Multiplex Genome Engineering Using CRISPR/Cas Systems

Le Cong\textsuperscript{1,2,*}, F. Ann Ran\textsuperscript{1,4,*}, David Cox\textsuperscript{1,3}, Shuailiang Lin\textsuperscript{1,5}, Robert Barretto\textsuperscript{6}, Naomi Habib\textsuperscript{1}, Patrick D. Hsu\textsuperscript{1,4}, Xuebing Wu\textsuperscript{7}, Wenyan Jiang\textsuperscript{8}, Luciano Marraffini\textsuperscript{8}, and Feng Zhang\textsuperscript{1,†}

\textsuperscript{1} Broad Institute of MIT and Harvard
    7 Cambridge Center
    Cambridge, MA 02142, USA

McGovern Institute for Brain Research
Department of Brain and Cognitive Sciences
Department of Biological Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

\textsuperscript{2} Program in Biological and Biomedical Sciences
Harvard Medical School
Boston, MA 02115, USA

\textsuperscript{3} Harvard-MIT Health Sciences and Technology
Harvard Medical School
Boston, MA 02115

\textsuperscript{4} Department of Molecular and Cellular Biology
Harvard University
Cambridge, MA 02138, USA

\textsuperscript{5} School of Life Sciences, Tsinghua University
Beijing, 100084, China

\textsuperscript{6} Department of Biochemistry and Molecular Biophysics
College of Physicians and Surgeons, Columbia University
New York, NY 10032, USA

\textsuperscript{7} Computational and Systems Biology Graduate Program, and
Koch Institute for Integrative Cancer Research
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

\textsuperscript{8} Laboratory of Bacteriology
The Rockefeller University
1230 York Ave.
New York, NY 10065, USA

*These authors contributed equally to this work.
†To whom correspondence should be addressed: zhang@broadinstitute.org.
SUPPLEMENTARY MATERIALS AND METHODS

Cell culture and transfection

Human embryonic kidney (HEK) cell line 293FT (Life Technologies) was maintained in Dulbecco’s modified Eagle’s Medium (DMEM) supplemented with 10% fetal bovine serum (HyClone), 2mM GlutaMAX (Life Technologies), 100U/mL penicillin, and 100µg/mL streptomycin at 37°C with 5% CO₂ incubation. Mouse neuro2A (N2A) cell line (ATCC) was maintained with DMEM supplemented with 5% fetal bovine serum (HyClone), 2mM GlutaMAX (Life Technologies), 100U/mL penicillin, and 100µg/mL streptomycin at 37°C with 5% CO₂.

293FT or N2A cells were seeded into 24-well plates (Corning) one day prior to transfection at a density of 200,000 cells per well. Cells were transfected using Lipofectamine 2000 (Life Technologies) following the manufacturer’s recommended protocol. For each well of a 24-well plate a total of 800ng plasmids was used.

Suveryor assay and sequencing analysis for genome modification

293FT or N2A cells were transfected with plasmid DNA as described above. Cells were incubated at 37°C for 72 hours post transfection before genomic DNA extraction. Genomic DNA was extracted using the QuickExtract DNA extraction kit (Epicentre) following the manufacturer’s protocol. Briefly, cells were resuspended in QuickExtract solution and incubated at 65°C for 15 minutes and 98°C for 10 minutes.

Genomic region surrounding the CRISPR target site for each gene was PCR amplified, and products were purified using QiaQuick Spin Column (Qiagen) following manufacturer’s protocol. A total of 400ng of the purified PCR products were mixed with 2µl 10X Taq polymerase PCR buffer (Enzymatics) and ultrapure water to a final volume of 20µl, and subjected to a re-annealing process to enable heteroduplex formation: 95°C for 10min, 95°C to 85°C ramping at – 2°C/s, 85°C to 25°C at – 0.25°C/s, and 25°C hold for 1 minute. After re-annealing, products were treated with SURVEYOR nuclease and SURVEYOR enhancer S (Transgenomics) following the manufacturer’s recommended protocol, and analyzed on 4-20% Novex TBE poly-acrylamide gels (Life Technologies). Gels were stained with SYBR Gold DNA
stain (Life Technologies) for 30 minutes and imaged with a Gel Doc gel imaging system (Bio-rad). Quantification was based on relative band intensities.

