

## **Improved vectors and genome-wide libraries for CRISPR screening**

Neville E. Sanjana<sup>1,2,3,4,5</sup>, Ophir Shalem<sup>1,2,3,4,5</sup> and Feng Zhang<sup>1,2,3,4</sup>

<sup>1</sup> Broad Institute of MIT and Harvard  
7 Cambridge Center  
Cambridge, MA 02142, USA

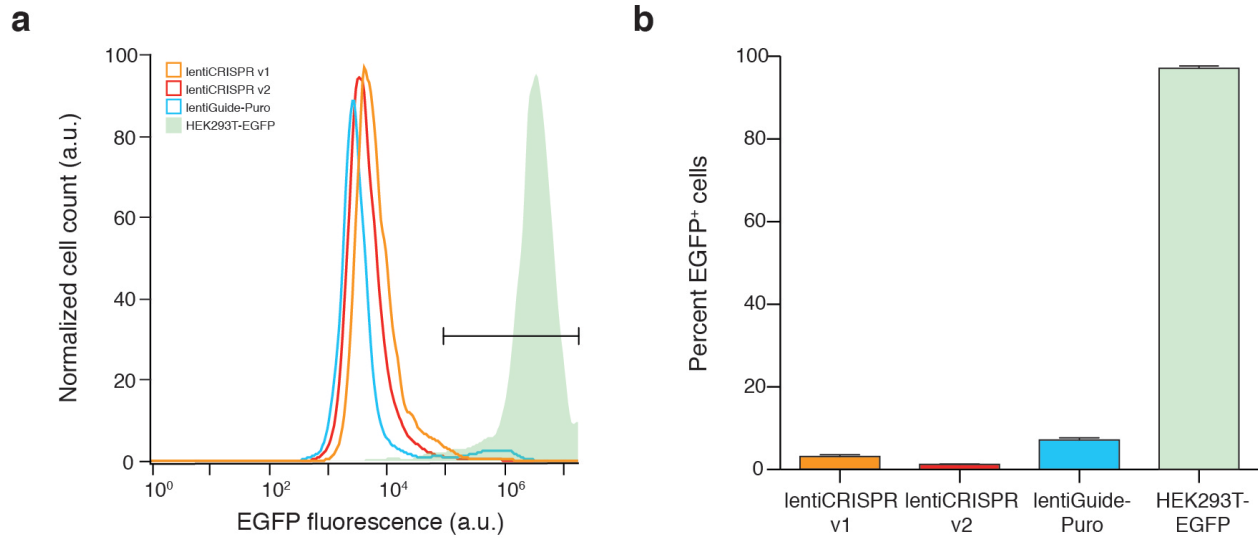
<sup>2</sup> McGovern Institute for Brain Research

<sup>3</sup> Department of Brain and Cognitive Sciences

<sup>4</sup> Department of Biological Engineering  
Massachusetts Institute of Technology  
Cambridge, MA 02139, USA

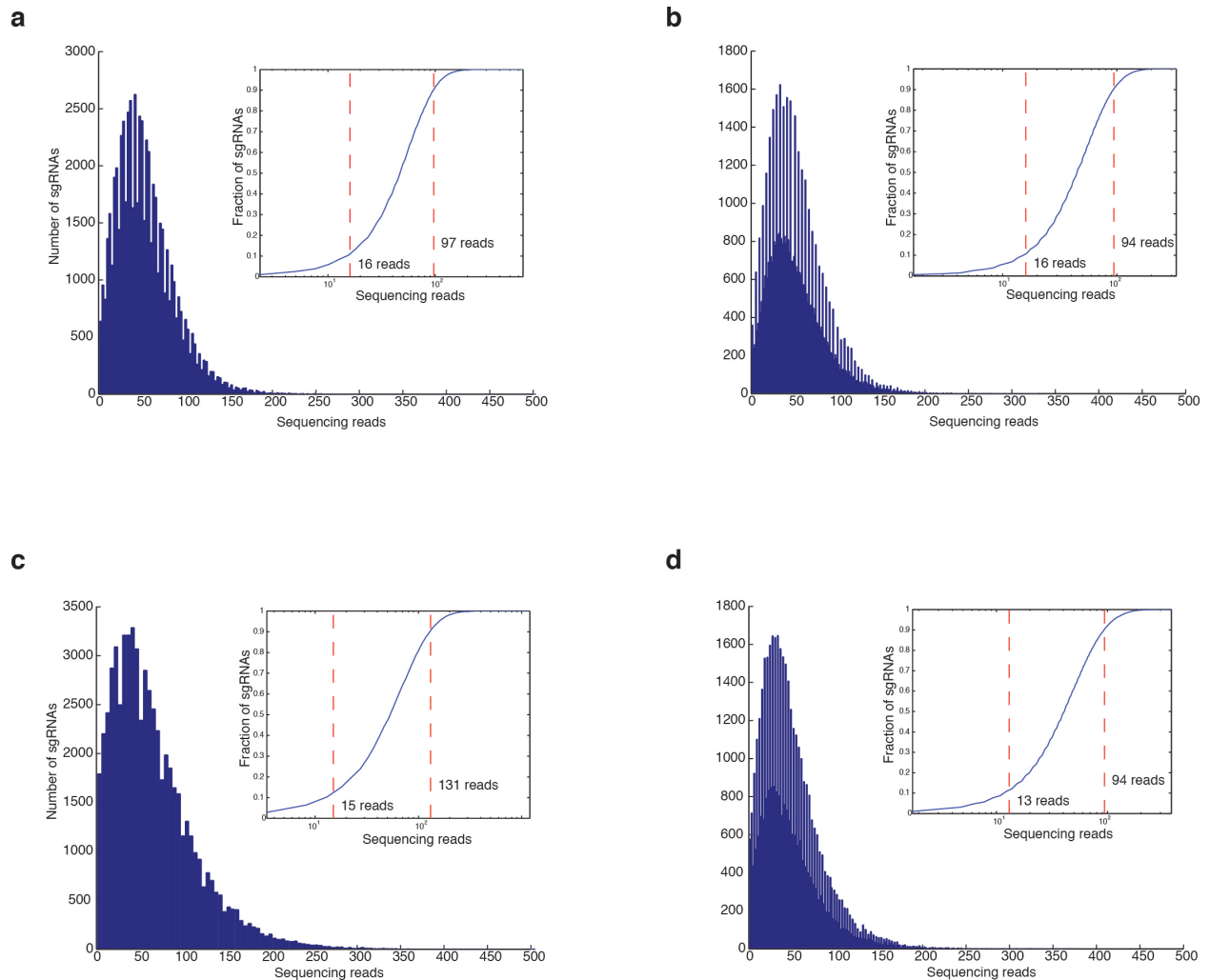
<sup>5</sup> These authors contributed equally to this work.  
email: [zhang@broadinstitute.org](mailto:zhang@broadinstitute.org).

**Supplementary Figure 1 | EGFP expression in HEK293T-EGFP cells after transduction with different lentiviral CRISPR constructs.**



