Supplementary Results

Titration of sgRNA- and Cas9-expressing plasmid amounts used for the EGFP disruption assay

Single guide RNAs (sgRNAs) were generated for three different sequences (Supplementary Fig. 2a) located upstream of EGFP nucleotide 502, a position at which the introduction of frameshift mutations via non-homologous end-joining can robustly disrupt expression of EGFP\textsuperscript{1,2}. For each of the three target sites, we initially transfected a range of sgRNA-expressing plasmid amounts (12.5 to 250 ng) together with 750 ng of a plasmid expressing a codon-optimized version of the Cas9 nuclease into our U2OS.EGFP reporter cells bearing a single copy, constitutively expressed EGFP-PEST reporter gene. All three RGENs efficiently disrupted EGFP expression at the highest concentration of sgRNA plasmid (250 ng) (Supplementary Fig. 2b). However, RGENs for target sites #1 and #3 exhibited equivalent levels of disruption when lower amounts of sgRNA-expressing plasmid were transfected whereas RGEN activity at target site #2 dropped immediately when the amount of sgRNA-expressing plasmid transfected was decreased (Supplementary Fig. 2b). We next titrated the amount of Cas9-encoding plasmid (range from 50 ng to 750 ng) transfected into our U2OS.EGFP reporter cells and assayed for EGFP disruption. As shown in Supplementary Fig. 2c, target site #1 tolerated a three-fold decrease in the amount of Cas9-encoding plasmid transfected without substantial loss of EGFP disruption activity. However, the activities of RGENs targeting target sites #2 and #3 decreased immediately with a three-fold reduction in the amount of Cas9 plasmid transfected (Supplementary Fig. 2c). Based on these results, we used 25ng/250ng, 250ng/750ng, and 200ng/750ng of sgRNA-/Cas9-expressing plasmids for EGFP target sites #1, #2, and #3, respectively, for the experiments described in the main text.
We do not understand the reasons why some sgRNA/Cas9 combinations work better than others in disrupting EGFP expression nor do we know why some of these combinations are more or less sensitive to the amount of plasmids used for transfection. Although it is possible that the range of off-target sites present in the genome for these three sgRNAs might influence each of their activities, we did not observe any differences in the numbers of genomic sites that differ by one to six bps for each of these particular target sites (Supplementary Table 3) that would account for the differential behavior of the three sgRNAs.
**Supplementary Fig. 1** Schematic illustrating a sgRNA/Cas9 nuclease complex bound to its target DNA site

Scissors indicate approximate cleavage points of the Cas9 nuclease on the genomic DNA target site.

Note the numbering of nucleotides on the guide RNA proceeds in an inverse fashion from 3' to 5'.
Supplementary Fig. 2  Titration of sgRNA- and Cas9-expressing plasmid amounts used for the human cell-based EGFP disruption assay

(a) EGFP Site 1 GGGCACGGGCAGCTTGCGATGG
EGFP Site 2 GATGCCGTTCTTCTGCTTGCG
EGFP Site 3 GGTGCTCAGATACTTCAGGG

(b) Percent EGFP negative cells vs. Amount of gRNA plasmid (Cas9 = 750ng)

(c) Percent EGFP negative cells vs. Amount of Cas9 plasmid (gRNA variable)

Supplementary Fig. 2 Titration of sgRNA- and Cas9-expressing plasmid amounts used for the human cell-based EGFP disruption assay
(a) Sequences for three RGEN target sites in *EGFP* used in this study. (b) Activities of RGENs in the EGFP disruption assay performed using varying amounts of sgRNA-expressing plasmids. (c) Activities of RGENs in the EGFP disruption assay performed using varying amounts of Cas9-expression plasmid.
**Target 1 (VEGFA Site 1):**

**OT1-3**  
AGACAGGACATCTCTGACACCCAGGAGAACCCTCCCTATCTCTCCAGGAGCAAAATGTCCTTGGAGTGTCA Wild-type x18

**OT1-6**  
GAGAGAGGCTCCCATCACGGGGGAGGGAGTTTGCTCCTGGGGAACCTGTGATCCCCACAGGGAACA Wild-type x87

**OT1-11**  
TGGACTCTACCCCTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Wild-type x27

**Target 2 (VEGFA Site 2):**

**OT2-2**  
ACCCACCTCTCTATCCCAAACCTTGCCAGGAGGAGGGGGCCCTAGGAGCGCAGCTCTGGGAGTGTGGGAGTGTGGGGGGTACTTTGTGGGCGTT Wild-type x71

**OT2-15**  
TGACTGTCGGTGCCCCACATGTGGCAGATGCCCAGAGGCGGGGTGTGGGGGGTACTTTGTGGGCGTT Wild-type x108

**AGACAGGACATCTCTGACACCCAGGAGAACCCTCCCTATCTCTCCAGGAGCAAAATGTCCTTGGAGTGTCA +2**
TGACTGTCGGTGCCCCACATGTGGCAGATGCCCAGAGCTT
GCGGGGTGTGGGGGGTACTTTGTGGGCGTT +2
ACAAGATGACTATGTGCTGCCCACATGTGGCAGATGCCCAGAGCTT
GCGGGGTGTGGGGGGTACTTTGTGGGCGTT +2
ACAAGATGACTATGTGCTGCCCACATGTGGCAGATGCCCAGAGCTT
GCGGGGTGTGGGGGGTACTTTGTGGGCGTT +2

\[ \text{Target 3 (VEGFA Site 3):} \]

\[ \text{Target 4 (EMX1):} \]
Supplementary Fig. 3  Sequences of off-target indel mutations induced by RGENs in human U2OS.EGFP cells