Restriction fragment length polymorphism assay for detection of homologous recombination
HEK 293FT and N2A cells were transfected with plasmid DNA, and incubated at 37°C for 72 hours before genomic DNA extraction as described above. The target genomic region was PCR amplified using primers outside the homology arms of the homologous recombination (HR) template. PCR products were separated on a 1% agarose gel and extracted with MinElute GelExtraction Kit (Qiagen). Purified products were digested with HindIII (Fermentas) and analyzed on a 6% Novex TBE poly-acrylamide gel (Life Technologies).

RNA extraction and purification
HEK 293FT cells were maintained and transfected as stated previously. Cells were harvested by trypsinization followed by washing in phosphate buffered saline (PBS). Total cell RNA was extracted with TRI reagent (Sigma) following manufacturer’s protocol. Extracted total RNA was quantified using Naonodrop (Thermo Scientific) and normalized to same concentration.

Northern blot analysis of crRNA and tracrRNA expression in mammalian cells
RNAs were mixed with equal volumes of 2X loading buffer (Ambion), heated to 95°C for 5 min, chilled on ice for 1 min and then loaded onto 8% denaturing polyacrylamide gels (SequaGel, National Diagnostics) after pre-running the gel for at least 30 minutes. The samples were electrophoresed for 1.5 hours at 40W limit. Afterwards, the RNA was transferred to Hybond N+ membrane (GE Healthcare) at 300 mA in a semi-dry transfer apparatus (Bio-rad) at room temperature for 1.5 hours. The RNA was crosslinked to the membrane using autocrosslink button on Stratagene UV Crosslinker the Stratalinker (Stratagene). The membrane was pre-hybridized in ULTRAhyb-Oligo Hybridization Buffer (Ambion) for 30 min with rotation at 42°C and then probes were added and hybridized overnight. Probes were ordered from IDT and labeled with [gamma-32P] ATP (Perkin Elmer) with T4 polynucleotide kinase (New England Biolabs). The membrane was washed once with pre-warmed (42°C) 2xSSC, 0.5% SDS for 1 min followed by two 30 minute washes at 42°C. The membrane was exposed to phosphor screen for one hour or overnight at room temperature and then scanned with phosphorimager (Typhoon).
**SUPPLEMENTARY FIGURES**

**Figure S1**

*Streptococcus pyogenes* SF370 type II CRISPR locus

**Fig. S1.** Schematic of the type II CRISPR-mediated DNA double-strand break. The type II CRISPR locus from *Streptococcus pyogenes* SF370 contains a cluster of four genes, *Cas9*, *Cas1*, *Cas2*, and *Csn1*, as well as two non-coding RNA elements, tracrRNA and a characteristic array of repetitive sequences (direct repeats) interspaced by short stretches of non-repetitive sequences (spacers, 30bp each) \(15-18, 30, 31\). Each spacer is typically derived from foreign genetic material (protospacer), and directs the specificity of CRISPR-mediated nucleic acid cleavage. In the target nucleic acid, each protospacer is associated with a protospacer adjacent motif (PAM) whose recognition is specific to individual CRISPR systems \(22, 23\). The Type II CRISPR system carries out targeted DNA double-strand break (DSB) in sequential steps \(12-14, 20, 21\). First, the pre-crRNA array and tracrRNA are transcribed from the CRISPR locus. Second, tracrRNA hybridizes to the direct repeats of pre-crRNA and associates with Cas9 as a duplex, which mediates the processing of the pre-crRNA into mature crRNAs containing individual, truncated spacer sequences. Third, the mature crRNA:tracrRNA duplex directs Cas9 to the DNA target consisting of the protospacer and the requisite PAM via heteroduplex formation between the spacer region of the crRNA and the protospacer DNA. Finally, Cas9 mediates cleavage of target DNA upstream of PAM to create a DSB within the protospacer.
Figure S2

(A) Schematic showing the design and sequences of two tracrRNA transcripts tested (short and long). Each transcript is driven by a U6 promoter. Transcription start site is marked as +1 and transcription terminator is as indicated. Blue line indicates the region whose reverse-complement sequence is used to generate northern blot probes for tracrRNA detection.

(B) SURVEYOR assay comparing the efficiency of SpCas9-mediated cleavage of the EMX1 locus. Two biological replicas are shown for each tracrRNA transcript.