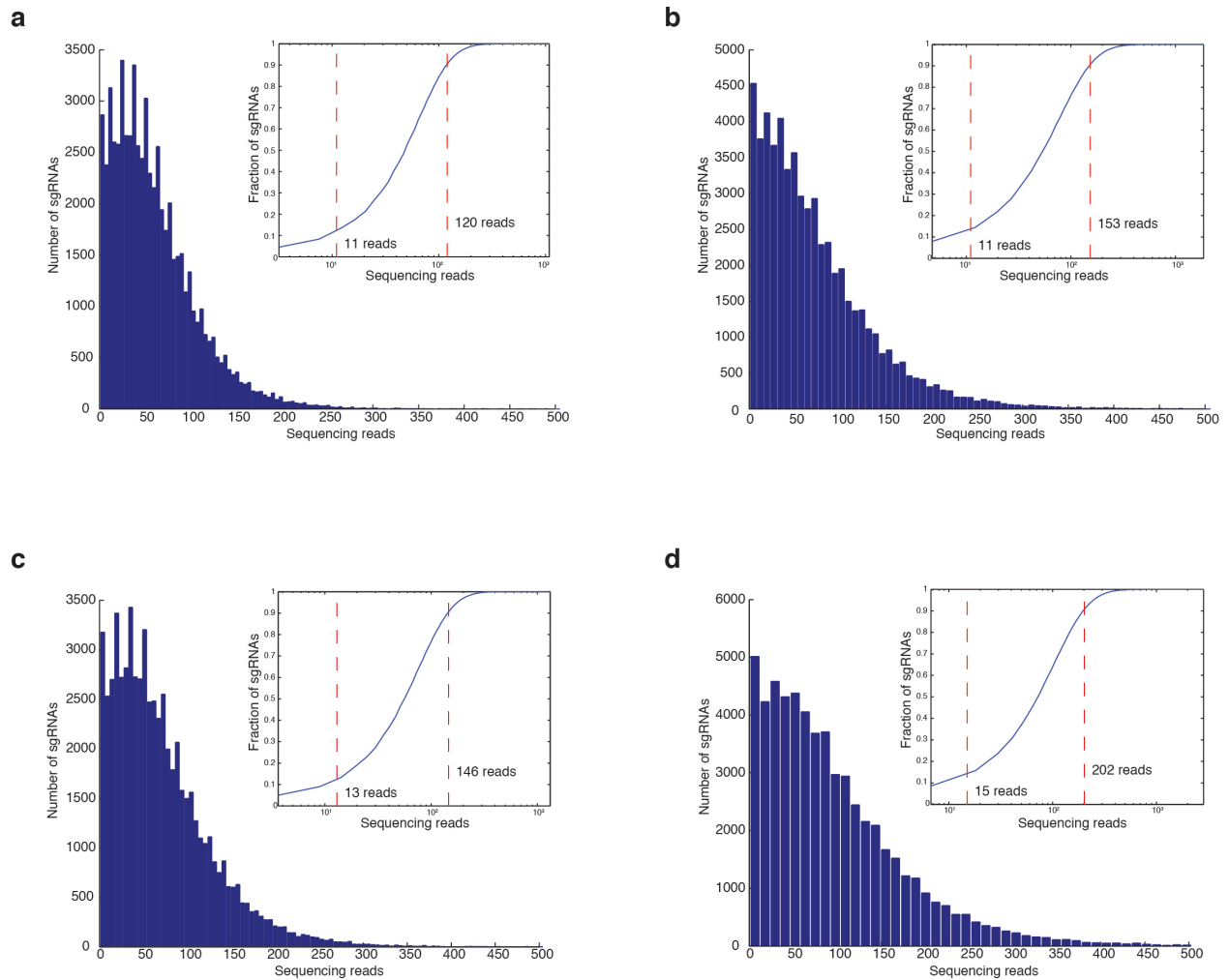
**(a)** Representative histograms of EGFP fluorescence from single transductions of HEK293T-EGFP cells with lentiCRISPR v1, lentiCRISPR v2, lentiGuide-Puro or no virus. For lentiGuide-Puro, HEK293T-EGFP cells had previously been transduced with lentiCas9-Blast and selected for 4 days with blasticidin. Twenty-four hours after transduction, cells were selected in puromycin and then analyzed by flow cytometry at 7 days after infection. **(b)** Percentage of EGFP positive cells (as given by gate drawn in **a**) per viral construct (error bars indicate s.e.m,  $n=3$  biological replicate transductions).

**Supplementary Figure 2 | Sequencing of human GeCKOv2 libraries cloned into lentiCRISPRv2 and lentiGuide-Puro.**



Histograms of sgRNA representation after cloning each half-library into sgRNA vectors. Inset: Cumulative distribution of sequencing reads. The number of sequencing reads for the 10<sup>th</sup> and 90<sup>th</sup> sgRNA percentile is indicated by the dashed red lines and text labels. All library sequencing was performed with PCR replicates; only one replicate is shown for each half-library. **(a,b)** Human GeCKOv2 library A **(a)** and library B **(b)** in lentiCRISPRv2. **(c,d)** Human GeCKOv2 library A **(c)** and B **(d)** in lentiGuide-Puro.

### Supplementary Figure 3 | Sequencing of mouse GeCKOv2 libraries cloned into lentiCRISPRv2 and lentiGuide-Puro.



Histograms of sgRNA representation after cloning each half-library into sgRNA vectors. Inset: Cumulative distribution of sequencing reads. The number of sequencing reads for the 10<sup>th</sup> and 90<sup>th</sup> sgRNA percentile is indicated by the dashed red lines and text labels. All library sequencing was performed with PCR replicates; only one replicate is shown for each half-library. **(a,b)** Mouse GeCKOv2 library A **(a)** and library B **(b)** in lentiCRISPRv2. **(c,d)** Mouse GeCKOv2 library A **(c)** and B **(d)** in lentiGuide-Puro.

## Supplementary Methods

**Lentiviral cloning and production.** For determination of lentiCRISPR v1, lentiCRISPR v2, and lentiGuide-Puro viral titers, the following sgRNA targeting EGFP (with no known targets in the human genome) was cloned into all 3 lentiviral transfer vectors:

EGFP sgRNA            GGGCGAGGAGCTGTTACCG

To clone the sgRNA guide sequence, plasmids were cut and dephosphorylated with FastDigest BsmBI and FastAP (Fermentas) at 37 °C for 2 hours. Oligonucleotides for the EGFP sgRNA guide sequence (Integrated DNA Technologies) were phosphorylated using polynucleotide kinase (Fermentas) at 37 °C for 30 minutes and then annealed by heating to 95 °C for 5 minutes and cooling to 25 °C at 1.5 °C/minute. Using T7 ligase (Enzymatics), annealed oligos were ligated into gel purified vectors (Qiagen) at 25 °C for 5 minutes. Cloned transfer plasmids were amplified using a endotoxin-free midi-prep kit (Qiagen).

To make lentivirus, the transfer plasmids were co-transfected with packaging plasmids pMD2.G and psPAX2 (Addgene plasmids 12259 and 12260), as described previously<sup>1</sup>. Briefly, for each virus, a T-75 flask of 80 % confluent HEK293T cells (ATCC CRL-3216) was transfected in OptiMEM (Life Technologies) using 10 ug of the transfer plasmid, 5 ug pMD2.G, 7.5 ug psPAX2, 100 ul of Plus Reagent (Life Technologies), and 50 ul of Lipofectamine 2000 (Life Technologies). After 6 hours, media was changed to D10 media, DMEM (Life Technologies) with 10 % fetal bovine serum (Hyclone), with 1 % bovine serum albumin (Sigma) added to improve virus stability. After 60 hours, viral supernatants were harvested and centrifuged at 3,000 rpm at 4 °C for 10 min to pellet cell debris. The supernatant was filtered through a 0.45 um low protein binding membrane (Millipore) and used immediately.

**Lentiviral functional titration.** Lentiviruses were titered in a functional assay by measuring puromycin resistance after transduction. All lentiviruses were produced in triplicate transfections. Relative titer measurements from separate viral production transfections performed weeks apart resulted in similar relative titers.

For each viral construct,  $2.5 \times 10^4$  HEK293T-EGFP cells were transduced in suspension (i.e. during plating) with 10, 100, or 1000  $\mu$ l of viral supernatant in wells of a 24-well plate. For lentiGuide-Puro transduction the HEK293T-EGFP cells also had a genomically-integrated copy of Cas9 from previous transduction with lentiCas9-Blast. Each transduction condition (construct and virus volume) was performed in triplicate. In each well, D10 culture media was added to make the final volume 1.5 ml. Cell without any virus added were also plated in six wells (3 wells for puromycin treatment, 3 wells as control).

At 24 hours post-transduction, media was changed to D10 with 1  $\mu$ g/ml puromycin (Sigma) for all wells except the uninfected controls without puromycin. At 3 days post-transduction, cells in all wells were split 1:5 to prevent any well from reaching confluence. Except for the uninfected controls without puromycin, new D10 media was supplemented with 1  $\mu$ g/ $\mu$ l puromycin. At 5 days post-transduction, all cells in the uninfected control wells treated with puromycin were floating/dead, which was verified using Trypan Blue exclusion (Sigma).

For the remaining wells, adherent cells were present and cell viability was measured using CellTiter Glo (Promega) following the manufacturer's protocol. After allowing cells to reach room temperature, media was aspirated from the cells and CellTiter Glo (diluted 1:1 in phosphate-buffered saline) was added. Plates were covered with foil, placed on an orbital shaker for 2 min, and then incubated for 10 minutes at room temperature. Luminescence was measured on an Synergy H4 plate imager (Biotek) using a 1 second integration time and auto-gain to utilize the full dynamic range of the detector. Positive controls (untransduced cells without puromycin) and negative controls (empty wells) were included in the assay.

