**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-3	CAAGCAGAAGACGGCATACGAGATA**A ATGAGCG**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-4	CAAGCAGAAGACGGCATACGAGAT**G GAATCTC**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-5	CAAGCAGAAGACGGCATACGAGAT**T TCTGAAT**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-6	CAAGCAGAAGACGGCATACGAGATA**A CGAATTC**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-7	CAAGCAGAAGACGGCATACGAGATA**A GCTTCAG**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-8	CAAGCAGAAGACGGCATACGAGAT**G CGCATTA**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-9	CAAGCAGAAGACGGCATACGAGAT**C ATAGCCG**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS

NGS-indel-R2-Rev-10	CAAGCAGAAGACGGCATACGAGAT**T** **TCGCGGA**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-11	CAAGCAGAAGACGGCATACGAGAT**G** **CGCGAGA**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS
NGS-indel-R2-Rev-12	CAAGCAGAAGACGGCATACGAGAT**C** **TATCGCT**GTGACTGGAGTTCAGACG TGTGCTCTTC	Universal second round amplification for NGS

Table 5. Troubleshooting table.

Step	Problem	Possible reason	Possible solution
28b	Insufficient number of colonies	Inefficient electroporation; plates dried out during bacteria spreading; salt contamination or low concentration of DNA for electroporation (custom library only).	Increase number of electroporations if feasible; follow the manufacturer's instructions for optimal efficiency (i.e. ensure cuvettes, tubes and cells are chilled thoroughly and recovery media is added immediately after electroporation); spread bacterial suspension evenly without letting plates dry; limit plate warming to <2h and add 1 volume of additional LB media before spreading as recommended; if the DNA for electroporation has a 260/230 ratio <1.8 or is <50 ng μl^{-1} by nanodrop, consider re-prepping (custom library only).
28c	Colonies	Incomplete digestion of	Increase the amount of restriction

	growing on negative control plate	library plasmid backbone	enzyme and phosphatase, ensure added DTT is fresh and stored at -20 ºC as recommended
41	Less than 70% of perfectly matched sgRNAs	NGS reads are low quality; Incorrect flag for presence of 'G' before guide spacer sequence; Low oligo synthesis quality (custom library only)	If <75% of clusters passed filter, consider resequencing the library with more optimal cluster density and cluster passing filter; Check for the proper flag indicating presence of 'G'; if both the negative gibson control and sequencing quality were good, contact your synthesis company or core to discuss your results
41	More than 0.5% of sgRNAs are undetected	Insufficient colonies during amplification process; Not enough cloning reactions during custom sgRNA library cloning	Increase the number of electroporations; scale up the number of reactions throughout cloning
41	Skew ratio of custom sgRNA library is greater than 10	Insufficient representation and potential bias during cloning process; Low oligo synthesis quality (custom library only)	Reduce the number of PCR cycles during oligo amplification; increase the number of reactions throughout cloning, ensure >500 colonies per guide during library amplification; contact your synthesis company or core to discuss your results.
45	Low survival following Lipofectamine	Cell density at transfection <75%; failure to replace media	Ensure even plating and a cell density of 80-90% during transfection. Replace media 4 h after transfection. The cell

	transfection	within 6 h post transfection	density should not be <80% at the time of lentivirus harvest.
45	Low Lipofectamine transfection efficiency	Incorrect amount or poor quality DNA used for transfection; poorly or unevenly seeded cells	Use low-passage-number cells (passage number <10) and do not let cells reach >90% confluency; reseed cells evenly at recommended density; prepare new DNA for transfection
47	Significant cell survival in the no virus, antibiotic-selected control	Incorrect antibiotic concentration	Construct a kill curve for the antibiotic and cell line of interest by titrating different concentrations of the antibiotic on cells seeded at 10% confluency. Choose the lowest concentration that kills all cells after 4-7 days. When using zeocin, only 90-95% of cells may be killed at the optimal concentration after 7 days, but this should not interfere with screening when using zeocin for selection of SAM sgRNAs.
47	Low lentivirus titer	Low lentivirus production or transduction efficiency	Concentrate the lentivirus using Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-100 membrane or ultracentrifugation at 88,000 × g for 2 h at 4 ºC.
53	No amplification of gDNA	Incorrect template or primer; incorrect template or primer concentration	Use appropriate NGS-Rev primer corresponding to the screen; titrate gDNA concentration to 1-3 μg for a 50 μl reaction; titrate primer concentration to a final concentration of 0.1-0.5 μM