(C) Northern blot analysis of total RNA extracted from 293FT cells transfected with U6 expression constructs carrying long or short tracrRNA, as well as SpCas9 and DR-EMX1(1)-DR. Left and right panels are from 293FT cells transfected without or with SpRNase III respectively. U6 indicate loading control blotted with a probe targeting human U6 snRNA. Transfection of the short tracrRNA expression construct led to abundant levels of the processed form of tracrRNA (~75bp) (19). Very low amounts of long tracrRNA are detected on the northern blot. As a result of these experiments, we chose to use short tracrRNA for application in mammalian cells.

**Fig S2. Comparison of different tracrRNA transcripts for Cas9-mediated gene targeting.**

(A) Schematic showing the design and sequences of two tracrRNA transcripts tested (short and long). Each transcript is driven by a U6 promoter. Transcription start site is marked as +1 and transcription terminator is as indicated. Blue line indicates the region whose reverse-complement sequence is used to generate northern blot probes for tracrRNA detection. (B) SURVEYOR assay comparing the efficiency of SpCas9-mediated cleavage of the EMX1 locus. Two biological replicas are shown for each tracrRNA transcript. (C) Northern blot analysis of total RNA extracted from 293FT cells transfected with U6 expression constructs carrying long or short tracrRNA, as well as SpCas9 and DR-EMX1(1)-DR. Left and right panels are from 293FT cells transfected without or with SpRNase III respectively. U6 indicate loading control blotted with a probe targeting human U6 snRNA. Transfection of the short tracrRNA expression construct led to abundant levels of the processed form of tracrRNA (~75bp) (19). Very low amounts of long tracrRNA are detected on the northern blot. As a result of these experiments, we chose to use short tracrRNA for application in mammalian cells.
Fig. S3. SURVEYOR assay for detection of double strand break-induced micro insertions and deletions (32). Schematic of the SURVEYOR assay used to determine Cas9-mediated cleavage efficiency. First, genomic PCR (gPCR) is used to amplify the Cas9 target region from a heterogeneous population of modified and unmodified cells, and the gPCR products are reannealed slowly to generate heteroduplexes. The reannealed heteroduplexes are cleaved by SURVEYOR nuclease, whereas homoduplexes are left intact. Cas9-mediated cleavage efficiency (% indel) is calculated based on the fraction of cleaved DNA.
Fig. S4. Northern blot analysis of crRNA processing in mammalian cells. (A) Schematic showing the expression vector for a single spacer flanked by two direct repeats (DR-EMX1(1)-DR). The 30bp spacer targeting the human EMX1 locus protospacer 1 (Table S1) is shown in blue and direct repeats are shown in gray. Orange line indicates the region whose reverse-complement sequence is used to generate northern blot probes for EMX1(1) crRNA detection. (B) Northern blot analysis of total RNA extracted from 293FT cells transfected with U6 expression constructs carrying DR-EMX1(1)-DR. Left and right panels are from 293FT cells transfected without or with SpRNase III respectively. DR-EMX1(1)-DR was processed into...
mature crRNAs only in the presence of SpCas9 and short tracrRNA, and was not dependent on the presence of SpRNase III. The mature crRNA detected from transfected 293FT total RNA is ~33bp and is shorter than the 39-42bp mature crRNA from *S. pyogenes* (19), suggesting that the processed mature crRNA in human 293FT cells is likely different from the bacterial mature crRNA in *S. pyogenes*. 
Fig. S5. Bicistronic expression vectors for pre-crRNA array or chimeric crRNA with Cas9.

(A) Schematic showing the design of an expression vector for the pre-crRNA array. Spacers can be inserted between two BbsI sites using annealed oligonucleotides. Sequence design for the oligonucleotides are shown below with the appropriate ligation adapters indicated. 

(B) Schematic of the expression vector for chimeric crRNA. The guide sequence can be inserted between two BbsI sites using annealed oligonucleotides. The vector already contains the partial direct repeat (gray) and partial tracrRNA (red) sequences. WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.
Fig. S6. Selection of protospacers in the human *PVALB* and mouse *Th* loci. Schematic of the human *PVALB* (A) and mouse *Th* (B) loci and the location of the three protospacers within the last exon of the *PVALB* and *Th* genes, respectively. The 30bp protospacers are indicated by black lines and the adjacent PAM sequences are indicated by the magenta bar. Protospacers on the sense and anti-sense strands are indicated above and below the DNA sequences respectively.
**Fig. S7. Occurrences of PAM sequences in the human genome.** Histograms of distances between adjacent *Streptococcus pyogenes* SF370 type II CRISPR PAM (NGG) (A) and *Streptococcus thermophilus* LMD-9 CRISPR1 PAM (NNAGAAW) (B) in the human genome. (C) Distances for each PAM by chromosome. Chr, chromosome. Putative targets were identified using both the plus and minus strands of human chromosomal sequences. Given that there may be chromatin, DNA methylation-, RNA structure, and other factors that may limit the cleavage activity at some protospacer targets, it is important to note that the actual targeting ability might be less than the result of this computational analysis.
Figure S8

(A) Schematic of CRISPR locus 2 from *Streptococcus thermophilus* LMD-9.