Fold differences in titer between viral constructs were calculated using luminescence values. Specifically, comparisons were made between pairs of viruses for the same volume of supernatant. Only viral volumes for which cell survival was greater than 1% and less than 20 % of control (untransduced) cells were directly compared. Assuming Poisson statistics, 20 % cell survival implies that approximately 90 % of cells surviving puromycin selection were infected by only a single virus.

Flow cytometry data was collected from the same set of infections using a BD Accuri C6 flow cytometer. Using FlowJo (TreeStar), single cells were distinguished from debris and doublets by gating in forward and side scatter area plots. EGFP fluorescence was measured in the gated

population from transduced and uninfected HEK293T-EGFP cells. For transductions with lentiGuide-Puro, HEK293T-EGFP cells had been transduced with lentiCas9-Blast and selected for 6 days with 5 ug/ml blasticidin (Life Technologies). For all antibiotic selections (puromycin and blasticidin), uninfected control cells were also exposed to the same concentration and duration of antibiotic selection to verify complete elimination of untransduced cells.

**Design of GeCKOv2 libraries.** Genome-wide sgRNA libraries for the human and mouse genomes were designed using the following steps:

(1) *Identification of constitutive exons:* For the human library, RNA sequencing data from the Illumina Human Body Map 2.0 (GEO accession number: GSE30611) was mapped to the reference human genome (hg19) using TopHat v1.0.14<sup>2</sup> and transcripts were reconstructed with Cufflinks v1.0.2<sup>3</sup>, as previously described<sup>4</sup>. Exons expressed in all tissues in the Human Body Map dataset were termed constitutive and chosen for sgRNA targeting. For the mouse library, we selected exons that are included in all of the NCBI RefSeq (<http://www.ncbi.nlm.nih.gov/refseq/>) transcripts for the same gene.

In both species, for each gene, we excluded the first and last exons and any exon that contained an alternative splicing site. Using these procedures, we chose 4 constitutive exons for each gene. For genes where there were less than 4 constitutive exons, we added exons starting from the second coding exon towards the end of the gene. The coding boundaries for each exon were identified using the NCBI Consensus CoDing Sequence (CCDS) database (<http://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi>).

(2) *Choice of sgRNA sequences:* For each gene, all *S. pyogenes* Cas9 sgRNA sequences of the form (N)<sub>20</sub>NGG in all constitutive exons (from the procedure in (1) above) were selected as candidate targets. Each 20mer candidate target sgRNA was mapped to a precompiled index containing all 20mer sequences in the human genome followed by either NGG or NAG. This mapping was done using the Bowtie short read aligner<sup>5</sup>, allowing up to 3 base mismatches (with parameters `-a --best -v 3`). For each potential off-target identified by the mapping, the following score was calculated<sup>6</sup>:

$$S_{\text{off-target}} = \prod_{e \in \mathcal{M}} (1 - W[e]) \times \frac{1}{\left(\frac{(19-\bar{d})}{19} \times 4 + 1\right)} \times \frac{1}{n_{mm}^2}$$

Within the first term,  $M$  is the set of mismatches between the potential off-target and the candidate target sgRNA. For each mismatch in  $M$ ,  $e$  denotes the position of the mismatch relative to the protospacer-adjacent motif (PAM). The values of  $e$  range between 1 (for the most PAM-distal base in the sgRNA) to 20 (for the most PAM-proximal base in the sgRNA). Using  $e$  as the index, a value from the look-up table  $W$  denotes the empirically-determined weight (relative to the other terms in the  $S_{\text{off-target}}$  metric) that penalizes mismatches based on their position in the sgRNA sequence<sup>6</sup>:

$e$	1	2	3	4	5	6	7	8	9	10
$W[e]$	0	0	0.014	0	0	0.395	0.317	0	0.389	0.079
$e$	11	12	13	14	15	16	17	18	19	20
$W[e]$	0.445	.508	0.613	0.851	0.732	0.828	0.615	0.804	0.685	0.583

$\bar{d}$  is the mean pairwise distance between all pairs of elements in  $M$  and  $n_{mm}$  is the number of elements in  $M$ .

Each of the individual off-targets ( $S_{\text{off-target}}$ ) for one candidate target sgRNA were integrated into an aggregate off-target score as follows:

$$S_{\text{guide}} = \frac{100}{1 + \sum_{i=1}^n S_{\text{off-target}}(i)}$$

$S_{\text{off-target}}$  is the score for a single potential off-target and  $S_{\text{guide}}$  is a score (0-100 scale) that provides a rating of how well a candidate sgRNA uniquely targets within the genome.

Candidate sgRNAs with higher values of  $S_{\text{guide}}$  are predicted to have less off-target activity.

For each gene, we chose 6 sgRNAs with the highest values of  $S_{\text{guide}}$  across all 4 constitutive exons (from step (1)) subject to constraint that no more than 2 sgRNAs could target a single exon. This procedure resulted in extremely uniform coverage of genes: For 99.4% of genes in the GeCKOv2 human library and 99.8% of genes in the GeCKOv2 mouse library, each gene was targeted with exactly 6 sgRNAs.



(3) *Targeting of mature miRNAs*: For the design of miRNA targeting guides, we used the pre-miRNA sequence coordinates from the mirBASE database ([www.mirbase.org](http://www.mirbase.org))<sup>7</sup>. For each pre-miRNA, we listed all the possible (N)<sub>20</sub>NGG sequences. In total, we chose up to 4 sgRNAs per miRNA or less if there were not 4 sgRNAs present in the pre-miRNA sequence.

**GeCKO library pooled synthesis and cloning.** DNA oligonucleotide library synthesis was completed on a programmable microarray using a B3 Synthesizer (CustomArray) and SAFC Proligo reagents (Sigma), as recommended by the manufacturer. The synthesis products were cleaved from the microarray and deprotected by overnight incubation in 28-30% ammonium hydroxide at 65 °C, dried, resuspended in 30 ul TE buffer and then purified using a QIAquick spin column (Qiagen). Full-length oligonucleotides (74 nt) were amplified by PCR using Phusion HS Flex (NEB). For the PCR reaction, the manufacturer's protocol was followed using 0.1 ul of synthesized oligonucleotide template, primers Array F and ArrayR (see below), an annealing temperature of 63 °C, an extension time of 15 s, and 20 cycles. After PCR, the 140 bp amplicon was size-selected using a 2% agarose E-Gel EX (Life Technologies, Qiagen).

ArrayF        TAACTTGAAAGTATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGAC  
                  GAAACACCG

ArrayR        ACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCT  
                  AGCTCTAAAAC

The vector backbone (lentiCRISPR v2 or lentiGuide-Puro) was digested with BsmBI (Fermentas) and treated with FastAP (Fermentas) at 37°C for 2 hours and gel-purified on a 1 % E-Gel EX (Life Technologies, Qiagen). A 20 ul Gibson ligation reaction (NEB) was performed using 10 ng of the gel-purified inserts and 25 ng of the vector (for lentiCRISPR v2) and using 5 ng of the gel-purified inserts and 25 ng of the vector (for lentiGuide-Puro). From the ligation, 0.5 ul of the reaction was transformed into 25 ul of electrocompetent cells (Lucigen) according to the manufacturer's protocol using a GenePulser (BioRad). To ensure no loss of representation, sufficient parallel transformations were performed using the same ligation reaction and plated onto 245 mm x 245 mm plates (Corning) with carbenicillin selection (50 ug/ml), which yielded

30-200 X library coverage. Colonies were scraped off plates and combined before plasmid DNA extraction using Endotoxin-Free Plasmid Maxiprep (Qiagen).

**Library sequencing and validation.** To check library representation, synthesis fidelity, and bias, libraries were amplified and then deep sequenced. First, libraries were PCR amplified for 16 cycles using Phusion Flash High-Fidelity (Thermo) with primers to add adaptors for Illumina sequencing. For all libraries, PCR reactions were performed in duplicate and barcoded to allow quantification of bias introduced by PCR. Samples were sequenced on a MiSeq following the manufacturer's protocol using a v3 150 cycle kit with 10% PhiX (Illumina).

PCR replicates were demultiplexed using FASTX-Toolkit (Hannon Lab, CSHL) and adaptors were trimmed using cutadapt to leave only the sgRNA guide sequence<sup>8</sup>. Alignment of the guide sequence to the appropriate GeCKO library index was done using Bowtie with parameters to tolerate up to a single nucleotide mismatch<sup>5</sup>. The Bowtie alignment was then read into Matlab for further analysis (Mathworks). For all libraries, greater than 90 % of sgRNAs were represented with at least one sequencing read and the difference in representation between the 90<sup>th</sup> and 10<sup>th</sup> percentile sgRNAs was always less than 15-fold. Histograms of sgRNA distributions and cumulative sgRNA read counts for libraries in both lentiCRISPRv2 and lentiGuide-Puro are shown in Supplementary Figure 2 (human) and Supplementary Figure 3 (mouse).