53	Primer dimer after gDNA amplification	Insufficient number of PCR cycles	Increase the number of cycles for PCR up to 25; lower gDNA input and perform more reactions
71	Very few or no colonies on plates	Incorrect template; Low transformation efficiency	Use an appropriate antibiotic for the resistance marker on the plasmid backbone; verify competency of *E. coli*
75	Low percentage of correctly cloned sgRNAs	Incorrect template or primer; incorrect Golden Gate Assembly reaction	Check that the primer design is correct and that the top and bottom strands are complementary; verify restriction enzyme site on plasmid backbone; ensure use of the suggested restriction enzyme (Fermentas Fast Digest) and DNA ligase (Enzymatics T7 ligase)
89	Multiple bands or no amplification during indel analysis	Inefficient or nonspecific priming; incorrect template concentration	Titrate annealing temperature (58-62 ºC); re-design primers to minimize primer dimer formation or nonspecific amplification; test a different high fidelity DNA polymerase; vary the magnesium concentration in the reaction. If none of the above steps solve the issue it may be helpful to clean up the Quickextract gDNA using Ampure beads to provide a cleaner PCR template.

Supplementary Tables

Table S1. Deoxyribonuclease I storage solution. Solution for resuspending and storing deoxyribonuclease I for fold activation analysis during validation. Prepared solution can be stored at -20 ºC for up to 2 years.

Component	Amount (µl)	Final concentration
Tris-HCl (1M, pH 7.5)	250	50 mM
CaCl$_2$ (1M)	50	10 mM
Glycerol	2500	50% (vol/vol)
UltraPure Water	2200	
Total	5000	

Table S2. RNA Lysis Buffer Setup. Buffer for lysing and harvesting RNA from cells for fold activation analysis during validation. Final pH of the solution should be approximately 7.8.

Component	Amount (ml)	Final concentration
Tris pH 8.0 (1M)	1.2	4.8 mM
Tris pH 7.5 (1M)	1.2	4.8 mM
MgCl$_2$ (1M)	0.125	0.5 mM
CaCl$_2$ (1M)	0.110	0.44 mM
Dtt (0.1M)	0.025	10 µM

Proteinase K (300 U ml^{-1})	2.5	3 U ml^{-1}
Triton X-114 (10% wt/vol)	2.5	0.1% (wt/vol)
UltraPure Water	242	
Total	250	

Table S3. EGTA (0.5M, pH 8.3) stock. EGTA stock solution used to prepare RNA lysis stop solution for fold activation analysis during validation. Adjust with NaOH to a final pH of 8.3 if necessary. Take aliquots to measure pH in order to keep main stock from being RNAse-contaminated by the pH probe.

Component	Amount
EGTA	9.5 g
Tris pH 8.0 (1M)	3.125 ml
NaOH 10N	6.1 ml
UltraPure Water	to 50 ml
Total	50 ml

Table S4. RNA lysis stop solution. Solution for terminating RNA lysis for fold activation analysis during validation. Prepare using EGTA (0.5M, pH 8.3) from **Table S3**.

Component	Amount (ml)	Final Concentration
Proteinase K inhibitor, 100 mM	0.150	1 mM