(B) Design of the expression system for the *S. thermophilus* CRISPR system. Human codon-optimized *StCas9* is expressed using a constitutive EF1α promoter. Mature versions of tracrRNA and crRNA are expressed using the U6 promoter to ensure precise transcription initiation. Sequences for the mature crRNA and tracrRNA are shown. A single base indicated by the lower case “a” in the crRNA sequence was used to remove the polyU sequence, which serves as a RNA Pol III transcriptional terminator. Sp, spacer.

(C) Schematic showing protospacer and direct repeats.

(D) Table showing indel (%).

Fig S8. Type II CRISPR from *Streptococcus thermophilus* LMD-9 can also function in eukaryotic cells. (A) Schematic of CRISPR locus 2 from *Streptococcus thermophilus* LMD-9. (B) Design of the expression system for the *S. thermophilus* CRISPR system. Human codon-optimized *StCas9* is expressed using a constitutive EF1α promoter. Mature versions of tracrRNA and crRNA are expressed using the U6 promoter to ensure precise transcription initiation. Sequences for the mature crRNA and tracrRNA are shown. A single base indicated by the lower case “a” in the crRNA sequence was used to remove the polyU sequence, which serves as a RNA Pol III transcriptional terminator. Sp, spacer. (C) Schematic showing protospacer and direct repeats.
corresponding PAM sequences targets in the human *EMX1* locus. Two protospacer sequences are highlighted and their corresponding PAM sequences satisfying the NNAGAAW motif are indicated by magenta lines. Both protospacers are targeting the anti-sense strand. (D) SURVEYOR assay showing StCas9-mediated cleavage in the target locus. RNA guide spacers 1 and 2 induced 14% and 6.4% respectively. Statistical analysis of cleavage activity across biological replica at these two protospacer sites can be found in Table S1.
Table S1. Protospacer sequences and modification efficiencies of mammalian genomic targets. Protospacer targets designed based on *Streptococcus pyogenes* type II CRISPR and *Streptococcus thermophilus* CRISPR1 loci with their requisite PAMs against three different genes in human and mouse genomes. Cells were transfected with Cas9 and either pre-crRNA/tracrRNA or chimeric RNA. Cells were analyzed 72 hours after transfection. Percent indels are calculated based on SURVEYOR assay results from indicated cell lines, N = 3 for all protospacer targets, errors are S.E.M. N.D., not detectable using the SURVEYOR assay; N.T., not tested in this study.