**Vector maps, library sequences and reagent distribution.** All lentiCRISPR plasmids and GeCKO libraries are available on Addgene. Please see <http://genome-engineering.org/gecko/> for downloadable vector maps. See labeled sequences in Supplementary Data for more details.

The sgRNAs in the human and mouse v2 GeCKO libraries are available for download on the GeCKO website or from Addgene.

## Supplementary Data

### LentiCRISPR v2

<b>Pink</b>	BsmBI TypeIIIs sites for cloning in 20bp sgRNA guide sequence after removing 1.8kb spacer sequence
<b>Green</b>	+85 chrRNA sequence for sgRNA
<b>Gray</b>	EFS promoter sequence
<b>Yellow</b>	Human codon-optimized SpCas9 coding sequence
<i>Italics</i>	FLAG peptide tag
<b>Red</b>	P2A bicistronic linker/self-cleaving peptide
<b>Blue</b>	Puromycin resistance coding sequence

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC  
CAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGGAGCAAAATTTAAGCTACAACAAGGCAAGG  
CTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATA  
CGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAG  
TTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAAT  
GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCC  
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCTTATTGACGTCAATGACGGTAAATGGCCCGCCTGG  
CATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT  
GGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA  
TTGACGTCAATGGGAGTTTGTGGCACAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATTTG  
ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTGCCTGTACTGGGTCTCTCTGGT  
TAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGA  
GTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTG  
GAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAGCGAAAAGGGAAACCCAGAGGAGCTCTCTCGACGCAGGAC  
TCGGCTTGTGAGCGCGCACGGCAAGAGCGAGGGCGGCGACTGGTGAGTACGCCAAAATTTTACTAGCGGGAG  
GCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGG  
TTAAGGCCAGGGGAAAGAAAAATATAAATTAACAATATAGTATGGGCAAGCAGGGAGCTAGAACGATTTCGAGT  
TAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGAT  
CAGAAGAATTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACC  
AAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCA  
GACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTA  
GGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCT  
TGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGT  
CTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGG  
GGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG  
TTGCTCTGGAAAACCTATTTGCACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTT  
GGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAA  
TCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTTGTGGAATTGGTTTAA  
CATAACAAATTTGGCTGTGGTATATAAATTTTTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAAGAATAGTTTTTG  
CTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTTCAGACCCACCTCCCAACCCCGAGG  
GGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGTGGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGG  
ATCGGCACTGCGTGCGCCAATTTGCGAGACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGATTGGG  
GGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTAC  
AAAAATTCAAAATTTTTCGGGTTTATTACAGGGACAGCAGAGATCCAGTTTTGGTTAATTAAGGTACCGAGGGCCTATT  
TCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACA  
CAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTTATGTTTT  
AAAATGGACTATCATATGCTTACCCTAATTGAAAGTATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACG  
AAACACCGGAGACGTTTGTAAATGAGCACACAAAATACACATGCTAAAATATTATATTCTATGACCTTTATAAAATC  
AACCAAAATCTTCTTTTAAATAACTTTAGTATCAATAATTAGAATTTTTATGTTCTTTTTGCAAACTTTTAAATAAA  
AATGAGCAAAATAAAAAACGCTAGTTTTAGTAACTCGCGTTGTTTTCTTACCTTTAATAATAGCTACTCCACCAC  
TTGTTCTTAAGCGGTACAGCTCCTGCTTCAATATTTTTGAGCATCTTCAAATGTTCTAACTCCACCAGCTGCTTTA  
ACTAAAGCATTGTCTTTAACAACCTGACTTCATTAGTTTAAACATCTTCAAATGTTGCACCTGATTTTGAAAATCCTGT

TGATGTTTTAACAAATTCTAATCCAGCTTCAACAGCTATTTACACAAGCTTTTCATGATTTCTTCTTTTGTAAATAAAC  
AATTTTCCATAATACATTTAACAAACATGTGATCCAGCTGCTTTTTTACAGCTTTTCATGTCTTCTAAAACATAATTCA  
TAATTTTTGTCTTTAATGCACCAATATTTAATACCATATCAATTTCTGTTGCACCATCTTTAATTGCTTCAGAAAC  
TTCGAATGCTTTTGTAGCTGTTGTGCATGCACCTAGAGGAAAACCTACAACATTTGTTATTTCTACATTTGTGCCTT  
TTAATAATTTCTTACAATAGCTTGTTCATATGAATTAACACAACTGTTGCAAAATCAAATTCATTTGCTTCATCA  
CATAATTTGTTAATTTTCAGCTTTTCGTAGCATCTTGTTTAATAATGTGTGATCTATATATTTGTTTAGTTTCATTTT  
TTCTCCTATATATTCATTTTTAATTTAATTTCTTTAATAATTTTCGTCTACTTTAACTTTAGCGTTTTGAACAGATTC  
ACCAACACCTATAAAATAAATTTTTAGTTTAGGTTTCCAGTTCCACTTGGGCGAACAGCAAATCATGACTTATCTTCTA  
AATAAAATTTTAGTAAGTCTTGTCTTGGCATATTATACATTTCCATTCGATGTAGTCTTCAACATTTAAACAATTTAAGT  
CCAGCAATTTGAGTTAAGGGTGTCTCTCAATGATTTTCAATGTTTCAATTTTTAATTTCTTTCTTCTTGTGTTT  
AAAATTTCAAGTTTAAAGTGAAGTGAATATGCACCCATTTCTTTAAATAAATCTTCTAAATAGTCTACTAATGTTT  
TATTTTGTTTTTTATAAAATCAAGCAGCCTCTGCTATTAATATAGAAGCTTGTATTCCATCTTTATCTCTAGCTGAG  
TCATCAATTACATATCCATAACTTTCTTATAAGCAAAAACAAAATTTAATCCGTTATCTTCTTCTTTAGCAATTTT  
TCTACCCATTCATTTAAATCCAGTTAAAGTTTTTACAATATTAACTCCATATTTTTTCATGAGCGATTCTATCACCCA  
AATCACTTGTACAAAACCTGAATATAGAGCCGGATTTTTTGAATGCTATTTAAGCGTTTTAGATTTGATAATTTT  
CAATCAATTAATTTGGTCTGTTTGTATTTCCATCTAATCTTACAAAATGACCATCATGTTTTATTGCCATTCCAAA  
TCTGTCAGCATCTGGGTCATTCATAATAATAATATCTGCATCATGTTTAAATACCATATTTCAAGCGGTATTTTTTCATG  
CAGGATCAAATTTCTGGATTTGGATTTACAACATTTTTAATGTTTCATCTTCAAATGCATGCTCTTCAACCTCAATA  
ACGTTATATCCTGATTCACGTAATTTTTTGGGGTAAATTTAGTTTCTGTTCCATTAACCTGCGCTAAAATAAATTTT  
TAAATCTTTTTTAGCTTCTTGTCTTTTTTGTACGCTCTGTTTAGAGCTAGAAAATAGCAAGTTAAAATAAGGCTA  
GTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGTCTTTTTGAATTCGCTAGCTAGGCTTGAAGGAGTGGG  
AATTGGCTCCGGTGCCTGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGGAGGGGTCCGGCA  
ATTGATCCGGTGCCTAGAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCC  
GAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAAC  
ACAGGACCGGTTCTAGAGCGCTGCCACCATGGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGG  
GCTGGGCGGTGATCACCGCAGTACAAGGTGCCAGCAAGAATCAAGGTGCTGGGCAACACCGGACCGGCACAGC  
ATCAAGAAGAACCTGATCGGAGCCCTGCTGTTTCGACAGCGGCAACAGCCGAGGCCACCGGCTGAAGAGAACCGC  
CAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG  
ACGACAGCTTCTTCCACAGACTGGAAGAGTCTTCTGTTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTC  
GGCAACATCGTGGACGAGGTGGCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAG  
CACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAGTTCCGGGGCCACTTCTGATCG  
AGGGCGACCTGAACCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTT  
GAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCTGTCTGCCAGACTGAGCAAGAGCAGACGGCT  
GGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGTGGCCTGTTTCGAAACCTGATTGCCCTGAGCCTGGGCC  
TGACCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGACGAC  
GACCTGGACAACCTGCTGGCCAGATCGGCGACCAAGTACGCCGACCTGTTTCTGGCCGCAAGAACCTGTCCGACGC  
CATCCTGCTGAGCGACATCTGAGAGTGAACACCGAGATACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGAT  
ACGACGAGCACACCAGGACCTGACCCCTGCTGAAAGCTCTCGTGCAGGAGCAGCTGCCTGAGAAGTACAAAGAGATT  
TTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTTCA  
CAAGCCCATCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGC  
AGCGGACCTTCGACAACCGCAGCATCCCCACAGATCCACCTGGGAGAGCTGCACGCCATTTCTGCGGCGGACAGGAA  
GATTTTTACCCATTTCTGAAGGACAACCGGAAAAGATCGAGAAGATCTGACCTTCCGCATCCCCACTACGTGGG  
CCCTCTGGCCAGGGGAAACAGCAGATTCGCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCTTGAACCTTCG  
AGGAAGTGGTGGACAAGGGCGCTTCCGCCAGAGCTTATCGAGCGGATGACCAACTTCGATAAGAACCTGCCAAC  
GAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTACCCTGTATAACGAGCTGACCAAAGTGAATACGT  
GACCGAGGGAATGAGAAAGCCCGCCTTCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCA  
ACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTCTTTCGACTCCGTGGAAATCTCC  
GGCGTGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCTT  
GGACAATGAGGAAAACGAGGACATTTCTGGAAGATATCGTGTGACCCCTGACACTGTTTGGAGACAGAGAGATGATCG  
AGGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTATGAAGCAGCTGAAGCGGCGGAGATACACCGGC  
TGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCTGGATTTCTGAA  
GTCCGACGGCTTCGCCAACAGAAAATTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGA  
AAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAG  
GGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAGTGTGAGGCGGCAAGCCCGAGAACATCGTGTAT  
CGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGG  
GCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTAC  
CTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAAGTGGACATCAACCGGCTGTCCGACTACGATGT

GGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACC  
GGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAATACTGGCGGCAGCTGCTGAACGCC  
AAGCTGATTACCCAGAGAAAGTTTCGACAATCTGACCAAGGCCGAGAGAGGCCGCTGAGCGAACTGGATAAAGGCCGG  
CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAGACAGTGGCACAGATCCTGGACTCCCGGATGAACA  
CTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTT  
CGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAATACTACCACCACGCCACGACGCCTACCTGAACGCCGT  
CGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACG  
TGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
AACTTTTTCAAGCCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGGCTGAGCATGCCCAAGTGAATATCG  
TGA AAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCAAGAGGAACAGCGATAAGCTGATC  
GCCAGAAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGT  
GGCAAAGTGGA AAAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAAGAGCTGCTGGGGATCACCATCATGGAAGAA  
GCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAGGACCTGATCATCAAG  
CTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGG  
AAACGAACTGGCCCTGCCCTCCAAATATGTGAACCTTCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCC  
CCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGC  
GAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAA  
GCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGT  
ACTTTGACACCACCATCGACCAGGAAGAGGTACACCAGCACCAAGAGGTTGCTGGACGCCACCCTGATCCACCAGAGC  
ATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAAGCGACCTGCCGCCACAAAGAAGGC  
TGGACAGGCTAAGAAGAAGAAA *GATTACAAAGACGATGACGATAAG* **GGATCCGGCGCAACAACTTCTCTCTGCTGA**  
**AACAAGCCGGAGATGTGGAAGAGAATCCTGGACCG** ACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGAC  
GTCCCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCCGCCACTACCCCGCCACGCGCCACACCGTCGATCCGGACCG  
CCACATCGAGCGGGTCACCGAGCTGCAAGAATCTTCTCACGCGCTCGGGCTCGACATCGGCAAGGTGTGGTTCG  
CGGACGACGGCGCGGTTGGCGGTCTGGACCACGCCGGAGAGCGTCAAGCGGGGGCGGTGTTCCGCCGAGATCGGC  
CCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGACCGGCC  
CAAGGAGCCCGCGTGGTTCTGGCCACCGTCGGAGTCTCGCCCGACCACAGGGCAAGGGTCTGGGCAGCGCCGTCG  
TGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCGGGGTGCCCGCTTCTGGAGACCTCCGCGCCCGCAACCTCCCC  
TTCTACGAGCGGCTCGGCTTACCCTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCG  
CAAGCCCGGTGCCTGAACGCGTTAAGTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTC  
TTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATG  
GCTTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTTTTATGAGGAGTTGTGGCCGTTGTGAGGCAACG  
TGGCGTGGTGTGCACTGTGTTTGTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCG  
GGACTTTTCGTTTTCCCTCCCTATTGCCACGGCGGAACCTCATCGCCGCTGCCTTGGCCGCTGCTGGACAGGGGCT  
CGGCTGTTGGGCACTGACAATCCGTGGTGTGTCGGGGAAATCATCGTCCTTTCTTGGCTGCTCGCCTGTGTTGC  
CACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCTTTCGGCCCTCAATCCAGCGGACCTTCTTCCCGCGGCC  
TGCTGCCGGCTCTGCGGCCTTTCGCGTCTTCGCCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCTCC  
CCGCGTCGACTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGACTG  
GAAGGGCTAATTCACCTCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTG  
AGCTGGGAGCTCTCTGGCTAAGTGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGGTGTCTCAAGTAG  
TGTGTGCCGCTGTTGTGTGACTCTGGTAAGTACAGATCCCTCAGACCTTTTATGTCAGTGTGGAAAATCTCTAGC  
AGGGCCCGTTTTAAACCCGCTGATCAGCCTCAGCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTTGCCCTCCCCG  
TGCTTCTTGGACCTGGAAGGTGCCACTCCCCTGTCCTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTG  
AGTAGGTGTCAATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA  
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGC  
CCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCAGCGCCCTAGCG  
CCCGCTCCTTTTCGTTTTCTTCCCTTCTTCTCGCCACGTTTCGCCGCTTTCCCGTCAAGCTCTAAATCGGGGGCT  
CCCTTTAGGGTTCGATTTAGTGTCTTACGGCACCTCGACCCAAAAAATTTGATTAGGGTGATGGTTCACGTAGTG  
GGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTCGAGTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTC  
ACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTATTATAAGGGATTTTGGCGATTTCCGGCCTATTGGTT  
AAAAAATGAGCTGATTTAACA AAAAATTTAACGCGAATTAATCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTC  
CCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAG  
GCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCC  
ATCCCGCCCTAACTCCGCCAGTTCGCCCATCTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGC  
CGAGGCGCCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGA AAAAGC  
TCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTTGACAATTAATCATCGGCATAGTATATCGGCA

TAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGGAC  
GTCGCCGGAGCGGTTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGACTTCGTGGAGGACGACTTCGCCGGTGT  
GGTCCGGGACGACGTGACCCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGT  
GGGTGCGCGGCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACCTCCGGGACGCCTCCGGGCCG  
GCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACCT  
CGTGGCCGAGGAGCAGGACTGACACGTGCTACGAGATTTCCGATTCACCCGCCGCTTCTATGAAAGGTTGGGCTTCG  
GAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCAAC  
TTGTTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACT  
GCATTTAGTTGTGGTTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCCGACCTCTAGCTAGAGCT  
TGGCGTAATCATGGTCATAGCTGTTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACAACATACGAGCCGGA  
AGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGTTTT  
CCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTGCGTATTGGGC  
GCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGCTGCGGCGAGCGGTATCAGCTCACTCAAAG  
GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG  
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCT  
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCT  
CCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC  
ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCCG  
ACCGCTGCGCCTTATCCGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC  
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCTAACTACGGCTA  
CACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT  
CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT  
CAAGAAGATCCTTTGATCTTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCAT  
GAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATG  
AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCC  
ATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGAT  
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTG  
GTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT  
AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTC  
CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGA  
TCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTTACTGTCATG  
CCATCCGTAAGATGCTTTTTCTGTGACTGGTGAAGTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAG  
TTGCTCTTGGCCGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTTAAAAGTGCTCATCATTGGAAAAC  
GTTCTTCCGGGGCGAAAACCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC  
TGATCTTCAGCATCTTTTACTTTACCAGCGTTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG  
AATAAGGGCGACACGGAAATGTTGAATACTCATACTTTCTTTTTCAATATTATTGAAGCATTATCAGGGTTATT  
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA  
GTGCCACCTGAC