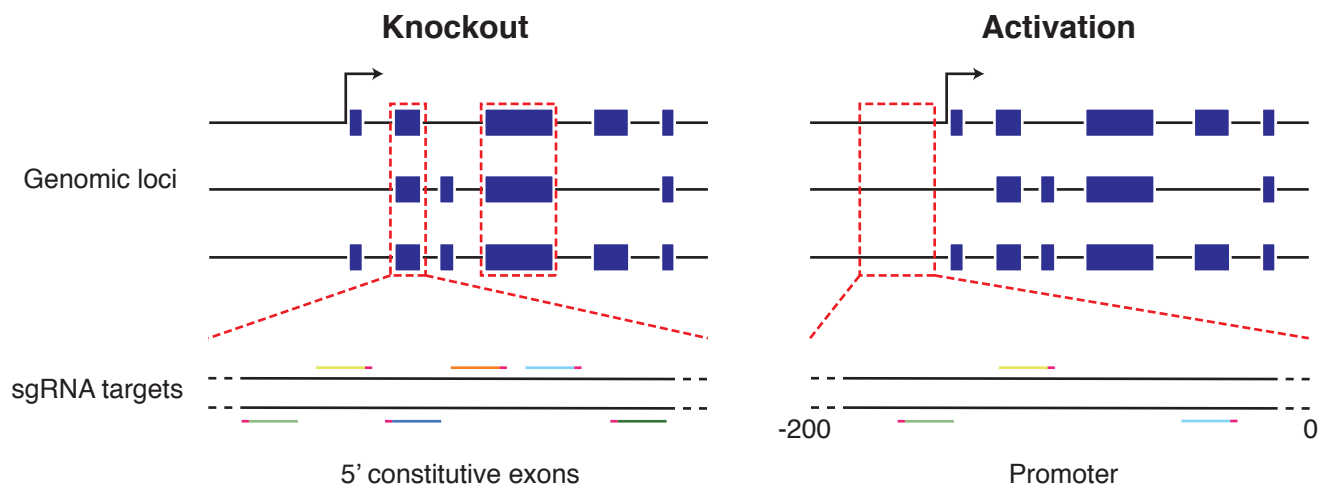|		
EGTA (0.5 M, pH 8.3)	2.694	90 mM
Dtt (0.1 M)	0.017	113 µM
UltraPure Water	12.14	
Total	15	
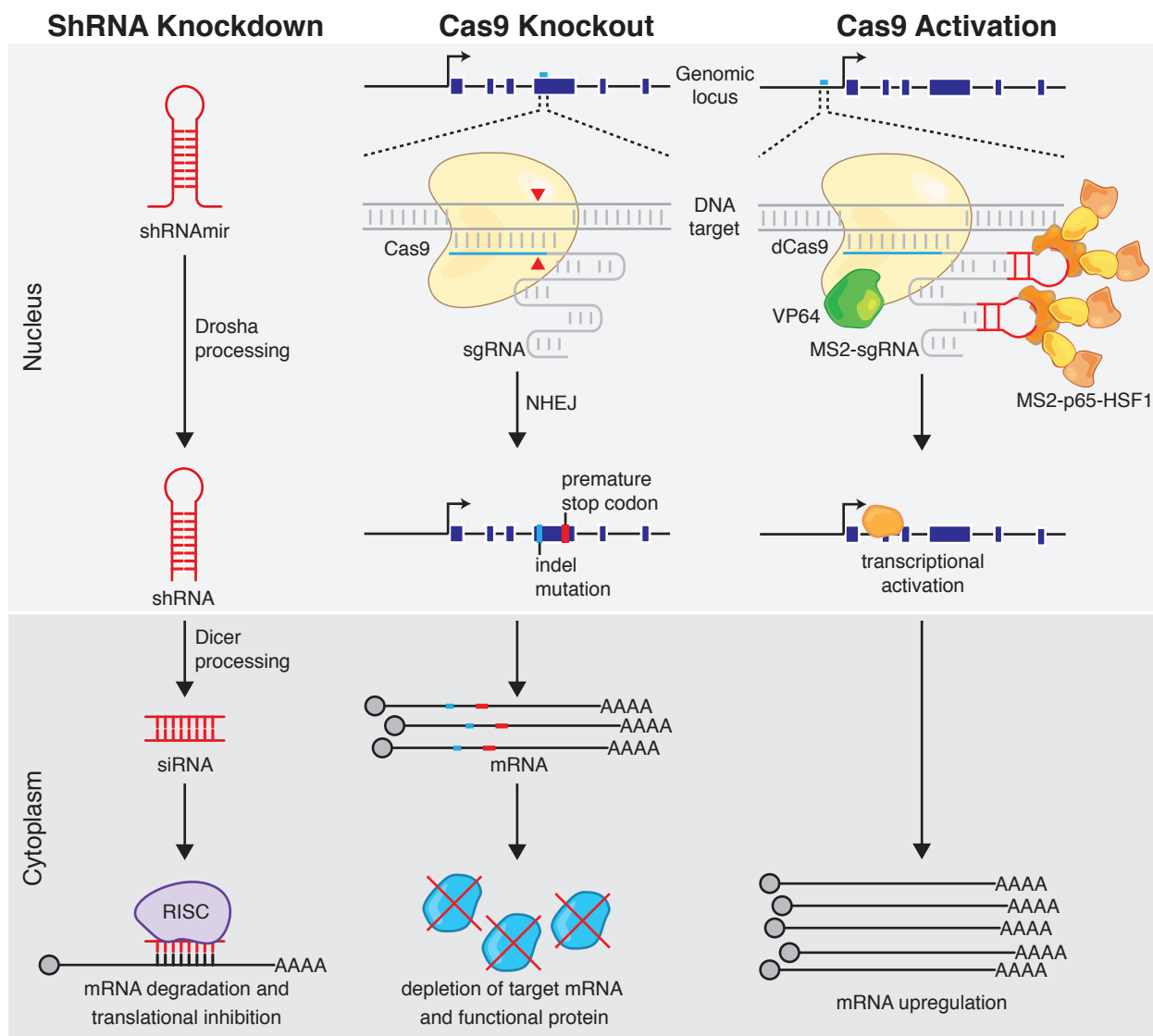
Figure 1
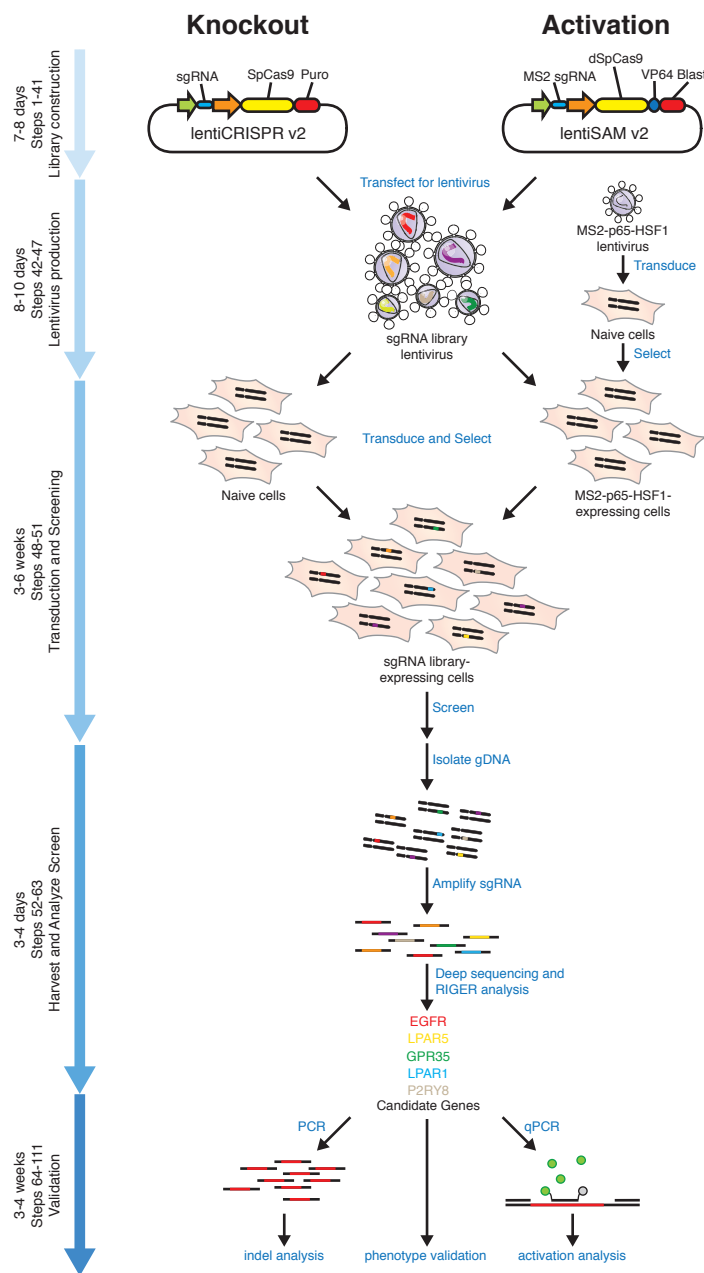
Figure 2

Figure 3

Figure 4

Box 1

Functional element by sgRNA target site

▽ Promoter
▽ 5′ or 3′ UTR
▽ Intron
▽ 5′ or 3′ distal region

SgRNA library

Genomic locus

$10^4 - 10^6$ bp

Box 4