FokI and Cas8 connected with a 30-aa linker (SEQ ID NO:455). Additional linker polypeptide sequences of varying length and amino acid composition were designed and used to connect FokI to the Cas8 protein in these vectors. The additional linker polypeptide sequences are listed in Table 43.

[00807]

Table 43
Amino Acid Linker Sequences

SEQ ID NO:	Linker length (amino acids)	Amino acid sequence
SEQ ID NO:817	5	GGGGS
SEQ ID NO:818	8	TGPGAAAR
SEQ ID NO:819	10	GGSGSSGGSG
SEQ ID NO:820	15	GGSGSSGGSGSSGGS
SEQ ID NO:821	17	ADPTNRAKGLEAVSVAS
SEQ ID NO:822	20	SGSETPGTSESATPESGGSG
SEQ ID NO:433	30	SGSETPGTSESATPESGGSGSSGGSGSSGG
SEQ ID NO:823	40	SGSETPGTSESATPESGGSGSSGGSGSSGGSGS SGGSGSS
SEQ ID NO:824	50	SGSETPGTSESATPESGGSGSSGGSGSSGGSGS SGGSGSSGGSGSSGGSG

[00808] B. Transfection of vectors encoding FokI-Cascade RNP complex components

[00809] Transfection conditions were essentially as described in Example 8B with the following modifications. Prior to nucleofection, 5 µl of plasmid vector solution was transferred to individual wells of a 96-well plate. Each well contained 2.4 µg of plasmid encoding FokI-Cascade RNP complex subunit protein components and ~1-2 µg of plasmid encoding the minimal CRISPR array.

[00810] C. Deep sequencing of gDNA from transfected cells

[00811] Deep sequencing was performed essentially as described in Example 8C with the following modifications. Instead of primers Y and Z from Table 36 of Example 8C, the target-specific primers were SEQ ID NO:825 to SEQ ID NO:1016.

[00812] D. Deep sequencing data analysis

[00813] Deep sequencing data analysis was performed essentially as described in Example 8D. FIG. 31A and FIG. 31B present the results of the data analysis. In FIG. 31A and FIG. 31B, percent genome editing is shown as a function of FokI-Cas8 linker type (FIG. 31A, FIG. 31B, ~~vertical axis 14-60~~ AA) and interspacer distance (n=1) (FIG. 31A, FIG. ~~32B~~, ~~horizontal axis, interspacer distance 5-50bp~~. In FIG. 31A, the grey scale vertical bar to the right

horizontal axis, 5-50
31B
vertical axis, 14-60 bp

31B

is percentage of indels. In FIG. ~~32B~~, values in the cells are the percent indels. An initial analysis of the data showed genome editing was highest with FokI-Cas8 linkers of 17 and 20 amino acids (SEQ ID NO:821 and SEQ ID NO:822, respectively) and with interspacer distances of ~26bp and ~30-32bp. The data was reprocessed and samples with less than a thousand sequences reads were removed as they may contain inflated editing values due to low coverage (sites were only retained if all the associated samples contained >1000 reads). This data, presented in FIG. 31A and FIG. 31B, showed that genome editing was highest with FokI-Cas8 linkers of 17 and 20 amino acids (SEQ ID NO:821 and SEQ ID NO:822, respectively) and with interspacer distances of ~30-32bp. Thus, efficient genome editing using Type I CRISPR-Cas complexes comprising Fok1-Cas8 fusion proteins was achieved by varying the interspacer distance and the linker polypeptide length of the FokI-Cas8 fusion protein. The amino acid composition of the linker polypeptides is discussed herein.

[00814]

Example 19

Identifying Cascade Homologs for Genome Editing

[00815] This Example illustrates the design and testing of multiple homolog Cascade complexes to evaluate the efficiency of genome editing.

[00816] A. Identification of sites for testing with homolog Cascade complexes

[00817] A panel of sites was identified for testing additional homolog Cascade complexes. Specifically, minimal CRISPR arrays were designed to target a set of loci in the human genome with 30-bp interspacer distances and that were flanked by either AAG or ATG PAM sequences. Guide polynucleotides containing "repeat-spacer-repeat-spacer-repeat" sequences were cloned following the method described in Example 9A with SEQ ID NO:454. The full set of oligonucleotide sequences used to generate the minimal CRISPR arrays are presented as SEQ ID NO:1017 to SEQ ID NO:1130 (Hsa33F, SEQ ID NO:1017, and Hsa33R, SEQ ID NO:1074, exemplify one pair). A positive control comprising guides targeting the TRAC locus was included (SEQ ID NO:454).

[00818] FokI-Cascade RNP subunit protein component-encoding genes were cloned into vectors comprising: CMV promoters to enable delivery and expression in mammalian cells; *cas* genes linked via 2A viral peptide "ribosome-skipping" sequences; a fusion protein comprising FokI and Cas8 connected with a 30-aa linker (SEQ ID NO:455).