*lentiCas9-Blast*

- Gray EFS promoter sequence
- Yellow Human codon-optimized SpCas9 coding sequence
- Italics FLAG peptide tag
- Red P2A bicistronic linker/self-cleaving peptide
- Blue Blast resistance coding sequence

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC  
CAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGC GCGAGCAAATTTAAGCTACAACAAGGCAAGG  
CTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATA  
CGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAG  
TTCCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAAT  
GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCC  
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGG  
CATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT  
GGTGTATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGTTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA  
TTGACGTCAATGGGAGTTTTGTTTTGGCACAAAATCAACGGGACTTCCAAAATGTCGTAACAACCTCCGCCCATTTG  
ACGCAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTGCCTGTACTGGGTCTCTCTGGT  
TAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGA  
GTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCTTTTAGTCAGTGTG  
GAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACGCAGGAC  
TCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCAGGGGGCGGCGACTGGTGTAGTACGCCAAAAATTTTACTAGCGGAG  
GCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGG  
TTAAGGCCAGGGGGAAAGAAAAAATATAAATTAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTTCGAGT  
TAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGAT  
CAGAAGAATTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACC  
AAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCA  
GACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAATTGAACCATTA  
GGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCT  
TGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGT  
CTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGG  
GGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAGGATCAACAGCTCCTGGGGATTTGGGG  
TTGCTCTGGAAAACCTATTTGCACCCTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTT  
GGAATCACACGACTGGATGGAGTGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAA  
TCGAAAACCCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTTGTTGAATTGGTTAA  
CATAACAAATTGGCTGTGGTATATAAAATTTTATCATAATGATAGTAGGAGGCTTGGTAGGTTAAGAATAGTTTTTG  
CTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGG  
GGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTTCGATTAGTGAACGG  
ATCGGCACTGCGTGCGCCAATTCTGCAGACAAATGGCAGTATTTCATCCACAATTTTAAAAGAAAAGGGGGATTGGG  
GGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAACTAAAGAATTACAAAAACAATTAC  
AAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCAGTTTGGTTAATTAAGCTAGCTAGGTCTTGAA  
AGGAGTGGGAATTTGGCTCCGGTGCCTGAGAGAAGGTGGCGCGGGGTAACCTGGGAAAGTGATGTGCTGTACTGGCTCCGC  
GGGTCCGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGGGGTAACCTGGGAAAGTGATGTGCTGTACTGGCTCCGC  
CTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGC  
CGCCAGAACACAGGACCGGTTCTAGAGCGCTGCCACCATTGGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCA  
ACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAGGTGCTGGGCAACACCCGAC  
CGGCACAGCATCAAGAAGAACCCTGATCGGAGCCCTGCTGTTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAA  
GAGAACCGCCAGAAGAAGTACACCAGACGGAAGAACCAGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGG  
CCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGACTCCTTCTGGTGGAAAGAGGATAAGAAGCAGCAGGCGGCAC  
CCCATCTTGGCAACACTCGTGGACGAGGTGGCTACCAGCAGAGAAGTACCCACCATCTACCACTGAGAAAGAAGAACT  
GGTGGACAGACCCGACAAGGCCGACCTGCGGCTGATCTATCTGAGCCCTGGCCACATGATCAAGTTCGGGGCCACT  
TCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAAC  
CAGCTGTTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAG  
CAGACGGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGAAACCTGATTGCCCTGA  
GCCTGGGCTGACCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACC  
TACGACGACGACCTGGACAACCTGCTGGCCAGATCGGCGACCAAGTACGCCGACCTGTTTCTGGCCGCAAGAACCT

GTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGA  
TCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC  
AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTA  
CAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAAGTCTCGTGAAGCTGAACAGAGAGGACCTGC  
TGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGG  
CGGCAGGAAGATTTTTACCCATTCTGAAGGACAACCGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTA  
CTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTGCGCTGGATGACCAGAAAAGAGCGGAGGAAACCATCACCCCT  
GGAACCTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAAC  
CTGCCAACGAGAAGGTGCTGCCAAGCACAGCTGTGTACGAGTACTTCACCGTGATAACGAGCTGACCAAAGT  
GAAATACGTGACCGAGGGAATGAGAAAAGCCCGCTTCTGAGCGGCAGCAGAAAAGGCCATCGTGGACCTGCTGT  
TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTG  
GAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATAACCAGATCTGCTGAAAATTATCAAGGACAA  
GGACTTCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAG  
AGATGATCGAGGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAAGTGAAGCAGCTGAAGCAGGCGGAGA  
TACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGA  
TTTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAGAGG  
ACATCCAGAAAAGCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCC  
ATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAGTGATGGGCCGGCACAAGCCCCGAGAA  
CATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGA  
TCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAG  
AAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAAGTGGACATCAACCGGCTGTCCGA  
CTACGATGTGGACCATATCGTGCCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCG  
ACAAGAACCAGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTGCTGAAGAAGATGAAGAACTACTGGCGGCAGCTG  
CTGAACGCCAAGCTGATTACCCAGAGAAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGGCTGAGCGAAGTGG  
TAAGGCGGCTTCATCAAGAGACAGCTGGTGGAAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCC  
GGATGAACACTAAGTACGACGAGAATGACAAGTATCGCGGAAGTGAAGTGATCACCTGAAGTCCAAGCTGGTGGT  
TCCGATTTCCGGAAGGATTTCCAGTTTTACAAAAGTGCAGGATCAACAACCTACCACCAGCCACGACCCCTACCT  
GAACGCCGTGCTGGGAACCGCCTGATCAAAAAGTACCCTAAGCTGGAAAAGCGAGTTCTGTACGGCGACTACAAGG  
TGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGC  
AACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAA  
CGGCGAAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGGCACCGTGCAGGAAAAGTGTGAGCATGCCCAAG  
TGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGAT  
AAGCTGATCGCCAGAAAAGAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGT  
GCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCA  
TGGAAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAAGAAGTGAAAAAGGACCTG  
ATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAAGT  
GCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCTGTACCTGGCCAGCCACTATGAGAAGCTGA  
AGGGCTCCCCCGAGGATAATGAGCAGAAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTGGACGAGATCATCGAG  
CAGATCAGCGAGTTCTCCAAGAGAGTGTATCCTGGCCGACGCTAATCTGGACAAAAGTGTGTCCGCCTACAACAAGCA  
CCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCCTGCCG  
CCTTCAAGTACTTTGACACCACCATCGACCCGGAAGAGGTACACCAGCACCAAGAGGTGCTGGACGCCACCCCTGATC  
CACCAGACATCACCGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAAGCGACCTGCCGAC  
AAAGAAGGCTGGACAGGCTAAGAAGAAGAAA *GATTACAAAGACGATGACGATAAGGGATCCGGCGCAACAAACTTCT*  
*CTCTGCTGAAAACAAGCCGGAGATGTGGAAGAGAATCCTGGACCGATGGCCAAGCCTTTGTCTCAAGAAGAATCCACC*  
*CTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTC*  
*TAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTTACTGGGGACCTTGTGCAGAAGTGTGGTGTGG*  
*GCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTGAGC*  
*CCCTGCGGACGGTGGCAGAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAGTGTGGACA*  
*GCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTAAGAATTCGATATCAAGCTTA*  
*TCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTACGCTA*  
*TGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGATAA*  
*ATCCTGGTTGCTGTCTTTTATGAGGAGTTGTGGCCCGTTGTGAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTG*  
*ACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATT*  
*GCCACGGCGGAACCTCATCGCCGCTGCCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGT*  
*GGTGTGTGCGGGGAAATCATCGTCTTTCTTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCTT*  
*TCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCTTCCCGCGGCTGCTGCCGGCTCTGCGGCCTCTTCCG*  
*CGTCTTCGCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCTCCCCGCATCGATACCGTCGACCTCGAGA*