Wild-type genomic off-target sites recognized by RGENs (including the PAM sequence) are highlighted in yellow and numbered as in Table 1 and Supplementary Table 2. Note that the complementary strand is shown for some sites. Deleted bases are shown as red dashes on a grey background. Inserted bases are italicized and highlighted in blue.
Target 1 (VEGFA Site 1):

OT1-3
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Wild-type x41
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ15
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ30
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ33
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ36

OT1-6
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ11
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ14
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ30
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ33
TCAGACGGACATTTCTGACACCCTCAGAGCCTCCCTCCCTACCTGCCCAAAATCCGGTCTTTAGATGG Δ36

OT1-11
AGCATCGCTGGACTCTACCCAGTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAA Wild-type x84
AGCATCGCTGGACTCTACCCAGTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAA Δ14
AGCATCGCTGGACTCTACCCAGTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAA Δ30
AGCATCGCTGGACTCTACCCAGTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAA Δ33
AGCATCGCTGGACTCTACCCAGTGAATGCCAGGAGCAAACTTCCCCTCCCCGAGTTGTGACAGCAAA Δ36

Target 2 (VEGFA Site 2):

OT2-2
CACCCTCTATCTCTCTTTAATCTGCAGGGGAGGCGGGAATTCGCTTCTGGGGGTGGAGGGGCCCCTAGGAGCGCCTTGGTGGGA Wild-type x74
CACCCTCTATCTCTCTTTAATCTGCAGGGGAGGCGGGAATTCGCTTCTGGGGGTGGAGGGGCCCCTAGGAGCGCCTTGGTGGGA Δ13
CACCCTCTATCTCTCTTTAATCTGCAGGGGAGGCGGGAATTCGCTTCTGGGGGTGGAGGGGCCCCTAGGAGCGCCTTGGTGGGA Δ30
CACCCTCTATCTCTCTTTAATCTGCAGGGGAGGCGGGAATTCGCTTCTGGGGGTGGAGGGGCCCCTAGGAGCGCCTTGGTGGGA Δ33
CACCCTCTATCTCTCTTTAATCTGCAGGGGAGGCGGGAATTCGCTTCTGGGGGTGGAGGGGCCCCTAGGAGCGCCTTGGTGGGA Δ36

GGGGTGAGGGGGCCCTAGGAGCGCCTTGGTGGGAG +150

OT2-15
GTCGTGGCCCACTATGCGAGATGCCTCACAGGCGGCCTGGGACTTCTTTGGAACGGTGATATCTTGCTGTTTGCGCTGATGCAAGCGTGTGAGTGAAT

OT2-24
GTCCCTCTGGGGCCCCATCCTGCCCTCCCTCCCCACCCCGCCTCAGGCTTGAAGAGGAAAGAAGAGCA

Target 3 (VEGFA Site 3):
OT3-2*
GAGAGCGAGTGAGTGAGTGAGTGAGTGTGTGTGGGGGGGACTCGGCTTGTTGTTGTCGGTGACTT Wild-type x26

OT3-9
TGGAGGTGTTGGGATGCGGGAGTGGGTGAGTGAGTGCGTGCGGGTGGCGATGCAAGCGTGTGAGTGAAT Wild-type x101

OT3-18
CAAAGACAGTAGATCTTTAAATGCTCACACACACACTACCCACACATATAAAAGGTGGTAACTGTGTGCTGACTT Wild-type x64

CAAAGACAGTAGATCTTTAAATGCTCACACACACACTACCCACACATATAAAAGGTGGTAACTGTGTGCTGACTT +71

CAAAGACAGTAGATCTTTAAATGCTCACACACACACTACCCACACATATAAAAGGTGGTAACTGTGTGCTGACTT +157

CAAAGACAGTAGATCTTTAAATGCTCACACACACACTACCCACACATATAAAAGGTGGTAACTGTGTGCTGACTT +190

CAAAGACAGTAGATCTTTAAATGCTCACACACACACTACCCACACATATAAAAGGTGGTAACTGTGTGCTGACTT +211 (∆16 +227)
**Target 4 (EMX1):**

OT4-1
GATTGCCTTTACTCCATGCTCTTCTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Wild-type x74
GATTGCCTTTACTCCATGCTCTTCTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA ∆9
GATTGCCTTTACTCCATGCTCTTCTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA x2
GATTGCCTTTACTCCATGCTCTTCTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA ∆6
GATTGCCTTTACTCCATGCTCTTCTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA ∆3 x3