<table>
<thead>
<tr>
<th>Cas9</th>
<th>target species</th>
<th>gene</th>
<th>protospacer ID</th>
<th>protospacer sequence (5' to 3')</th>
<th>PAM strand</th>
<th>cell line tested</th>
<th>% indel (pre-crRNA + tracrRNA)</th>
<th>% indel (chimeric RNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. pyogenes</td>
<td>Homo sapiens</td>
<td>EMX1</td>
<td>1</td>
<td>GGAAAGGCCCCTGACCGAGAAGAAGACAA</td>
<td>TGG</td>
<td>293FT</td>
<td>20 ± 1.8</td>
<td>6.7 ± 0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>2</td>
<td>CATTGGAGGTGGCCTGCTGCTCCCTCCCAT</td>
<td>TGG</td>
<td>293FT</td>
<td>2 ± 0.31</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>3</td>
<td>GGACGTAGCTGCTCCTCCAATGACTAGG</td>
<td>TGG</td>
<td>293FT</td>
<td>14 ± 1.1</td>
<td>2.1 ± 0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>4</td>
<td>CATCGATGCTCCCTTACTGGCCTGTCGG</td>
<td>TGG</td>
<td>293FT</td>
<td>11 ± 1.7</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>5</td>
<td>TCTGTCGCCACCTGCCACCGGTGAGTCTA</td>
<td>TGG</td>
<td>293FT</td>
<td>4.3 ± 0.46</td>
<td>2.1 ± 0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>6</td>
<td>TCGTGGCACATGCCACCGGTGTCGATGAT</td>
<td>TGG</td>
<td>293FT</td>
<td>4.0 ± 0.66</td>
<td>0.41 ± 0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>7</td>
<td>TCAAGCTTTGCGTTTGTACCTTGTCCTTC</td>
<td>TGG</td>
<td>293FT</td>
<td>1.5 ± 0.12</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>8</td>
<td>GGAGGAGGGCCAGATAGAGAAGCCTAGG</td>
<td>AGG</td>
<td>293FT</td>
<td>7.8 ± 0.83</td>
<td>2.3 ± 1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td>PVALB</td>
<td>9</td>
<td>AAGGGCCGAGATTGGGTTCAGGGCAGAG</td>
<td>AGG</td>
<td>293FT</td>
<td>21 ± 2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td>PVALB</td>
<td>10</td>
<td>ATGCAGAGGGTGCCAGAGGGCCGAGAT</td>
<td>TGG</td>
<td>293FT</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td>PVALB</td>
<td>11</td>
<td>GGGGCGAGGGGCGAGGATGCTGGTCCTCC</td>
<td>AGG</td>
<td>293FT</td>
<td>N.D.</td>
</tr>
<tr>
<td>S. thermophilus</td>
<td>Mus musculus</td>
<td>EMX1</td>
<td>12</td>
<td>CAGGCATGAGTGGCCATTAGCTAAAACTGAT</td>
<td>AGG</td>
<td>Neuro2A</td>
<td>27 ± 4.3</td>
<td>4.1 ± 2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>13</td>
<td>ATGCATAGGAGTACACCCACAGGGTCCAG</td>
<td>CAG</td>
<td>Neuro2A</td>
<td>4.8 ± 1.2</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>14</td>
<td>ACAACATGAGGAGGCGAGCCAGCCAGAA</td>
<td>AGG</td>
<td>Neuro2A</td>
<td>11.3 ± 1.3</td>
<td>N.D.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td>EMX1</td>
<td>15</td>
<td>GGAGGAGGGTAGTATACAGAAACACAAGA</td>
<td>GTAAGAT</td>
<td>293FT</td>
<td>14 ± 0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMX1</td>
<td>16</td>
<td>AGAATGAGGAGTACAGAAACACAGCA</td>
<td>GATAGA</td>
<td>293FT</td>
<td>7.8 ± 0.77</td>
<td>N.T.</td>
</tr>
</tbody>
</table>

*Note: Percent indels are calculated based on SURVEYOR assay results.*
Table S2. Sequences for primers and probes used for SURVEYOR assay, RFLP assay, genomic sequencing, and Northern blot.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Assay</th>
<th>Genomic Target</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp-EMX1-F</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>EMX1</em></td>
<td>AAAACCACCTTCTCTCTGGC</td>
</tr>
<tr>
<td>Sp-EMX1-R</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>EMX1</em></td>
<td>GGAGATTGGAGACACGGAGAG</td>
</tr>
<tr>
<td>Sp-PVALB-F</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>PVALB</em></td>
<td>CTGGAAGCCAATGCCCTGAC</td>
</tr>
<tr>
<td>Sp-PVALB-R</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>PVALB</em></td>
<td>GGCAGCAAACTCCTTGCTCT</td>
</tr>
<tr>
<td>Sp-Th-F</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>Th</em></td>
<td>GTGCTTTGCAGAGGCCTACC</td>
</tr>
<tr>
<td>Sp-Th-R</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>Th</em></td>
<td>CCTGGAGCGCATGCAGTAGT</td>
</tr>
<tr>
<td>St-EMX1-F</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>EMX1</em></td>
<td>ACCTTCTGTGTTTCCACCATTC</td>
</tr>
<tr>
<td>St-EMX1-R</td>
<td>SURVEYOR assay, sequencing</td>
<td><em>EMX1</em></td>
<td>TTGGGGAGTGCACAGACTTC</td>
</tr>
<tr>
<td>Sp-EMX1-RFLP-F</td>
<td>RFLP, sequencing</td>
<td><em>EMX1</em></td>
<td>GGCTCCCTGGTTCAAAGTA</td>
</tr>
<tr>
<td>Sp-EMX1-RFLP-R</td>
<td>RFLP, sequencing</td>
<td><em>EMX1</em></td>
<td>AGAGGGGTCTGGATGTCGTA</td>
</tr>
<tr>
<td>Pb_EMX1_sp1</td>
<td>Northern Blot Probe</td>
<td>Not applicable</td>
<td>TAGCTCTAAAAACTTTTCTTCTGCTCGGAC</td>
</tr>
<tr>
<td>Pb_tracrRNA</td>
<td>Northern Blot Probe</td>
<td>Not applicable</td>
<td>CTAGCCTTTTTAAGCTTGTGCTGCTGTT</td>
</tr>
</tbody>
</table>
SUPPLEMENTARY SEQUENCES

> U6-short tracrRNA (Streptococcus pyogenes SF370)
GAGGGCCTATTTCCCATGATATCCCTTATATTTGCAATACGATACAAAGGCTGTTAGAGATAA
TTGGAATTAATTTGAGCTGAAACACAAAGATATTAGTACAAAAATACGTGACGAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> U6-long tracrRNA (Streptococcus pyogenes SF370)
GAGGGCCTATTTCCCATGATATCCCTTATATTTGCAATACGATACAAAGGCTGTTAGAGATAA
TTGGAATTAATTTGAGCTGAAACACAAAGATATTAGTACAAAATACGTGACGTAAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> U6-DR-BbsI backbone-DR (Streptococcus pyogenes SF370)
GAGGGCCTATTTCCCATGATATCCCTTATATTTGCAATACGATACAAAGGCTGTTAGAGATAA
TTGGAATTAATTTGAGCTGAAACACAAAGATATTAGTACAAAATACGTGACGTAAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> U6-chimeric RNA-BbsI backbone (Streptococcus pyogenes SF370)
GAGGGCCTATTTCCCATGATATCCCTTATATTTGCAATACGATACAAAGGCTGTTAGAGATAA
TTGGAATTAATTTGAGCTGAAACACAAAGATATTAGTACAAAATACGTGACGTAAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> 3xFLAG-NLS-SpCas9-NLS
ATGGACATTAAGGACCAGCAGGAGACGACTACAAGGATCATGATATTGATTACAAGAGCATGAGC
ATACAAGATGCCCAAAGCGAAGCAGAGATATAGCTACAAAAATACGTGACGAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> 3xFLAG-NLS-SpCas9-NGS
ATGGACATTAAGGACCAGCAGGAGACGACTACAAGGATCATGATATTGATTACAAGAGCATGAGC
ATACAAGATGCCCAAAGCGAAGCAGAGATATAGCTACAAAAATACGTGACGAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT

> 3xFLAG-NLS-SpCas9-NGS
ATGGACATTAAGGACCAGCAGGAGACGACTACAAGGATCATGATATTGATTACAAGAGCATGAGC
ATACAAGATGCCCAAAGCGAAGCAGAGATATAGCTACAAAAATACGTGACGAGAAATTAATA
ATTCTTGGGTAGTTGCGAGTTTTAAAAATATGCTTTTTAAAAATGCATACATATCTCCGAGTA
ACTTTGAAGGTATTTGACATTTTGGCTTTTATATATCTCTGAGAAGACGAAACACCCGAACCA
TTCAAAAAACAGCATAAGCAAGTTAAATAAAGGCTAGTGCTTTATCAAATGAAAAGTTGGCACCAG
GTTCGGTCTTTTTTT
S\textsuperscript{pRNase3-mCherry-NLS}

ATGAAGCAGCTGGAGGAGTTACTTTCTACCTCTTTCGACATCCAGTTTAATGACCTGACCCCCTGC
TGAAACCCGCCTTTCACACTCTTACCGGAAATGACCCAGCCCTACTGAATGTGACCAAAGCAACT
TGCACTGCTTGCAGCTACTCTAAGAAAGCTGGT

> 3xFLAG-NLS-\textsuperscript{SpCas9n-NLS} (the D10A nickase mutation is labeled in red)

ATGGACTATAAGGACCACGACGGAAGATCACAAGGATCATGATATTGATTACAAAGACGATGACG
ATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAA
GTACAGCATCGCGCTGCCATCGCCACAAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTAC
AAGGTGCCACGCAAGAAGATTCAAGGATCTGGGCTGGGCCACCGACCGTAAGAAGAAACC
TGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGC
CAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATA
AGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTA
CCCCACCATCACAGCGCTTGGCAAGAAGAGAATGGCCTGTTCGACCGTGACGGGCTTGAGCAGG

ATGGAAGCTGAGGGCTTCCCGGAGGATAATGAGCAGAAACACAGCTGTGTTTGTGGAACAGCACA
AGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGAG
CGCTAATCTGGACAAAGTGCTGTCCGC

CTACAACAAGCACGGGATAAGCCCATCAGAGAGCAG
GCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACT
TTGACACCACCACCAGGCTAAGAGGTAGTACCTGCTAGTCTTCCAGGAGGCTGGGAGGCGACAAG
CGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA

3xFLAG-NLS-\textsuperscript{SpCas9n-NLS} (the D10A nickase mutation is labeled in red)

ATGGACTATAAGGACCACGACGGAAGATCACAAGGATCATGATATTGATTACAAAGACGATGACG
ATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAA
GTACAGCATCGCGCTGCCATCGCCACAAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTAC
AAGGTGCCACGCAAGAAGATTCAAGGATCTGGGCTGGGCCACCGACCGTAAGAAGAAACC
TGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGC
CAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATA
AGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTA
CCCCACCATCACAGCGCTTGGCAAGAAGAGAATGGCCTGTTCGACCGTGACGGGCTTGAGCAGG

ATGGAAGCTGAGGGCTTCCCGGAGGATAATGAGCAGAAACACAGCTGTGTTTGTGGAACAGCACA
AGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGAG
CGCTAATCTGGACAAAGTGCTGTCCGC

CTACAACAAGCACGGGATAAGCCCATCAGAGAGCAG
GCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACT
TTGACACCACCACCAGGCTAAGAGGTAGTACCTGCTAGTCTTCCAGGAGGCTGGGAGGCGACAAG
CGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA

S\textsuperscript{pRNase3-mCherry-NLS}

ATGAAGCAGCTGGAGGAGTTACTTTCTACCTCTTTCGACATCCAGTTTAATGACCTGACCCCCTGC
TGAAACCCGCCTTTCACACTCTTACCGGAAATGACCCAGCCCTACTGAATGTGACCAAAGCAACT
TGCACTGCTTGCAGCTACTCTAAGAAAGCTGGT
AGCACTACCTGGACGAGATCATCGACGAGATCAAGGAGGAGTGATCCTGGCCGA
CGTAATCTGGACAAAGTGCTGTCGCCCTACAACAAGACCCGGGATAAGCCCACTAGAGACGAG
GCCGAATATCATCCACCTTGTGATACTTGGAGACCCTTGTGGGCCCTTCAAGTACT
TTGACACCAAATCGACCGGAGAAGTGACACACCACAAAGAGGTTGCTGGAGCGCCCACCCTTAT
CCACCGAGAGCAGTCAGGAGCTCAGGCTCTTCGACTGCTGCGGAGGCACAAG
CTGCTCTGCTGACTAAAGATAGCTGCTGCTTCCAAGGAAAGTCTGGTCAAGCTAAGAAAA

> hEMX1-HRTemplate-HindIII-NheI
GAATGCTGCCCTAGACCAGCGCCCTTTGCTTCTCTTGTGCTTCAAGGAGGAATGAGGTCTC
AGAAGCTGAACTCTGCTAACTGCTGACATTCACATTACACGACCAGGACCAAGGCCAA
CTGACTCCACCTTGTGCTTCTGTGGCTGAGAAGAGCTGAGTGAAGATCGATCGAGA
TGAACTGCTTTGTTAGAGGAAACGAAACCTCCCTTCTCTTCTCTTCTCTTCTCTCT
TGCAACCCCTTCTGCTTGAGAAGGGAACGAAAGAAACCTCCCTTCTCTTCTCTTCTCT
GCCCTGCACTCCTCTGCCAGGCTTGAGAAGGGAAGGACTGCTGTCGAGGAAGGAGGA
CAGGGCCATAGGGGAGACACATCGAGTTACACCTACCTAGTGAACAACTTGGTGCGGAGAACG
AATAGGCAGGGCAGGCTGCTCTTGGCTGAGACAGAGACAGGGCTTAAGGCTGAGCCTGCA
ACCAGTCCCCAGTGACTCAGGGCCTCCTCAGCCAAAGAAAGACGAACTGGCAGG

> NLS-StCsn1-NLS
ATGAAAAAGGCCGGCCCGACGAGAAGGAAAGGCCGGCCGAGCAGAAAGAAGAGGTTCCTGACTAG
TTGAGCTGAACTCTGCTGACATTCACATTACACGACCAGGACCAAGGCCAA
CTGACTCCACCTTGTGCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCT
GCCCTGCACTCCTCTGCCAGGCTTGAGAAGGGAAGGACTGCTGTCGAGGAAGGAGGA
CAGGGCCATAGGGGAGACACATCGAGTTACACCTACCTAGTGAACAACTTGGTGCGGAGAACG
AATAGGCAGGGCAGGCTGCTCTTGGCTGAGACAGAGACAGGGCTTAAGGCTGAGCCTGCA
ACCAGTCCCCAGTGACTCAGGGCCTCCTCAGCCAAAGAAAGACGAACTGGCAGG
> U6-St_tracrRNA(7-97)
GAGGGGCTATTTTCCATGATTTTCTTCATATTTGCTATATACGATACAAGGCTGTCTTGAGAGATAA
TTGGAATTAAATTTGACGTGAAACAAAAGATATTAGTACAAAAATACGTGAGCTAAAAAGTTAA
ATTTCTTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTTAAAAATGGACTATCATATATGCTTACGTA
ACTTGAAAGATTTTCCATTTGCTTTATATATTTGTTGGAAAGGAGCAAACACCGTTTTACTTT
AAATCTTTCGACAAGCTCAAAAGATAAAGGCTGTCTTGCCGAAATCAACACCTGTCATTTTTATGGC
AGGGTTTGTCTTTATTAAAA

> EMX1_TALEN_Left
ATGGCAGTATAAGGACCACGCCAGGAGACTACAAAGGATCATGATATTGATTTCAAAGGACGATGACG
ATAAGATGCTGCCCCAAAAGAAAGAAACGAAAGGCTGGTCTACACGGCTTCTGTTGCCCAGGGAAAG
CTGTTGACGCAGTCGTTCCGCAATTCCGAGAGGAGGGCTTTTTACGAGAGAGGCTTTTTACGAGAGGCT
TTTAAAATTATGTTTTAAAAATGGACTATCATATATGCTTACGTA
ACTTGAAAGATTTTCCATTTGCTTTATATATTTGTTGGAAAGGAGCAAACACCGTTTTACTTT
AAATCTTTCGACAAGCTCAAAAGATAAAGGCTGTCTTGCCGAAATCAACACCTGTCATTTTTATGGC
AGGGTTTGTCTTTATTAAAA

TAAACCCAAACTCGATTTCAAGCGTCTCCTGCTGCTACTAAGAAAGCTGGTCTAGCTAAAGAAAAAG
AAATTTAA
T>
GATGAGATGCAACGTATATGTCGAAAGATACAGACCAGAAACACACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTTAAGTTTTTGTTTGTCTCTGGGCATTTCAAGGGCAACTATAAGGCCCAGCTCACACGGTTGAATCACATTACGAACTGCAATGGTGCGGTTTTGCTCCGTAGAGGAACTGCTCATTGGTGGAGAAATGATCAAAGCGGGAACTCTGACACTGGAAGAAGTCAGACGCAAGT
TTAACAATGGCGAGATCAATTTCCGCTCA

>EMX1_TALEN_Right
ATGGACCTATAAGGCCACGCAGCGAGACTCAAAGGATCATGATATTGATTACAAAGACGATGACGATAAGATGGCTGCGAATAGGATCCTTCATCAGTGACCGAGTTTAAGTTTTTGTTTGTCTCTGGGCATTTCAAGGGCAACTATAAGGCCCAGCTCACACGGTTGAATCACATTACGAACTGCAATGGTGCGGTTTTGCTCCGTAGAGGAACTGCTCATTGGTGGAGAAATGATCAAAGCGGGAACTCTGACACTGGAAGAAGTCAGACGCAAGT

GATGAGATGCAACGTATATGTCGAAAGATACAGACCAGAAACACACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTTAAGTTTTTGTTTGTCTCTGGGCATTTCAAGGGCAACTATAAGGCCCAGCTCACACGGTTGAATCACATTACGAACTGCAATGGTGCGGTTTTGCTCCGTAGAGGAACTGCTCATTGGTGGAGAAATGATCAAAGCGGGAACTCTGACACTGGAAGAAGTCAGACGCAAGT

TTAACAATGGCGAGATCAATTTCCGCTCA
References and Notes