CCTAGAAAAACATGGAGCAATCACAAAGTAGCAATACAGCAGCTACCAATGCTGATTGTGCCTGGCTAGAAGCACAAAG  
AGGAGGAGGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT  
AGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCAACGAAGACAAGATATCCTTGATCTGTG  
GATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCT  
TTGGATGGTGTACAAGCTAGTACCAGTTGAGCAAGAGAAGGTAGAAGAAGCCAATGAAGGAGAGAACACCCGCTTG  
TTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTATTAGAGTGGAGGTTTGACAGCCGCCTAGC  
ATTTTCATCACATGGCCCGAGAGCTGCATCCGGACTGTAAGGTTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGC  
TCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAAGTAGTGTGTGCCCGT  
CTGTTGTGTGACTCTGGTAAGTACTAGAGTCCCTAGACCCCTTTTAGTCAGTGTGAAAAATCTCAGCAGGGCCCGCTT  
AAACCCGCTGATCAGCTGCCTTCTAGTCCCTGAGCCATCTGTTGTTTGGCCCTCCCCGCTCCCTTCTCTTG  
ACCCTGGAAGGTGCCACTCCACTGTCCTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA  
TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG  
CGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGC  
GCATTAAGCGCGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTT  
CGCTTTCTCCCTTCTCTTCTCGCCACGTTCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGT  
TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGTGTTTACGTAAGTGGGCCATCGCCC  
TGATAGACGGTTTTTCCGCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTCCAAACTGGAACAAC  
ACTCAACCCTATCTCGGTCTATTCTTTTGAATTAATAAGGGATTTTGGCGATTTCCGCTATTGGTTAAAAAATGAGC  
TGATTTAACAATAATTAACGCGAATTAATCTGTGGAATGTGTGTGAGTTAGGGTGTGGAAAGTCCCAGGCTCCC  
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCA  
GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCAGCCCTAACTCCGCCATCCCGCCCT  
AACTCCGCCAGTTCGCCCCATTCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCT  
CTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGAAAAAGCTCCCGGGAGCT  
TGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATAC  
GACAAGGTGAGGAACAAACCATGGCCAAGTTGACCAGTCCGCTCCGGTGTCCAGCGCGCAGCTCGCCGGAGC  
GGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCGGGACTTCTGAGGAGCAGACTTCGCCGGTGTGGTCCGGGACG  
ACGTGACCCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTCCGGAACAACACCCTGGCCTGGGTGTGGGTGCGCGGC  
CTGGACGAGCTGTACGCCGAGTGGTCCGAGGTCGTGTCCACGAACCTCCGGGACGCTCCGGGCCGGCCATGACCGA  
GATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGCAACTGCGTGCCTTCTGAGCCGAGG  
AGCAGGACTGACACGTGCTACGAGATTTGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTT  
CGGGACGCCGGCTGGATGATCTCCAGCGCGGGGATCTCATGCTGGAGTCTTCCGCCACCCCAACTTGTATTGCT  
AGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTACTGCATTCTAGTT  
GTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCA  
TGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAAGTG  
TAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTCGCTTCCGCTCACTGCCCGCTTCCAGTCGGGAA  
ACCTGTCTGTCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGGCTATTGGGCGCTCTTCCGCT  
TCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCCGCTGCGGCGAGCGGTATCAGCTCAAGGCGGTAATACG  
GTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA  
AGGCCGCGTGTGTCGGGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGG  
TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGAC  
CCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCCTTTCTCATAGCTCAGCTGTAGGT  
ATCTCAGTTCCGTTGAGTGTGCTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTCCAGCCGACCGCTGCGCC  
TTATCCGGTAACTATCGTCTTGTGATCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG  
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA  
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA  
ACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC  
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCAAGATTATCAA  
AAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAATCAATCTAAAGTATATATGAGTAAACTTGG  
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCTTCCATCCATAGTTGCCTG  
ACTCCCGTCTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCAAGTGTGCAATGATACCGCGAGACC  
CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACT  
TTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGC  
CGTGTGGCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCCAAC  
GATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCCGGTCTCCGATCGTTGTGAGA  
AGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTTCTTACTGTGATGCCATCCGTAAG  
ATGCTTTTCTGTGACTGGTGTGACTCAACCAAGTCAATCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC  
CGCGTCAATACGGGATAATACCGGCCACATAGCAGAACCTTAAAAGTGTCTCATCATTGAAAAACGTTCTTCCGGG

CGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC  
ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA  
CACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGC  
GGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGA  
C

*lentiGuide-Puro*

- Pink** BsmBI TypeII sites for cloning in 20bp sgRNA guide sequence after removing 1.8kb spacer sequence
- Green** +85 chrRNA sequence for sgRNA
- Gray** EF1a promoter sequence
- Blue** Puromycin resistance coding sequence

```
CCCGGGTGCAAAGATGGATAAAGTTTTAAACAGAGAGGAATCTTTGCAGCTAATGGACCTTCTAGGTCTTGAAAGGA
GTGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGT
CGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTT
TTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGC
AGAACACAGGTAAGTGCCGTGTGTGGTCCCGCGGGCCTGGCCTCTTACGGGTTATGGCCCTTGCCTGCCTTGAAT
TACTTCCACTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGTGGGAGAGTTCGAGGCCTTG
CGCTTAAGGAGCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGCCCGCGTGCATCTGTT
GGCACCTTCGCGCTGTCTCGTCTTTGATAAGTCTTAGCCATTTAAAATTTTTGATGACCTGTGCGACGCTT
TTTTTCTGGCAAGATAGTCTTGAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGCGGGCGG
CGACGGGGCCCGTGCCTCCAGCGCACATGTTGCGCGAGGCGGGCCTGCGAGCGCGGCCACCGAGAATCGGACGGG
GGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCGTGTATCGCCCCGCCCTGGGCGGCAAGG
CTGGCCCGGTGCGCACCAAGTTCGCTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAG
GACGCGGCGCTCGGGAGAGCGGGCGGGTGTGAGTACCCACACAAAGGAAAAGGGCCTTTCGCTCTCAGCCGTGCTT
CATGTGACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTGCTCTTTA
GGTTGGGGGGAGGGTTTTATGCGATGGAGTTTCCACACTGAGTGGTGGAGACTGAAGTTAGGCCAGCTTGGA
CTTGATGTAATTTCTCTTGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTC
AAAGTTTTTTCTTCCATTTCAAGGTGTCGTGACGTACGGCCACCATGACCGAGTACAAGCCCACGGTGCCTCGCC
ACCCGCGACGACGTCCCCAGGGCCGTACGCACCCTCGCCGCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGT
CGATCCGGACCGCCACATCGAGCGGGTACCGAGCTGCAAGAACTCTTCTCACGCGCGTTCGGGCTCGACATCGGCA
AGGTGTGGGTGCGGACGACGGCGCCGCGTGGCGGTCTGGACCACGCCGGAGAGCGTCAAGCGGGGGCGGTGTTT
GCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCTGGC
GCCGACCCGGCCAAAGGAGCCCGCTGGTTCCTGGCCACCGTCGGAGTCTGCCCGACCACAGGGCAAGGCTCGG
GCAGCCCGCTGTGCTCCCGGAGTGGAGCGCCGAGCGCGCGGGGTGCCCGCTTCTGGAGACCTCCGCGCCC
CGAACCTCCCCTTCTACGAGCGGCTCGGCTTACCCTGTCACCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTG
GTGCATGACCCGCAAGCCCGGTGCTGAACGCGTTAAGTGCACAATCAACCTCTGGATTACAAAATTTGTGAAAGAT
TGACTGGTATTCTTAACTATGTTGCTCTTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATT
GCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCGT
TGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTC
AGCTCCTTTCCGGGACTTTTCGCTTTCCCTCCTATTGCCACGGCGGAACCTCATCGCCGCTGCCTTGGCCGCTGC
TGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGTTGGTGTGTCGGGGAAATCATCGTCTTTCTTGGCTGCT
CGCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTC
CTTCCCGCGGCTGCTGCCGGCTCTGCGGCCTTCCCGCTTTCGCTTTCGCCCTCAGACGAGTCGGATCTCCCTT
TGGGCCGCTCCCGCGCTCGACTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAA
AAGGGGGGACTGGAAGGGCTAATCACTCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTTGTT
AGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCTTGGAG
TGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGG
AAAATCTCTAGCAGTACGTATAGTAGTTTATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGAAATGAATATC
AGAGAGTGAGAGGAACCTGTTTATTGCAGCTTATAATGTTTCAAATAAAGCAATAGCATCACAATTTACAAATA
AAGCATTTTTTCACTGCATTTAGTTGTGGTTTTGTCCAAATCATCAATGTATCTTATCATGTCTGGCTTAGCTA
TCCCGCCCCTAATCCGCCATCCCGCCCCTAATCCGCCAGTTCGCCCATTTCCGCCCTAGGCTGACTAATT
TTTTTTATTTATGCAGAGGCCGAGGCCCTCGGCCCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGG
CCTAGGGACGTACCCAATTCGCCCTATAGTGAGTCGATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGA
CTGGGAAAACCCTGGCGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAG
AGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATT
AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTT
CTTCCCTTCTTTCTCGCCACGTTCCCGGCTTTCCCGTCAAGCTCTAATCGGGGGCTCCCTTTAGGGTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAAATTGATTAGGGTGTGTTTACGTAAGTGGGCCATCGCCCTGATAG
ACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCAAAATGGAACAACACTCAA
CCCTATCTCGGTCTATTCTTTGATTTATAAGGGATTTGCCGATTTGGCCTATTGGTTAAAAAATGAGCTGATTT
```

AACAAAAATTTAACGCGAATTTTAAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCG  
GAACCCCTATTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTT  
CAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCCTTTTTTTCGGGCATTTTG  
CCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT  
ACATCGAAGTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCGAAGAACGTTTTCCAATGATGAGCACT  
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCCTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTA  
TTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT  
GCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTA  
ACCGTTTTTTTTGCACAACATGGGGGATCATGTAACCTGCCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACC  
AAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTTGCAGAACTATTAACCTGGCGAATCTTAA  
CTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTT  
CCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC  
AGATGGTAAGCCCTCCCCTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA  
TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT  
TTAAAACCTTATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTAACG  
TGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCG  
TAATCTGCTGCTTCAAACAACAAAAACCACCGCTACCAGCGGTGGTTTTGTTTCCCGGATCAAGAGCTACCAACTCTT  
TTTCCGAAGGTAACCTGGCTTTCAGCAGAGCGCAGATACCAAACTACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCA  
CTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA  
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGGTTCCG  
TGACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAC  
GCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGGAGCTTC  
CAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGTTTTCCGCACCTCTGACTTGAGCGTCGATTTTTGTGATGC  
TCGTGAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGTGCGCCTTT  
TGCTCACATGTTCTTTCTGCGTTATCCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGTGATACCG  
CTCGCCGACCGAAGCGACGCGCAGCGAGTCAGTGCAGGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCCCT  
CTCCCCGCGCTTGGCCGATTCTTAATGCAGCTGGCACGACAGTTTTCCCGACTGGAAGCGGGCAGTGAGCGCAA  
CGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG  
GAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAATTAACCCCTCAC  
TAAAGGGAACAAAAGCTGGAGCTGCAAGCTTAATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGAAAGTAAGGTGGTACGATCG  
TGCCCTTATTAGGAAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGCCGCATTGCAGAGATATTG  
TATTTAAGTGCCTAGCTCGATACATAAACGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAA  
CTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGGAGTCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGA  
CTCTGGTAACCTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCGAACAGGGGACTT  
GAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCGAGG  
GGCGGCGACTGGTGAGTACGCCAAAAATTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCACT  
ATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTTAA  
ACATATAGTATGGGCAAGCAGGGAGCTAGAACCATTGCGAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTA  
GACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAACAGTAGCA  
ACCCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAAGACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAA  
CAAAAGTAAGACCACCGCACAGCAAGCGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTTGGAGA  
AGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAGAGAAAGAGTGGT  
GCAGAGAGAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCG  
CAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGCTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGG  
GCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGT  
GGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTGTCTGGAATACTATTTGCACCACTGCTGTGC  
CTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAA  
ATTAACAATTACACAAGCTTAATACACTCCTTAATTTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATT  
ATTGGAATTAGATAAATGGGCAAGTTTTGTGGAATTTGGTTAACATAACAAATTTGGCTGTGGTATATAAATTTATTCA  
TAATGATAGTAGGAGGCTTGGTAGGTTTTAAGAATAGTTTTTGTGCTGACTTTCTATAGTGAATAGAGTTAGGCAGGGA  
TATTCACCATTATCGTTTTAGACCCACCTCCCAACCCCGAGGGGACCCAGAGAGGGCCTATTTCCCATGATTCCTTC  
ATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC  
AAAATACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCAT  
ATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGAGACGG  
TTGTAATGAGCACACAAAATACACATGCTAAAATATTATATTCTATGACCTTTATAAAATCAACCAAAATCTTCTT  
TTAATAACTTTAGTATCAATAATTAGAATTTTTATGTTCTTTTTGCAAACTTTTAATAAAAATGAGCAAAATAAA

AAAACGCTAGTTTTAGTAACTCGCGTTGTTTTCTTCACCTTTAATAATAGCTACTCCACCACTTGTTCCCTAAGCGGT  
CAGCTCCTGCTTCAATCATTTTTGAGCATCTTCAAATGTTCTAACTCCACCAGCTGCTTTAACTAAAGCATTGTCT  
TTAACAACGACTTCATTAGTTAACATCTTCAAATGTTGCACCTGATTTTGAAAATCCTGTTGATGTTTTAACAAA  
TTCTAATCCAGCTTCAACAGCTATTTACAAGCTTTCATGATTTCTTCTTTTGTTAATAAACAATTTTCCATAATAC  
ATTTAACAAACATGTGATCCAGCTGCTTTTTTACAGCTTTCATGTCTTCTAAAACCTAATTCATAATTTTTGTCTTT  
AATGCACCAATATTTAATACCATATCAATTTCTGTTGCACCATCTTTAATTGCTTCAGAACTTCGAATGCTTTTGT  
AGCTGTTGTGCATGCACCTAGAGGAAAACCTACAACATTTGTTATTCCTACATTTGTGCCTTTTAATAATTCTTTAC  
AATAGCTTGTTCATATGAATTAACACAAACTGTTGCAAAATCAAATTCATTGCTTCATCACATAATTGTTAATT  
TCAGTTTTCGTAGCATCTTGTAAAAATAATGTGTGATCTATATATTTGTTAGTTTCATTTTTCTCTATATATTC  
ATTTTTAATTTAATTCTTTAATAATTTCTGCTACTTTAACTTTAGCGTTTTGAACAGATTCACCAACACCTATAAA  
ATAAATTTTTAGTTTAGGTTTCAGTTCCACTTGGGCGAACAGCAAATCATGACTTATCTTCTAAATAAAATTTTAGTA  
AGTCTTGCTCCTGGCATATTATACATTCATCGATGTAGTCTTCAACATTAACAACCTTTAAGTCCAGCAATTTGAGTT  
AAGGGTGTGCTCTCAATGATTTCAATTAATGGTTCAATTTTTAATTTCTTTCTTCTGTTTTAAAATTCAAGTTTAA  
AGTGAAAGTGTAATATGCACCCATTTCTTTAAATAAATCTTCTAAATAGTCTACTAATGTTTTATTTGTTTTTAT  
AAAATCAAGCAGCCTCTGCTATTAATATAGAAGCTTGTATTCCATCTTTATCTCTAGCTGAGTCATCAATTACATAT  
CCATAACTTTCTTCAAGCAAAAACAAAATTTAATCCGTTATCTTCTTTAGCAATTTCTTACCCATTCATTT  
AAATCCAGTTAAAGTTTTTACAATATTAACCTCCATATTTTTCATGAGCGATTCTATCACCCAAATCACTTGTTACAA  
AAGTTGAATATAGAGCCGGATTTTTTGAATGCTATTTAAGCGTTTTAGATTTGATAATTTTCAATCAATTTAAAT  
GGTCTGTTTGATTTCCATCTAATCTTACAAAATGACCATCATGTTTTATTGCCATTCAAATCTGTCAGCATCTGG  
GTCATTCATAATAATAATATCTGCATCATGTTAATACCATATTCAGCGGTATTTTTCATGCAGGATCAAATTCCTG  
GATTTGGATTTACAACATTTTTAAATGTTTCATCTTCAAATGCATGCTTTC AACCTCAATAACGTTATATCCTGAT  
TCACGTAATATTTTTGGGGTAAATTTAGTTCCTGTTCCATTAAGTGCAGCTAAAATAATTTTTAAATCTTTTTTAGC  
TTCTTGCTCTTTTTTGTA **CGTCTC** **GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTT**  
**GAAAAAGTGGCACCGAGTCGGTGC**TTTTTTAAGCTTGGCGTAACTAGATCTTGAGACAAATGGCAGTATTCATCCAC  
AATTTTTAAAAGAAAAGGGGGGATTGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACA  
AACTAAAGAATTACAAAACAAATTACAAAATTTCAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTT  
GGCGCCGGCTCGAGGGG

## References

1. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
2. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
3. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562–578 (2012).
4. Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**, 1593–1599 (2012).
5. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
6. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832 (2013).
7. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**, D68–73 (2014).
8. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).