**Supplementary Fig. 4  Sequences of off-target indel mutations induced by RGENs in human HEK293 cells**

Wild-type genomic off-target sites recognized by RGENs (including the PAM sequence) are highlighted in yellow and numbered as in Table 1 and Supplementary Table 2. Note that the complementary strand is shown for some sites. Deleted bases are shown as red dashes on a grey background. Inserted bases are italicized and highlighted in blue. *Yielded a large number of single bp indels*
Supplementary Table 1  Sequences of oligonucleotides used to generate expression plasmids

encoding sgRNAs/variant sgRNAs targeted to sites in the EGFP reporter gene and sgRNAs targeted
to six endogenous human gene targets

See accompanying file
Supplementary Table 2  Sequences and characteristics of genomic on- and off-target sites for six RGENs targeted to endogenous human genes and primers and PCR conditions used to amplify these sites

See accompanying file
Supplementary Table 3  Numbers of off-target sites in the human genome for six RGENs targeted to endogenous human genes and three RGENs targeted to the *EGFP* reporter gene

<table>
<thead>
<tr>
<th>Target Site</th>
<th>Number of mismatches to on-target site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Target 1 (VEGFA Site 1)</td>
<td>1</td>
</tr>
<tr>
<td>Target 2 (VEGFA Site 2)</td>
<td>1</td>
</tr>
<tr>
<td>Target 3 (VEGFA Site 3)</td>
<td>1</td>
</tr>
<tr>
<td>Target 4 (EMX)</td>
<td>1</td>
</tr>
<tr>
<td>Target 5 (RNF2)</td>
<td>1</td>
</tr>
<tr>
<td>Target 6 (FANCF)</td>
<td>1</td>
</tr>
<tr>
<td>EGFP Target Site #1</td>
<td>0</td>
</tr>
<tr>
<td>EGFP Target Site #2</td>
<td>0</td>
</tr>
<tr>
<td>EGFP Target Site #3</td>
<td>0</td>
</tr>
</tbody>
</table>

Off-target sites for each of the six RGENs targeted to the *VEGFA*, *RNF2*, *FANCF*, and *EMX1* genes and the three RGENs targeted to EGFP Target Sites #1, #2 and #3 were identified in human genome sequence build GRCh37. Mismatches were only allowed for the 20 nt region to which the sgRNA anneals and not to the PAM sequence.
Supplementary Table 4  Indel mutation frequencies at on- and off-target genomic sites induced by different amounts of Cas9- and sgRNA-expressing plasmids for the RGEN targeted to VEGFA Target Site 3

<table>
<thead>
<tr>
<th>Site</th>
<th>Sequence</th>
<th>250ng sgRNA/750 ng Cas9 Mean indel frequency (%) ± SEM</th>
<th>12.5ng sgRNA/250 ng Cas9 Mean indel frequency (%) ± SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>T3 (On-target)</td>
<td>GGTGAGTGAGTGTTGGCGTG</td>
<td>49.4 ± 3.8</td>
<td>33.0 ± 3.7</td>
</tr>
<tr>
<td>OT3-1</td>
<td>GGTGAGTGAGTGGTGGAGG</td>
<td>7.4 ± 3.4</td>
<td>N.D.</td>
</tr>
<tr>
<td>OT3-2</td>
<td>AGTGAGTGAGTGTTGGGG</td>
<td>24.3 ± 9.2</td>
<td>9.8 ± 4.2</td>
</tr>
<tr>
<td>OT3-4</td>
<td>GCTGAGTGAGTGGTGCGTG</td>
<td>20.9 ± 11.8</td>
<td>4.2 ± 1.2</td>
</tr>
<tr>
<td>OT3-9</td>
<td>GGTGAGTGAGTGGCGGGTGG</td>
<td>3.2 ± 0.3</td>
<td>N.D.</td>
</tr>
<tr>
<td>OT3-17</td>
<td>GTGAGTGAGTGGTGCGGAGG</td>
<td>2.9 ± 0.2</td>
<td>N.D.</td>
</tr>
<tr>
<td>OT3-18</td>
<td>TGTGCGTGAGTGGTGCGGAGG</td>
<td>13.4 ± 4.2</td>
<td>4.9 ± 0.0</td>
</tr>
<tr>
<td>OT3-20</td>
<td>AGAGTGAGTGAGTGGTAGG</td>
<td>16.7 ± 3.5</td>
<td>7.9 ± 2.4</td>
</tr>
</tbody>
</table>

Amounts of sgRNA- and Cas9-expressing plasmids transfected into U2OS.EGFP cells for these assays are shown at the top of each column. (Note that data for 250 ng sgRNA/750 ng Cas9 are the same as those presented in Table 1.) Mean indel frequencies were determined using the T7EI assay from replicate samples as described in Methods. OT = Off-target sites, numbered as in Table 1 and Supplementary Table 2. N.D. = none detected
Supplementary References:
