Supplementary Information for

Template-based genome editing directed by *Svi*Cas3

Authors:

Wang-Yu Tong*, Zhong-Fan Yu, Yong Li, Shou-Dong Ye, De-Xiang Yong, An-Jing Wang, Yan-Yan Tang, Xin Xu, Cai-Hua Qiu, Xing-Wang Yang, Ting-Ting Xia, Mei-Li Li, Qing-Yang Liu, Si-Qi Zhu, Yan Zhang, Su-Li Cao, Yan Sun and Xue Li

Affiliations:

Integrated Biotechnology Laboratory, School of Life Sciences, Anhui University, 111 Jiulong Road, Hefei 230601, China

* *Corresponding author: tongwy@ahu.edu.cn*

Tel.: +86-551-63861282

Fax: +86-551-63861282

Table of contents

Supplementary discussion

1. The subtype I-B-*Svi* CRISPR-Cas system

The three confirmed CRISPRs: The genome of *S. virginiae* IBL14 contains three confirmed *Svi*CRISPRs, i.e., *Svi*CRISPRs 1, 2 and 3 (Extended Data Table 4). All 26 spacers in the three *Svi*CRISPRs are confined to a length range of 34 nt to 44 nt, a GC content range of 51% to 86% and a Tm range of 66 ºC to 84 ºC. A total of 29 repeats in the three *Svi*CRISPRs are composed of 30 nucleotides (nt) and share a GC content range of 57% to 70% (lower than the average GC content of 72.98% in the *S. virginiae* IBL14 genome) and a melting temperature (Tm) range of 67 ºC to 74ºC. *Svi*CRISPRs 1, 2 and 3 have 13, 5 and 11 repeats, respectively. All 13 repeats in *Svi*CRISPR 1 have the same base sequences, but the base sequences in the 5 repeats of *Svi*CRISPR 2 are not the same. Two and four out of the 11 repeats in *Svi*CRISPR 3 have the same base sequences, respectively. In addition, there is one repeat in *Svi*CRISPR 2 that shares the same base sequences with two repeats in *Svi*CRISPR 3. In comparison to the repeat architecture (gcttcAACCCcacaaGGGTTcgtctgaaac, a total of 30 nt) of the subtype I-B CRISPR-Cas system in *Haloferax volcanii* [1], 26 out of the 29 repeats have a pentameric palindrome composed of two inverted repeats interspaced by 3 nt with the following stoichiometry: gt-5 nt-CCCCT-3 nt-AGGGG-10 nt or gt-8 nt-CCCCT-3 nt-AGGGG-7 nt, implying that in the 30 nt repeats, a 5 nt-palindrome with a higher GC content is probably important for subtype I-B CRISPR-Cas-system-directed adaptive immunity to invader DNA. All 29 repeats begin with the base “g” in the 5’-terminal and 27 out of

the 29 repeats end with the two bases “ac” in the 3’-terminal, similar to the repeats in the I-B CRISPR-Cas system in *Haloferax volcanii* [1]. It is worth noting that the same sequence motif “CCC/GGG” in the repeats of the two subtype I-B CRISPR-Cas systems also exists in the repeats of the subtype I-E (gtgtTCCCCGCgccaGCGGGGAtaaacc-28 nt, in *E. coli*) [2] and type VI (ccaCCCCaatatcgaaGGGGactaaaac-28 nt, in *L. shahii*) [3] CRISPR-Cas systems, suggesting that in adaptive immunity to invasive genetic factors (DNA or RNA) directed by CRISPR-Cas systems, the complementary motif is probably necessary for maintaining the stability of crRNA secondary structure.

It was reported that approximately 0.4% of spacers (100 of 23550 spacers) were self-targeting spacers (fully matching a portion of endogenous genomic sequences that is not part of a CRISPR array, termed self-proto-spacer) and that such self-targeting was a mode of autoimmunity rather than gene regulation [4]. Although there are no full self-targeting spacers in the three *Svi*CRISPR arrays (Extended Data Table 4), partial sequences (18 nt to 19 nt) of five spacers match a portion of five encoding gene sequences in the strain IBL14 genome (the five genes *gvgl*005738, *gvgl*005245, *gvgl*002878, *gvgl*001077 and *gvgl*006237 encode squalene-hopene cyclase/EC5.4.99.17, thymidine kinase/EC 2.7.1.21, ATP-binding cassette/ABC transporter integral membrane protein, TetR family transcriptional regulator and putative dehydrogenase, in proper sequence) (http://www.kegg.jp/; http://www.ncbi.nlm.nih.gov; http://www.uniprot.org/help/uniprotkb). It is worth noting that the nucleotide sequences of *Svi*CRISPR 3 fall within the sequences of the

coding gene *gvgl*007771 (accession number in the NCBI database: KY243079) that shares the highest identity (77.34%) with the DUF721 domain-containing protein of *Streptomyces subrutilus* (WP_069924566.1). The gene *gvgl*007771 encodes an unknown functional protein, GVGL007771, which is widely present in bacteria (http://www.ncbi.nlm.nih.gov).

The subtype I-B-Svi Cas system: The architecture of the CRISPR-Cas system in the *S. virginiae* IBL14 genome was delineated as shown in Extended Data Fig. 1A (*gvgl*007770-*cas*7-*cas*5-*cas*3-*cas*4-*cas*1-*cas*2 in transcriptional direction). In contrast to the typical subtype I-B Cas system containing eight *cas* genes with the arrangement *cas*6-*cas*8b-*cas*7-*cas*5-*cas*3-*cas*4-*cas*1-*cas*2 in the transcriptional direction found in *Lactobacillus crispatus* VMC3, *Thermotoga neapolitana* DSM 4359, *Haloarcula marismortui* str. ATCC 43049 and *Clostridium kluyveri* CKL_2758-CKL_2751 [5-7], the *Svi*Cas system lacks the common gene *cas*6 and the signature gene *cas*8 used for subtype classification in type I Cas systems, implying that Gvgl007770 may play the roles of Cas6 or Cas8 or other Cas proteins in the *Svi*Cas system. Based on the architectural similarity of *cas* genes between the *Svi*Cas system and subtype I-B Cas systems [5,7], the *Svi*Cas system was temporarily named the subtype I-B-*Svi* Cas system.

In the sequences of the seven *cas* genes, there are two sets of genes (*cas*7 with *cas*5 and *cas*3 with *cas*4) with the four overlapping bases “gtga”, among which “gtg” (valine) serves as a start codon in *Svi*Cas5 and *Svi*Cas4 translation and “tga” as a stop codon in *Svi*Cas7 and *Svi*Cas3 translation, and four superficially redundant sequences, i.e., 29 nt between *gvgl*007770 and *cas*7, 161 nt between *cas* 5 and *cas* 3, 1 nt between *cas* 4 and

cas 1 and 7 nt between *cas* 1 and *cas* 2. As expected, the number of bases in the overlapping and redundant sequences is not a multiple of three, to prevent misreading of the coding sequences. In addition, the gaps between *Svi*CRISPR 3 and *Svi*CRISPR 2, *Svi*CRISPR 2 and *gvgl*00770 as well as *cas*2 and *Svi*CRISPR 1 in the system are 645 (226+419) nt, 1240 nt and 139 nt, respectively.

Furthermore, the biochemical properties of the seven *Svi*Cas proteins in the subtype I-B-*Svi cas* gene operon were investigated (Extended Data Table 5) [7] (http://web.expasy.org/protparam). All the *Svi*Cas proteins are hydrophilic, as expected, i.e., their grand average of hydropathicity (GRAVY) values are negative. In the six definite *Svi*Cas proteins (*Svi*Cas1, 2, 3, 4, 5 and 7), the numbers of cysteine residues, respectively, are 2, 1, 2, 5, 1 and 0, and the numbers of methionine residues are 6, 0, 4, 0, 2 and 2. That is, there are no disulfide bonds in *Svi*Cas proteins 2, 5 and 7, which are generally responsible for RNA processing [5,7-9]. Strikingly, among the seven *Svi*Cas proteins, the signature protein *Svi*Cas3 has the largest theoretical molecular weight (MW) of 84352.40 Da (771 aa), the lowest theoretical isoelectric point (pI) of 5.78, the lowest instability index (I-ind) of 35.49 and the highest GRAVY of -0.153.

The results of Nucleotide BLAST for the seven *cas* genes showed that *cas*1, *cas*3, *cas*4 and *cas*7 had the highest identity with the four genes (74.51%, 74.18%, 75.09% and 70.78%, in proper order): Sequence ID-KM527027.1 [*Salinispora pacifica* strain DSM 45549 Cas1 protein I-B (*cas*1) gene, complete cds], Sequence ID-CP054919.1 [*Kitasatospora sp*. NA04385 chromosome, complete genome], Sequence ID-CP061007.1 [*Saccharopolyspora spinosa* strain CCTCC M206084 chromosome,

complete genome] and Sequence ID-CP016076.1 [*Actinoalloteichus fjordicus* strain ADI127-7 chromosome, complete genome], while the results of Nucleotide BLAST for the three genes *cas*2, *cas*5 and *gvgl*007770 were "No significant similarity found" (http://www.ncbi.nlm.nih.gov). Furthermore, we performed Protein BLAST on the seven *Svi*Cases and found that *Svi*Cas1, *Svi*Cas2, *Svi*Cas3, *Svi*Cas4, *Svi*Cas5, *Svi*Cas7 and Gvgl007770 shared the highest identity with the seven proteins (85.89%, 85%, 64.77%, 84.85%, 75.71%, 78.92% and 59.23%, respectively): Sequence ID-WP_123471944.1 [type I-B CRISPR-associated endonuclease Cas1, *Streptomyces sp.* CEV 2-1], Sequence ID-WP_065001230.1 [CRISPR-associated endonuclease Cas2, *Streptomyces sp*. H-KF8], Sequence ID-WP_121433360.1 [CRISPR-associated helicase Cas3', *Actinomadura pelletieri*], Sequence ID-WP_055591556.1 [CRISPR-associated protein Cas4, *Streptomyces sp*. H-KF8], Sequence ID-WP_158013542.1 [CRISPR-associated protein Cas5, partial, *Streptomyces thermoautotrophicus*], Sequence ID-WP_045691965.1 [type I-B CRISPR-associated protein Cas7/Cst2/DevR, *Streptomyces rubellomurinus*] and Sequence ID-WP_147449255.1 [hypothetical protein, *Actinomadura pelletieri*] (http://www.ncbi.nlm.nih.gov) [10]. In the subtype I-B-*Svi* Cas system, the unknown protein Gvgl7770 is neither Cas6 nor Cas8. What is the function of Gvgl7770? The classic *Sp*Cas9 from *Streptococcus pyogenes* has the highest identity (85.84%) to Cas9 (Sequence ID: WP_138125798.1) from *Streptococcus dysgalactiae* and the second highest (85.09%) to Cas9 (Sequence ID: WP_159583349.1) from *Streptococcus halichoeri*, indicating that *Svi*Cas3 (64.77%) is more distinctive

than *Sp*Cas9 and probably performs some functions that neither *Sp*Cas9 nor other Cas3 proteins share. In summary, the subtype I-B-*Svi* Cas system has unique features.

Cas3 superfamily proteins typically contain an N-terminal histidine-aspartate (HD) domain (belonging to a superfamily of metal-dependent phosphohydrolases) and a C-terminal ATP-dependent superfamily 2 helicase (Hel) domain [9,11-13]. The secondary structure of *Svi*Cas3 was predicted (Extended Data Fig. 1B) by online software (GOR4) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html). *Svi*Cas3 includes 49.42% (381 aa) alpha helix (h) (55.92% in *Sp*Cas9), 9.99% (77 aa) extended strand (e) (8.55% in *Sp*Cas9) and 40.60% (313 aa) random coil (c) (35.53% in *Sp*Cas9). Notably, it contains seven repeat motifs, consisting of four (AALL, AHQL, GLPA and LEEI) or three (GRL, RFG and VAT) amino acid residues. The motifs AALL, AHQL and LEEI are only present in the alpha helix, and the motif GLPA is present in the random coil. However, the functions of these repeat motifs are unclear.

Cas9 in type II systems contains two nuclease domains, i.e., an HNH-like nuclease domain in the middle of the protein that cleaves the crRNA complementary strand and a RuvC-like nuclease domain near the amino (N) terminus that cleaves the noncomplementary strand, which generates blunt-ended double-strand breaks (DSB), typically 3 nt in front of the 3’ end of a proto-spacer (a sequence in the target genomic DNA that is consistent with a spacer) adjacent motif (PAM) [7,8,14]. Cas9 is also involved in the trimming of pre-crRNA to form crRNA via a Cas9-crRNA-tracrRNA ternary complex and in the silencing of invasive DNA [10,15]. Unlike Cas9 cutting both the target DNA strand (TS) and the nontarget DNA strand (NTS) using its two nuclease domains,

Cas3 can perform crRNA-guided, PAM-dependent cleavage sequentially on both the NTS and TS using its HD domain with the help of its Hel domain in the presence of ATP [9,10,16-19], implying that it probably produces a local single strand broken (SSB) DNA or DNAs with sticky-ends. Analytical results show that *Svi*Cas3 exhibits an HD-like nuclease (HD domain) at the N-terminus (1-186 aa) and a DNA helicase (Hel domain) in the middle of the protein (232-591 aa) (http://www.ebi.ac.uk/interpro/sequencesearch/iprscan5-S20170430-040754-0890-51239778-oy), implying that the carboxyl termini of Cas3 and Cas9 are both nonconserved.

2. Self-targeting genome editing in *S. virginiae*

Self-targeting genome editing refers to endogenous genome editing that is triggered by delivering an engineered gene editing plasmid into a cell and further implemented by the Cas system(s) and HDR or the nonhomologous end joining (NHEJ) of the cell itself, in which the spacer designed in the guide DNA (g-DNA) fragment fully matches a portion of the complete genomic sequences of the cell itself. Since 2015, we have repeatedly demonstrated that the endogenous subtype I-B-*Svi* Cas system can be programmed for self-targeting genome editing by transforming a gene editing plasmid into the cells of *S. virginiae* IBL14 [20-25].

To develop self-targeting gene editing tools and investigate the mechanism of antibiotic metabolism in *S. virginiae* IBL14, we selected seven endogenous genes participating in penicillin metabolism as targets (*sviipe* with 1143 nt: encoding

isopenicillin-N epimerase, EC 5.1.1.17; *svipam*1 with 2901 nt and *svipam*2 with 2049 nt: encoding penicillin amidase, EC 3.5.1.11; *svibla*1 with 2082 nt, *svibla*2 with 1542 nt , *svibla*3 with 1245 nt and *svibla*4 with 921 nt: encoding beta-lactamase, EC 3.5.2.6; corresponding accession numbers in the NCBI database: KY243071, KY243072, KY243073, KY243074, KY243075, KY243076 and KY243077) for self-targeting genome editing, and chose plasmid pKC1139 [26] as a vector to construct gene editing plasmids (Extended Data Fig. 2A, Extended Data Tables 1 and 2). Additionally, we arbitrarily designed the fragments of template DNA donor (t-DNA) (upstream homologous arm/UHA plus downstream homologous arm/DHA plus insertion sequence if present: 201+391+807=1399 nt, 945+1102=2047 nt, 414+616=1030 nt, 219+322=541 nt, 418+648=1066 nt, 414+424=838 nt and 727+832=1559 nt for the genes *sviipe*, *svipam*1, *svipam*2, *svibla*1, *svibla*2, *svibla*3 and *svibla*4, respectively), the insertion fragment (807 nt for chloramphenicol acetyltransferase biosynthesis/*sviipe*) and the deletion fragments (327 nt/*svipam*1, 447 nt/*svipam*2, 577 nt/*svibla*1, 402 nt *svibla*2, 318 nt/*svibla*3 and 513 nt/*svibla*4) (Extended Data Fig. 2C and Extended Data Table 2).

Experimentally, we first extracted the *S. virginiae* IBL14 genome (~8.4 Mbp) and plasmid pKC1139 (6308 bp), and chemically synthesized g-DNA fragments (essential elements: promoter-repeat-spacer-repeat-terminator) (Extended Data Fig. 2A) (http://www.addgene.org/; http://www.ncbi.nlm.nih.gov). After preparing the t-DNA fragments (Extended Data Fig. 2B) of the seven genes of interest and constructing the seven gene editing plasmids (Extended Data Figs. 2C and 2D, Extended Data Tables 1

and 2), we transformed these gene editing plasmids into the protoplasts of *S. virginiae* IBL14 following the procedures described in the section "Construction of gene-edited mutants of *S. virginiae* IBL14". For each gene editing, three randomly picked single colonies on MR2YE plates were verified as potential gene-edited mutants (*strain name-edited gene abbreviation*) by DNA electrophoresis (Extended Data Fig. 2E, Extended Data Table 2) following the procedures described in Materials and Methods [23,27]. Next, we constructed two double gene deletion mutants (*S. virginiae* IBL14-Δ*svipam*1Δ*sviipe*::*cat* and *S. virginiae* IBL14-Δ*svipam*1Δ*svibla*1) using the plasmid-removed mutant *S. virginiae* IBL14-Δ*svipam*1 as a host according to the same procedures as described above. The results demonstrated that the strategy of iterative gene editing in self-targeting genome editing of *S. virginiae* IBL14 is convenient and practicable (Extended Data Figs. 2F and 2G, Extended Data Tables 1 and 2).

In self-targeting genome editing, including the deletion and insertion of DNA sequences of the genes of interest, all the experimental results (Extended Data Fig. 2, Extended Data Table 2) were correct, as we anticipated, implying that it is not impossible for other class 1 (types I, III and IV) CRISPR-Cas systems to be utilized for self-targeting genome editing. In particular, all tested mutants were the correctly gene-edited mutants, motivating us to further design and construct gene editing tools based on the subtype I-B-*Svi* CRISPR-Cas system for non-self-targeting genome editing.

3. RNA-guided genome editing in *E. coli*

To test whether the subtype I-B-*Svi* CRISPR-Cas system was effective in non-self-targeting genome editing, we selected the gene *lacZ* (3075 nt) encoding beta-D-galactosidase as a target in *E. coli* JM109 (DE3) genome for its advantages of blue-white selection and a known CRISPR-Cas system, which is incapable of self-targeting genome editing (http://crispr.i2bc.paris-saclay.fr/) [28,29]. Following the procedures described in both Construction of gene editing plasmids and Construction of Cas expression plasmids, we tentatively designed and constructed a set of gene editing tools composed of a gene editing plasmid pKC1139-t/g-Δ*lacZ* and a Cas expression plasmid pCas-*cas*7-5-3-4-1-2 bearing all six definite *cas* genes of the subtype I-B-*Svi* Cas system (Extended Data Figs. 3A-E, Extended Data Tables 1 and 2). The arbitrarily designed fragment sizes of t-DNA (UHA plus DHA) and deletion in the gene editing of *lacZ* were 1144 nt (423+721 nt) and 420 nt, respectively, and the selected PAM was ttc (Extended Data Fig. 3H and Extended Data Table 2). After transformation of the Cas expression plasmid pCas-*cas*7-5-3-4-1-2 and the gene editing plasmid pKC1139-t/g-Δ*lacZ* into *E. coli* JM109 (DE3) competent cells, three randomly selected white single colonies on LBPETs as the candidate gene-edited mutant *E. coli* JM109(DE3)-Δ*lacZ* (Extended Data Fig. 3F) were verified by DNA electrophoresis (Fig. 3G) and DNA sequencing analysis (Fig. 3H). The results in Extended Data Fig. 3 suggested that the subtype I-B-*Svi* CRISPR-Cas system could also be developed as a gene editing tool for genome editing in other prokaryotic species.

It has been reported [7,8,14,30] that proteins Cas1, Cas2 and Cas4 are, respectively, a metal-dependent deoxyribonuclease with a unique fold consisting of an N-terminal β strand domain and a C-terminal α-helical domain, a small endoribonuclease specific to

U-rich regions and a RecB-like nuclease with a three-cysteine C-terminal cluster (the three-cysteine C-terminal cluster of *Svi*Cas4 with 165 aa is composed of Cys153, Cys156 and Cys162), and all of them are responsible for new spacer acquisition in the adaptation phase in CRISPR-Cas systems that direct microbial immunity to exogenous DNA. Accordingly, we designed and constructed the Cas expression plasmid: pCas-*cas*7-5-3, harbouring the three genes *cas*7, *cas*5 and *cas*3 (Extended Data Figs. 3A-E). Together with the plasmid pKC1139-t/g-Δ*lacZ*, the plasmid pCas-*cas*7-5-3 was delivered into *E. coli* JM109 (DE3) competent cells. White single colonies on LBPETs, potentially gene-edited mutant *E. coli* JM109(DE3)-Δ*lacZ*, were also obtained (Extended Data Fig. 3F) and further confirmed by DNA electrophoresis and sequencing analysis (Figs. 3G and 3H). These results demonstrated that Cas1, Cas2 and Cas4 are not required in non-self-targeting genome editing based on the subtype I-B-*Svi* CRISPR-Cas system and suggested that the gene editing tools constructed based on *cas*7-5-3 can be simply and easily used for the genome editing of other prokaryotic microorganisms because of their monocistron and relatively low molecular weight advantages (a total of 143813.69 Da/1323 aa for *Svi*Cas7, *Svi*Cas5 and *Svi*Cas3; 158441.41 Da/1368 aa for the *Sp*Cas9).

Cas5, Cas6, Cas7 and Cas8 are subunits of the Cascade/Csy complex in the type I CRISPR-Cas system, which are responsible for recognizing the repeat-derived 5' handle of mature crRNA, yielding mature crRNA, forming the backbone of the Cascade complex, and recognizing the PAM sequence and recruiting Cas3 (or Cas2/3) [5,7,8,30-33], respectively. Therefore, we further simplified and constructed a Cas expression plasmid

pCas-*cas*3 harbouring only the single gene *Svicas*3 due to the lack of definite Cas6 and Cas8 proteins in this subtype I-B-*Svi*Cas system, and a gene editing plasmid pKC1139-t/g-Δ*lacZ*::*cat* (the same as pKC1139-t/g-Δ*lacZ* except for t-Δ*lacZ*::*cat* instead of t-Δ*lacZ*) (Extended Data Fig. 3, Extended Data Tables 1 and 2). After transforming the plasmids pCas-*cas*3 plus pKC1139-t/g-Δ*lacZ* or pCas-*cas*3 plus pKC1139-t/g-Δ*lacZ*::*cat* (t-DNA=UHA+*cat*+DHA: 1400 bp=278 bp+807 bp+315 bp; deletion fragment=441 bp; PAM: ttc) into *E. coli* JM109 (DE3) competent cells, we obtained white gene-edited mutants of *E. coli* JM109(DE3)-Δ*lacZ* and *E. coli* JM109(DE3)-Δ*lacZ*::*cat*, respectively (Extended Data Fig. 3F), and verified them by DNA electrophoresis and DNA sequencing analysis. Luckily, the band size of DNA electrophoresis and the result of DNA sequencing were exactly as designed (Extended Data Figs. 3G and 3H, Extended Data Table 2), demonstrating that in the non-self-targeting genome editing of *E. coli* JM109(DE3), *Svi*Cas3 is necessary while both *Svi*Cas7 and *Svi*Cas5 are auxiliary. This finding breaks the view that Cascade is necessary in genome editing directed by type I CRISPR-Cas systems [17,34,35] but is consistent with the view that a full R-loop formation (the R-loop–forming Cascade) enables Cas3 to bind to the NTS bulge and thus reduces off-target effects [17,18,32]. Notably, during CRISPR interference in type I-C CRISPR-Cas systems, Cas3-mediated target DNA recognition and cleavage are independent of the composition and architecture of the Cascade surveillance complex [19], which also indirectly supports our results.

The main disadvantage in the genome editing of *E. coli* JM109(DE3) was that many blue or blue-white colonies occurred in addition to white colonies on LBPETs,

and many small white colonies switched to blue-white colonies and even to blue colonies (Extended Data Fig. 3F), which resulted in two bands in the DNA electrophoresis of colony PCR. The reason is that the cells of the correctly gene-edited mutant *E. coli* JM109 (DE3)-Δ*lacZ* initially hydrolysed the antibiotics apramycin and kanamycin, and then the host *E. coli* JM109 (DE3) cells grew, which was validated by our experiments that small white colonies grown early were correctly *lacZ*-edited mutants, but blue colonies later grown were not.

4. RNA-guided and DNA-guided genome editing in *C. glutamicum*

RNA-guided genome editing in C. glutamicum: To further verify the effectiveness of the single *Svi*Cas3 in genome editing of prokaryotic microorganisms, an important industrial (producing diverse amino acids) strain *Corynebacterium glutamicum* ATCC 13032 lacking the CRISPR system (Because Cas9 is toxic to *C. glutamicum* ATCC 13032 cells, gene editing of the genome of *C. glutamicum* ATCC 13032 is difficult) (http://crispr.i2bc.paris-saclay.fr/) [36] was selected as the host of interest and the 945 nt gene *ldh* (encoding lactate dehydrogenase, which catalyses the reversible conversion of lactic acid to pyruvate) was selected as the gene of interest. In the genome editing of the strain *C. glutamicum* ATCC 13032, we designed and constructed the Cas expression plasmid pEC-XK99E-*cas*3 (Extended Data Figs. 4B and 4C, Extended Data Tables 1 and 2) and the gene editing plasmid pXMJ19-t/g-Δ*ldh*::*egfp* (*egfp* encoding an enhanced green fluorescent protein) (Extended Data Figs. 4A and 4C). The arbitrarily designed fragment sizes of t-DNA (UHA plus *egfp* plus DHA) and

deletion in the gene editing of *ldh* were 2402 nt (776+726+900 nt) and 1998 nt, respectively, and the selected PAM was ttc (Extended Data Fig. 4E and Extended Data Table 2). After delivering the two plasmids pXMJ19-t/g-Δ*ldh*::*egfp* and pEC-XK99E-*cas*3 into *C. glutamicum* ATCC 13032 competent cells, we obtained potential gene-edited mutants on LBHIS plates with 25 μg/ml chloramphenicol and 50 μg/ml kanamycin and tested them by basic PCR (primers: *ldh*-vF/vR and *ldh*-*egfp*-vF/*ldh*-vR, respectively) and subsequent sequencing analysis (Extended Data Fig. 4) following the procedures described in the section "Construction of gene-edited mutants of *C. glutamicum* ATCC 13032". Unfortunately, transformants of both *C. glutamicum* ATCC 13032-pXMJ19-t/g-Δ*ldh*::*egfp* and *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* showed green fluorescence, suggesting that the arbitrarily designed 776 bp-HUA fragment in pXMJ19-t/g-Δ*ldh*::*egfp* may have a promoter (as there is no promoter designed for *egfp* gene transcription in the plasmid pXMJ19-t/g-Δ*ldh*::*egfp*.).

Nevertheless, we observed from Extended Data Fig. 4D that two of three lanes of the DNA gel electrophoresis of the PCR products of edited sequences in the gene-edited mutant *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* genome had an inconspicuous band of 2498 bp (gene-edited mutant *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp*) in addition to an obvious band of 3770 bp (wild-type host *C. glutamicum* ATCC 13032) when the primer pair *ldh*-vF/vR was used (template: genomic DNA), implying that the gene editing tools developed based on the subtype I-B-*Svi* CRISPR-Cas system were effective although the efficiency was not very high in the genome editing of the strain *C. glutamicum* ATCC 13032. The conclusion was further verified by DNA gel electrophoresis of the

PCR products of the edited sequences when the primer pair *ldh*-vF/vR (Extended Data Fig. 4D) was replaced with the primer pair *ldh*-*egfp*-vF/*ldh*-vR (Extended Data Fig. 4D) and subsequently subjected to DNA sequencing analysis (Fig. 4E). In addition, dual plasmid transformation with pXMJ19-t/g-Δ*ldh::egfp* plus pEC-XK99E-*cas*9 was carried out for comparison in the genome editing of *C. glutamicum* ATCC 13032, and unfortunately, none of the colonies grew on LBHIS plates, suggesting that the biocompatibility between *Svi*Cas3 and the host cells of *C. glutamicum* ATCC 13032 is much better than that between *Sp*Cas9 and the host cells [36] and indicating that it is possible for *Svi*Cas3 to be applied in the genome editing of other prokaryotic species where Cas9 is difficult to use.

In the RNA-guided genome editing of *E. coli* JM109(DE3) and *C. glutamicum* ATCC 13032, the crRNA fragment composed of R-S-R can search for both the PAM sequence and spacer's base-pairing sequence (equivalent to proto-spacer) in a complementary strand of the endogenous DNA, guide *Svi*Cas3 (*Svicas*3 without codon optimization) to perform a site-specific cleavage of the target gene and hence genome editing mediated by the HDR of the cells themselves, indicating that both the architecture of crRNA and the codon optimization of *Svicas*3 (GC%: *Svicas*3-69%, *E. coli* genome-50.8%; *C. glutamicum* genome-53.8%) may not be required in the genome editing of prokaryotes conducted by the single *Svi*Cas3.

DNA-guided gene editing in C. glutamicum: In the DNA-guided genome editing of the strain *C. glutamicum* ATCC 13032, all construction procedures of the gene editing tools and gene-edited mutants were the same as those described in the RNA-

guided genome editing of the strain *C. glutamicum* ATCC 13032 above, except as mentioned in the text. To verify the effectiveness of the DNA-guided genome editing conducted by *Svi*Cas3, we first designed and constructed the gene editing plasmid-pXMJ19-t-Δ*ldh*::*egfp* (without g-DNA compared to the plasmid-pXMJ19-t/g-Δ*ldh*::*egfp*). In contrast to the designed fragment sizes of t-DNA (UHA+*egfp*+DHA=776+726+900=2402 nt) and deletion (1998 nt) in plasmid pXMJ19-t/g-Δ*ldh*::*egfp*, the designed fragment sizes of t-DNA and deletion in plasmid pXMJ19-t-Δ*ldh*::*egfp* were 2001 nt (375+726+900 nt) and 1623 nt, respectively (Extended Data Fig. 5A and 5C). After successively transforming the two plasmids pEC-XK99E-*cas*3 and pXMJ19-t-Δ*ldh*::*egfp* into *C. glutamicum* ATCC 13032 competent cells, we obtained potential gene-edited mutants on LBHIS plates containing 25 μg/ml chloramphenicol and 50 μg/ml kanamycin (Extended Data Table 1). The results of DNA gel electrophoresis and DNA sequencing of the target sequence (Extended Data Fig. 5B) were consistent with the designed DNA fragment size and sequence (Extended Data Fig. 5C and Extended Data Table 2), demonstrating that the potential gene-edited mutants were the correctly gene-edited mutants *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp*).

Furthermore, we designed and constructed an all-in-one gene editing tool- pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp* (pEC-XK99E-*cas*3 instead of pPXMJ19 as a starting plasmid, compared to the plasmid-pXMJ19-t-Δ*ldh*::*egfp*) (Extended Data Fig. 5A). After transforming the all-in-one plasmid pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp* into *C. glutamicum* ATCC 13032 competent cells, we obtained potential Δ*ldh*::*egfp*-edited

colonies of *C. glutamicum* ATCC 13032 on LBHIS plates and verified them through DNA gel electrophoresis and DNA sequencing analysis of the PCR products of the edited sequences. In the gene-edited mutants (*C. glutamicum* ATCC 13032-Δ*ldh*::*egfp*) obtained by single-plasmid transformation (pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp*) and double-plasmid transformation (pEC-XK99E-*cas*3 plus pXMJ19-t-Δ*ldh*::*egfp*), the results of DNA gel electrophoresis analysis (Extended Data Fig. 5B) and DNA sequencing of the target sequence (Extended Data Fig. 5C) were consistent with the designed fragment size and the designed sequence (Extended Data Fig. 5C and Extended Data Table 2), indicating that the all-in-one gene editing tool pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp* is practical in the DNA-guided, template-based genome editing of *C. glutamicum* ATCC 13032, as is the dual-plasmid gene editing tool pEC-XK99E-*cas*3 plus pXMJ19-t-Δ*ldh*::*egfp*. The success of DNA-guided *Svi*Cas3-based genome editing in *C. glutamicum* provides further support for the initial finding that *Svi*Cas3 can direct DNA-guided genome editing in *S. cerevisiae*. The results also support the idea that if there is no biocompatibility problem between an enzyme and host cells, the enzyme effective for a substrate in prokaryotes should also be effective for the substrate in eukaryotes, and vice versa.

5. Genome editing efficiency

Lethal DSBs in the host genome can often be repaired by HDR mechanisms (with lower frequencies and fewer errors) of cells themselves when homologous DNA sequences are present (achieved by introducing an engineered t-DNA with its two end

sequences complementary to the terminal sequences around the two DSB products), which is the foundation of template-based genome editing. NHEJ (with higher frequencies and more errors) often occurs in the absence of homologous DNA sequences and hence leads to frame shift mutations through nucleotide insertions or deletions (indels), especially in genome with blunt-ended DSBs [10,14].

In genome editing directed by CRISPR-Cas systems, the molecular events after a set of gene editing tools enters host cells mainly include (i) duplication, transcription and translation of the *cas* gene, (ii) transcription and processing of g-DNA to form crRNA, (iii) site-specific cleavage of genomic DNA by the Cas enzyme and (iv) homologous recombination between target DNA and t-DNA by HDR. Obviously, events (i) and (iv) are performed by the molecular components of host cells and plasmids, and only the events (ii) and (iii) should be accomplished by Cas(es)[10,14]. In fact, event (ii) is easily achieved through a vector with engineered g-DNA after crRNA optimization. Accordingly, event (iii) in genome editing becomes a key step, requiring a restriction endonuclease such as *Sp*Cas9 or *Svi*Cas3 to perform RNA-guided site-specific genomic DNA cleavage (protospacer: a fragment of genomic DNA complementary to a spacer).

According to the molecular principles and operating processes of genome editing directed by the CRISPR-Cas system discussed above, we know that genome editing efficiency (GEE) with high fidelity directed by the CRISPR-Cas system consists of the transformation or transformation/transfection efficiency (TE) of gene editing tools and the homologous recombination efficiency (HRE) of t-DNA. The former (easy to

optimize) mainly depends on the characteristics of plasmids (e.g., genetic markers) and physiological status of cells (e.g., competent state of recipient cells), while the latter (difficult to optimize) depends on the characteristics of the HDR mechanisms of the host cells themselves, the features of DNA products (e.g., cohesive-ended DSB, blunt-ended DSB or local SSB) of digestion by the Cas effector complex, and the biocompatibility between the Cas effector complex and host cells [37,38]. To test the GEEs based on the subtype I-B-*Svi* CRISPR-Cas system, many transformation experiments were carried out. Moreover, RNA-guided mammalian genome editing (two genes: *DROSHA* and *Vegfa*) was performed with *Sp*Cas9 as a control (Extended Data Fig. 11).

In the *Svi*Cas3-based, RNA-guided genome editing of six different species, the GEE values varied greatly by up to five orders of magnitude (respectively: 1.0×10^{-2} for *S. virginiae* IBL14, 9.0×10^{-5} for *E. coli* JM109 (DE3), 8.6×10^{-7} for *C. glutamicum* ATCC 13032, 6.4×10^{-2} for *S. cerevisiae* LYC4, 4.4×10^{-2} for HEK293T and 7.3×10^{-2} for NIH3T3) (Extended Data Table 3). Further analysis shows that the great change in GEEs (from 8.6×10^{-7} to 7.3%) is obviously dependent on the change of TEs (ranging from 1.8×10^{-6} to 12.2×10^{-2}, up to five orders of magnitude; respectively: 1.1×10^{-2} for *S. virginiae* IBL14, 3.3×10^{-2} for *E. coli* JM109 (DE3), 1.8×10^{-6} for *C. glutamicum* ATCC 13032, 6.8×10^{-2} for *S. cerevisiae* LYC4, 12.2×10^{-2} for HEK293T, and 11.9×10^{-2} for NIH3T3) but not on the change in HREs (ranging from 27.3% to 94%, on the same order of magnitude; respectively: 92% for *S. virginiae* IBL14, 27.3% for *E. coli* JM109 (DE3), 48.0% for *C. glutamicum* ATCC 13032, 94% for *S. cerevisiae* LYC4, 36.0% for HEK293T, and 61.3% for NIH3T3). It is worth noting that the HRE changes in the

genome editing of six different species based on the *Svi*Cas system were of the same order of magnitude, all higher than 27.3% (Note: in the genome editing of *E. coli* JM109 (DE3), the plasmid pCas harbours the genes *exo*, *bet* and *gam* that encode λ Red recombinases to enhance HRE), indicating that in the template-based genome editing directed by *Svi*Cas3, the critical factor affecting GEE is not HRE but TE. In particular, we found that in different batches of genome-editing experiments with the same strain, the effect of experimental strategies (e.g., genetic marker) and physiological status of cells (e.g., competent cells) on GEE was very significant, suggesting that in template-based genome editing, HDR-system engineering in cells of different species is not a priority as long as exogenous endonucleases (e.g., Cas3, Cas9, zinc finger nucleases, etc.) and their HDR machineries (including allelic crossing-over mechanisms) work properly. In short, to improve GEE, the optimization of experimental strategies and operational conditions is indispensable.

To further verify this view, we conducted a series of *Svi*Cas3-based, DNA-guided eukaryotic genome editing experiments (Fig. 3 and Fig. 4). As seen in Extended Data Table 3, the GEE values in the DNA-guided genome editing directed by the *Svi*Cas3 (6.6% on average: 6.5% for *S. cerevisiae* LYC4, 7.8% for HEK293T, and 5.5% for NIH3T3) differed by approximately 10% from those in the RNA-guided genome editing (6.0% on average). The average of 6.6% GEE in the DNA-guided mode is not lower than the average of 6.0% GEE in the RNA-guide mode, implying that in the template-based genome editing conducted by *Svi*Cas3 in the presence of t/g-DNA, R-loops derived from g-DNA and D-loops from t-DNA (UHA and DHA) may coexist in

transformed cells, thus affecting GEEs. In the RNA-guided, template-based genome editing of mammalian cells, we found that the GEE values (4.4-7.3%) directed by *Svi*Cas3 were approximately 10 times higher than those directed by *Sp*Cas9 (0.5-0.7%) (Extended Data Table 3).

6. Off-target analysis in RNA-guided genome editing

Off-target mutations are a major concern in genome editing in mammalian cells, because they can lead to genomic instability and the functional failure of some normal genes, especially in biomedical and clinical applications [39-41]. Many strategies for minimizing off-target effects in CRISPR/Cas9-mediated genome engineering have been reported, including optimization of the crRNA sequence, concentration control of the Cas9-sgRNA complex, mutation of *cas*9 and so on, but potential off-target cleavage could still occur, especially at the genomic DNA sites consistent with a seed sequence of 6-11 nt (corresponding to the terminal sequence of a spacer directly adjacent to PAM) [40,41].

To investigate the off-target effects of *Svi*Cas3-based genome editing, we stochastically chose four potential off-target sites for each of the six gene-edited mutants (7 potential off-target sites in *E. coli* JM109(DE3)-Δ*lacZ* based on the motif: PAM+seed=3+8=TTC+CCGCCCGG; 18 potential off-target sites in *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* based on the motif: PAM+seed=3+7= TTC+CATTCCA, 87 potential off-target sites in *S. cerevisiae* LYC4-Δ*crtE* based on the motif: PAM+seed=3+6=TTC+tagtgc; 315 potential off-target sites in HEK293T-

Δ*DROSHA*::*egfp* based on the motif: PAM+seed=3+11=TTC+tttcaacagtg; 11 potential off-target sites in HEK293T-Δ*CAMKMT*::*egfp* based on the motif: PAM+seed=3+12=TCC+ttgaatgttgaa; and 17 potential off-target sites in NIH-3T3-Δ*Lepr*::*egfp* based on the motif: PAM+seed=3+10=TAC+gttcctgagt, respectively), designed and chemically synthesized 24 primer pairs (Extended Data Table 2) for off-target-analysis based on the principle that the targeting specificity of Cas9 in genome editing could be tightly controlled by suitable selection of both spacer and PAM[39,42]. Fortunately, no off-target changes or indel-formation were found in the results of DNA gel electrophoresis or further DNA sequencing of the PCR products of potential off-target sites in the non-self-targeting genome editing conducted by *Svi*Cas3 (Extended Data Fig. 12). This may be because: (i) Cas9 lacks an R-loop locking to verify a complete R-loop formation to trigger cleavage, while Cas3 requires a locked full R-loop to trigger cleavage [32,43]; (ii) even in the presence of t-DNA, Cas9 still produces blunt-ended DSBs (an HNH-like nuclease domain to cleave TS and a RuvC-like nuclease domain to cleave NTS) [7,8,14], while *Svi*Cas3 probably generates SSBs (HD domain cleave NTS and TS in sequence) [10,16,17] and (iii) *Svi*Cas3 is also a DNA-guided endonuclease (locating D-loop), so the specificity of *Svi*Cas3 can be further enhanced by forming two large stable D-loops or multiple dynamic intermediate D-loops between t-DNA (because both UHA and DHA can be much longer than a spacer of approximately 40 bp) and target sequences [32,43]. In addition, these two facts that (i) Cascade is not needed in the RNA-guided genome editing conducted by *Svi*Cas3 alone and (ii) the subtype I-B-*Svi* CRISPR-Cas system lacks Cas8 (the PAM recognition

component) indirectly support the results that “no off-target effects or indel-formation were detected in target sequencing”.

Furthermore, as a comparison, we conducted a study on the off-target effects of genome editing mediated by *Sp*Cas9. From Extended Data Fig. 12, we can see that in the RNA-guided genome editing mediated by *Svi*Cas3, no off-target changes or indel-formation were detected, while in the RNA-guided genome editing mediated by *Sp*Cas9, two out of four randomly selected potential off-target sites were found to show off-target changes (8468 potential off-target sites in HEK293T-Δ*DROSHA*::*egfp*-Cas9 based on the motifs: seed+PAM=10+3=agactttgta+TGG, seed+PAM=12+3=ccagactttgta+AGG, seed+PAM=13+3=tacaagactttgt+AGG, and seed+PAM=14+3=gaccagactttgta+GGG). It is worth noting that in addition to the off-target effects of CRISPR-Cas editing per se (the specific genome editing protocol used), off-target effects might be cell-type specific, depending on the integration of DSBs in the genome by the HDR pathway of a particular cell type[42,44].

7. Signature Cas proteins in different types of CRISPR-Cas systems

The main biochemical characteristics of the typical signature proteins (all for DNA excision except Cas13a for RNA excision; note: Ngago-GDNA can also cleave RNA rather than DNA[45]) in six types of CRISPR-Cas systems excepting type IV are summarized in Extended Data Table 6 [3,46,47] (http://web.expasy.org/protparam; http://www.addgene.org/; http://www.ncbi.nlm.nih.gov;). We can see that the contents of cysteine in all the proteins are lower than 1% (0.1% to 0.8%; 2 to 8 cysteines) and

that the contents of negatively charged aspartic acid (5.5% to 7.5%) and glutamic acid (6.4% to 10.4%) in all the proteins vary by less than 40%. However, the contents of positively charged arginine (3.9% to 9.2%) and lysine (1.3% to 15.3%) in all the proteins vary greatly (more than 235%). To be more precise, the arginine content in *Svi*Cas3 (9.2%) is significantly higher than those in *Sp*Cas9 (5.6%) from *Streptococcus pyogenes*, *Cj*Cas9 (4.3%) from *Campylobacter jejuni*, *Ll*Cas10 (5.3%) from *Lactococcus lactis* subsp. Lactis, *As*Cas12a (4.4%) from *Acidaminococcus sp*. BV3L6 and *Ls*Cas13a (3.9%) from *Leptotrichia shahii* and the lysine content in *Svi*Cas3 (1.3%) is much lower than those in *Sp*Cas9 (11.0%), *Cj*Cas9 (14.7%), *Ll*Cas10 (7.5%), *As*Cas12a (9.0%) and *Ls*Cas13a (15.3%). Additionally, we can see that the asparaginate content in *Svi*Cas3 (1.2%) is much lower than those in *Sp*Cas9 (5.1%), *Cj*Cas9 (6.8%), *Ll*Cas10 (6.7%), *As*Cas12a (5.7%) and *Ls*Cas13a (10.7%). In addition, among the six proteins, *Svi*Cas3 has the highest contents of Ala (15.4%), Arg (9.2), Gly (7.7%), His (3.8), Leu (12.3%), Pro (5.7%), Trp (1.7%) and Val (7.1%) and the lowest contents of Asn (1.2), Glu (6.4%), Ile (3.2%), Lys (1.3%), Met (0.5%), Phe (3.4%), Ser (4.0%) and Tyr (2.2%), suggesting that *Svi*Cas3 probably has distinct functions from the other Cas signature proteins.

In particular, compared to *Sp*Cas9 (1368 aa, type II Cas system) (http://www.addgene.org/) and *As*Cas12a (1307 aa, type V Cas system) [47] that have been used for genome editing, we can see that the significant differences between *Svi*Cas3 and them are their MWs (*Svi*Cas3: 84352.40 Da, *Sp*Cas9: 158441.41 Da and *As*Cas12a: 151206.55 Da) and pIs (*Svi*Cas3: 5.78, *Sp*Cas9: 8.98 and *As*Cas12a: 8.01).

The MW of *Svi*Cas3 is slightly over half of the MWs of *Sp*Cas9 (53.2%) and *As*Cas12a (55.8%), even lower than the MW (73.4%) of the smallest *Cj*Cas9 (*Cj*Cas9: 984 aa, 114896.12 Da, pI 9.37) [46], which suggests that *Svi*Cas3 alone may be more easily loaded into vectors than *Sp*Cas9, *Cj*Cas9 and *As*Cas12a and is therefore suitable for genome editing in eukaryotic cells.

8. Main features of typical gene editing techniques

The targeting approaches of genome engineering based on site-specific recombination processes mainly include: recombinases (e.g., phi C31 integrase) [48,49], transposons (e.g., the ISY100 transposase) [50], homologous recombination HR [51,52] and NHEJ [53], among which the homologous gene targeting reported in yeast is one of the first methods for rational genome engineering (with an efficiency of only 10^{-6} to 10^{-9}) [54,55]. Generally, endonucleases used in genome engineering can be divided into five groups [56,57]: meganucleases (also known as homing endonucleases) [56,58,59], chemical nucleases (synthetic endonucleases resulting from the fusion of DNA-reactive agents to a DNA-binding polymer) [60], zinc finger nucleases (ZFNs) [61-65], transcription activator like effector nucleases (TALENs) [37,66,67] and CRISPR-Cas enzymes [5,19,68-70], which mainly introduce DSB or SSB at the target DNA site to stimulate the cell's own DNA repair pathways to make DNA modifications (HDR: conservative homology-directed repair in the presence of t-DNA; NHEJ: error-prone non-homologous end joining in the absence of t-DNA).

Meganucleases (I-SceI meganuclease, the first reported for meganuclease used in mammalian cells) [58,71] can be used as a tool for targeted genome engineering, but their disadvantages, such as site-sequence degeneracy, homodimeric nature, the conversion of sequences adjacent to a DNA break, and the need to first introduce a known cleavage site into the region of interest, hinder their widespread use in genome editing [56,58,71]. Although its off-target effect and relatively large molecular weight limit its practical application, the most commonly used genome editing tool is still the Cas9-based gene editing system based on RNA-guided mechanisms [72,73]. The main features of typical gene editing techniques in three classes of genome editing based on the principle of DNA-orientation are summarized in Extended Data Table 7.

As seen from Extended Data Table 7, genome editing technology based on *Svi*Cas3 has obvious advantages over other technologies, such as free selection of target sequences, simple operation and no off-target changes detected. In particular, t-DNA without the PAM requirement acts as both a template and a guide, suggesting that CRISPR, as the basis of crRNA design, is unnecessary in genome editing directed by *Svi*Cas3. Furthermore, we demonstrated that the DNA-guided *Svi*Cas3 enzyme can simultaneously perform both gene editing and base editing in the genomic DNA of mammalian cells.

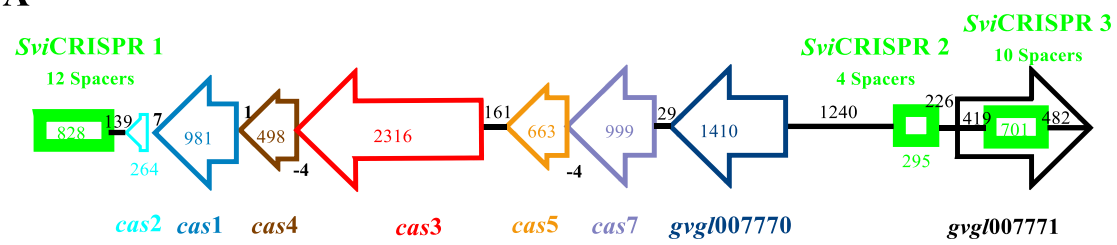
In RNA-guided genome editing directed by *Svi*Cas3, a gene editing plasmid, in addition to the essential components of the plasmid itself, should contain at least one *Svicas*3 gene, one g-DNA fragment and one t-DNA fragment for HDR. However, in DNA-guided genome editing conducted by *Svi*Cas3, g-DNA is not needed because t-

DNA serves as both an editing template and a DNA recognition guide. After a gene editing plasmid (carrying *Svicas*3 and t-DNA) enters host cells, the consequent molecular events that occur are: (i) duplication, transcription and translation of the *Svicas*3 gene, (ii) site-specific cleavage of genomic DNA by *Svi*Cas3 and (iii) homologous recombination between target DNA and t-DNA by HDR machinery. Events (i) and (iii) are wholly controlled by molecular components of plasmid(s) and host cells, and only event (ii) should be performed by *Svi*Cas3.

In addition, compared with commercial *Sp*Cas9, *Svi*Cas3 has the following advantages: (i) a small molecular weight suitable for vector carrying (*Svicas*3/2316 bp, *Spcas*9/4107 bp) and protein expression (*Svi*Cas3/84352.40 Da, *Sp*Cas9/158441.41 Da); (ii) similar to DNA and RNA, an acidic isoelectric point (*Svi*Cas3/pI5.78, *Sp*Cas9/pI8.98) carrying negative charges at normal physiological pH; (iii) a bacterial protein of non-animal origin with a low possibility of eliciting the *Svi*Cas3 antibody; and (iv) in particular, no off-target changes or indel (insertion and deletion) formation detected in the analytical experiments of potential off-target sites.

Extended Data Figs. 1 to 12

A



*Svi*CRISPR1: 96665-97492=828 bp; ***cas2***: 97632-97895=264 bp; ***cas1***: 97903-98883=981 bp; ***cas4***: 98885-99382=498 bp; ***cas3***: 99379-101694=2316 bp; ***cas5***:101856-102518=663 bp; ***cas7***: 102515-103513=999 bp; ***gvgl007770/cas6***: 103543:104952=1410 bp; *Svi*CRISPR2: 106193-106487=295 bp; *gvgl*L007771: 106714:108315=1602 bp; *Svi*CRISPR3: 107133-107833=701 bp; arrow: transcriptional direction.

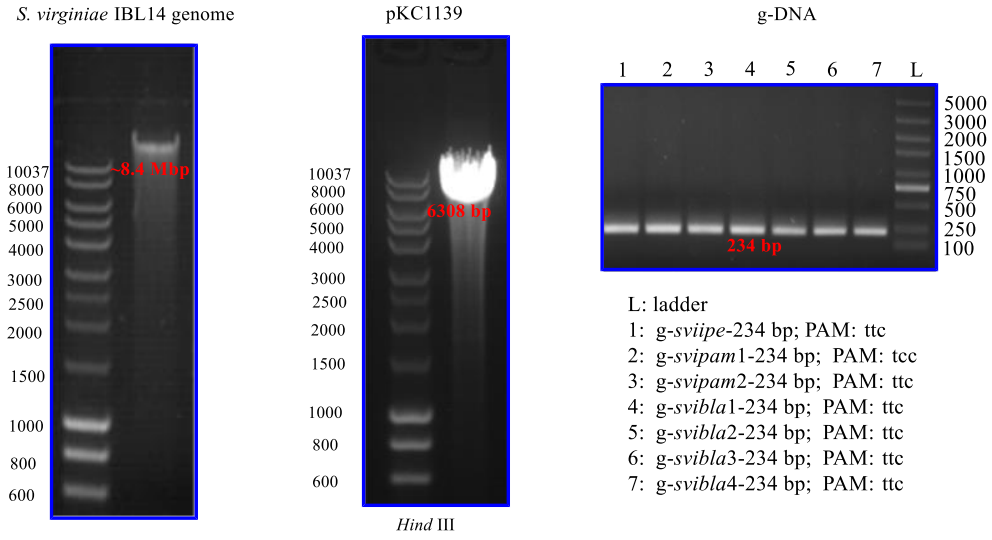
B

```
1         10        20        30        40        50        60        70
VGRLDAVEDVFGGRFWPVVELAGLTHDAGKIPEGFQRMLAGYSRAWGERHEVASLGFLPALIGDPDVLLW
ccccccceeeccccccceeeeeccccccccchhhhhhhhhhhhhhhccchhhhhhhccccccccccceeeee
VATAVATHHRPLTGQNGRDLQTLYSGVTITELAHRFGPFDPRAVPALEAWLRASAIRVGLPAAAVPDDGT
eeeeeeccccccccccccceeeeeccceeehhhhccccccccccchhhhhhhhhhhhhhccccccccccccc
LTDTGVVAGAHQLLEEILDRWADRVRPEVGLAAVLLQGAVTLADHLSSAHQALPTVQPLGAGFRSRLEKE
ecccchhhhhhhhhhhhhhhhhhhcccchhhhhhhhhhhhhhhhhhhhhhhhhhhccccccccccchhhhhhh
FAERGRTLRAHQLEAATVTGHLLLRGPTGSGKTEAALLWAASQVEALKAEGRGVPRVFFTLPYLASINAM
hhhhhhhhhhhhhhhhhhceeeeccccccchhhhhhhhhhhhhhhhhhhhcccccceeeeeeccchhhhhh
ATRLGDTLGDGEAVGVAHSRAASYHLAQAIAPQDGDEEDEHGAPCRVDAAAKALSRAAATKLFRESVRVA
hhhhccccccceeeehhhhhhhhhhhhhhhhcccccccccccccchhhhhhhhhhhhhhhhhhhhhhcccccc
TPYQLLRAALAGPAHSGILIDAANSVFILDELHAYDARRLGYILASARLWERLGGRITVLSATLPRALAD
chhhhhhhhhccccccchhhhhhchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhccceeeeeeccchhhhhh
LFESTLTAPITFLDTPDLGLPARHLLHTRGHHLTDPATLEEIRLRLSRDESVLVIANNVSQAIALYEQLA
hhhccccccceeeccccccccchhhhhhhhccccccchhhhhhhhhhhccccceeeeeecchhhhhhhhhhcc
PDVCERFGQDAALLLHSRFRRMDRSRIEQKIADRFATVAPDAQNSRKPGLVVATQVVEVSLDVDFDVLFT
cchhhhhccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhccccccccccceeeeeeeeeeccccchhhhhc
GAAPLEALLQRFGRTNRVGARPPADVIVHHPAWTTRRRQPGEYADGIYPREPVESAWHILTRNHGRVIDE
ccchhhhhhhhccccccccccccceeeccccccccccccccccccccccccchhhhhhccccchhhh
ADATAWLDEVYATDWGRQWHREVLERRERFDRAFLQFRYPFEDRTDLADTFDELFDGSEAILAEDQDAYS
hhhhhhhceeeeccccchhhhhhhhhhhhhhhhhhhcccccccccchhhhhhhhccchhhhhhhhhhhhh
AALAAPDGDHPGAGRLLAEEYLIPVPHWASPLSRYEKQLKVRVINGDYHPDHGLMAVRGLPQPAYRAGEV
hhhhccccccccchhhhhhhhhhhcccccccchhhhhhhheeeeeccccccccchhhhhcccccccccccee
L
e
```

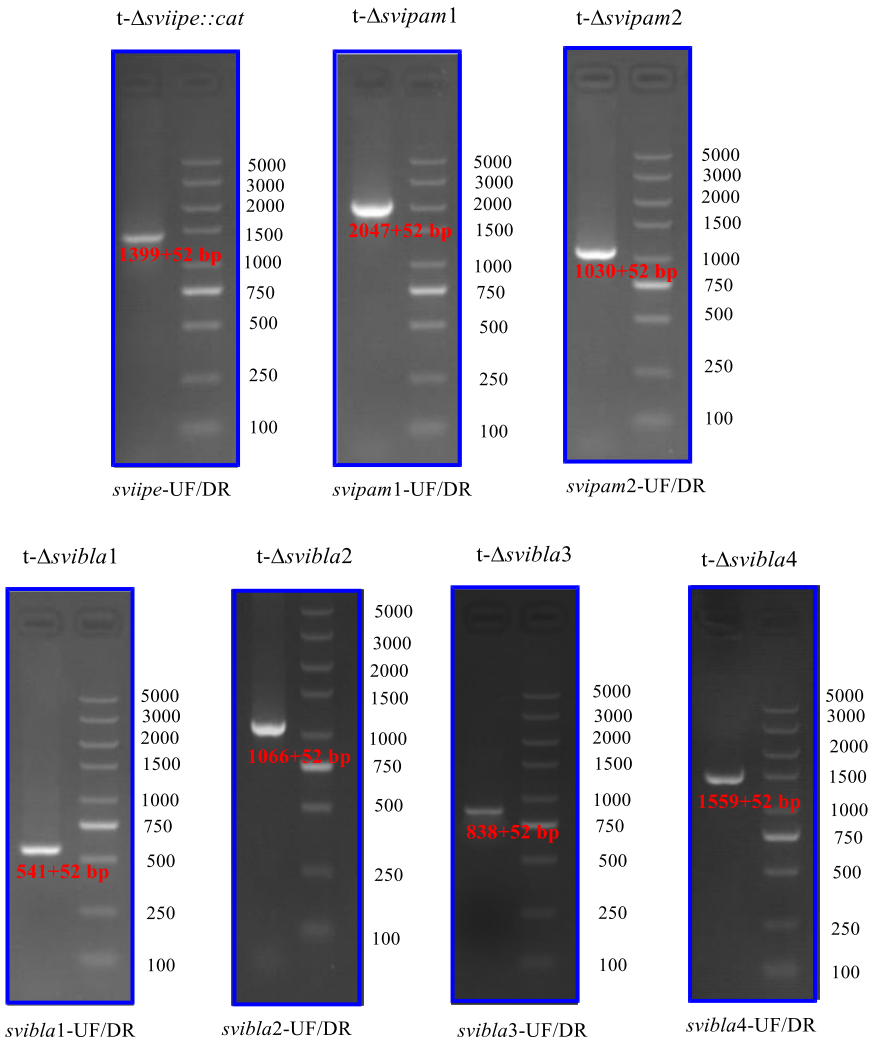
c: random coil, e: extended strand, h: alpha helix, letters with red underline represent HD domain, letters with red dash-line represent Hel domain

Extended Data Fig. 1. The architecture of the subtype I-B-*Svi* CRISPR-Cas system in strain *S. virginiae* IBL14 and the predicted secondary structure of *Svi*Cas3 in the strain IBL14. (A) The architectural diagram of the subtype I-B-*Svi* CRISPR-Cas system loci. **(B)** The predicted secondary structure of *Svi*Cas3.

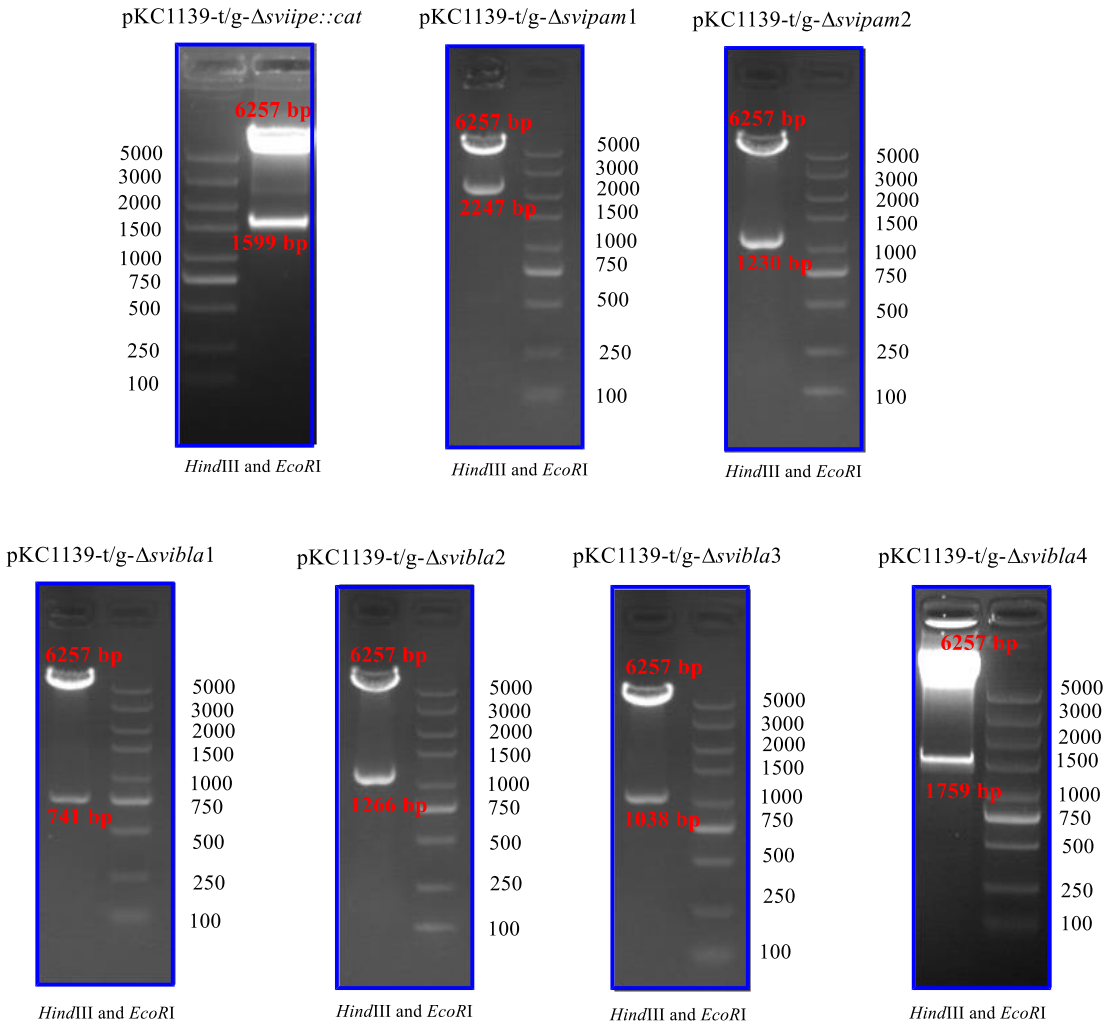
A


S. virginiae IBL14 genome
~8.4 Mbp
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
pKC1139
6308 bp
Hind III
g-DNA
1 2 3 4 5 6 7 L
234 bp
5000
3000
2000
1500
1000
750
500
250
100
L: ladder
1: g-sviipe-234 bp; PAM: ttc
2: g-svipam1-234 bp; PAM: tcc
3: g-svipam2-234 bp; PAM: ttc
4: g-svibla1-234 bp; PAM: ttc
5: g-svibla2-234 bp; PAM: ttc
6: g-svibla3-234 bp; PAM: ttc
7: g-svibla4-234 bp; PAM: ttc

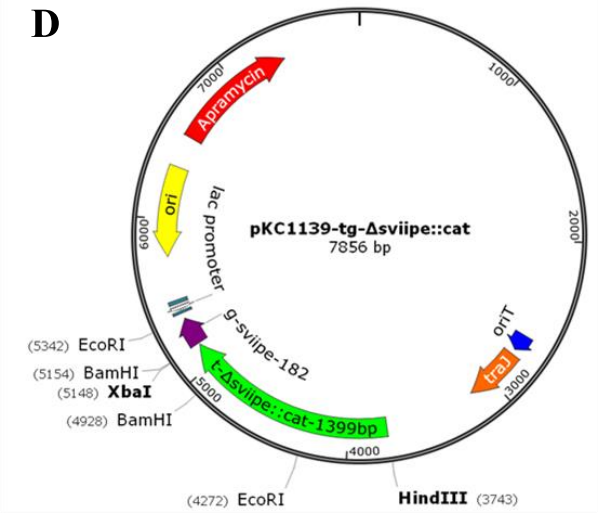
B


t-Δsviipe::cat
1399+52 bp
sviipe-UF/DR
t-Δsvipam1
2047+52 bp
svipam1-UF/DR
t-Δsvipam2
1030+52 bp
svipam2-UF/DR
t-Δsvibla1
541+52 bp
svibla1-UF/DR
t-Δsvibla2
1066+52 bp
svibla2-UF/DR
t-Δsvibla3
838+52 bp
svibla3-UF/DR
t-Δsvibla4
1559+52 bp
svibla4-UF/DR
5000
3000
2000
1500
1000
750
500
250
100

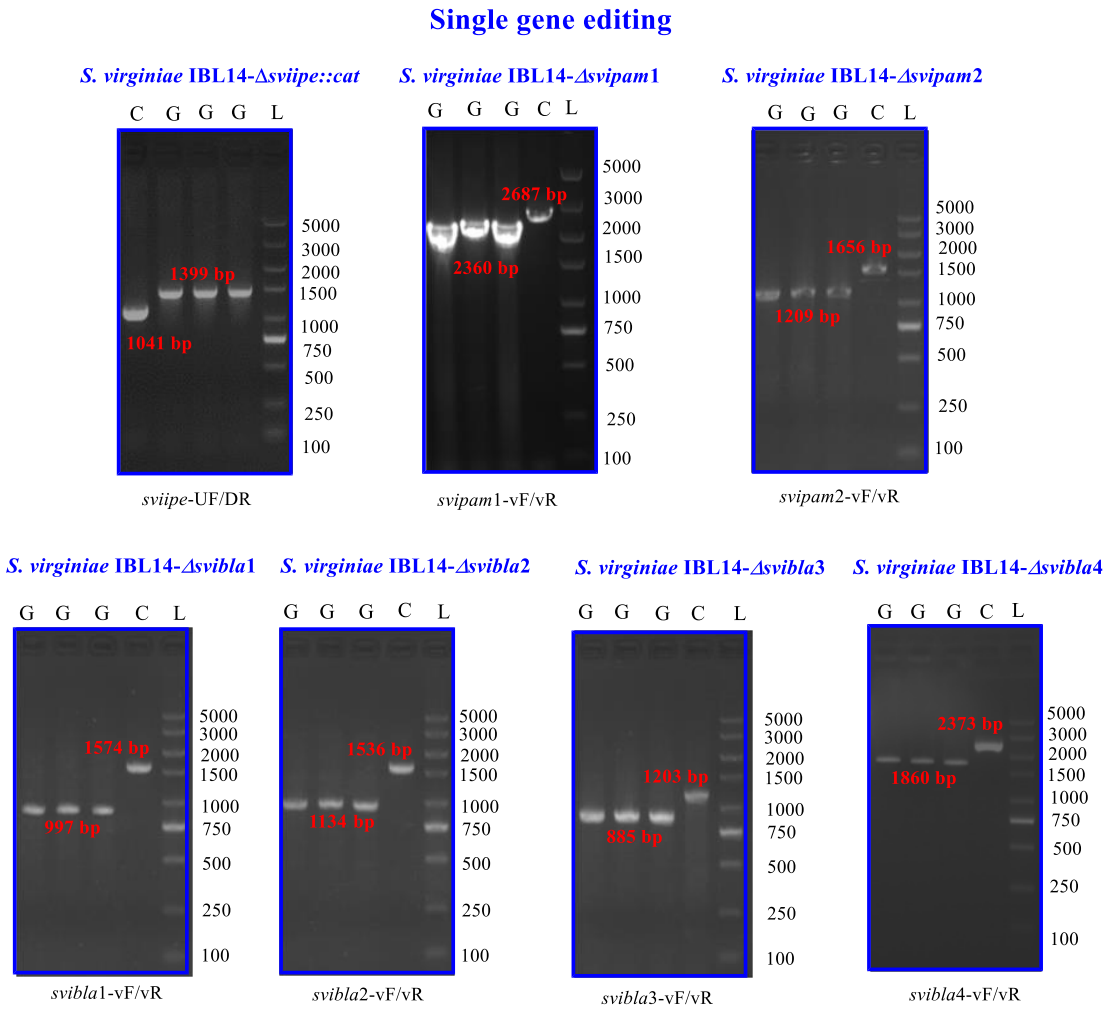
C


pKC1139-t/g-Δsviipe::cat
6257 bp
1599 bp
5000
3000
2000
1500
1000
750
500
250
100
HindIII and EcoRI
pKC1139-t/g-Δsvipam1
6257 bp
2247 bp
HindIII and EcoRI
pKC1139-t/g-Δsvipam2
6257 bp
1230 bp
HindIII and EcoRI
pKC1139-t/g-Δsvibla1
6257 bp
741 bp
HindIII and EcoRI
pKC1139-t/g-Δsvibla2
6257 bp
1266 bp
HindIII and EcoRI
pKC1139-t/g-Δsvibla3
6257 bp
1038 bp
HindIII and EcoRI
pKC1139-t/g-Δsvibla4
6257 bp
1759 bp
HindIII and EcoRI

D

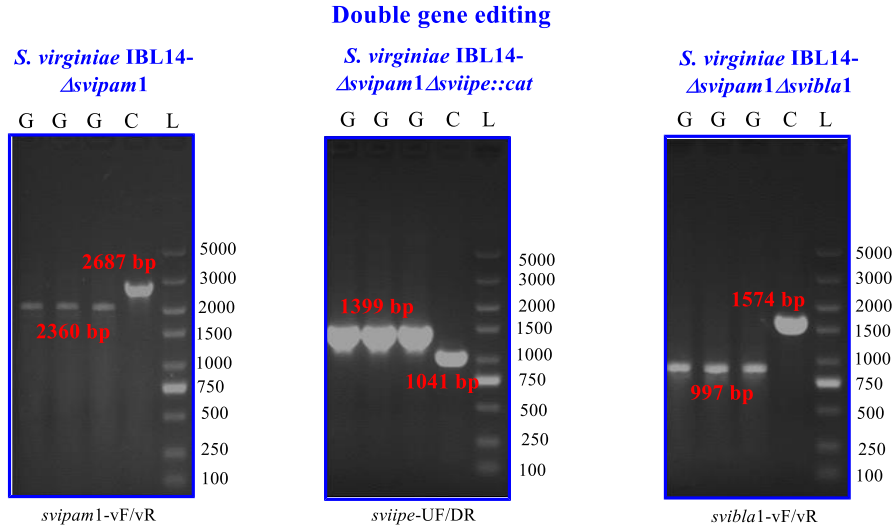

pKC1139-tg-Δsviipe::cat
7856 bp
Apramycin
ori
lac promoter
g-sviipe-182
t-Δsviipe::cat-1399bp
oriT
traJ
(5342) EcoRI
(5154) BamHI
(5148) XbaI
(4928) BamHI
(4272) EcoRI
HindIII (3743)

E

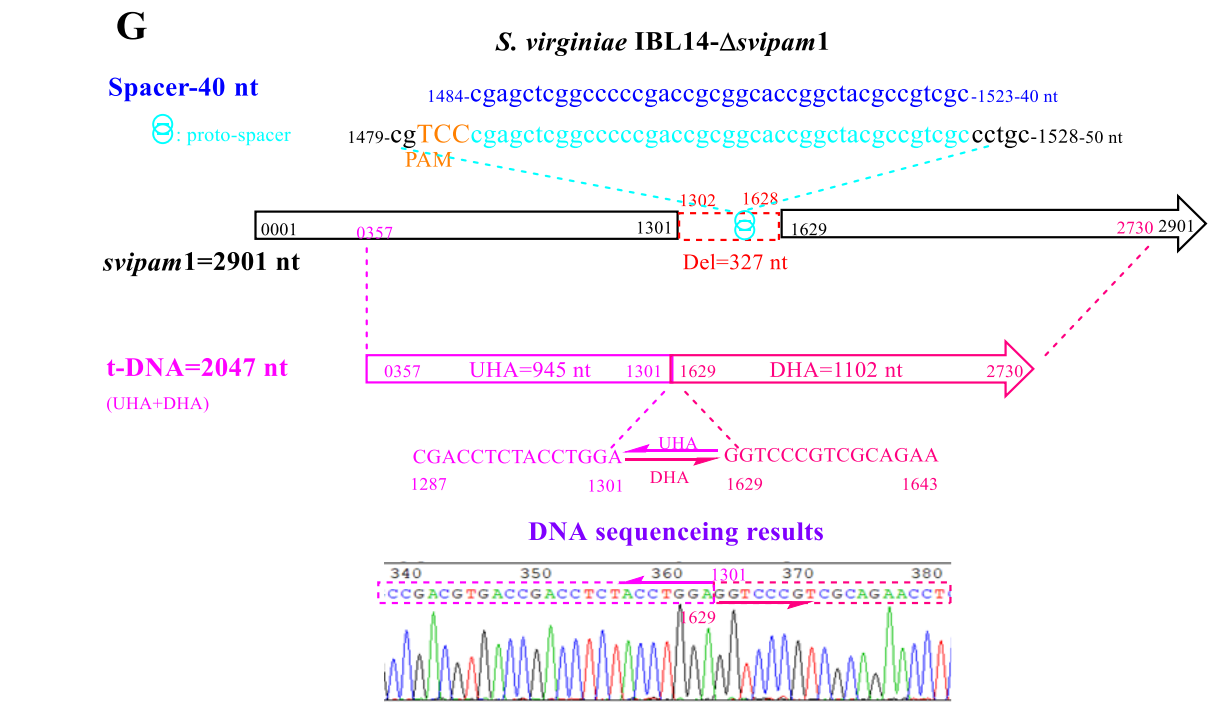


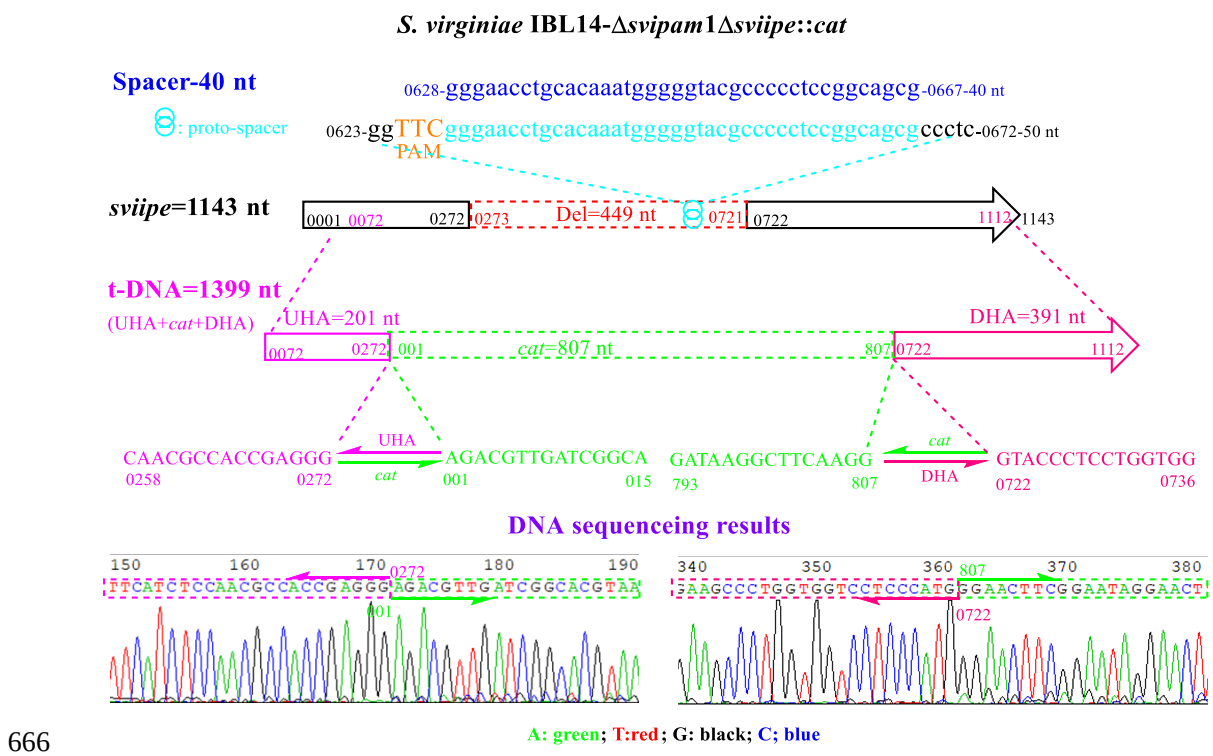
C: control- *S. virginiae* IBL14-pKC1139; G: gene-edited mutant; L: ladder

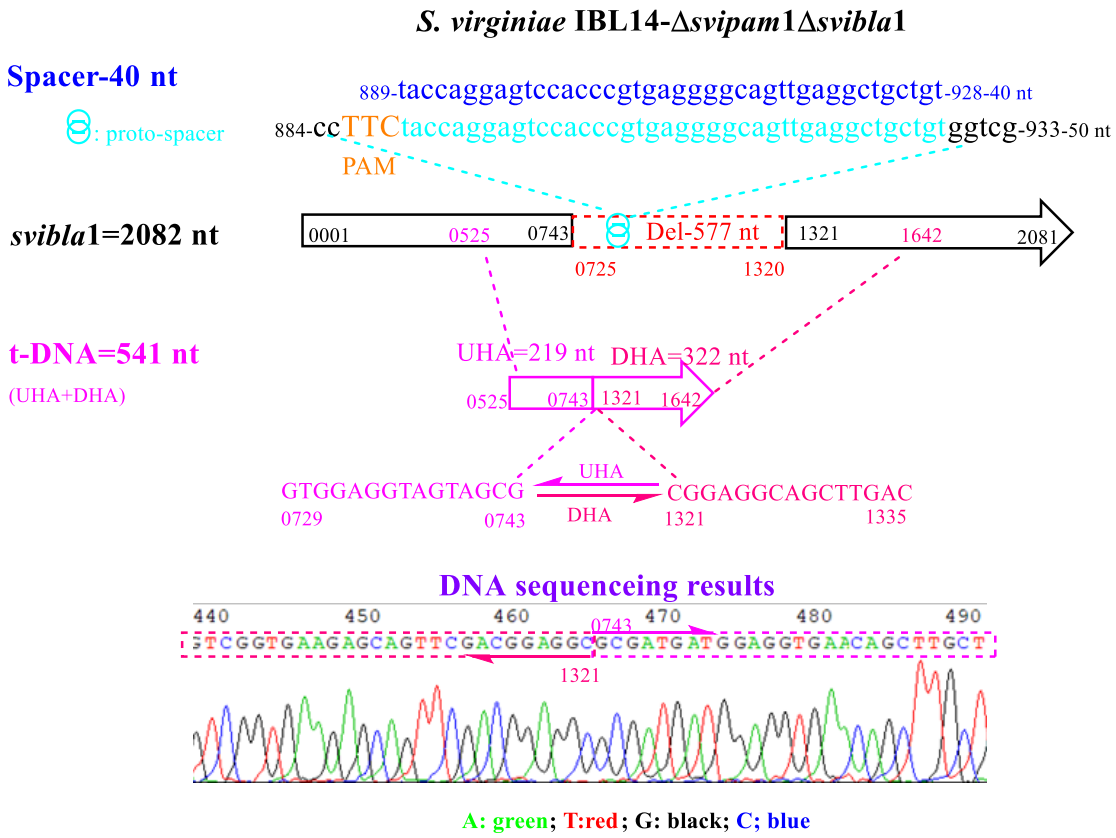
F



C: control-*S. virginiae* IBL14-pKC1139; G: gene-edited mutant; L: ladder


G
S. virginiae IBL14-Δsvipam1
Spacer-40 nt
: proto-spacer
1484-cgagctcggcccccgaccgcggcaccggctacgccgtcgc-1523-40 nt
1479-cgTCCcgagctcggcccccgaccgcggcaccggctacgccgtcgccctgc-1528-50 nt
PAM
svipam1=2901 nt
Del=327 nt
t-DNA=2047 nt
(UHA+DHA)
UHA=945 nt
DHA=1102 nt
CGACCTCTACCTGGA
GGTCCCGTCGCAGAA
UHA
DHA
DNA sequenceing results
A: green; T:red; G: black; C; blue


S. virginiae IBL14-Δsvipam1Δsviipe::cat
Spacer-40 nt
: proto-spacer
0628-gggaacctgcacaaatgggggtacgccccctccggcagcg-0667-40 nt
0623-ggTTCgggaacctgcacaaatgggggtacgccccctccggcagcgccctc-0672-50 nt
PAM
sviipe=1143 nt
Del=449 nt
t-DNA=1399 nt
(UHA+cat+DHA)
UHA=201 nt
cat=807 nt
DHA=391 nt
CAACGCCACCGAGGG
AGACGTTGATCGGCA
GATAAGGCTTCAAGG
GTACCCTCCTGGTGG
UHA
cat
DHA
DNA sequenceing results
A: green; T:red; G: black; C; blue



668

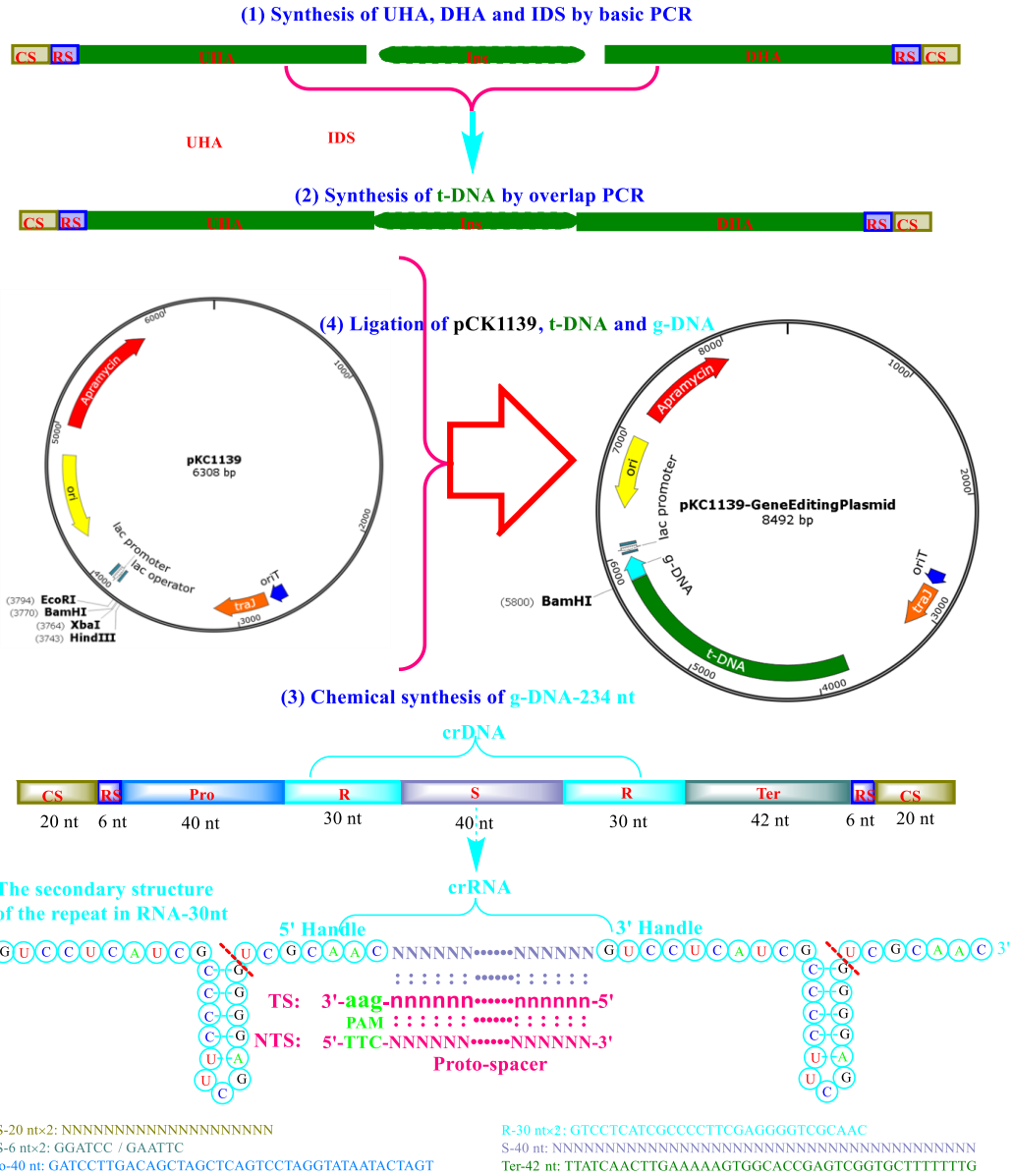
669

670

671 **Extended Data Fig. 2. Construction of gene editing plasmids and the RNA-guided**
672 **self-targeting genome editing of *S. virginiae* IBL14. (A)** The extracted *S. virginiae*
673 IBL14 genome, linearized pKC1139 by *Hind*III and chemically synthesized g-DNA
674 fragments. **(B)** The t-DNA fragments amplified by overlap PCR. **(C)** The results of the
675 constructed gene editing plasmids digested by *Hind*III and *Eco*RI. **(D)** The map of gene
676 editing plasmid pKC1139-t/g-Δ*sviipe*::*cat* (In the self-targeting senome editing of *S.*
677 *virginiae* IBL14, other six gene editing plasmids except the corresponding g-DNAs and
678 t-DNAs were the same as this map). **(E)** DNA gel electrophoresis of the PCR products
679 of edited sequences in single gene-edited mutants. **(F)** DNA gel electrophoresis of the
680 PCR products of edited sequences in double gene-edited mutants. **(G)** The selected
681 proto-spacers, the engineered t-DNAs as well as the DNA sequencing results of the

682 double gene-edited mutants (The junction sequences between UHA and DHA, UHA
683 and *cat*, and *cat* and DHA are highlighted in Extended Data Fig. 2G).
684

A ***Gene editing plasmid construction process***



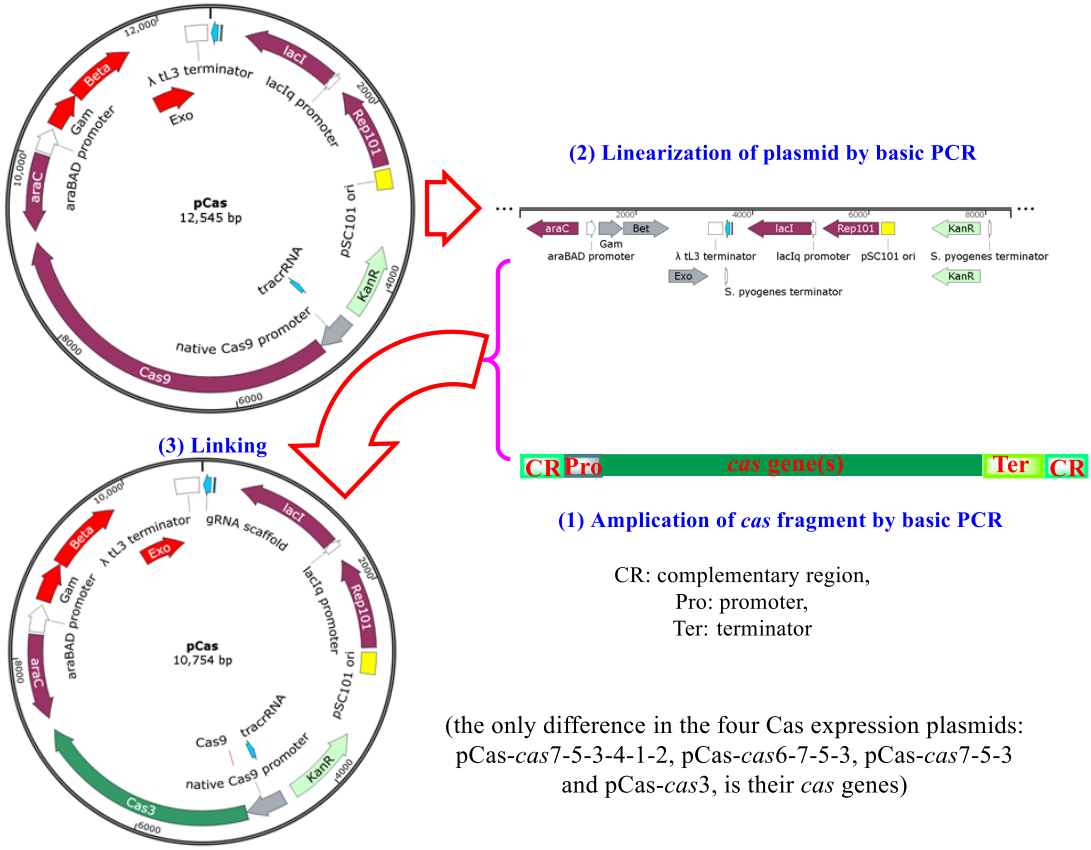
CS: complementary sequence
DHA: downstream homologous arm
Ins: Inserted DNA sequence
NTS: non-target strand DNA
Pro: promoter-pJ23119
R: repeat-the CRISPR3
RS: restriction site
S: spacer
TS: target strand DNA
Ter: terminator-*S. pyogenes*
UHA: upstream homologous arm

685

686

B

Cas expression plasmid construction process

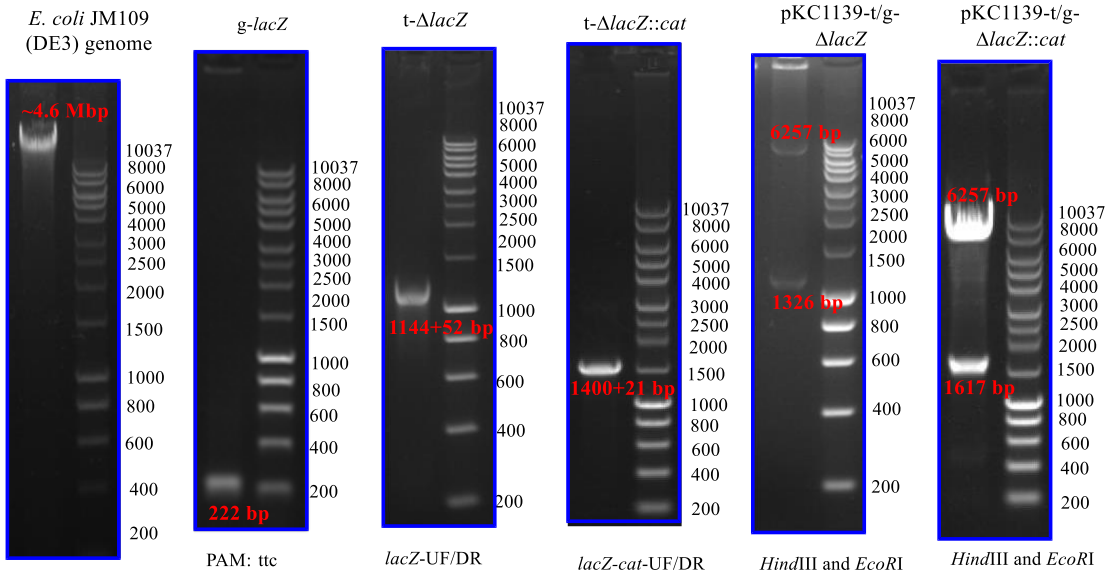


CR: complementary region,
Pro: promoter,
Ter: terminator

(the only difference in the four Cas expression plasmids: pCas-*cas*7-5-3-4-1-2, pCas-*cas*6-7-5-3, pCas-*cas*7-5-3 and pCas-*cas*3, is their *cas* genes)

687

688 **C**

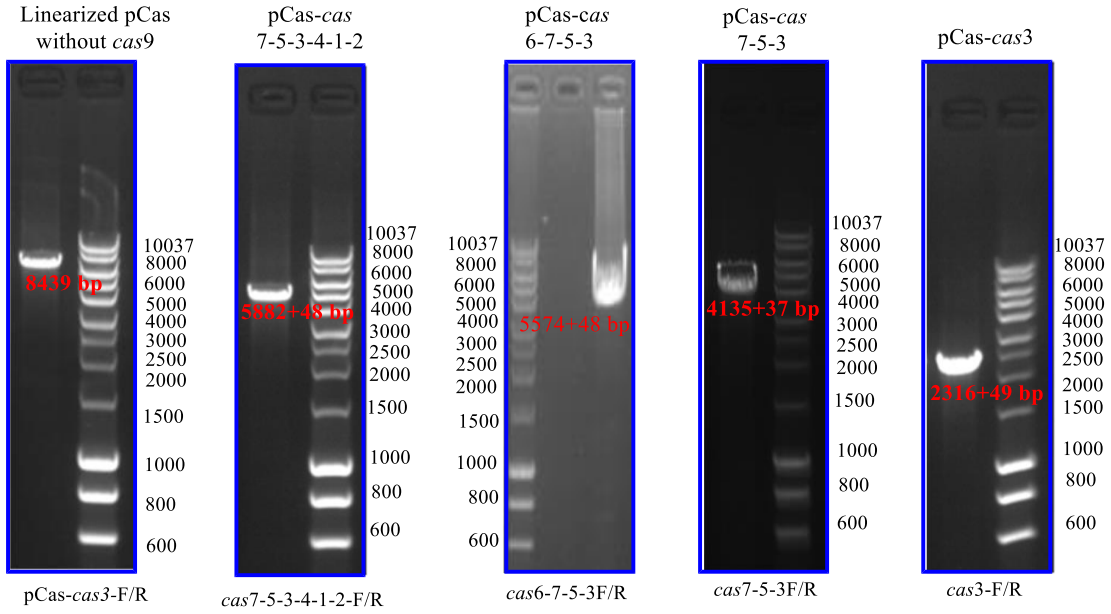


689

690

691

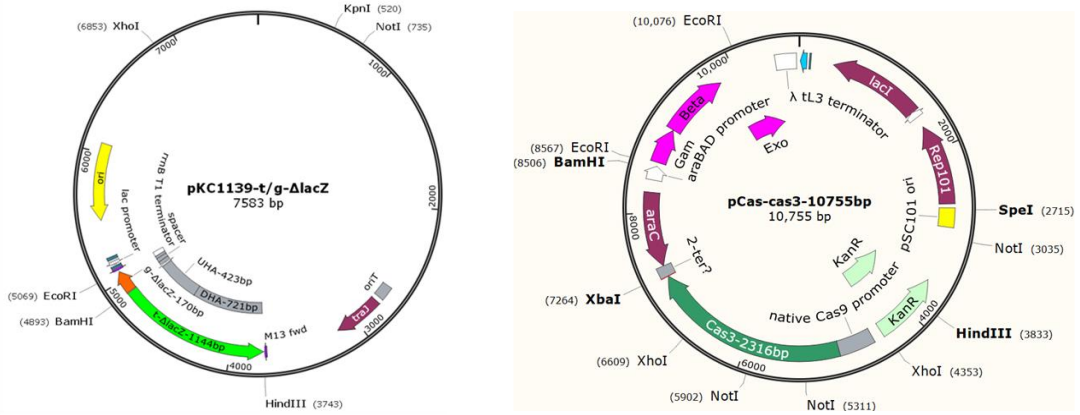
692 D



693

694

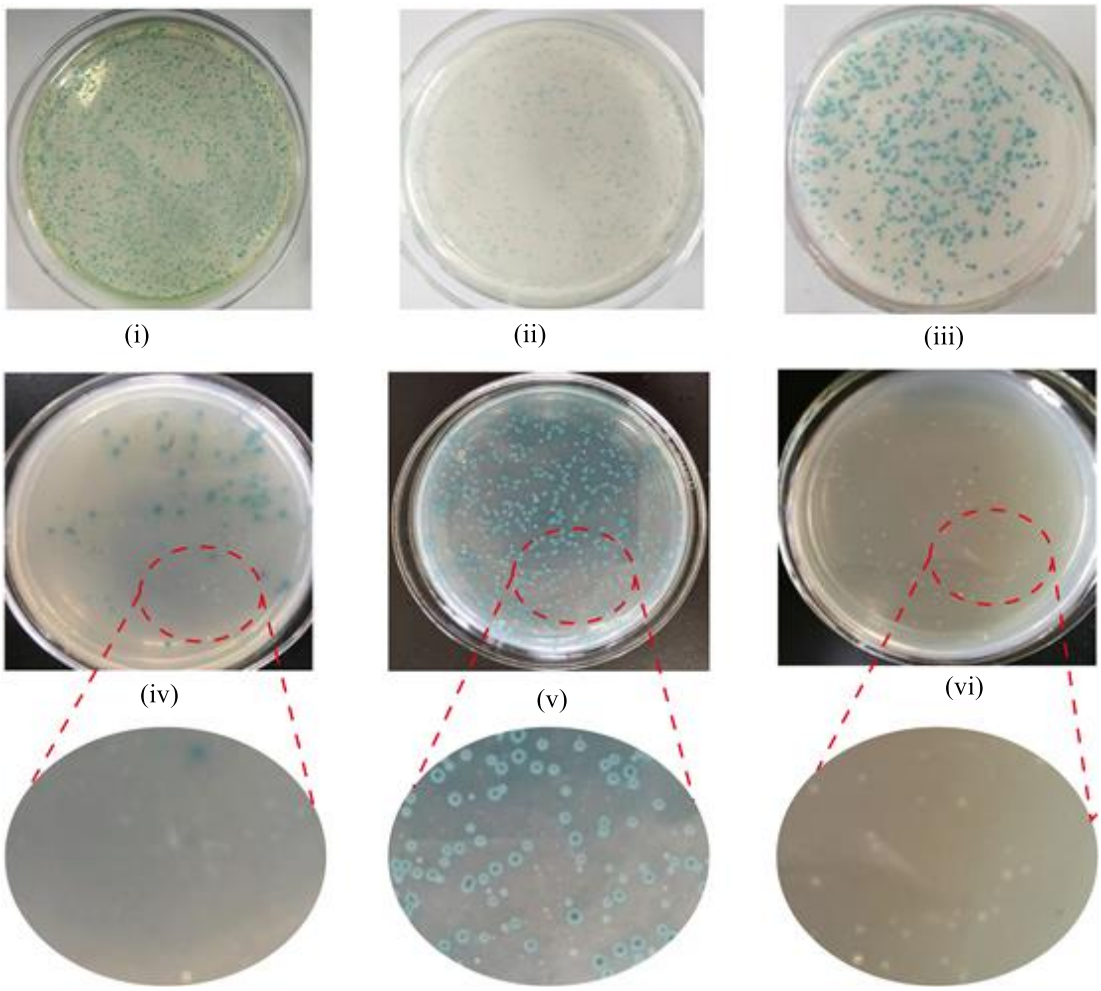
E



695

696

F



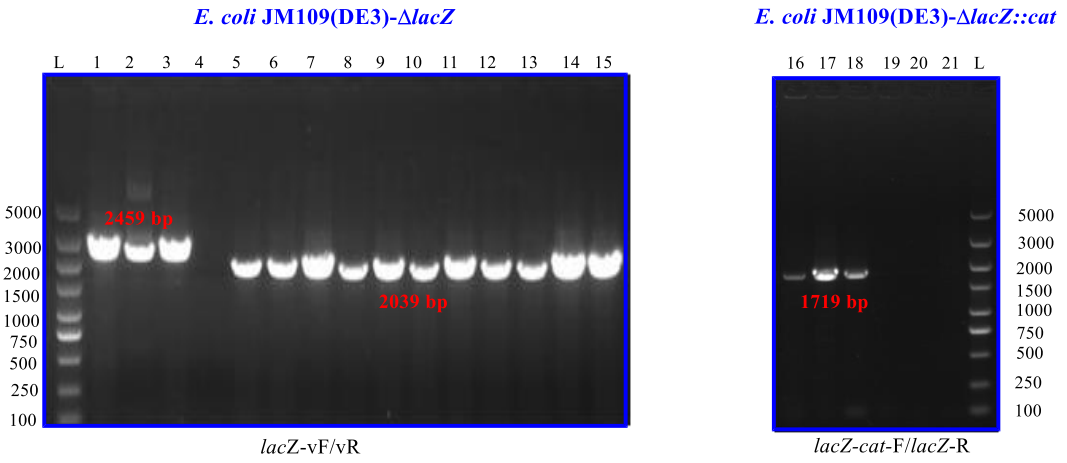
i-iii: *E. coli* JM109 (DE3) or its single plasmid transformants by pKC1139-t/g-Δ*lacZ*, pKC1139-t/g-Δ*lacZ*::*cat*, pCas-*cas*7-5-3-4-1-2, pCas-*cas*6-7-5-3, pCas-*cas*7-5-3 or pCas-*cas*3.
iv-vi: double plasmid transformants by pKC1139-t/g-Δ*lac*Z plus pCas-*cas*7-5-3-4-1-2, pCas-*cas*6-7-5-3, pCas-*cas*7-5-3 or pCas-*cas*3, respectively.

697

698

699

700 G



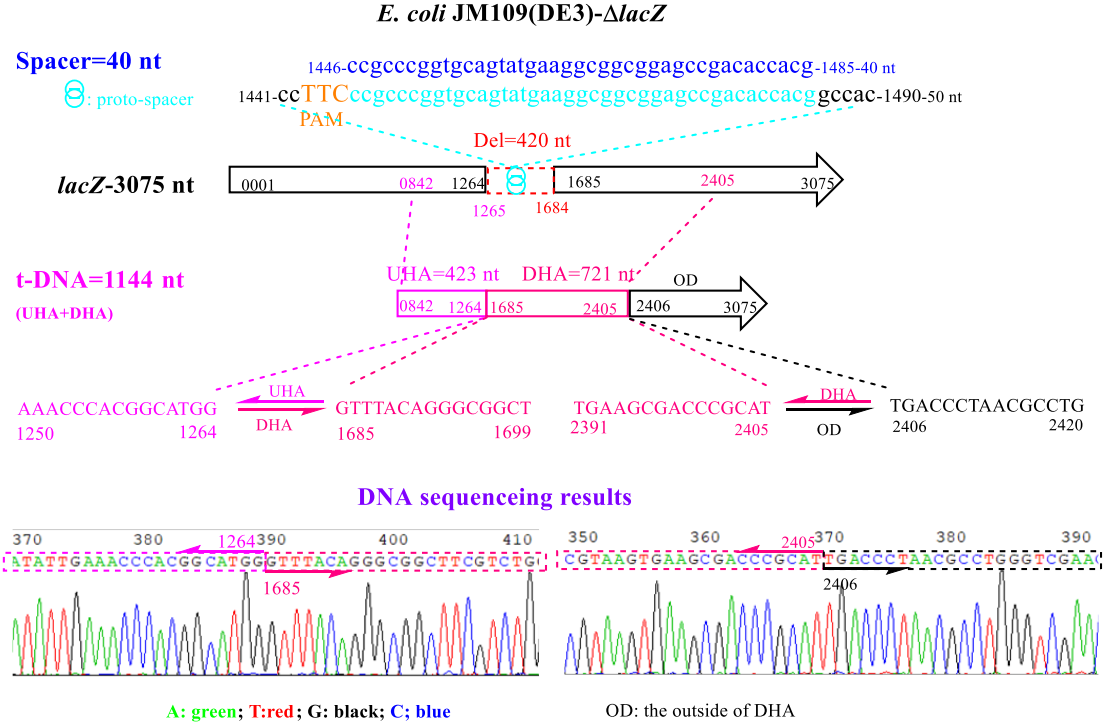
L: ladder
1, 21: *E. coli* JM109(DE3)
2: *E. coli* JM109(DE3) transformant by pKC1139-t/g-Δ*lacZ*
3, 20: *E. coli* JM109(DE3) transformant by pCas-*cas*3
4: ddH_2O
5-7: *E. coli* JM109(DE3)-Δ*lacZ* by pCas-*cas*3+pKC1139-t/g-Δ*lacZ*
8-10: *E. coli* JM109(DE3)-Δ*lacZ* by pCas-*cas*7-5-3+pKC1139-t/g-Δ*lacZ*
11-13:*E. coli* JM109(DE3)-Δ*lacZ* by pCas-*cas*7-5-3-4-1-2+pKC1139-t/g-Δ*lacZ*
14-15: *E. coli* JM109(DE3)-Δ*lacZ* by pCas-*cas*6-7-5-3+pKC1139-t/g-Δ*lacZ*
16-18: *E. coli* JM109(DE3)-Δ*lacZ*::*cat* by pCas-*cas*3+pKC1139-t/g-Δ*lacZ*::*cat*
19: *E. coli* JM109(DE3) transformant by pKC1139-t/g-Δ*lacZ*::*cat*

* 14-15: *E. coli* JM109(DE3)-Δ*lacZ* by pCas-*cas*6-7-5-3+pKC1139-t/g-Δ*lacZ*, where *cas*6 temporarily stands for *gvgl*007770, indicating that the product Gvgl007770 of gene *gvgl*007770 has no side effect in the genome editing of strain *E. coli* JM109(DE3).

701

702

H

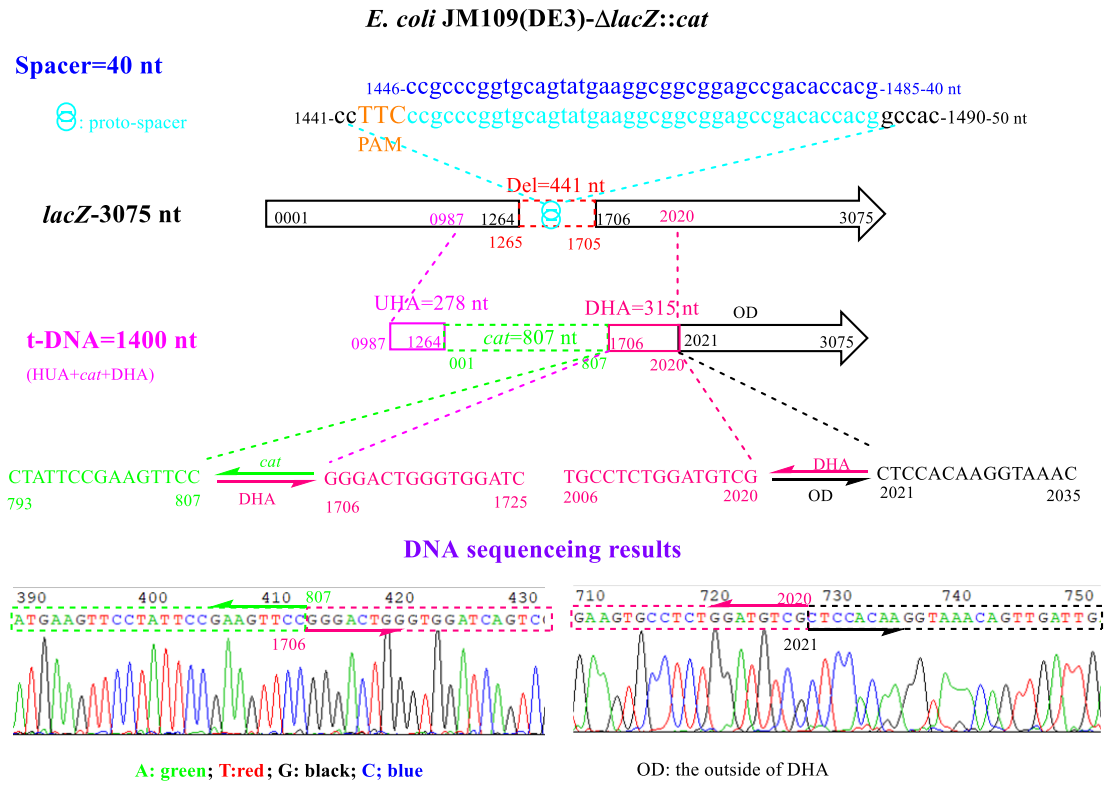


A: green; T:red; G: black; C; blue

OD: the outside of DHA

703

704



705

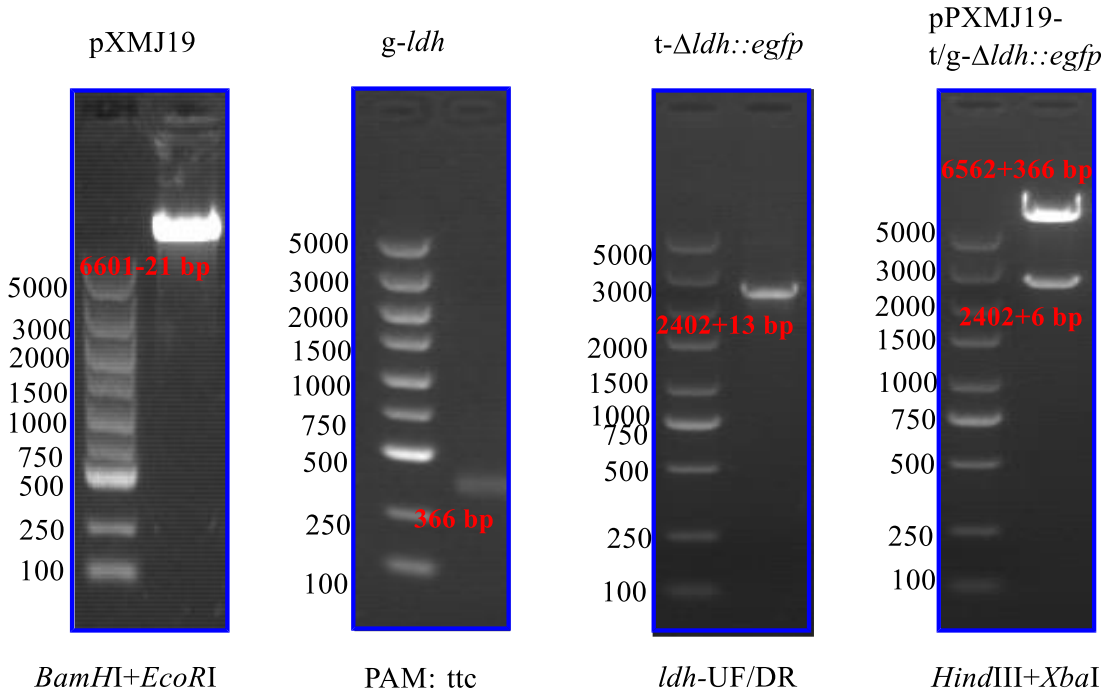
706

707

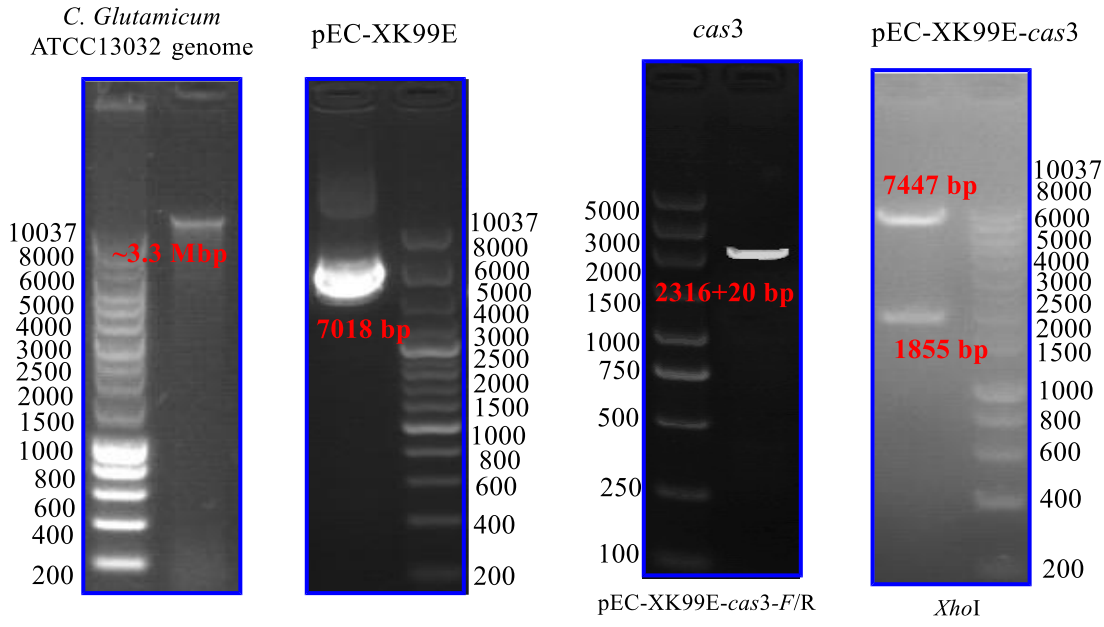
708 **Extended Data Fig. 3. Construction of gene editing tools and the RNA-guided**
709 **genome editing of *E. coli* JM109 (DE3) based on the subtype I-B-*Svi* CRISPR-CAS**
710 **system. (A)** A schematic diagram of gene editing plasmid construction process. **(B)** A
711 schematic diagram of Cas expression plasmid construction process. **(C)** The extracted
712 *E. coli* JM109 (DE3) genome, chemically synthesized g-*lacZ*, overlap PCR amplified
713 t-Δ*lacZ* and t-Δ*lacZ*::*cat* and the results of gene editing plasmids pKC1139-t/g-Δ*lacZ*
714 and pKC1139-t/g-Δ*lacZ*::*cat* digested by *Hind*Ⅲ and *Eco*RI, respectively. **(D)** The
715 linearized pCas without *cas*9 by PCR and the fragments of genes *cas*7-5-3-4-1-2, *cas*6-
716 7-5-3, *cas*7-5-3 and *cas*3 amplified by PCR from pCas-*cas*7-5-3-4-1-2, pCas-*cas*6-7-5-
717 3, pCas-*cas*7-5-3 and pCas-*cas*3, respectively. **(E)** The maps of gene editing plasmid
718 pKC1139-*lac*Z and Cas expression plasmid pCas-*cas*3 (In the genome editing of *E. coli*
719 JM109 (DE3), other gene editing tools are the same as these two maps except for the

corresponding t/g-DNAs and the *cas* genes). **(F)** Typical features of both host *E. coli* JM109(DE3) and its transformants on LBPETs. **(G)** DNA gel electrophoresis of the PCR products of edited sequences in the genome editing of *E. coli* JM109(DE3) (Lanes 16-21: PCR products occur only while the genome of the gene-edited mutant *E. coli* JM109(DE3)-Δ*lacZ*::*cat* is utilized as template because the primer *lac*Z-*cat*-F is a portion of the gene *cat* in t-Δ*lacZ*::*cat* and the primer *lacZ*-R is part of the gene *lacZ* but is not included in t-Δ*lacZ*::*cat*.). **(H)** The selected proto-spacers, the engineered t-DNAs as well as the DNA sequencing results of the PCR products of edited sequences from the two gene-edited mutants *E. coli* JM109 (DE3)-Δ*lacZ* and *E. coli* JM109 (DE3)-Δ*lacZ*::*cat* genomes, respectively (The junction sequences between UHA and DHA, *cat* and DHA, and DHA and the outside of DHA are highlighted in Extended Data Fig. 3H).

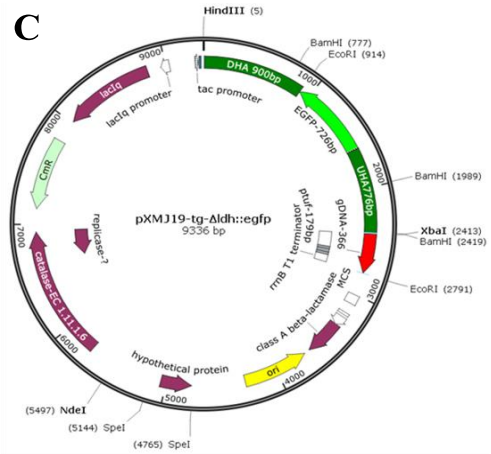
A

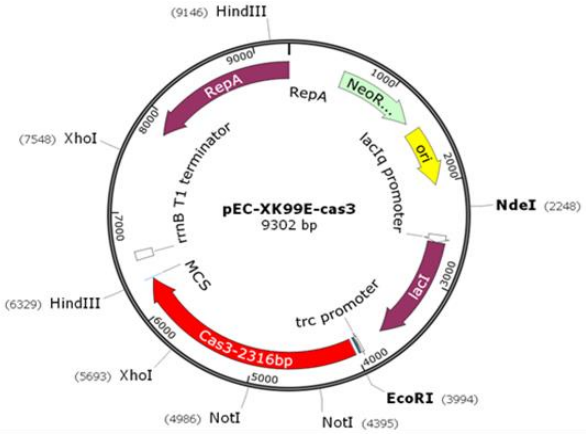

pXMJ19
g-*ldh*
t-Δ*ldh::egfp*
pPXMJ19-t/g-Δ*ldh::egfp*
6601-21 bp
366 bp
2402+13 bp
6562+366 bp
2402+6 bp
5000
3000
2000
1500
1000
750
500
250
100
*BamH*I+*EcoR*I
PAM: ttc
ldh-UF/DR
*Hind*III+*Xba*I

B


C. Glutamicum ATCC13032 genome
pEC-XK99E
*cas*3
pEC-XK99E-*cas*3
~3.3 Mbp
7018 bp
2316+20 bp
7447 bp
1855 bp
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
pEC-XK99E-*cas*3-*F*/R
*Xho*I

C

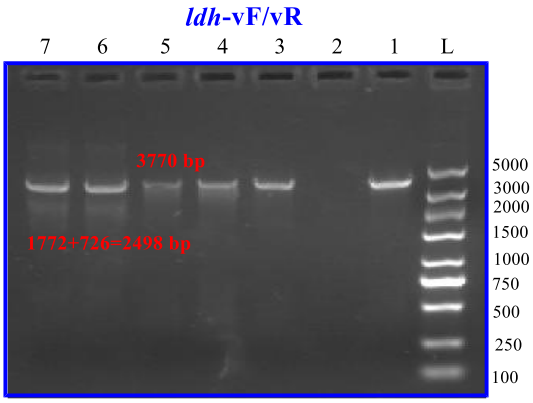




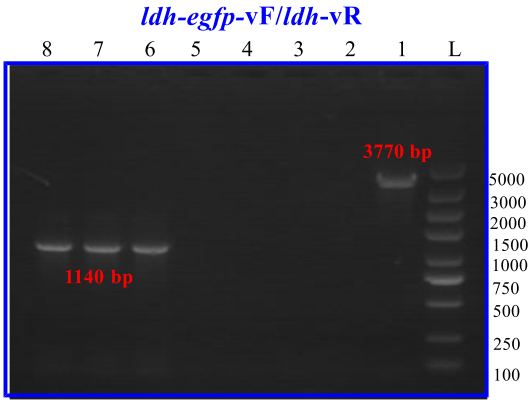
739

740

741 **D**



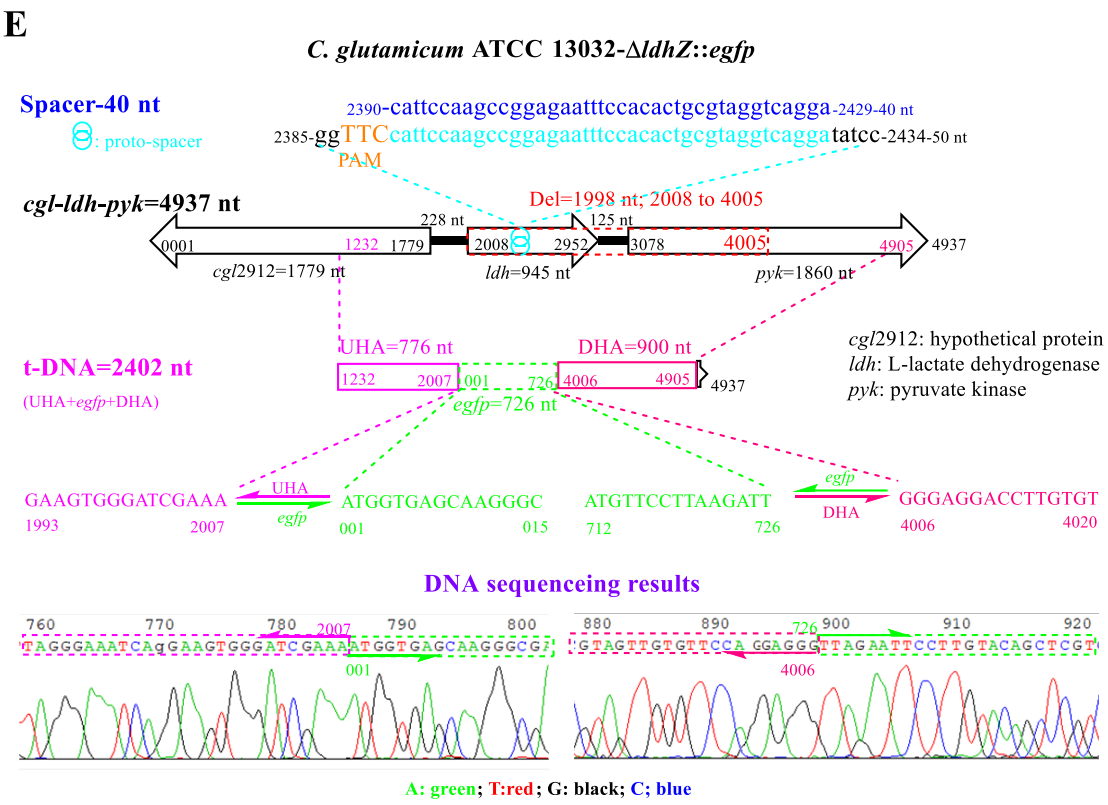
L: ladder
1: *C Glutamicum* ATCC13032
2: water
3: pXMJ19-t/g-*ldh*::*egfp*
4: pEC-XK99E-*cas*3
5-7: *C Glutamicum* ATCC13032-Δ*ldh*::*egfp*



L: ladder
1: *C Glutamicum* ATCC13032 (*ldh*-vF/vR)
2: *C Glutamicum* ATCC13032
3: water
4: pXMJ19-t/g-*ldh*::*egfp*
5: pEC-XK99E-*cas*3
6-8: *C Glutamicum* ATCC13032-Δ*ldh*::*egfp*

742

743



744

745

746

747 **Extended Data Fig. 4. Construction of gene editing tools and the RNA-guided**

748 **genome editing of *C. glutamicum* ATCC 13032. (A)** The plasmid pXMJ19 digested

749 by *BamH*I and *EcoR*I, chemically synthesized g-*ldh*, overlap PCR amplified t-

750 Δ*ldh*::*egfp* and the result of the gene editing plasmid pXMJ19-t/g-Δ*ldh*::*egfp* digested

751 by *Hind*III and *Xba*I. **(B)** The extracted *C. glutamicum* ATCC 13032 genome, extracted

752 plasmid pEC-XK99E, PCR amplified *cas*3 from pEC-XK99E-*cas*3 and the result of

753 pEC-XK99E-*cas*3 digested by *Xho*I (two *Xho*I restriction sites in this plasmid),

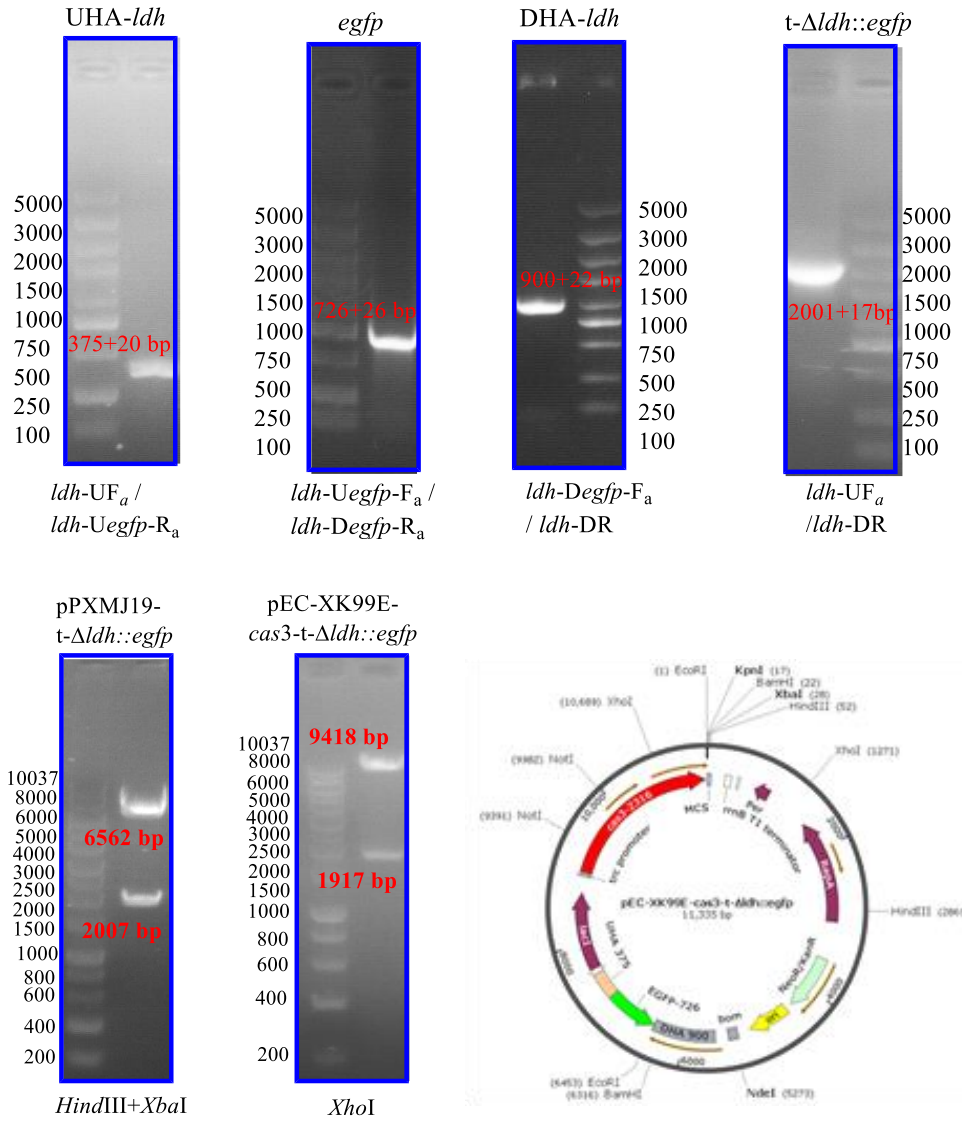
754 respectively. **(C)** The maps of gene editing plasmid pXMJ19-t/g-Δ*ldh*::*egfp* and Cas

755 expression plasmid pEC-XK99E-*cas*3.**(D)** DNA gel electrophoresis of the PCR

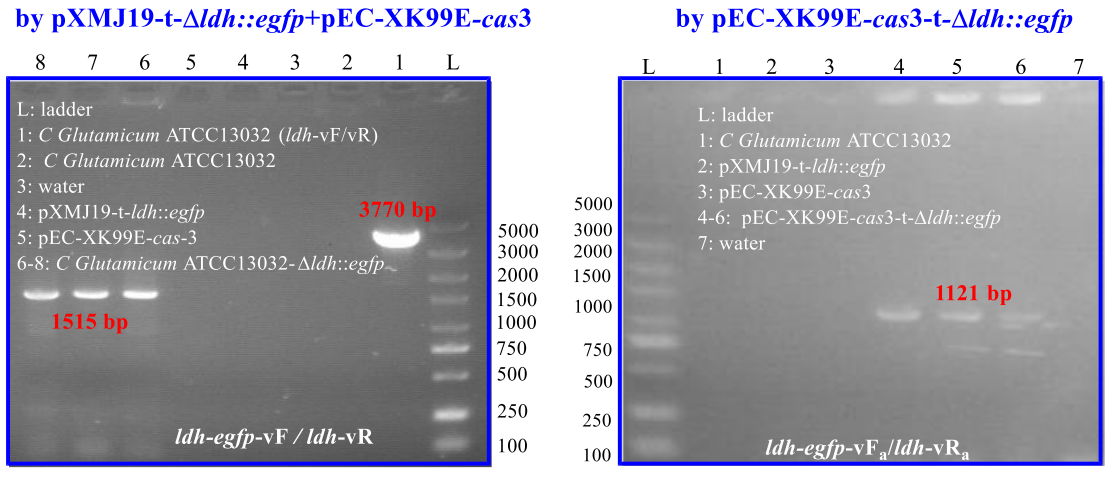
756 products of edited sequences in the genome editing of *C. glutamicum* ATCC 13032

(The primers are respectively *ldh*-vF / vR and *ldh-egfp*-vF / *ldh*-vR, wherein, the primer *ldh-egfp*-vF is a portion of the gene *egfp* in the t-Δ*ldh*::*egfp* fragment and the primer *ldh*-vR is part of the gene *ldh* but is not contained in the t-Δ*ldh*::*egfp*. Therefore, the PCR products can occur only when the genome of the correctly gene-edited mutant *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* is used as template). **(E)** The selected proto-spacer, the engineered t-DNA as well as the DNA sequencing results of the PCR products of the edited sequences of the gene-edited mutant *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* genome (The junction sequences between UHA and *egfp* and between *egfp* and DHA are highlighted in Extended Data Fig. 4E).

A


UHA-ldh
egfp
DHA-ldh
t-Δldh::egfp
5000
3000
2000
1500
1000
750
500
250
100
375+20 bp
726+26 bp
900+22 bp
2001+17bp
ldh-UF_a / ldh-Uegfp-R_a
ldh-Uegfp-F_a / ldh-Degfp-R_a
ldh-Degfp-F_a / ldh-DR
ldh-UF_a /ldh-DR
pPXMJ19-t-Δldh::egfp
pEC-XK99E-cas3-t-Δldh::egfp
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
6562 bp
2007 bp
9418 bp
1917 bp
HindIII+XbaI
XhoI

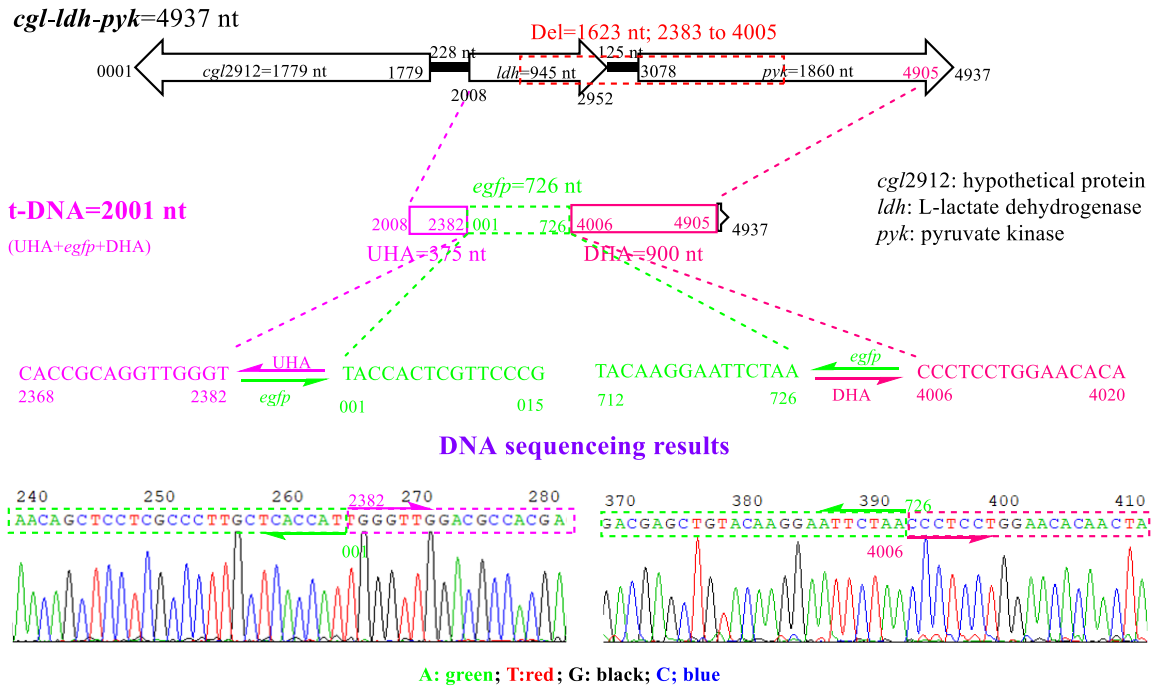
B


by pXMJ19-t-Δldh::egfp+pEC-XK99E-cas3
8 7 6 5 4 3 2 1 L
L: ladder
1: C Glutamicum ATCC13032 (ldh-vF/vR)
2: C Glutamicum ATCC13032
3: water
4: pXMJ19-t-ldh::egfp
5: pEC-XK99E-cas-3
6-8: C Glutamicum ATCC13032-Δldh::egfp
3770 bp
1515 bp
ldh-egfp-vF / ldh-vR
5000
3000
2000
1500
1000
750
500
250
100
by pEC-XK99E-cas3-t-Δldh::egfp
L 1 2 3 4 5 6 7
L: ladder
1: C Glutamicum ATCC13032
2: pXMJ19-t-ldh::egfp
3: pEC-XK99E-cas3
4-6: pEC-XK99E-cas3-t-Δldh::egfp
7: water
1121 bp
ldh-egfp-vF_a/ldh-vR_a

771

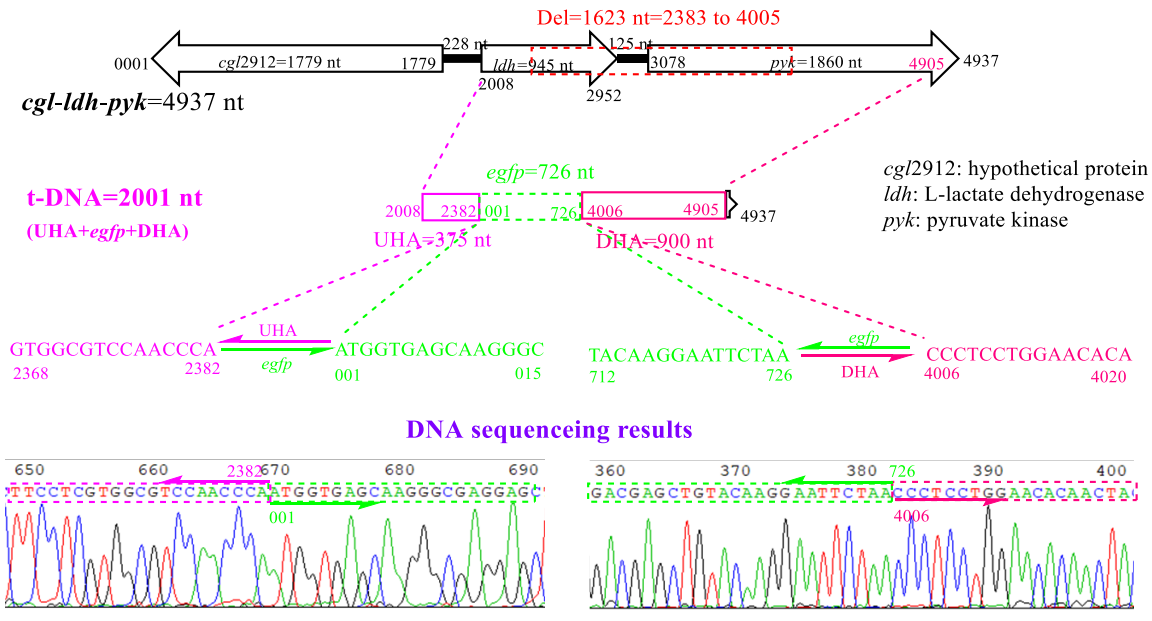
C

by pXMJ19-t-Δ*ldh::egfp*+pEC-XK99E-*cas*3



772 **A: green; T:red; G: black; C; blue**

by pEC-XK99E-*cas*3-t-Δ*ldh::egfp*



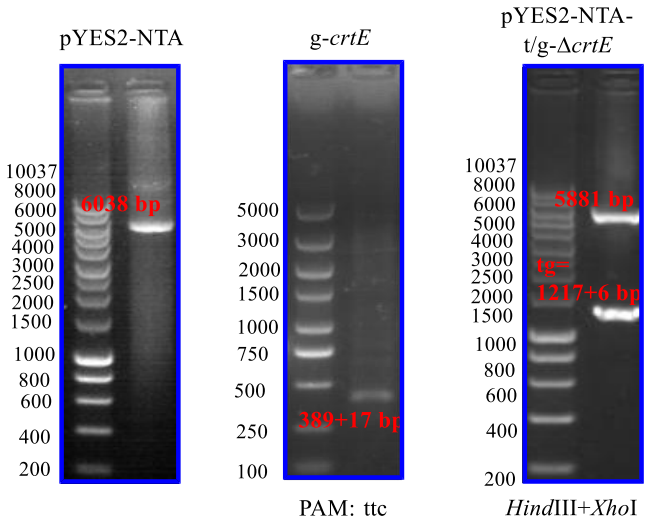
773 **A: green; T:red; G: black; C; blue**

774

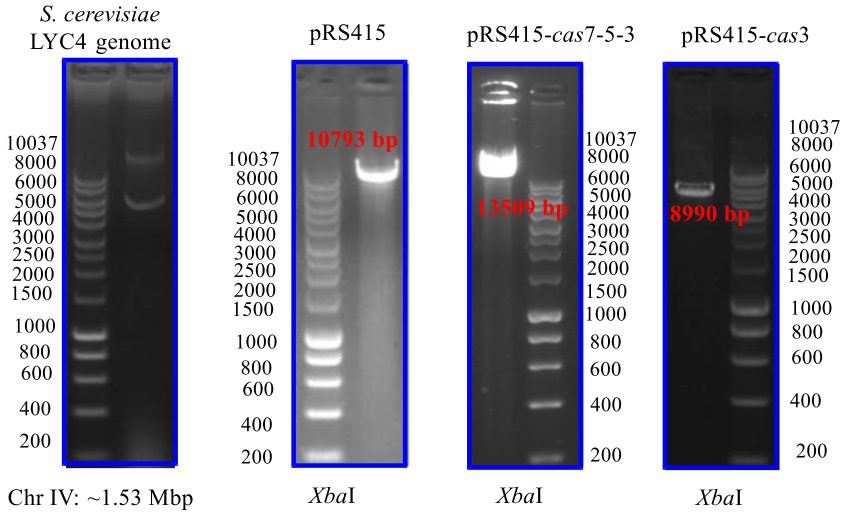
775

Extended Data Fig. 5. Construction of gene editing tools and the DNA-guided genome editing of *C. glutamicum* ATCC 13032. (A) The construction of DNA-guided genome editing tools (The PCR amplified UHA-*ldh*, *egfp*, DHA-*ldh* and the overlap PCR amplified t-Δ*ldh*::*egfp*, the results of plasmids pXMJ19-t-Δ*ldh*::*egfp* and pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp* respectively digested by *Hind*III+*Xba*I and *Xho*I, as well as the map of plasmid pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp*). **(B)** The results of DNA gel electrophoresis of the PCR product of target sequences in the DNA-guided genome editing of *S C. glutamicum* ATCC 13032 by pEC-XK99E-*cas*3 plus pXMJ19-t-Δ*ldh*::*egfp* and pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp*, respectively. **(C)** The engineered t-DNA structure diagrams and the results of DNA sequencing of the PCR product of target sequences in the DNA-guided genome editing of *S C. glutamicum* ATCC 13032 by pEC-XK99E-*cas*3-t-Δ*ldh*::*egfp* (PCR products only occur in gene-edited mutants *C. glutamicum* ATCC 13032-Δ*ldh*::*egfp* because the primers *ldh-egfp*-vF and *ldh-egfp*-vF_a are a portion of gene *egfp* in t-Δ*ldh*::*egfp* and the primers *ldh*-vR and *ldh*-vR_a are a portion of gene *ldh* that is not contained in t-Δ*ldh*::*egfp*.).

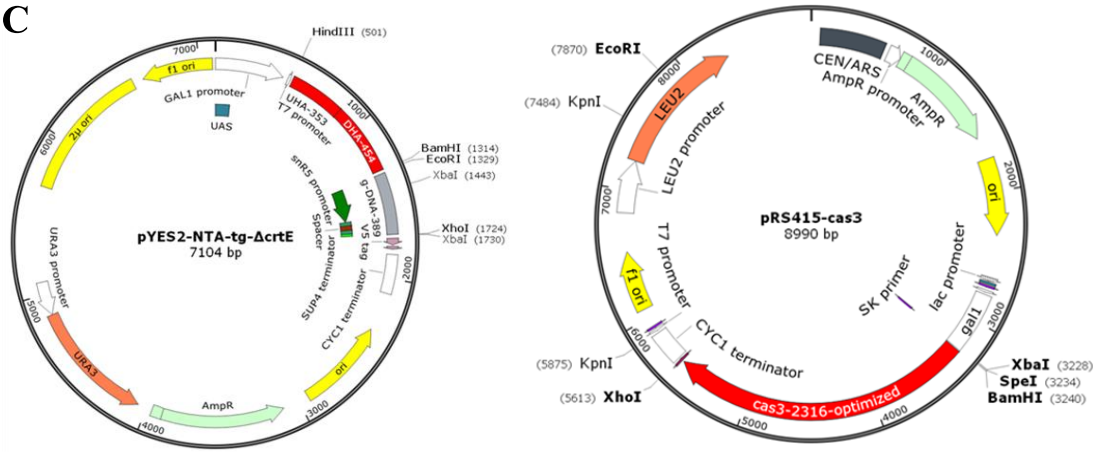
A


pYES2-NTA
g-crtE
pYES2-NTA-t/g-ΔcrtE
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
6038 bp
5000
3000
2000
1500
1000
750
500
250
100
389+17 bp
5881 bp
tg=
1217+6 bp
PAM: ttc
HindIII+XhoI

B

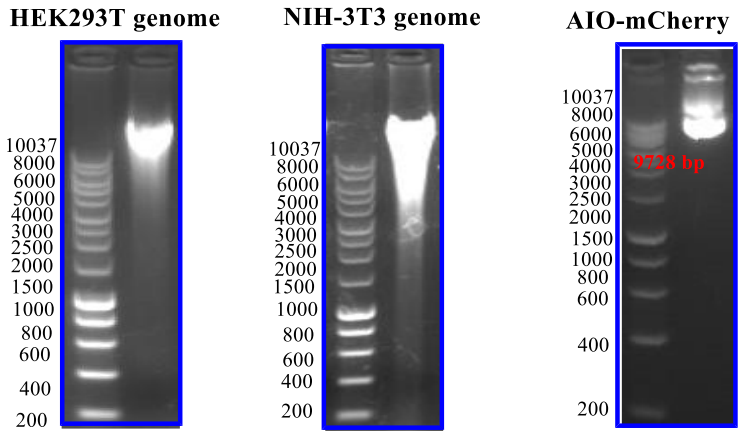

S. cerevisiae
LYC4 genome
pRS415
pRS415-cas7-5-3
pRS415-cas3
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
10793 bp
13509 bp
8990 bp
Chr IV: ~1.53 Mbp
XbaI
XbaI
XbaI

C

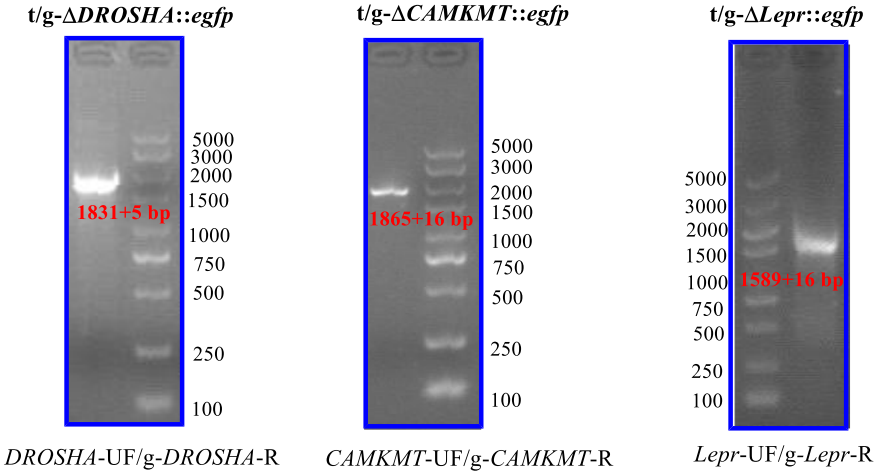

pYES2-NTA-tg-ΔcrtE
7104 bp
HindIII (501)
f1 ori
GAL1 promoter
UAS
UHA-353
T7 promoter
DHA-452
BamHI (1314)
EcoRI (1329)
XbaI (1443)
snR5 promoter
Spacer
g-DNA-389
V5 tag
XhoI (1724)
XbaI (1730)
SUP4 terminator
CYC1 terminator
ori
AmpR
URA3
URA3 promoter
2μ ori
pRS415-cas3
8990 bp
(7870) EcoRI
(7484) KpnI
CEN/ARS
AmpR promoter
AmpR
LEU2
LEU2 promoter
ori
T7 promoter
f1 ori
SK primer
lac promoter
gal1
CYC1 terminator
(5875) KpnI
(5613) XhoI
cas3-2316-optimized
XbaI (3228)
SpeI (3234)
BamHI (3240)

Extended Data Fig. 6. Construction of gene editing tools in the RNA-guided genome editing of *S. cerevisiae* LYC4. (A) The extracted pYES2-NTA, chemically synthesized g-*crtE* and the result of pYES2-NTA-t/g-Δ*crtE* digested by *Hind*III and *Xho*I. **(B)** The extracted *S. cerevisiae* LYC4 genome and the linearized results of plasmids pRS415, pRS415-*cas*7-5-3 and pRS415-*cas*3 digested by *Xba*I. **(C)** The maps of gene editing plasmid pYES2-NTA-t/g-Δ*crt*E and Cas expression plasmid pRS415-*cas*3 (The map of plasmid pRS415-*cas*7-5-3 was the same as this map of plasmid pRS415-*cas*3, except for the codon optimized genes of *cas*5 and *cas*7).

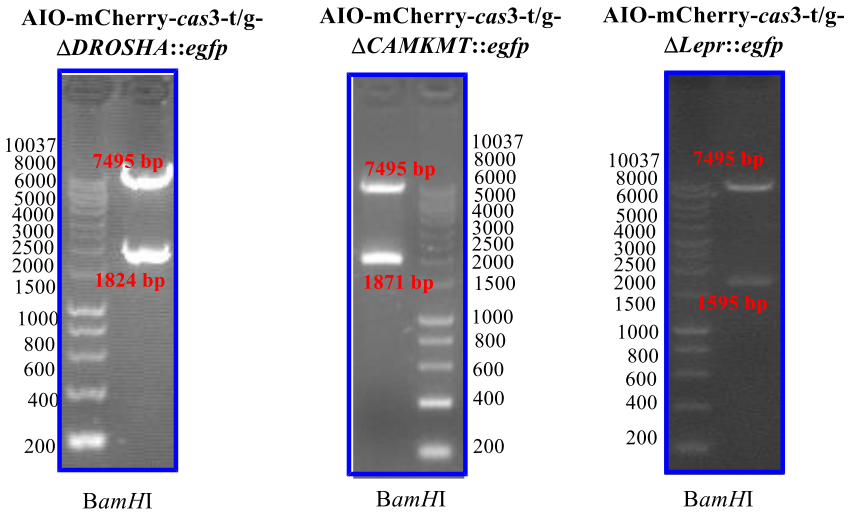
A

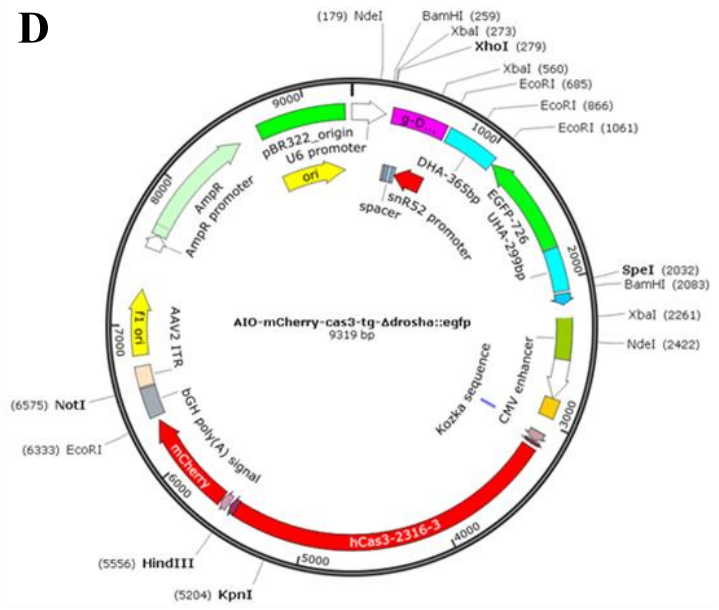

HEK293T genome
NIH-3T3 genome
AIO-mCherry
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
9728 bp

B


t/g-ΔDROSHA::egfp
t/g-ΔCAMKMT::egfp
t/g-ΔLepr::egfp
5000
3000
2000
1500
1000
750
500
250
100
1831+5 bp
1865+16 bp
1589+16 bp
DROSHA-UF/g-DROSHA-R
CAMKMT-UF/g-CAMKMT-R
Lepr-UF/g-Lepr-R

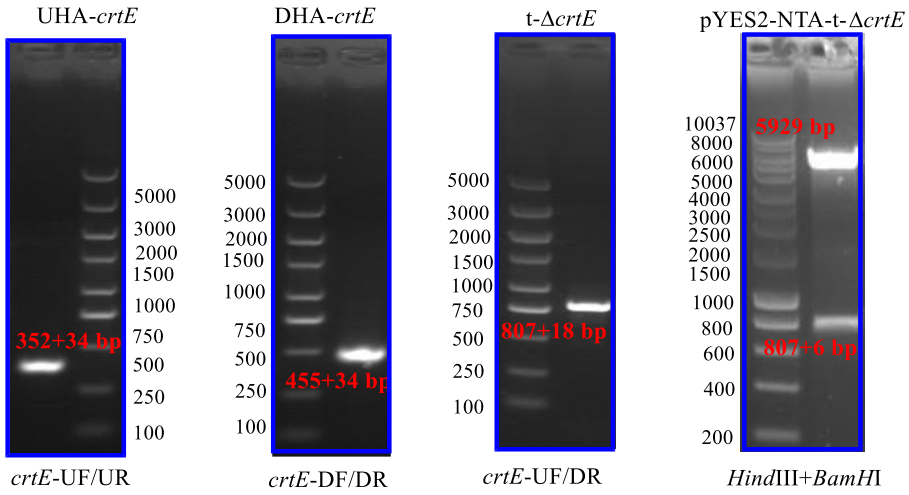
C


AIO-mCherry-cas3-t/g-ΔDROSHA::egfp
AIO-mCherry-cas3-t/g-ΔCAMKMT::egfp
AIO-mCherry-cas3-t/g-ΔLepr::egfp
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200
7495 bp
1824 bp
1871 bp
1595 bp
BamHI
BamHI
BamHI

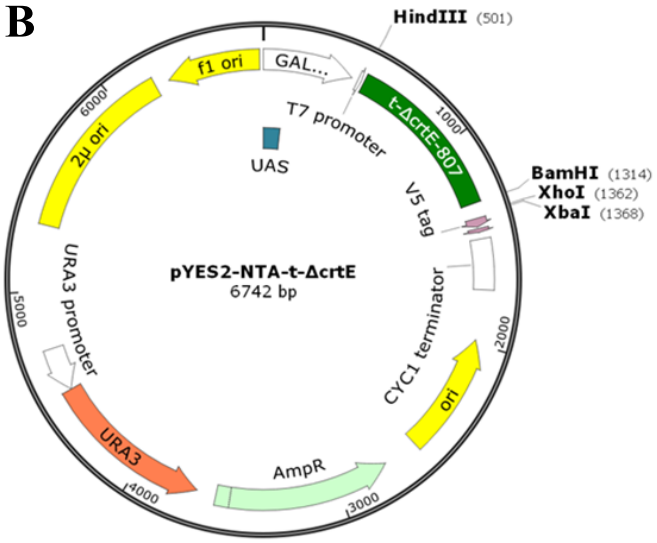


Extended Data Fig. 7. Construction of gene editing vector in the RNA-guided genome editing of HEK293T and NIH-3T3 cells. (A) The extracted HEK293T cell genome, extracted NIH-3T3 cell genome and extracted vector AIO-mCherry. **(B)** The overlap PCR amplified t/g-Δ*DROSHA*::*egfp*, overlap PCR amplified t/g-Δ*CAMKMT*::*egfp*, overlap PCR amplified t/g-Δ*Lepr*::*egfp*. **(C)** The linearized results of the three gene editing vectors AIO-mCherrym-*cas*3-t/g-Δ*DROSHA*::*egfp*, AIO-mCherrym-*cas*3-t/g-Δ*CAMKMT*::*egfp* and AIO-mCherrym-*cas*3-t/g-Δ*Lepr*::*egfp* digested by *Bam*HI (each vector harboring two *Bam*HI restriction sites). **(D)** The map of gene editing plasmid AIO-mCherry-*cas*3-t/g-Δ*DROSHA*::*egfp* (In the RNA-guided genome editing of mammalian cells, the maps of the other two genome editing vectors AIO-mCherry-*cas*3-t/g-Δ*CAMKMT*::*egfp* and AIO-mCherry-*cas*3-t/g-Δ*Lepr*::*egfp* were the same as this map, except for the corresponding g-DNAs and t-DNAs).

A

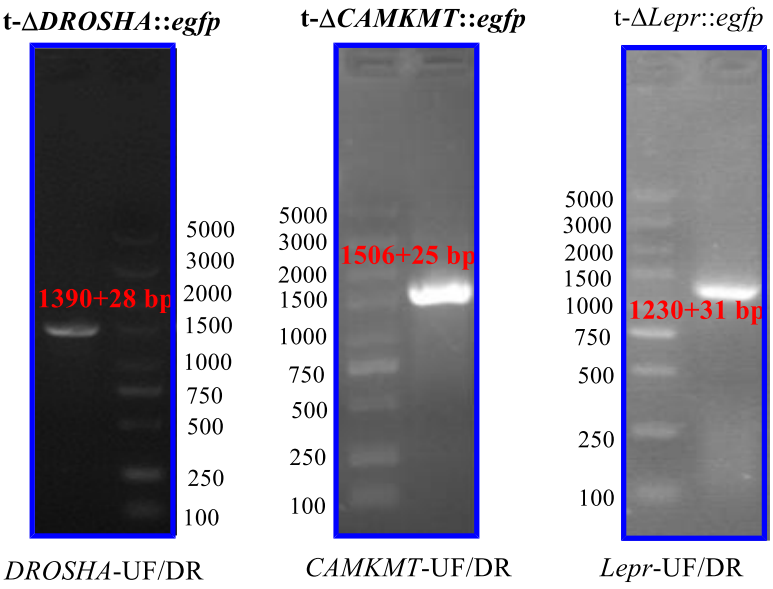


B



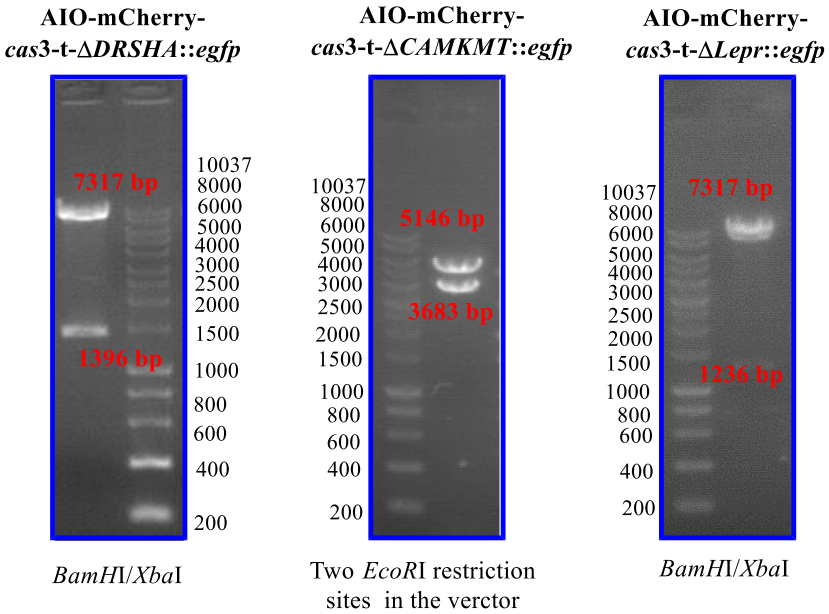
Extended Data Fig. 8. Construction of gene editing tools in the DNA-guided genome editing of *S. cerevisiae* LYC4. (A) The PCR amplified UHA-*crtE* and DHA-*crtE*, overlap PCR amplified t-*crtE* and the result of gene editing plasmid pYES2-NTA-t-Δ*crtE* digested by *Hind*III+*BamH*I, respectively (UHA+DHA=353+454=807 nt, deletion=252 nt). **(B)** The map of gene editing plasmid pYES2-NTA-t-Δ*crt*E.

841 **A**



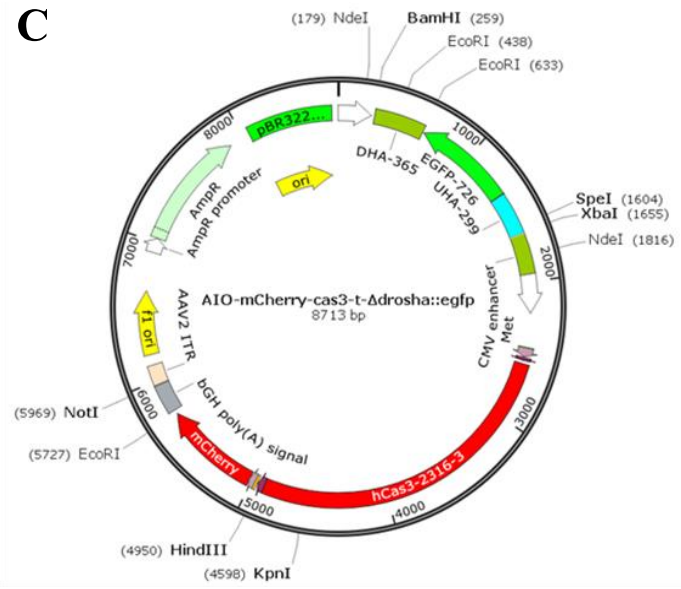
842

843 **B**



844

845



846

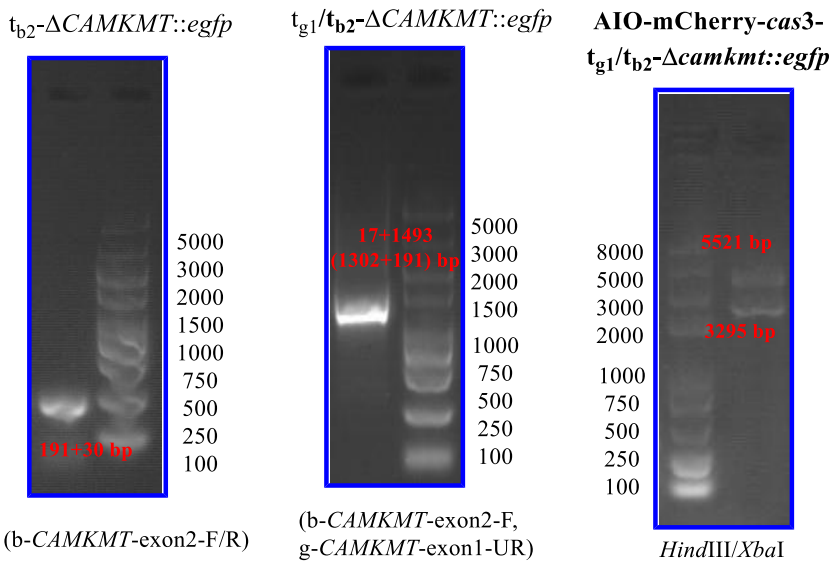
847

848

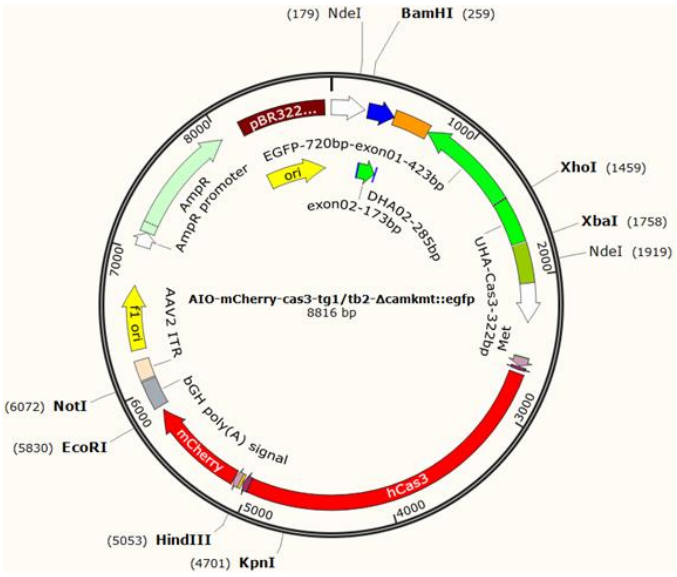
849 **Extended Data Fig. 9. Construction of gene editing tools in the DNA-guided**
850 **genome editing of mammalian cells. (A)** The PCR amplified fragments of t-
851 Δ*DROSHA*::*egfp*, t-Δ*CAMKMT*::*egfp* and t-Δ*Lepr*::*egfp*. **(B)** The results of the three
852 vectors (AIO-mCherry-*cas*3-t-Δ*DROSHA*::*egfp*, AIO-mCherry-*cas*3-t-
853 Δ*CAMKMT*::*egfp* and AIO-mCherry-*cas*3-t-Δ*Lepr*::*egfp*) digested by *Bam*HI+*Xba*I,
854 *Eco*RI and *Bam*HI+*Xba*I, respectively. **(C)** The map of gene editing plasmid AIO-
855 mCherry-*cas*3-t-Δ*DROSHA*::*egfp* (In the DNA-guided genome editing of mammalian
856 cells, the maps of the other two genome editing vectors AIO-mCherry-*cas*3-t-
857 Δ*CAMKMT*::*egfp* and AIO-mCherry-*cas*3-t-Δ*Lepr*::*egfp* were identical to this map,
858 except for the corresponding t-DNAs).

859

A



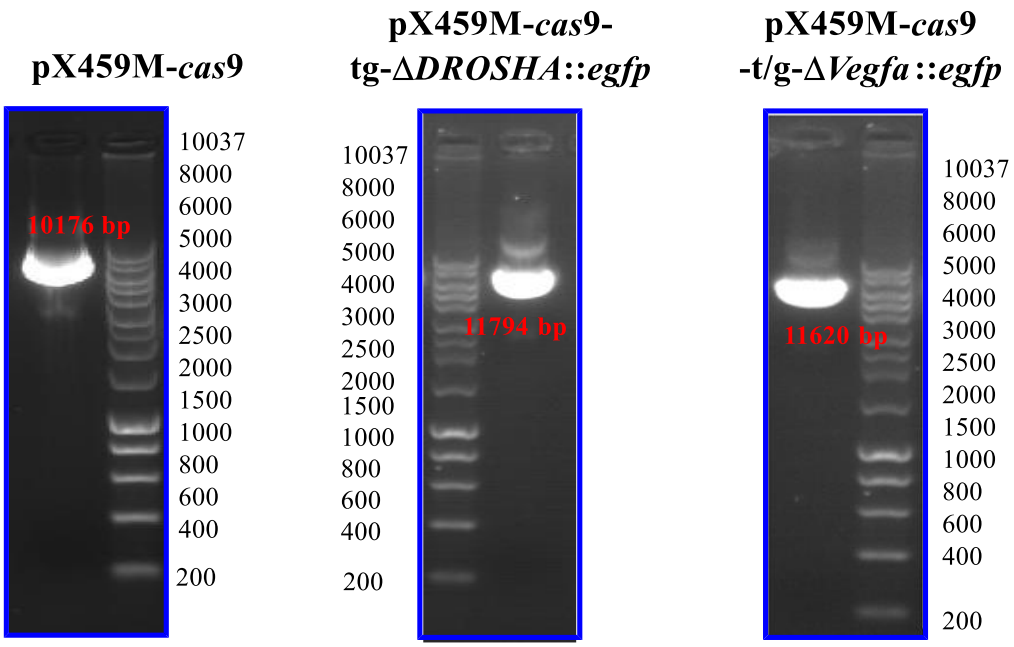
B

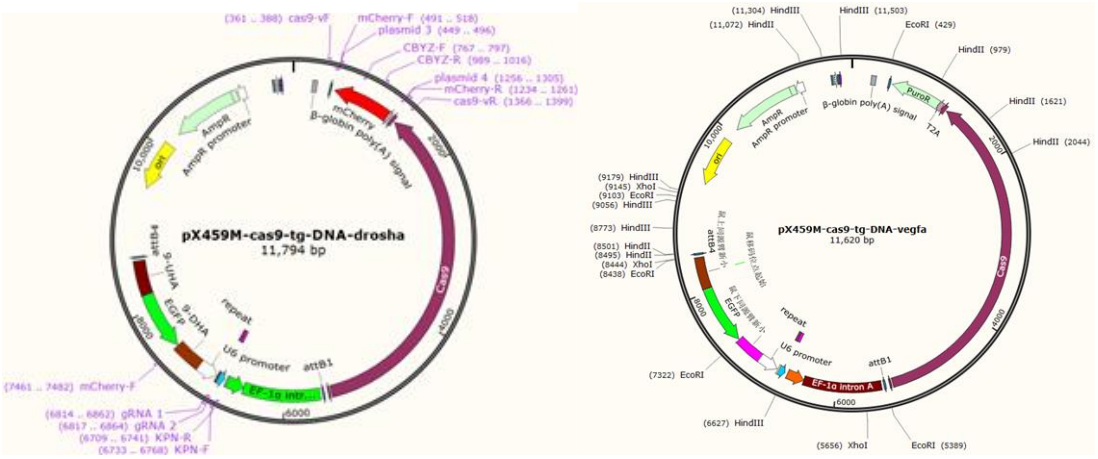


Extended Data Fig. 10. Construction of the DNA-guided gene and base editing vector AIO-mCherry-*cas*3-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp* for the *CAMKMT* gene in HEK293T cells. (A) The PCR amplified fragments of t_{b2}-Δ*CAMKMT*::*egfp* and t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*, and the result of the vector AIO-mCherry-*cas*3-t_{g1}/t_{b2}-

870 Δ*CAMKMT*::*egfp* digested by *Hind*III plus *Xba*I. **(B)** The map of the vector AIO-
871 mCherry-*cas*3-t_{g1}/t_{b2}- Δ*CAMKMT*::*egfp* ("b- or b2-" representing base editing of the
872 exon2, "g- or g1-" representing gene editing of the exon1).
873

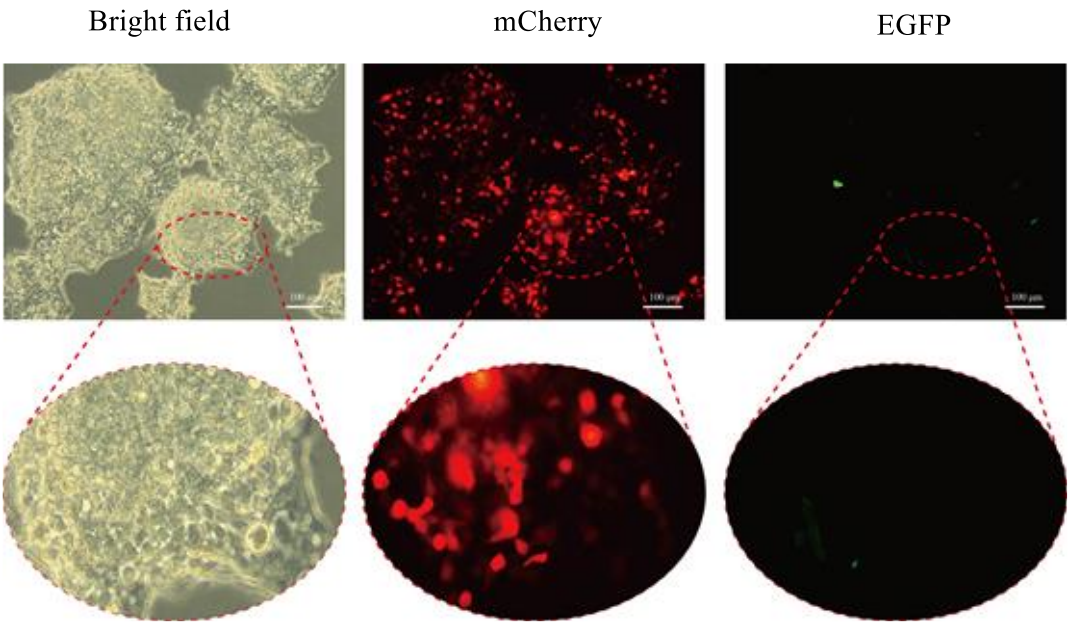
A


pX459M-*cas*9
10176 bp
pX459M-*cas*9-tg-Δ*DROSHA*::*egfp*
11794 bp
pX459M-*cas*9-t/g-Δ*Vegfa*::*egfp*
11620 bp
10037
8000
6000
5000
4000
3000
2500
2000
1500
1000
800
600
400
200


pX459M-cas9-tg-DNA-drosha
11,794 bp
pX459M-cas9-tg-DNA-vegfa
11,620 bp
AmpR promoter
mCherry
β-globin poly(A) signal
U6 promoter
attB1
repeat
Cas9
EGFP
puroR
T2A
ori
HindIII
EcoRI
XhoI
HindII

B

pX459M-*cas*9-t/g-Δ*DROSHA*::*egfp* (with mCherry)



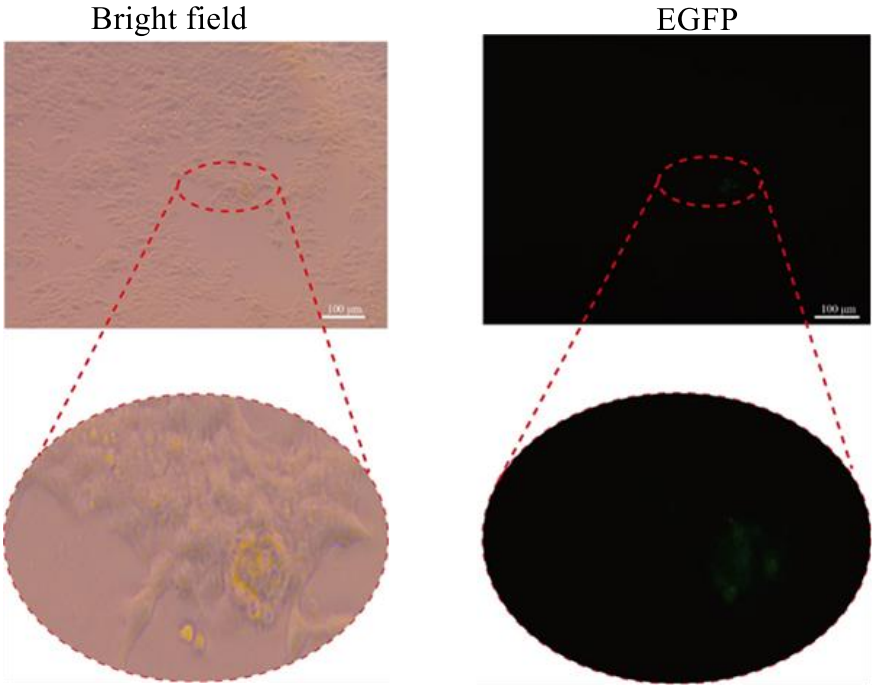
HEK293T
HEK293T-t/g-Δ*DROSHA*::*egfp*
HEK293T-Δ*DROSHA*::*egfp*

HEK293T-t/g-Δ*DROSHA*::*egfp*

HEK293T-Δ*DROSHA*::*egfp*

879
880

pX459M-cas9-Δ*Vegfa*::*egfp* (without mCherry)

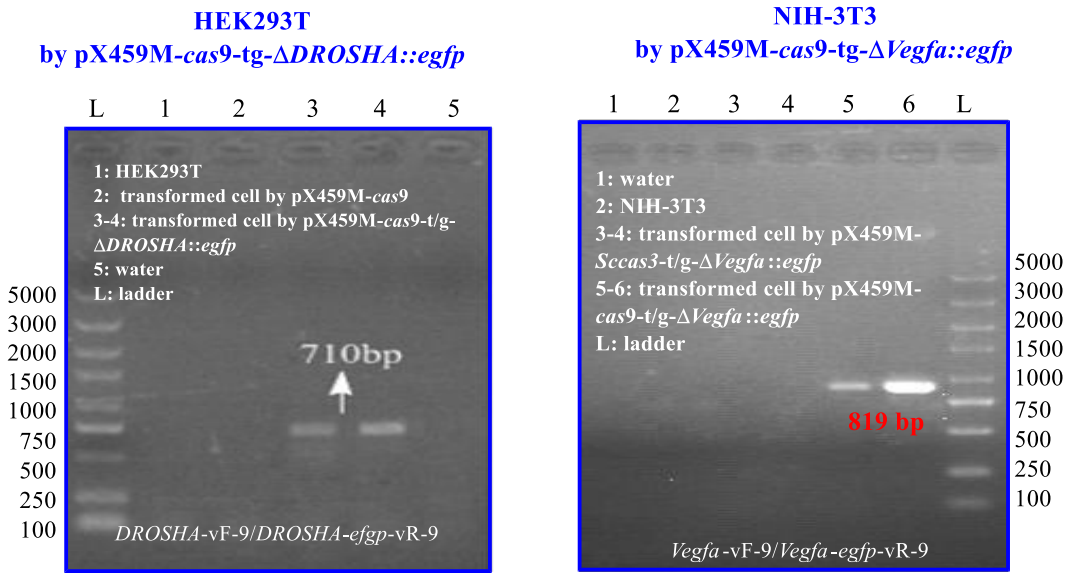


NIH-3T3
NIH-3T3-t/g-Δ*Vegfa*::*egfp*
NIH-3T3-Δ*Vegfa*::*egfp*

NIH-3T3-Δ*Vegfa*::*egfp*

881
882

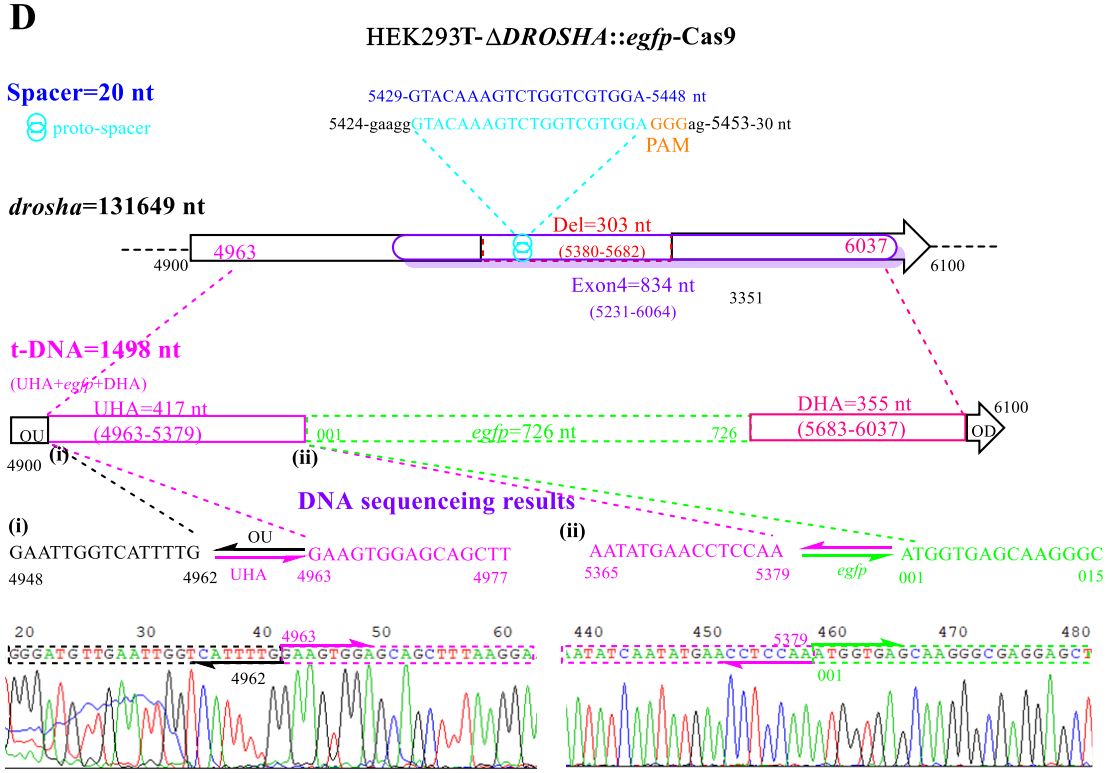
883 **C**



884 ***Sccas3*: codon optimized gene *Svicas3* based on the base preference of *Saccharomyces cerevisiae***

885

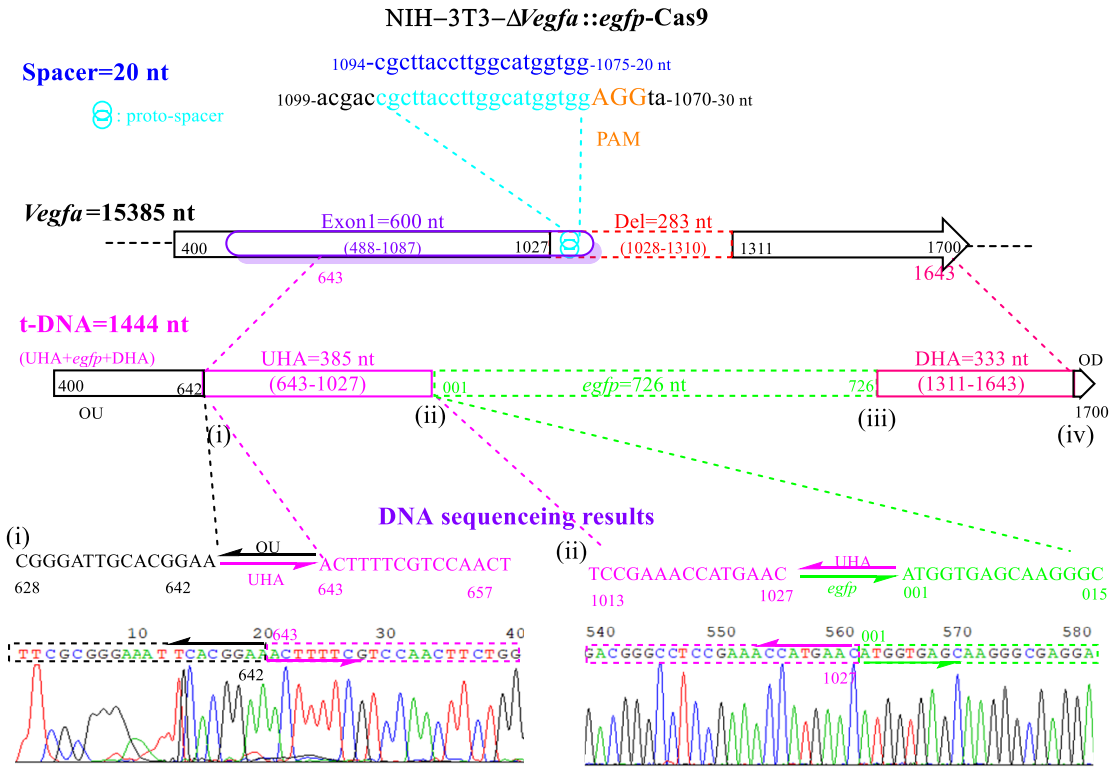
D



886 OD: the outside of DHA, OU: the outside of UHA

A: green; T:red; G: black; C; blue

887



888 OD: the outside of DHA, OU: the outside of UHA A: green; T:red; G: black; C; blue

889

890

891 **Extended Data Fig. 11. Construction of Cas9-based gene editing tools and mutant**

892 **verification in mammalian cell genome editing mediated by Cas9. (A)** The

893 construction of Cas9-based genome editing tools (The original plasmid pX459M-*cas*9,

894 the two constructed Cas9-based gene editing plasmids pX459M-*cas*9-tg-

895 Δ*DROSHA*::*egfp* with mCherry gene and pX459M-*cas*9-tg-Δ*Vegfa*::*egfp* without

896 mCherry gene as well as their maps). **(B)** Typical micrographs of *Sp*Cas9-based genome

897 editing process in mammal cells. **(C)** The results of DNA gel electrophoresis of the PCR

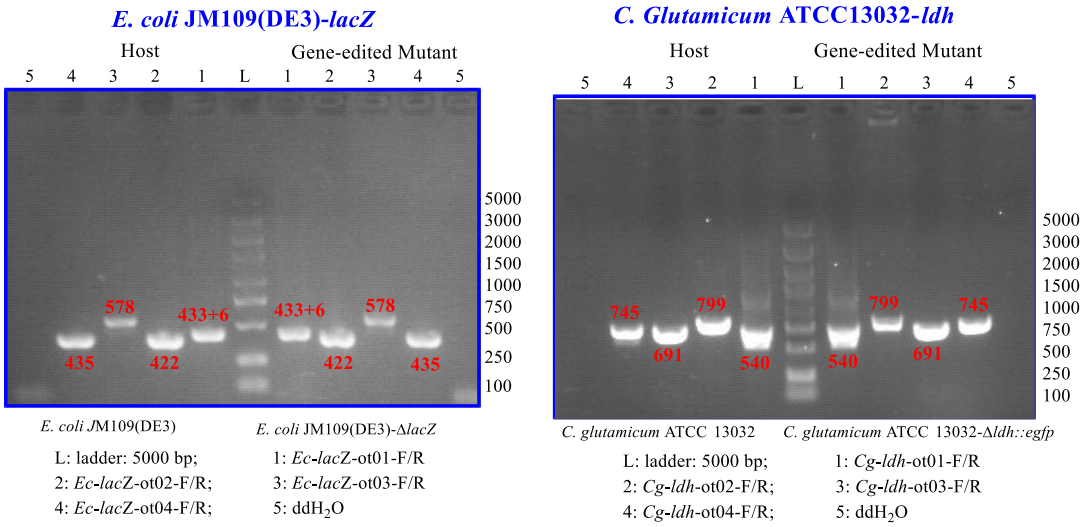
898 product of target sequences in the Cas9-based mammalian cell genome editing by

899 pX459M-*cas*9-tg-Δ*DROSHA*::*egfp* and pX459M-*cas*9-tg-Δ*Vegfa*::*egfp*, respectively.

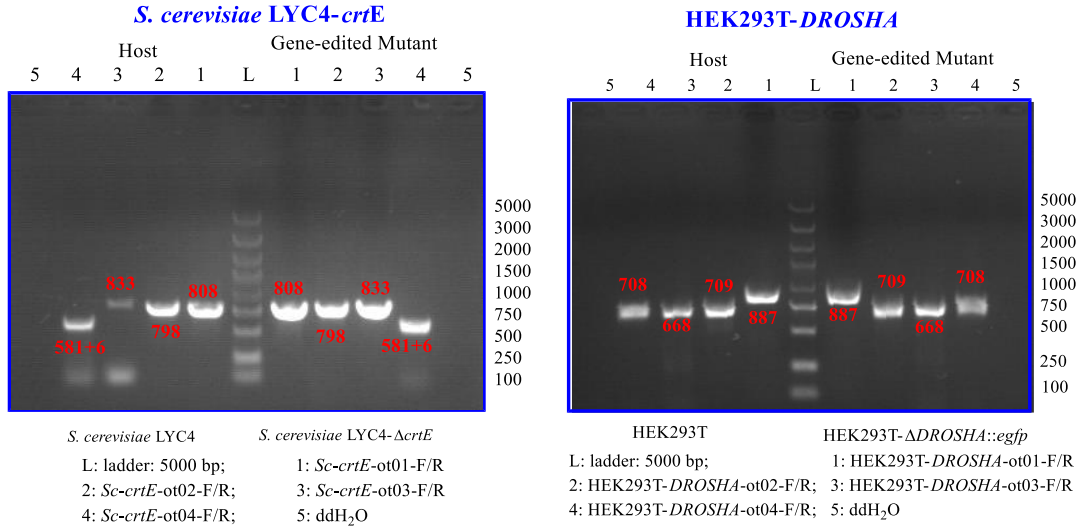
900 **(D)** The engineered t-DNA structure diagrams and the results of DNA sequencing of

the PCR product of target sequences in the Cas9-based mammalian cell genome editing by pX459M-*cas*9-tg-Δ*DROSHA*::*egfp* and pX459M-*cas*9-tg-Δ*Vegfa*::*egfp* (PCR products only occur in gene-edited variants HEK293T-Δ*DROSHA*::*egfp*-9 and NIH-3T3-Δ*Vegfa*::*egfp*-9 because the primers *DROSHA*-*egfp*-vR-9 and *Vegfa*-*egfp*-vR-9 are a portion of gene *egfp* in t-Δ*DROSHA*::*egfp* and t-Δ*Vegfa*::*egfp* and the primers *DROSHA*-vF-9 and *Vegfa*-vF-9 are a portion of genes *DROSHA* and *Vegfa* that are not contained in t-Δ*DROSHA*::*egfp* and t-Δ*Vegfa*::*egfp*, respectively.).

909 **A**

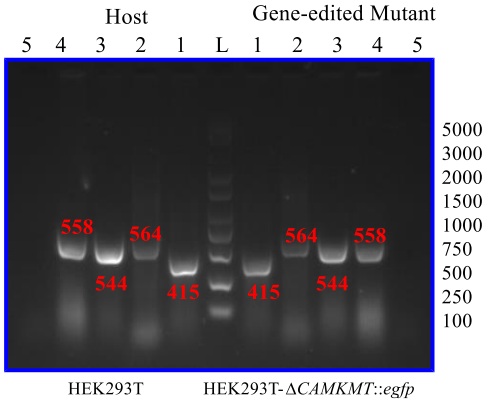

E. coli JM109(DE3)-lacZ
Host
Gene-edited Mutant
5 4 3 2 1 L 1 2 3 4 5
578
433+6
433+6
578
435
422
422
435
5000
3000
2000
1500
1000
750
500
250
100
E. coli JM109(DE3)
E. coli JM109(DE3)-ΔlacZ
L: ladder: 5000 bp;
2: Ec-lacZ-ot02-F/R;
4: Ec-lacZ-ot04-F/R;
1: Ec-lacZ-ot01-F/R
3: Ec-lacZ-ot03-F/R
5: ddH2O
C. Glutamicum ATCC13032-ldh
Host
Gene-edited Mutant
5 4 3 2 1 L 1 2 3 4 5
745
799
799
745
691
540
540
691
5000
3000
2000
1500
1000
750
500
250
100
C. glutamicum ATCC 13032
C. glutamicum ATCC 13032-Δldh::egfp
L: ladder: 5000 bp;
2: Cg-ldh-ot02-F/R;
4: Cg-ldh-ot04-F/R;
1: Cg-ldh-ot01-F/R
3: Cg-ldh-ot03-F/R
5: ddH2O

910


S. cerevisiae LYC4-crtE
Host
Gene-edited Mutant
5 4 3 2 1 L 1 2 3 4 5
833
808
808
833
798
798
581+6
581+6
5000
3000
2000
1500
1000
750
500
250
100
S. cerevisiae LYC4
S. cerevisiae LYC4-ΔcrtE
L: ladder: 5000 bp;
2: Sc-crtE-ot02-F/R;
4: Sc-crtE-ot04-F/R;
1: Sc-crtE-ot01-F/R
3: Sc-crtE-ot03-F/R
5: ddH2O
HEK293T-DROSHA
Host
Gene-edited Mutant
5 4 3 2 1 L 1 2 3 4 5
708
709
709
708
668
887
887
668
5000
3000
2000
1500
1000
750
500
250
100
HEK293T
HEK293T-ΔDROSHA::egfp
L: ladder: 5000 bp;
2: HEK293T-DROSHA-ot02-F/R;
4: HEK293T-DROSHA-ot04-F/R;
1: HEK293T-DROSHA-ot01-F/R
3: HEK293T-DROSHA-ot03-F/R
5: ddH2O

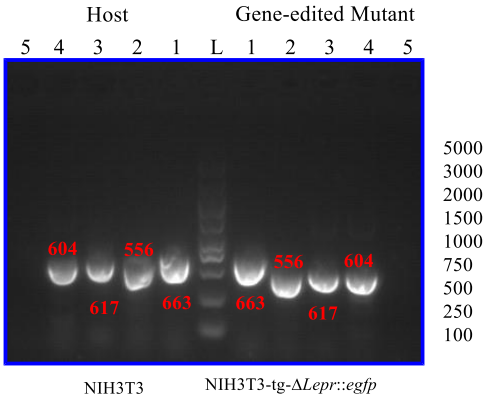
911

HEK293T-*CAMKMT*



L: ladder: 5000 bp; 1: HEK293T-CAMKMT-ot01-F/R
2: HEK293T-CAMKMT-ot02-F/R; 3: HEK293T-AMKMT-ot03-F/R
4: HEK293T-CAMKMT-ot04-F/R; 5: ddH_2O

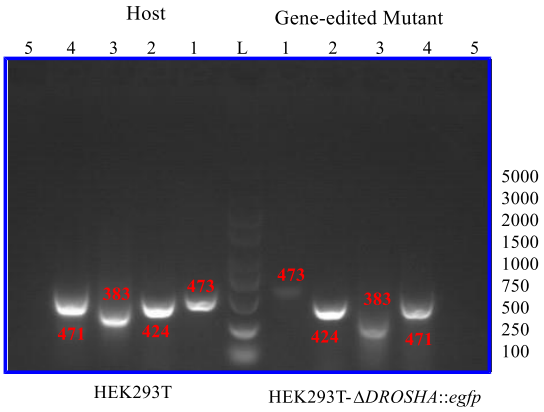
NIH3T3-*Lepr*



L: ladder: 5000 bp; 1: NIH3T3-*Lepr*-ot01-F/R
2: NIH3T3-*Lepr*-ot02F/R; 3: NIH3T3-*Lepr*-ot03F/R
4:NIH3T3-*Lepr*-ot04F/R; 5: ddH_2O

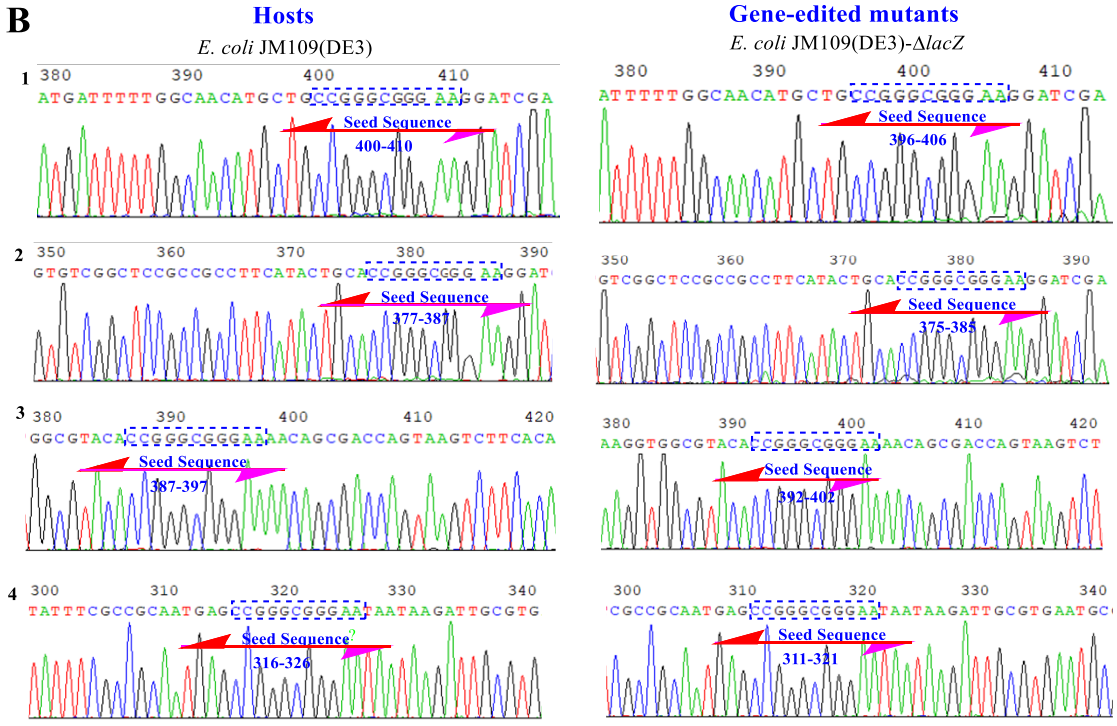
912

HEK293T-*DROSHA*-Cas9



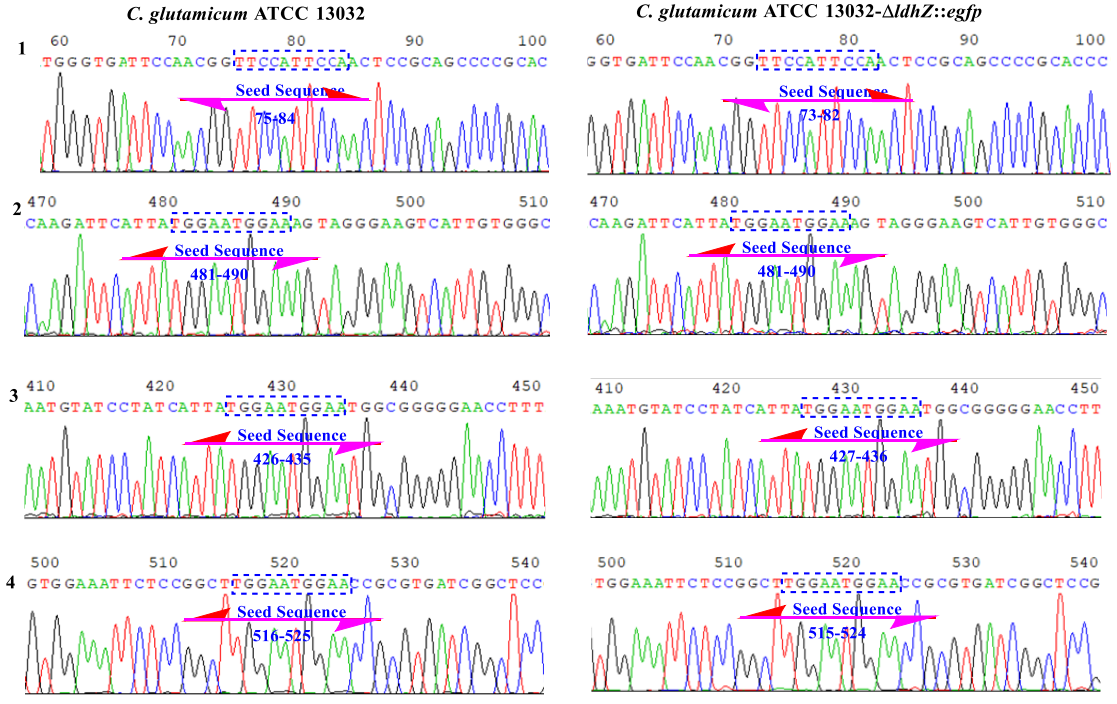
L: ladder: 5000 bp; 1: HEK293T-*DROSHA*-9-ot01-F/R
2: HEK293T-*DROSHA*-9-ot02-F/R; 3: HEK293T-*DROSHA*-9-ot03-F/R
4: HEK293T-*DROSHA*-9-ot04-F/R; 5: ddH_2O

913



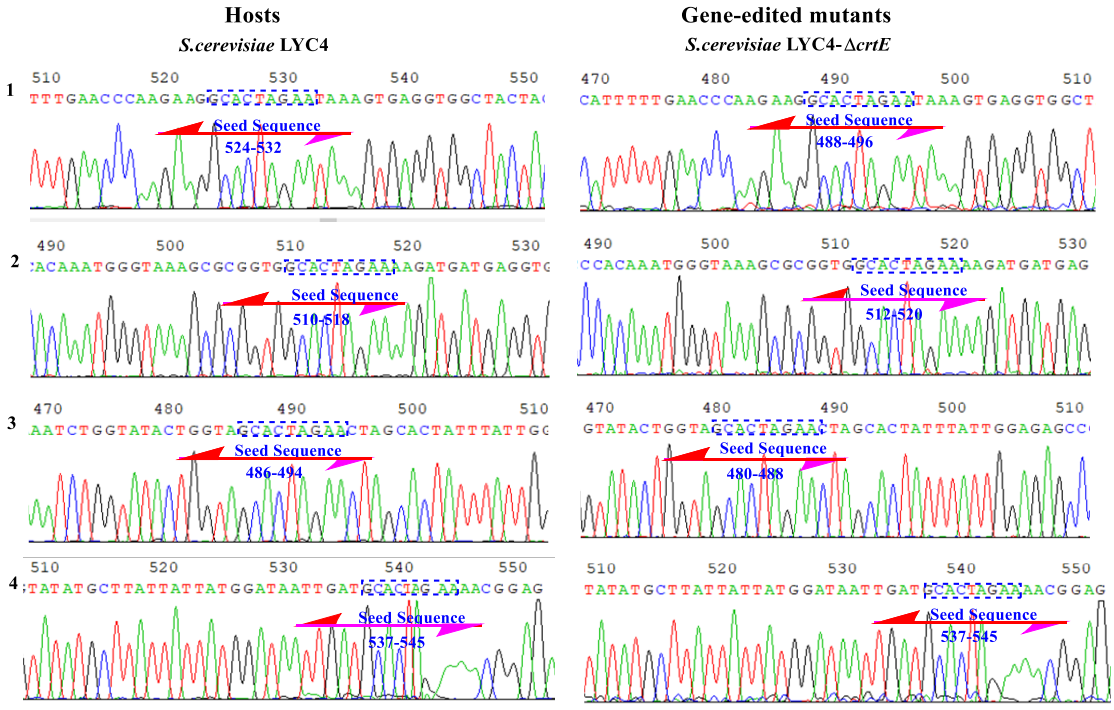
914 A: green; T:red; G: black; C; blue

915

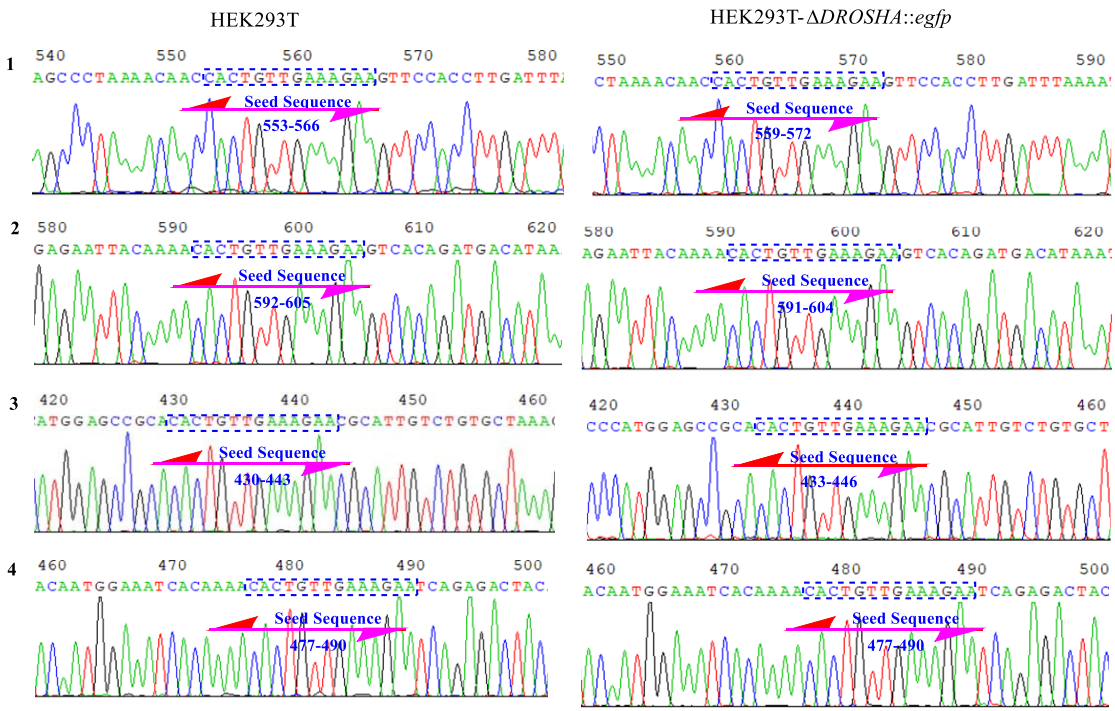


916 A: green; T:red; G: black; C; blue

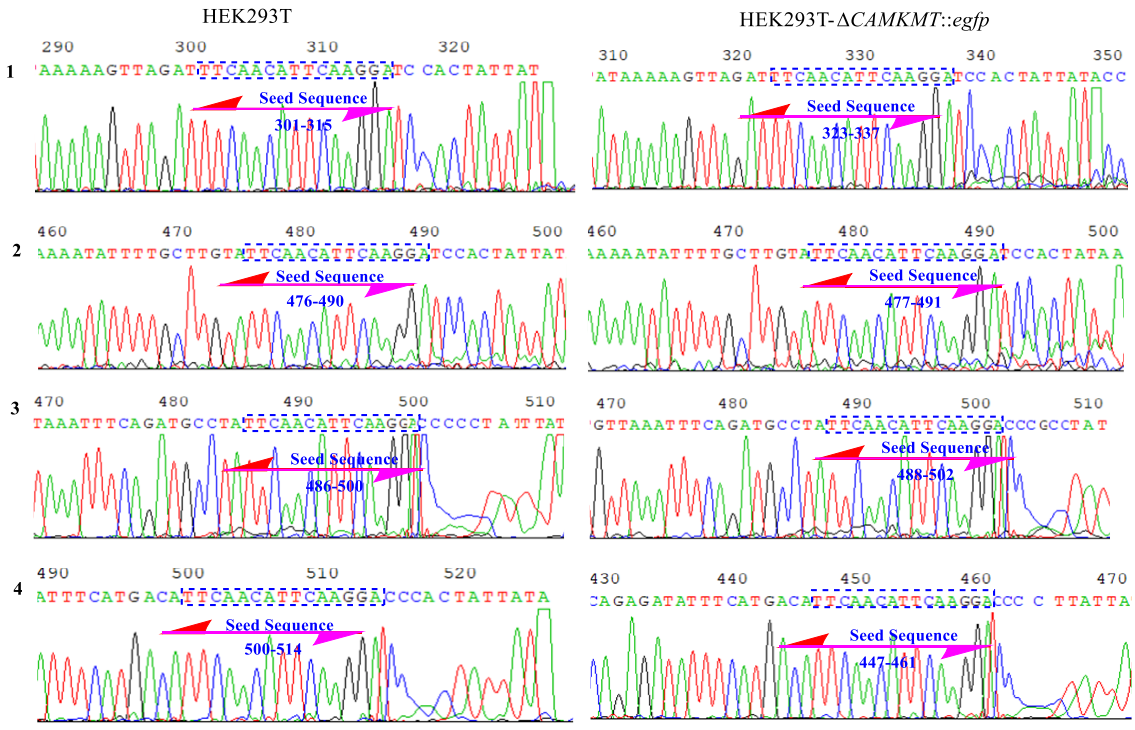
917



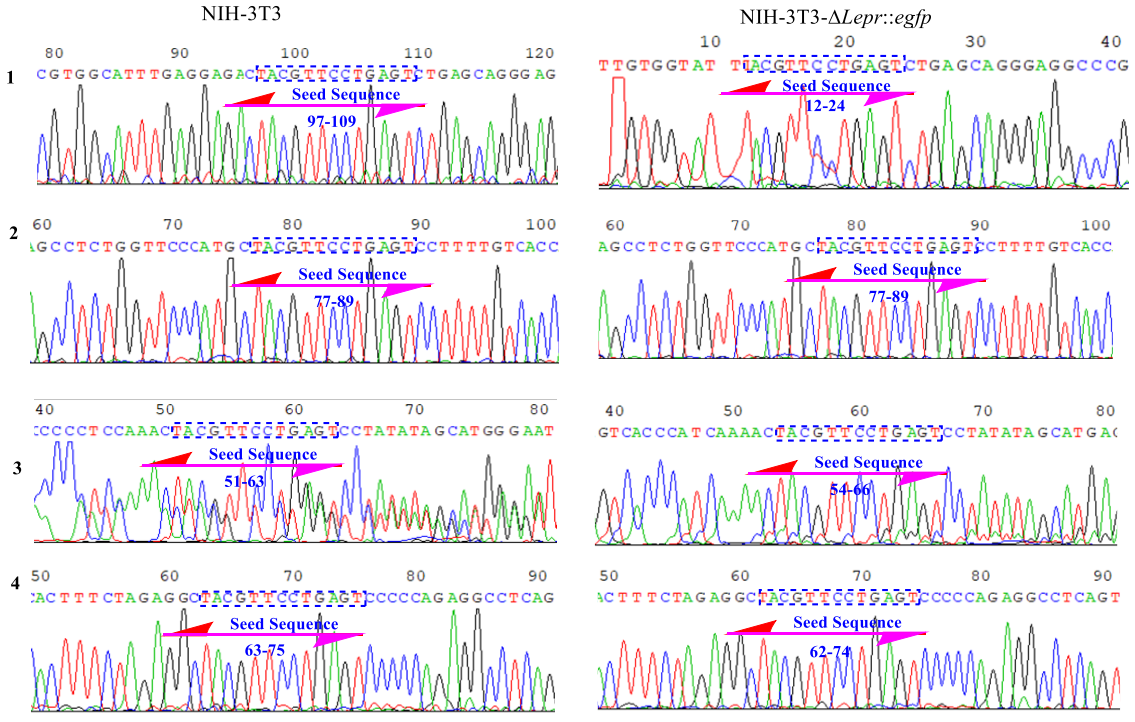
A: green; T:red; G: black; C; blue



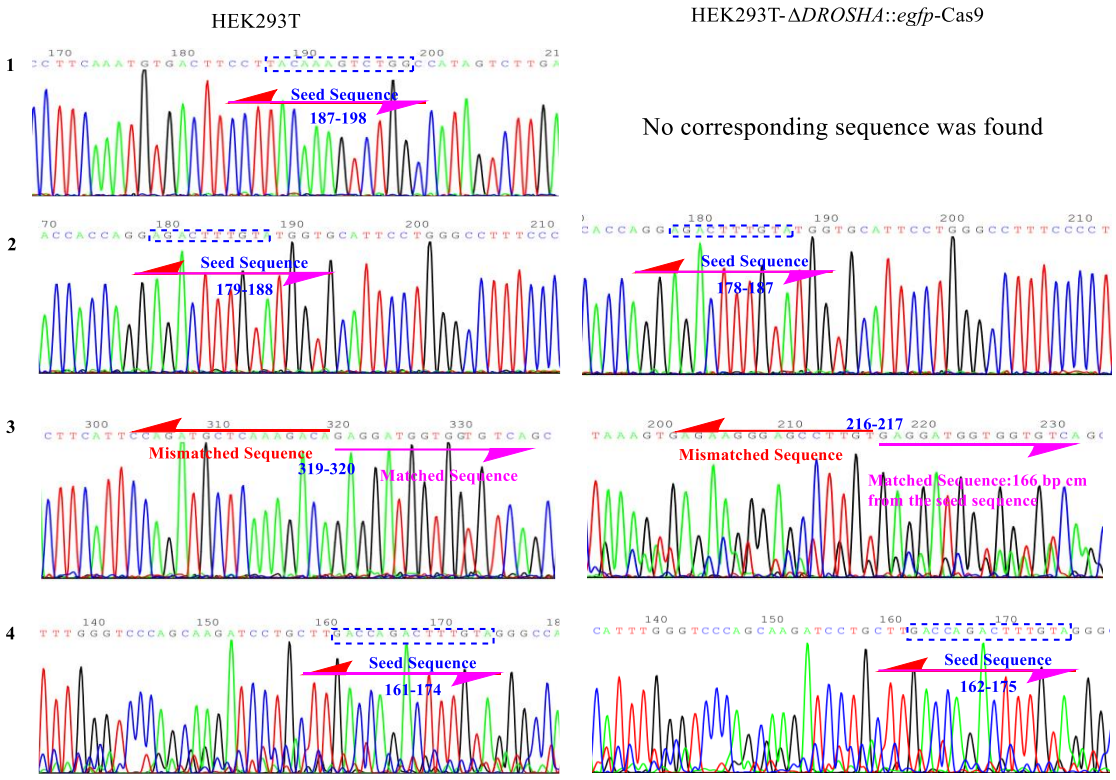
A: green; T:red; G: black; C; blue



A: green; T:red; G: black; C; blue



A: green; T:red; G: black; C; blue



Extended Data Fig. 12. Off-target analysis of the RNA-guided genome editing conducted by the *Svi*Cas3 in prokaryotic and eukaryotic cells. (A) DNA gel electrophoresis of the PCR products of potential off-target sites in the genomes of gene-edited mutants. **(B)** The DNA sequencing results of the PCR products of potential off-target sites in the gene-edited mutants (As can be seen from Extended Data Fig. 12, in the RNA-guided genome editing mediated by the *Svi*Cas3, no off-target cases or indel formation were detected, while in the RNA-guided genome editing mediated by *Sp*Cas9, off-target cases were detected in HEK293T-Δ*DROSHA*::*egfp*-Cas9 cells).

Extended Data Tables. 1 to 7

Extended Data Table 1. Strains and plasmids used in this study (http://www.addgene.org/; http://www.ncbi.nlm.nih.gov)

Item	Description	Application	Source
Strain			
C. glutamicum ATCC 13032	wild-type	*C. glutamicum* ATCC 13032	Our lab
C. glutamicum ATCC 13032-pXMJ19	pXMJ19	*C. glutamicum* ATCC 13032	This study
C. glutamicum ATCC 13032-pEC-XK99E	pEC-XK99E	*C. glutamicum* ATCC 13032	This study
C. glutamicum ATCC 13032-pXMJ19-t/g-Δ*ldh::egfp*	pXMJ19-t/g-Δ*ldh::egfp*	*C. glutamicum* ATCC 13032	This study
C. glutamicum ATCC 13032-pEC-XK99E-*cas*3	pEC-XK99E-*cas*3	*C. glutamicum* ATCC 13032	This study
C. glutamicum ATCC 13032-Δ*ldh::egfp*	Δ*ldh::egfp*	*C. glutamicum* ATCC 13032	This study
E. coli DH5α	F^-，φ80d*lac*ZΔM15，Δ(*lac*ZYA-*arg*F)U169，*deo*R，*rec*A1，*end*A1，*hsd*R17(rk^-，mk^+)，*pho*A，*sup*E44，λ-，*thi*-1，*gyr*A96，*rel*A1		Our lab
E. coli DH5α-pCas	pCas	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pCas-*cas*7-5-3-4-1-2	pCas-*cas*7-5-3-4-1-2	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pCas-*cas*6-7-5-3	pCas-*cas*6-7-5-3	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pCas-*cas*7-5-3	pCas-*cas*7-5-3	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pCas-*cas*3	pCas-*cas*3	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pEC-XK99E	pEC-XK99E	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pEC-XK99E-*cas*3	pEC-XK99E-*cas*3	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pXMJ19	pXMJ19	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pXMJ19-t-Δ*ldh::egfp*	pXMJ19-t-Δ*ldh::egfp*	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pEC-XK99E-*cas*3-t-Δ*ldh::egfp*	pEC-XK99E-*cas*3-t-Δ*ldh::egfp*	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pKC1139	pKC1139	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*sviipe*::*cat*	pKC1139-t/g-Δ*sviipe*::*cat*	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*svipam*1	pKC1139-t/g-Δ*svipam*1	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*svipam*2	pKC1139-t/g-Δ*svipam*2	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*svibla*1	pKC1139-t/g-Δ*svibla*1	*S. virginiae* IBL14	This study

E. coli DH5α-pKC1139-Δ*svibla*2	pKC1139-t/g-Δ*svibla*2	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*svibla*3	pKC1139-t/g-Δ*svibla*3	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-Δ*svibla*4	pKC1139-t/g-Δ*svibla*4	*S. virginiae* IBL14	This study
E. coli DH5α-pKC1139-t/g-Δ*lacZ*	pKC1139-t/g-*lacZ*	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pKC1139-t/g-Δ*lacZ*::*cat*	pKC1139-t/g-Δ*lacZ*::*cat*	*E. coli* JM109 (DE3)	This study
E. coli DH5α-pRS415	pRS415	*S. cerevisiae* LYC4	This study
E. coli DH5α-pRS415-*cas*3	pRS415-*cas*3	*S. cerevisiae* LYC4	This study
E. coli DH5α-pRS415-*cas*7-5-3	pRS415-*cas*7-5-3	*S. cerevisiae* LYC4	This study
E. coli DH5α-pYES2-NTA	pYES2-NTA	*S. cerevisiae* LYC4	This study
E. coli DH5α-pYES2-NTA-t/g-Δ*crtE*	pYES2-NTA-t/g-Δ*crtE*	*S. cerevisiae* LYC4	This study
E. coli DH5α-pYES2-NTA-t-Δ*crtE*	pYES2-NTA-t-Δ*crtE*	*S. cerevisiae* LYC4	This study
E. coli DH5α-pYES2-NTA-g-Δ*crtE*	pYES2-NTA-g-*crtE*	*S. cerevisiae* LYC4	This study
E. coli DH5α-pXMJ19	pXMJ19	*C. glutamicum* ATCC 13032	This study
E. coli DH5α-pXMJ19-t/g-Δ*ldh*::*egfp*	pXMJ19-t/g-Δ*ldh*::*egfp*	*C. glutamicum* ATCC 13032	This study
E. coli JM109 (DE3)	*end*A1, *rec*A1, *gyr*A96, *thi*-1, *hsd*R17 (rk^-, mk^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB)/[F', *tra*D36, *pro*AB, *lac*I^qZΔM15], λ–, lDE3	*E. coli* JM109 (DE3)	Our lab
E. coli JM109 (DE3)-Δ*lacZ*	Δ*lacZ*	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-Δ*lacZ*::*cat*	Δ*lacZ*::*cat*	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pKC1139	pKC1139	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pCas	pCas	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pCas-*cas*7-5-3-4-1-2	pCas-*cas*7-5-3-4-1-2	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pCas-*cas*6-7-5-3	pCas-*cas*6-7-5-3	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pCas-*cas*7-5-3	pCas-*cas*7-5-3	*E. coli* JM109 (DE3)	This study
E. coli JM109 (DE3)-pCas-*cas*3	pCas-*cas*3	*E. coli* JM109 (DE3)	This study
E. coli TOP10	F^-, *mcr*AΔ(*mrr-hsd* RMS-*mcr*BC), *φ*80, *lac*ZΔM15, Δ*lac*X74, *rec*A1, *ara*Δ139Δ(*ara-leu*)7697, *gal*U, *gal*K, *rps*L, (*Strr*) *end*A1, *nup*G	HEK293T / NIH-3T3	Our lab

E. coli TOP10-AIO-mCherry	AIO-mCherry	HEK293T / NIH-3T3	This study
E. coli TOP10-AIO-mCherry-*cas*3	AIO-mCherry-*cas*3	HEK293T / NIH-3T3	This study
E. coli TOP10-AIO-mCherry-t/g-Δ*DROSHA*::*egfp*	AIO-mCherry-t/g-Δ*DROSHA*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-*cas*3-t/g-Δ*DROSHA*::*egfp*	AIO-mCherry-*cas*3-t/g-Δ*DROSHA*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-t/g-Δ*CAMKMT*::*egfp*	AIO-mCherry-t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-*cas*3-t/g-Δ*CAMKMT*::*egfp*	AIO-mCherry-*cas*3-t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-t/g-Δ*Lepr*::*egfp*	AIO-mCherry-t/g-Δ*Lepr*::*egfp*	NIH-3T3	This study
E. coli TOP10-AIO-mCherry-*cas*3-t/g-Δ*Lepr*::*egfp*	AIO-mCherry-*cas*3-t/g-Δ*Lepr*::*egfp*	NIH-3T3	This study
E. coli TOP10-AIO-mCherry-t-Δ*DROSHA*::*egfp*	AIO-mCherry-t-Δ*DROSHA*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-*cas*3-t-Δ*DROSHA*::*egfp*	AIO-mCherry-*cas*3-t-Δ*DROSHA*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-t-Δ*CAMKMT*::*egfp*	AIO-mCherry-t-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-*cas*3-t-Δ*CAMKMT*::*egfp*	AIO-mCherry-*cas*3-t-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	AIO-mCherry-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-*cas*3-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	AIO-mCherry-*cas*3-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	HEK293T	This study
E. coli TOP10-AIO-mCherry-t-Δ*Lepr*::*egfp*	AIO-mCherry-t-Δ *Lepr*::*egfp*	NIH-3T3	This study
E. coli TOP10-AIO-mCherry-*cas*3-t-Δ*Lepr*::*egfp*	AIO-mCherry-*cas*3-t-Δ*Lepr*::*egfp*	NIH-3T3	This study
E. coli TOP10-pX459M-*cas*9	pX459M-*cas*9	HEK293T / NIH-3T3	This study
E. coli TOP10-pX459M-*cas*9-t/g-Δ*DROSHA*::*egfp*-9	pX459M-*cas*9-t/g-Δ*DROSHA*::*egfp*-9	HEK293T	This study
E. coli TOP10-pX459M-*cas*9-t/g-Δ*Vegfa*::*egfp*-9	pX459M-*cas*9-t/g-Δ*Vegfa*::*egfp*-9	NIH-3T3	This study

human embryonic kidney 293 /HEK293T	wild-type	HEK293T	Our lab
HEK293T-AIO-mCherry	AIO-mCherry	HEK293T	This study
HEK293T-AIO-mCherry-*cas*3	AIO-mCherry-*cas*3	HEK293T	This study
HEK293T-AIO-mCherry-t/g-Δ*DROSHA*::*egfp*	AIO-mCherry-t/g-Δ*DROSHA*::*egfp*	HEK293T	This study
HEK293T-AIO-mCherry-t-Δ*DROSHA*::*egfp*	AIO-mCherry-t-Δ*DROSHA*::*egfp*	HEK293T	This study
HEK293T-Δ*DROSHA*::*egfp*	Δ*DROSHA*::*egfp*	HEK293T	This study
HEK293T-pX459M-Cas9-t/g-Δ*DROSHA*::*egfp*-9	pX459M-Cas9-t/g-Δ*DROSHA*::*egfp*-9	NIH3T3	This study
HEK293T-Δ*DROSHA*::*egfp*-Cas9	Δ*DROSHA*::*egfp*-9	HEK293T	This study
HEK293T-AIO-mCherry-t/g-Δ*CAMKMT*::*egfp*	AIO-mCherry-t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
HEK293T-AIO-mCherry-t-Δ*CAMKMT*::*egfp*	AIO-mCherry-t-Δ*CAMKMT*::*egfp*	HEK293T	This study
HEK293T-AIO-mCherry-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	AIO-mCherry-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	HEK293T	This study
HEK293T-Δ*CAMKMT*::*egfp*	Δ*CAMKMT*::*egfp*	HEK293T	This study
HEK293T-Δ*CAMKMT*::*egfp*-g1/b2	Δ*CAMKMT*::*egfp*-g1/b2	HEK293T	This study
NIH-3T3 mouse fetus fibroblast cells/NIH-3T3	wild	NIH-3T3	Our lab
NIH-3T3-AIO-mCherry	AIO-mCherry	NIH-3T3	This study
NIH-3T3-AIO-mCherry-*cas*3	AIO-mCherry-*cas*3	NIH-3T3	This study
NIH-3T3-AIO-mCherry-t/g-Δ*Lepr*::*egfp*	AIO-mCherry-t/g-Δ*Lepr*::*egfp*	NIH-3T3	This study
NIH-3T3-AIO-mCherry-t-Δ*Lepr*::*egfp*	AIO-mCherry-t-Δ*Lepr*::*egfp*	NIH-3T3	This study
NIH-3T3-Δ*Lepr*::*egfp*	Δ*Lepr*::*egfp*	NIH-3T3	This study
NIH3T3-pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*-9	pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*-9	NIH3T3	This study
NIH-3T3-Δ*Vegfa*::*egfp*-Cas9	Δ*Vegfa*::*egfp*-9	NIH-3T3	This study
S. cerevisiae LYC4	carrying genes *crtE*, *crtB*, *crtB* and *zds* in chromosome IV in *S. cerevisiae* S288C	*S. cerevisiae* LYC4 (Original abbreviation of the strain: SyBE_Sc14D14)	Our lab
S. cerevisiae LYC4-Δ*crtE*	Δ*crtE*	*S. cerevisiae* LYC4	This study
S. cerevisiae LYC4-pYES2-NTA-t/g-Δ*crtE*	pYES2-NTA-t/g-Δ*crtE*	*S. cerevisiae* LYC4	This study
S. cerevisiae LYC4-pRS415-*cas*3	pRS415-*cas*3	*S. cerevisiae* LYC4	This study
S. cerevisiae LYC4-pRS415-*cas*7-5-3	pRS415-*cas*7-5-3	*S. cerevisiae* LYC4	This study

S. cerevisiae LYC4-pYES2-NTA-t-Δ*crtE*	pYES2-NTA-t-Δ*crtE*	*S. cerevisiae* LYC4	This study
S. cerevisiae LYC4-pYES2-NTA-g-Δ*crtE*	pYES2-NTA-g-Δ*crtE*	*S. cerevisiae* LYC4	This study
S. virginiae IBL14	wild-type	*S. virginiae* IBL14	Our lab
S. virginiae IBL14-pKC1139	pKC1139	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*sviipe*::*cat*	Δ*sviipe*::*cat*	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svipam*1	Δ*svipam*1	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svipam*2	Δ*svipam*2	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svibla*1	Δ*svibla*1	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svibla*2	Δ*svibla*2	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svibla*3	Δ*svibla*3	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svibla*4	Δ*svibla*4	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svipam*1Δ*sviipe*::*cat*	Δ*svibla*1Δ*sviipe*::*cat*	*S. virginiae* IBL14	This study
S. virginiae IBL14-Δ*svipam*1Δ*svibla*1	Δ*svibla*1Δ*svipam*1	*S. virginiae* IBL14	This study
Plasmid			
AIO-mCherry (9728 6bp)	Mammalian Expression (high copy), CRISPR, AmpR, f1 ori, U6 promoter, mCherry, Cas9n	HEK293T / NIH-3T3	Addgene
AIO-mCherry-*cas*3	codon optimized *cas*3	HEK293T / NIH-3T3	This study
AIO-mCherry-t/g-Δ*DROSHA*::*egfp*	t/g-Δ*DROSHA*::*egfp*	HEK293T	This study
AIO-mCherry-t-Δ*DROSHA*::*egfp*	t-Δ*DROSHA*::*egfp*	HEK293T	This study
AIO-mCherry-*cas*3-t/g-Δ*DROSHA*::*egfp*	*cas*3-t/g-Δ*DROSHA*::*egfp*	HEK293T	This study
AIO-mCherry-*cas*3-t-Δ*DROSHA*::*egfp*	*cas*3-t-Δ*DROSHA*::*egfp*	HEK293T	This study
AIO-mCherry-t/g-Δ*CAMKMT*::*egfp*	t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-t-Δ*CAMKMT*::*egfp*	t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-*cas*3-t/g-Δ*CAMKMT*::*egfp*	*cas*3-t/g-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-*cas*3-t-Δ*CAMKMT*::*egfp*	*cas*3-t-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-*cas*3-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	*cas*3-t_{g1}/t_{b2}-Δ*CAMKMT*::*egfp*	HEK293T	This study
AIO-mCherry-t/g-Δ*Lepr*::*egfp*	t/g-Δ*Lepr*::*egfp*	NIH-3T3	This study
AIO-mCherry-t-Δ*Lepr*::*egfp*	t-Δ*Lepr*::*egfp*	NIH-3T3	This study

AIO-mCherry-*cas*3-t/g-Δ*Lepr*::*egfp*	*cas*3-t/g-Δ*Lepr*::*egfp*	NIH-3T3	This study
AIO-mCherry-*cas*3-t-Δ*Lepr*::*egfp*	*cas*3-t-Δ*Lepr*::*egfp*	NIH-3T3	This study
pCas(12545 bp)	*cas*9, lambda-Red recombinase expression plasmid, Kan^R, ParaB promoter	*E. coli* JM109 (DE3)	Addgene
pCas-*cas*7-5-3-4-1-2	*cas*7-5-3-4-1-2	*E. coli* JM109 (DE3)	This study
pCas-*cas*6-7-5-3	*Cas*6-7-5-3	*E. coli* JM109 (DE3)	This study
pCas-*cas*7-5-3	*cas*7-5-3	*E. coli* JM109 (DE3)	This study
pCas-*cas*3	*cas*3	*E. coli* JM109 (DE3)	This study
pEC-XK99E (7018bp)	Shuttle vector (*C. glutamicum* / *E. coli*), ori pUC, KanR, lacIq, *rep*A	*C. glutamicum* ATCC 13032	Biovector Inc
pEC-XK99E-*cas*3	*cas*3	*C. glutamicum* ATCC 13032	This study
pEC-XK99E-*cas*3-t-Δ*ldh::egfp*	*cas*3-t-Δ*ldh::egfp*	*C. glutamicum* ATCC 13032	This study
pKC1139 (6308 bp)	Expression vector in *Streptomyces*, $Apramycin^R$, oriT, lac promoter	*S. virginiae* IBL14 / *E. coli* JM109 (DE3)	Our lab
pKC1139-t/g-Δ*sviipe*::*cat*	t/g-Δ*sviipe*::*cat*	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svipam*1	t/g-Δ*svipam*1	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svipam*2	t/g-Δ*svipam*2	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svibla*1	t/g-Δ*svibla*1	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svibla*2	t/g-Δ*svibla*2	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svibla*3	t/g-Δ*svibla*3	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*svibla*4	t/g-Δ*svibla*4	*S. virginiae* IBL14	This study
pKC1139-t/g-Δ*lacZ*	t/g-Δ*lacZ*	*E. coli* JM109 (DE3)	This study
pKC1139-t/g-Δ*lacZ*::*cat*	t/g-Δ*lacZ*::*cat*	*E. coli* JM109 (DE3)	This study
pRS415 (rooted from pRS415_pGal-nCas9,10793 bp)	Yeast Expression, $Leu2^+$, , Amp^R, *cas*9	*S. cerevisiae* LYC4	Addgene
pRS415-*cas*3	codon optimized *cas*3	*S. cerevisiae* LYC4	This study
pRS415-*cas*7-5-3	codon optimized *cas*7-5-3	*S. cerevisiae* LYC4	This study
pX459M-*cas*9 (10176 bp)	codon optimized *cas*9, without *mCheery* gene	HEK293T / NIH-3T3	This study

pX459M-*cas*9-t/g-Δ*DROSHA*::*egfp*	t/g-Δ*DROSHA*::*egfp*-9, with *mCheery* gene	HEK293T	This study
pX459M-*cas*9-t/g-Δ*Vegfa*::*egfp*	t/g-Δ*Vegfa*::*egfp*-9, without *mCheery* gene	NIH-3T3	This study
pXMJ19 (6601 bp)	Shuttle vector (*C. glutamicum* / *E. coli*), CmR, ori pUC, lacIq, *cat* encoding Catalase	*C. glutamicum* ATCC 13032	Biovector Inc
pXMJ19-t/g-Δ*ldh*::*egfp*	t/g-Δ*ldh*::*egfp*	*C. glutamicum* ATCC 13032	This study
pYES2-NTA (6038 bp)	Galactose-inducible expression plasmid in Yeast, Ura3^{+}, Amp^{R}, *Dcr*1	*S. cerevisiae* LYC4	Addgene
pYES2-NTA-t/g-Δ*crtE*	t/g-Δ*crtE*	*S. cerevisiae* LYC4	This study
pYES2-NTA-t-Δ*crtE*	t-Δ*crtE*	*S. cerevisiae* LYC4	This study

941

Extended Data Table 2. Primers, g-DNAs and t-DNAs used in this study*

Items	Sequence (5´ to 3´)	Note
Primers		
***E. coli* JM109 (DE3)**		
lacZ-F	tacccaacttaatcgccttgcagcaca	*lacZ*
lacZ-R	ccgtcgatattcagccatgtgccttctt	*lacZ*
lacZ-vF	cgtttcatctgtggtgcaac	Δ*lacZ*
lacZ-vR	catcagttgctgttgactgtag	Δ*lacZ*
lacZ-UF	ggctgcaggtcgactctagaGGATCCatgagcgtggtggttatgcc	UHA
lacZ-UR	acgaagccgccctgtaaacccatgccgtgggtttcaata	UHA
lacZ-DF	tattgaaacccacggcatgggtttacagggcggcttcgt	DHA
lacZ-DR	ctagagtcgacctgcagcccAAGCTTatgcgggtcgcttcacttac	DHA
lacZ-cat-UF	cgcGGATCCcgatgtcggtttccgcgag	UHA
lacZ-cat-DR	ccccccAAGCTTcgacatccagaggcacttcacc	DHA
lacZ-U*cat*-F	attgaaacccacggcatggagacgttgatcggcacgt	UHA-*cat*
lacZ-U*cat*-R	acgtgccgatcaacgtctccatgccgtgggtttcaat	UHA-*cat*
lacZ-D*cat*-F	aaatgaagttcctattccgaagttccgggactgggtggatcagtcg	DHA-*cat*
lacZ-D*cat*-R	cgactgatccacccagtcccggaacttcggaataggaacttcattt	DHA-*cat*
lacZ-cat-F	gagctggtgatatgggatagtgttcaccct	t-Δ*lacZ::cat*
*cas*7-5-3-4-1-2-F	cggagctcgaattcggatccgtggtcgccggtgccccg	*cas*7-5-3-4-1-2
*cas*7-5-3-4-1-2-R	gtttaactttaagaaggagatataccatggtcacaggatgtcactgggggcgg g	*cas*7-5-3-4-1-2
pCas-*cas*7-5-3-4-1-2-F	ttttaggaggcaaaagtggtcgccggtgcc	linearization of pCas without *cas*9
pCas-*cas*7-5-3-4-1-2-R	catctaaaatatacttcacaggatgtcactggggg	linearization of pCas without *cas*9
*cas*6-7-5-3-F	acttttttattttaggaggcaaaattggtgctgacggcgcatccg	*cas*6-7-5-3
*cas*6-7-5-3-R	aaataatcttcatctaaaatatacttcacaagacctccccggcgcggta	*cas*6-7-5-3
pCas-*cas*6-7-5-3-F	ttttaggaggcaaaatcacaagacctccccggc	linearization of pCas without *cas*9
pCas-*cas*6-7-5-3-R	catctaaaatatactttggtgctgacggcgcat	linearization of pCas without *cas*9
*cas*7-5-3-F	ttattttaggaggcaaaagtggtcgccggtgccccgaac	*cas*7-5-3
*cas*7-5-3-R	tcttcatctaaaatatacttcacaagacctccccggcgc	*cas*7-5-3

pCas-*cas*7-5-3-F	ggaggtcttgtgaagtatattttagatgaagattatttc	linearization of pCas without *cas*9
pCas-*cas*7-5-3-R	ggggcaccggcgaccacttttgcctcctaaaataaaaag	linearization of pCas without *cas*9
*cas*3-F	acttttattttaggaggcaaaagtgggccgtctggacgcggtgga	*cas*3
*cas*3-R	aaataatcttcatctaaaatatactttcacaagacctccccggcgcggta	*cas*3
pCas-*cas*3-F	taccgcgccggggaggtcttgtgaaagtatattttagatgaagattattt	linearization of pCas without *cas*9
pCas-*cas*3-R	tccaccgcgtccagacggcccacttttgcctcctaaaataaaaagt	linearization of pCas without *cas*9
Ec-lacZ-ot-F	cgatcc+TTC+CCGCCCGG (xN+PAM+seed: 6+3+8=17 nt)	number of potential off-target site: 7
Ec-lacZ-ot01-F/R	cgatccttcccgcccgg / cgcgcagggcgacaattttg	off-target analysis
Ec-lacZ-ot02-F/R	cgatccttcccgcccgg / ccgtcagcgctggatgcg	off-target analysis
Ec-lacZ-ot03-F/R	gaagtcaccggcgaaccg / gacgcaggatcatcaatgg	off-target analysis
Ec-lacZ-ot04-F/R	gcagagagttcctgttcggaag / cacacttgcggatcgcatgatg	off-target analysis
***C. glutamicum* ATCC 13032**		
ldh-F	agtgggatcgaaaatgaaagaaaccgt	*ldh*
ldh-R	actaggcgccaaagatttagaagaact	*ldh*
ldh-vF	gcctttgtgagttactcttcccag	Δ*ldh-egfp*
ldh-vR	cgatgaaacttatccacggcg	Δ*ldh-egfp*
ldh-UF	gcTCTAGAttaatactgtgctgggttaattcgc	UHA
ldh-U*egfp*-R	gctcaccattttcgatcccacttcctgatttc	UHA-*egfp*
ldh-U*egfp*-F	gtgggatcgaaaatggtgagcaagggcgag	UHA-*egfp*
ldh-D*egfp*-R	ttccaggagggttaGAATTCcttgtacagctcgtcc	DHA-*egfp*
ldh-D*egfp*-F	caagGAATTCtaaccctcctggaacacaactac	DHA-*egfp*
ldh-DR	cccAAGCTTgcagcctactctttcgttgg	DHA
ldh-egfp-vF	tcgtccttgaagaagatggtgc	*egfp*
pEC-XK99E-F	gaggtcttgtgatcgacaaggagcaagcttggctg	linearization of pEC-XK99E
pEC-XK99E-R	cgaattccttggtctgtttcctgtgtg	linearization of pEC-XK99E
pEC-XK99E-*cas*3-F	aggaattcggtgggccgtctggacgcggtg	linearization of pEC-XK99E-*cas*3-F
pEC-XK99E-*cas*3-R	ctccttgtcgatcacaagacctccccggcg	linearization of pEC-XK99E-*cas*3-R
pXMJ19-vF	tcagcttggctgttttggcggat	pXMJ19
pXMJ19-vR	ttgagatccttttttctgcgcgtaatctg	pXMJ19

pEC-XK99E-*cas*3-vF	acagaccatggaattcggtgggcc	*cas*3
pEC-XK99E-*cas*3-vR	gccaagcttgctccttgtcgatca	*cas*3
Cg-ldh-ot-F	cacgcgg+TTC+CATTCCA (xN+PAM+seed: 7+3+7=17 nt)	number of potential off-target site: 18
Cg-ldh-ot01-F/R	aggaaggcaagttcgctgac / cttcaaagcgcacaccgatg	off-target analysis
Cg-ldh-ot02-F/R	ctaagaagaatgcgaagccgag / catctagttgggcgagtctg	off-target analysis
Cg-ldh-ot03-F/R	gcacaaccacatcagacggatc / gaagaggttaacgtccagctg	off-target analysis
Cg-ldh-ot04-F/R	ttgtccagcattcggctaag /cttggtctgaccacattttc	off-target analysis
ldh-UF$_a$	gcTCTAGAatgaaagaaaccgtcggtaacaagattg	UHA
ldh-U*egfp*-R$_a$	ctcctcgcccttgctcaccattgggttggacgccacgaggaag	UHA-*egfp*
ldh-U*egfp*-F$_a$	cttcctcgtggcgtccaacccaatggtgagcaagggcgaggag	UHA-*egfp*
*ld*h-D*egfp*-R$_a$	gtgttccaggagggttaGAATTCcttgtacagctcgtcc	DHA-*egfp*
ldh-D*egfp*-F$_a$	caaggaattctaaccctcctggaacacaactacg	DHA-*egfp*
ldh-egfp-vF$_a$	cgatgcccttcagctcgatg	*egfp*
ldh-vR$_a$	ggatccccggaactagctc	Δ*ldh-egfp*
pEC-XK99E-*cas*3-F$_a$	cggcatgcatttacgttgcagcctactctttcgttggg	linearization of pEC-XK99E-*cas*3
pEC-XK99E-*cas*3-R$_a$	ccattcgatggtgtcatgaaagaaaccgtcggtaacaag	linearization of pEC-XK99E-*cas*3
HEK293T		
AIO-mCherry-F	cccAAGCTTggctccggagccacgaacttc	AIO-Mcherry
AIO-mCherry-R	gcTCTAGAgccatttgtctgcagaattgg	AIO-Mcherry
g-*DROSHA*-F	aggaattcggcaggaagagtctttgaaaagataatgtatg	g-*DROSHA*
g-*DROSHA*-R	cgcGGATCCaagggccctctagactcgagagac	g-*DROSHA*
DROSHA-vF	ctgtttttaccatgaaggaaacgagagtatgaag	Δ*DROSHA-egfp*
DROSHA-vR	catttgctcttttcctagtatttcctatctc	Δ*DROSHA-egfp*
DROSHA-egfp-vF	cacaacgtctatatcatggccgacaagc	*egfp*
DROSHA-egfp-vR	acttgaagaagtcgtgctgcttcatgtggtcg	*egfp*
DROSHA-UF	cgcGGATCCatgcctataagtctttttagctgctg	UHA
DROSHA-U*egfp*-R	ctcgcccttgctcaccatcatgatgttccgcctggatatgtcac	UHA-*egfp*
DROSHA-U*egfp*-F	gtgacatatccaggcggaacatcatgatggtgagcaagggcgag	UHA-*egfp*
DROSHA-D*egfp*-R	ccaccttctaaggacctaattccacctttaatttagaattccttgtacagctc	DHA-*egfp*

DROSHA-De*gfp*-F	gagctgtacaaggaattctaaattaaaggtggaattaggtccttagaaggtgg	DHA-*egfp*
DROSHA-DR	ctcttcctgccgaattcctgacaccattaatagtccttaagacatgtaacagg	DHA
DROSHA-egfp-vF_a	gcaccatcttcttcaaggacga	*egfp*
HEK293T-*DROSHA*-ot-F	tgctca+TTC+tttcaacagtg (xN+PAM+seed: 6+3+11=20 nt)	number of potential off-target site: 315
HEK293T-*DROSHA*-ot01-F/R	agatgtagcctcactcctgagatc / cattaagagctcaggtcaagtgtatg	ChrVI, off-target analysis
HEK293T-*DROSHA*-ot02-F/R	gatattgattcttccaagtcatgagc / ccacatacacgaatcaataagtg	ChrV, off-target analysis
HEK293T-*DROSHA*-ot03-F/R	caaagccaggctttggtcatcatgc / gttgtattgtcctcagatggttc	ChrV, off-target analysis
HEK293T-*DROSHA*-ot04-F/R	tattggttccatatgaaatgtggaatag / ctccaaataataagagctagctatg	ChrV, off-target analysis
g-*DROSHA*-F-9	gagggcctatttcccatgattc	g-*DROSHA*-9
g-*DROSHA*-R-9	ccggagccaagcttaaaaaaaagcaccgactcggtgccac	g-*DROSHA*-9
DROSHA-vF-9	tggtgaacaatgggatgttgaattggtc	Δ*DROSHA-egfp*-9
DROSHA-vR-9	gtttacctgctccgttcgtagc	Δ*DROSHA-egfp*-9
DROSHA-UF-9	ctcgaggaattcggcaggaagaagtggagcagctttaaggaatggtcg	UHA-9
DROSHA-U*egfp*-R-9	cgcccttgctcaccatttggaggttcatattgatattgcacagg	UHA-*egfp*-9
DROSHA-U*egfp*-F-9	cctgtgcaatatcaatatgaacctccaaatggtgagcaagggcg	UHA-*egfp*-9
DROSHA-D*egfp*-R-9	ctgataattaacctgctgcggcatgactgttagaattccttgtacag	DHA-*egfp*-9
DROSHA-D*egfp*-F-9	ctgtacaaggaattctaacagtcatgccgcagcaggttaattatcag	DHA-*egfp*-9
DROSHA-DR-9	tcatgggaaataggccctcgatggtgttctccctcggtcataatcag	DHA-9
DROSHA-egfp-vF-9	cacaacgtctatatcatggccgacaagc	*egfp*
DROSHA-egfp-vR-9	cttgccggtggtgcagatgaacttc	*egfp*
pX459M-*cas*9-*DROSHA*-F	gctctagaagatttctttttcttagcttgaccagctttcttagtagcagcag	linearization of pX459M-*cas*9-F
pX459M-*cas*9-*DROSHA*-R	cgcggatcctgcatcgagactagataactgatctacccagctttcttgtac	linearization of pX459M-*cas*9-R
pX459M-*cas*9-*mCherry*-F	cgcggatccttagaattccttgtacagctcg	*mCherry*
pX459M-*cas*9-*mCherry*-R	gctctagatctctgttaaagcaagcaggag	*mCherry*
HEK293T-*DROSHA*-9-ot	seed+PAM=10+3=agactttgta+ TGG, seed+PAM=12+3=ccagactttgta+AGG, seed+PAM=13+3=tacaagactttgt+AGG, seed+PAM=14+3=gaccagactttgta+GGG	Potential off-target sites: 8468, PAM: NGG
HEK293T-*DROSHA*-9-ot01-F/R	ccaataatgtcagttcagatattgtg / gttaacatttatatttggttaactag	ChrII, off-target analysis

HEK293T-*DROSHA*-9-ot02-F/R	cacagaacaggaaggatttcaatg / cgaaaatacaggtctccttg	ChrVIII, off-target analysis
HEK293T-*DROSHA*-9-ot03-F/R	caaggcaggcagagtatacaag / cccgctgacaccaccatcctc	ChrVIII, off-target analysis
HEK293T-*DROSHA*-9-ot04-F/R	ctcactcccaggaaatctgcc / ctttgaacctcattttccgtg	ChrX, off-target analysis
g-*CAMKMT*-F	ccatcataggagctatggactctttgaaaagataatgtatgattatg	g-*CAMKMT*
g-*CAMKMT*-R	cgGGATCCagacataaaaaacaaaaaaaccctcg	g-*CAMKMT*
CAMKMT-vF	ctgtagatggaatagcagacctgaag	Δ*CAMKMT*::*egfp*
CAMKMT-vR	gctcatatttcattgacaattcactatcttaagag	Δ*CAMKMT*::*egfp*
CAMKMT-*egfp*-vF	gatcactctcggcatggacgag	*egfp*
CAMKMT-UF	cgGGATCCtcagcttgattctctaag	UHA
CAMKMT-U*egfp*-R	tcgcccttgctcaccattccactattatgcctagaaaggag	UHA-*egfp*
CAMKMT-U*egfp*-F	atggtgagcaagggcgag	UHA-*egfp*
CAMKMT-D*egfp*-R	ttagaattccttgtacagctcgtcc	DHA-*egfp*
CAMKMT-D*egfp*-F	ctgtacaaggaattctaagagatggagagagtaggtcagaaag	DHA-*egfp*
CAMKMT-DR	cattatcttttcaaagagtccatagctcctatgatggtatg	DHA
CAMKMT-vF_a	tgctgttatagagtaaatttatgtattgttgagtgtgg	Δ*CAMKMT*::*egfp*
CAMKMT-vR_a	cctaaattcagccaaaagttaacagttctg	Δ*CAMKMT*::*egfp*
g-*CAMKMT*-exon1-vF	cacccccatacatacgcaaattgag	Δ*CAMKMT*::*egfp*-exon1
g-*CAMKMT*-exon1-vR	cagagactggcgagaaggggaaaag	Δ*CAMKMT*::*egfp*-exon1
g-*CAMKMT*-exon1-*egfp*-vR	gaacttgtggccgtttacgtcg	*egfp*
g-*CAMKMT*-exon1-UR	gctctagaggctctggacacaaaccagtg	g-exon1-UHA
g-*CAMKMT*-exon1-U*egfp*-R	aagcacagaaggctaaaacaatggtgagcaagggcgaggag	g-exon1-UHA-*egfp*
g-*CAMKMT*-exon1-U*egfp*-F	tcgcccttgctcaccattgttttagccttctgtgcttcatctc	g-exon1-UHA-*egfp*
g-*CAMKMT*-exon1-D*egfp*-R	tggacgagctgtacaagtaagtaagggagaacctgctcgcc	g-exon1-DHA-*egfp*
g-*CAMKMT*-exon1-D*egfp*-F	gcgagcaggttctcccttacttacttgtacagctcgtccatgcc	g-exon1-DHA-*egfp*
b-*CAMKMT*-exon1-exon2-R	gaagtagaagaacccctctaaatggttaccttGaTgagatactgtattc	g-b-exon1-exon2-DHA
b-*CAMKMT*-exon1-exon2-F	atctcAtCaaggtaaccatttagaaggggttcttctacttcgatctg	g-b-exon1-exon2-DHA
b-*CAMKMT*-exon2-F	cgcggatccatttcaaaggtActgaaAcaaaaacac	b-exon2
b-*CAMKMT*-exon2-F’	cgcggatccatttcaaaggtTctgaaGcaaaaacac	b-exon2
b-*CAMKMT*-exon2-vF	tatttggctttatttcaaaggtActgaaAc	b-exon2
b-*CAMKMT*-exon2-vR	atacagcaagacccagttaaaatagcag	b-exon2

b-*CAMKMT*-exon2-vF$_a$	gtctttgctctctgatgtcttcaagtaatg	b-exon2
b-*CAMKMT*-exon2-vR$_a$	tagaataatggttaccttGaTgagatactg	b-exon2
HEK293T-*CAMKMT*-ot-F	ataatagtgga+TCC+ttgaatgttgaa (xN+PAM+seed: 6+3+11=20 nt)	number of potential off-target site: 11
HEK293T-*CAMKMT*-ot01-F/R	ataatagtggaTCCttgaatgttgaa / cctttgaggtagctgtctatgaaag	ChrI, off-target analysis
HEK293T-*CAMKMT*-ot02-F/R	ataatagtggaTCCttgaatgttgaa / gtacaatagtatccacagtgcac	ChrI, off-target analysis
HEK293T-*CAMKMT*-ot03-F/R	ataatagtggaTCCttgaatgttgaa / agaccaaaaagactgactgttatg	ChrI, off-target analysis
HEK293T-*CAMKMT*-ot04-F/R	ataatagtggaTCCttgaatgttgaa / cattctgtatccttgagtaggg	ChrIII, off-target analysis
NIH-3T3		
g-*Lepr*-F	caatacatgttatgctgagtgatctttgaaaagataatgtatgattatg	g-*Lepr*
g-*Lepr*-R	cgGGATCCagacataaaaaacaaaaaaacccctcg	g-*Lepr*
Lepr-vF	tgaggaaaattgatgcaaggatttccag	Δ*Lepr*::*egfp*
Lepr-vR	catctaattgactcaattaagtcttctctgtcc	Δ*Lepr*::*egfp*
Lepr-egfp-vR	gaacttgtggccgtttacgtcg (reverse complement)	*egfp*
Lepr-egfp-vR$_a$	cttgccggtggtgcagatgaacttc (reverse complement)	*egfp*
Lepr-UF	cgggatccgaaggtagacgctcagggttg	UHA
Lepr-U*egfp*-R	tcctcgcccttgctcaccatgataccacttgaattaaatttag	UHA-*egfp*
Lepr-U*egfp*-F	atggtgagcaagggcgag	UHA-*egfp*
Lepr-D*egfp*-R	ttagaattccttgtacagctcgtcc	DHA-e*gfp*
Lepr-D*egfp*-F	ggacgagctgtacaaggaattctaaactgaagggaagacactggc	DHA-*egfp*
Lepr-DR	atcatacattatcttttcaaagatcactcagcataacatgtattgattg	DHA
NIH-3T3-*Lepr*-ot-F	gtggtatc+TAC+gttcctgagt (xN+PAM+seed: 8+3+10=21nt)	number of potential off-target site: 17
NIH-3T3-*Lepr*-ot01-F/R	ctgcccacctcttaccatag/agtggtagcacaatccttaaatc	ChrI, off-target analysis
NIH-3T3-*Lepr*-ot02-F/R	gatgaagggtgtgaatggtg/catgatagaaagacacgtggaatac	ChrI, off-target analysis
NIH-3T3-*Lepr*-ot03-F/R	cacctttacatcagctacgg/gtatgctaggtgctgttgaag	ChrI, off-target analysis
NIH-3T3-*Lepr*-ot04-F/R	aagtagacatggtggcagag/gccatggaagtgttagctttc	ChrII, off-target analysis
g-*Vegfa*-F-9	gagggcctatttcccatgattc	g-*Vegfa*-9
g-*Vegfa*-R-9	ccggagccaagcttaaaaaaaagcaccgactcggtgccac	g-*Vegfa*-9
Vegfa-vF-9	ttgcatcggaccagtcgcgctgacggac	Δ*Vegfa*::*egfp*-9

Vegfa-vR-9	gtgcaggtggcccaacaagctagagcggtg	Δ*Vegfa*::*egfp*-9
Vegfa-UF-9	gttggctccgaatttctcgacttttcgtccaacttctgggctcttctcgctc	UHA-9
Vegfa-U*egfp*-R-9	cgcccttgctcaccatgttcatggtttcggaggccc	UHA-*egfp*-9
Vegfa-U*egfp*-F-9	gggcctccgaaaccatgaacatggtgagcaagggcg	UHA-*egfp*-9
Vegfa-D*egfp*-R-9	gatcgtacgtgcggtgactctggtgggtggttagaattccttgtacag	DHA-*egfp*-9
Vegfa-D*egfp*-F-9	ctgtacaaggaattctaaccacccaccagagtcaccgcacgtacgatc	DHA-*egfp*-9
Vegfa-DR-9	ctcttcctgccgaattcctcagcaatccatcctaaaactttccccaaactcactgc	DHA-9
*Vegfa-egfp-v*R-9	acttgaagaagtcgtgctgcttcatgtggtcg	*egfp*-9
pX459M-*cas*9-F	gaccttccgcttcttctttggcatgg	linearization of pX459M-*cas*9-F
pX459M-*cas*9-R	actagtggtggcggctcaaagcgtc	linearization of pX459M-*cas*9-R
pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*-F	ttcctgccgaattcctcgag	linearization of pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*-F
pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*-R	ttttttttaagcttggctccgg	linearization of pX459M-Cas9-t/g-Δ*Vegfa*::*egfp*–R
NIH3T3-*Vegfa*-ot-F	atgaact+TTC+tgctctcttg (xN+PAM+seed: 7+3+10=20 nt)	number of potential off-target site: 341
NIH3T3-*Vegfa*-ot01-F/R	ggattacctcctggtatgttattc / cacacaaactccaaacaaaagac	ChrXVII, off-target analysis
NIH3T3-*Vegfa*-ot02-F/R	gccattcagatatactgctttctag / cactaaccaggaagaaatgtactg	ChrXV, off-target analysis
NIH3T3-*Vegfa*-ot03-F/R	cagccctgttctgctaatctc / gccagatgtggtaacacatatg	ChrXV, off-target analysis
NIH3T3-*Vegfa*-ot04-F/R	gtaacttgcggattattgactcg / tagaacagcctcagactttagc	ChrXVII, off-target analysis
***S. cerevisiae* LYC4**		
crtE-F	tccattggagtttactccacaag	*crtE*
crtE-R	ttttccctgtacgaccttcggatt	*crtE*
crtE-UF	cccAAGCTTcagccattccattggagtttactcc	UHA
crtE-UR	aacggcaatacgaaacaaaccaccaactggcataggaataggtgttggac	UHA
crtE-DF	gtccaacacctattcctatgccagttggtggtttgtttcgtattgccgtt	DHA
crtE-DR	cgcGGATCCttttccctgtacgaccttcggatt	DHA
crtE-vF	cagcatacacctcactagggtag	Δ*crtE*
crtE-vR	ccccttccagtgcattatgcaa	Δ*crtE*
*cas*3-Fsc	aaaaaaatgggtagattggatgccgttg	codon *cas*3

*cas*3-Rsc	gtcagccctgcttaaaacttcaccggccctatatgcag	codon optimized *cas*3
pRS415-Fsc	tatagggccggtgaagttttaagcagggctgaccccaag	pRS415 without *cas*9
pRS415-Rsc	caacggcatccaatctacccatttttttcccgggggatccactagttc	pRS415 without *cas*9
pRS415-*cas*3-Fsc	ccggattctagaactagtggatc	pRS415 with codon optimized *cas*3
pRS415-*cas*3-Rsc	cttttcggttagagcggatgtgg	pRS415 with codon optimized *cas*3
*cas*7-Fsc	ataaaacatcatcacagaattcatggttgccggtgcacc	codon optimized *cas*7
*cas*7-Rsc	aaaccttggatggatcacaccttcctcttcttcttgggg	codon optimized *cas*7
pRS415-*cas*7-Fsc	gaagaagaggaaggtgtgatccatccaaggtttcaaggccg	pRS415 with codon optimized *cas*7
pRS415-*cas*7-Rsc	ggcaaccatgaattctgtgatgatgttttatttgttttgattgg	pRS415 with codon optimized *cas*7
*cas*5-Fsc	atacaataatagaattcatgacaggtactgaagttactgcc	codon optimized *cas*5
*cas*5-Rsc	aattcttagttaaaagcacttcacaccttcctcttcttcttgggg	codon optimized *cas*5
pRS415-*cas*5-Fsc	aagaggaaggtgtgaagtgcttttaactaagaattattagtcttttctgc	pRS415-*cas*7-3
pRS415-*cas*5-Rsc	tcagtacctgtcatgaattctattattgtatgttatagtattagttgcttgg	pRS415-*cas*7-3
pRS415-*cas*7-5-Fsc	tgattgtttatttacaggatccctctccccgcgcgttg	pRS415 with codon optimized *cas*7-5
pRS415-*cas*7-5-Rsc	gtagtggaatcaaggatccgcggtttgcgtattgggcg	pRS415 with codon optimized *cas*7-5
pRS415-*cas*7-5-3-Fsc	gctccggtgcccgtcagtggatccttgattccactacagttac	pRS415-*cas*7-5-3
pRS415-*cas*7-5-3-Rsc	gacgctttgagccgccaccactagtgcaaattaaagccttcgagcgtcc	pRS415-*cas*7-5-3
Sc-crtE-ot-F	ctccgt+TTC+tagtgc (xN+PAM+seed: 6+3+6=15 nt)	number of potential off-target site: 87
Sc-crtE-ot01-F/R	cctatcgaacagaagattagactac / cagccgtatgttcatcg	ChrII, off-target analysis
Sc-crtE-ot02-F/R	tttactacccatgcgacttgtc / gtacaaatagctacaccac	ChrII, off-target analysis
Sc-crtE-ot03-F/R	gttgatgatgggtcactattgaag / ggatttgcattgttcaatcc	ChrVII, off-target analysis
Sc-crtE-ot04-F/R	ctccgtttctagtgc / cacacttaaacgtatacgag	ChrXVI, off-target analysis
***S. virginiae* IBL14**		
sviipe-vF	ccaggccaatcccgcccaagcggccctttt	Δ*sviipe*::*cat*
sviipe-vR	cccgcaagggctggatcatggacacctat	Δ*sviipe*::*cat*
sviipe-UF	gtaaaacgacggccagtgccAAGCTTgcacctcaaccacggctcctac gg	UHA
sviipe-U*cat*-R	tcttacgtgccgatcaacgtctccctcggtggcgttggagatgaa	UHA-*cat*
sviipe-U*cat*-F	ttcatctccaacgccaccgagggagacgttgatcggcacgtaaga	UHA-*cat*
sviipe-D*cat*-R	agccctggtggtcctcccatgggaacttcggaataggaacttcattta	DHA-*cat*

sviipe-D*cat*-F	taaatgaagttcctattccgaagttcccatgggaggaccaccagggct	DHA-*cat*
sviipe-DR	tcgcgcgcggccgcggatccTCTAGAaggcgttcgtactcctcgggc	DHA
*svipam*1-vF	tcgttcccgcagaccaccggctccctcaa	Δ*svipam*1
*svipam*1-vR	ttggcccacagggtcgtctggtcggtgtagtt	Δ*svipam*1
*svipam*1-UF	gtaaaacgacggccagtgccAAGCTTcgagaccgacgccttcct	UHA
*svipam*1-UR	gatgaggttctgcgacgggacctccaggtagaggtcggtcacg	UHA
*svipam*1-DF	cgtgaccgacctctacctggaggtcccgtcgcagaacctcatc	DHA
*svipam*1-DR	tcgcgcgcggccgcggatccTCTAGAgttgaggttgacgaccatccg	DHA
*svipam*2-vF	gtgaccaccgaggcctacga	Δ*svipam*2
*svipam*2-vR	gaggaggggcgtggtgagga	Δ*svipam*2
*svipam*2-UF	gtaaaacgacggccagtgccAAGCTTaccaccgaggcctacgaggtct	UHA
*svipam*2-UR	tctccaggatctcgacctcgacgaagaggatgtgggtgccgat	UHA
*svipam*2-DF	atcggcacccacatcctcttcgtcgaggtcgagatcctggaga	DHA
*svipam*2-DR	tcgcgcgcggccgcggatccTCTAGAtgctgtcggcctccatgtgc	DHA
*svibla*1-vF	acggcgacgagggactgtg	Δ*svibla*1
*svibla*1-vR	ccgcccgtggtgcggaccgtcgagaagggga	Δ*svibla*1
*svibla*1-UF	gtaaaacgacggccagtgccAAGCTTctcccccgaacacccctactac	UHA
*svibla*1-UR	aagagcagttcgacggaggcgcgatgatggaggtgaacagc	UHA
*svibla*1-DF	gctgttcacctccatcatcgcgcctccgtcgaactgctctt	DHA
*svibla*1-DR	tcgcgcgcggccgcggatccTCTAGAtgtccatcccggagaaccag	DHA
*svibla*2-vF	atgcgcacgcaccgcgccgc	Δ*svibla*2
*svibla*2-vR	ccgggtgaaggtgcccaggtgttc	Δ*svibla*2
*svibla*2-UF	gtaaaacgacggccagtgccAAGCTTatcctgctcacagccgggacgg	UHA
*svibla*2-UR	tgatccgtttcccgtcgaggaccccgctgcggtgggagaacaggtc	UHA
*svibla*2-DF	gacctgttctcccaccgcagcggggtcctcgacgggaaacggatca	DHA
*svibla*2-DR	tcgcgcgcggccgcggatccTCTAGAccaggtactcgatgcggaccttgc	DHA
*svibla*3-vF	gtgcctccaggggctgcgcac	Δ*svibla*3
*svibla*3-vR	ctccgcccggccctgtttgacca	Δ*svibla*3

*svibla*3-UF	gtaaaacgacggccagtgccAAGCTTctctcaccgatgtcgcagaactgat	UHA
*svibla*3-UR	caggtgcaggcagacgtggccggggtggtggtgggtgatg	UHA
*svibla*3-DF	catcacccaccaccaccccggccacgtctgcctgcacctg	DHA
*svibla*3-DR	tcgcgcgcggccgcggatccTCTAGAcgccgcagatgggcttccgcttcc	DHA
*svibla*4-vF	gacggagttgaccgcccgagcc	Δ*svibla*4
*svibla*4-vR	ggcactcgcctggctgttcctc	Δ*svibla*4
*svibla*4-UF	gtaaaacgacggccagtgccAAGCTTgctggtgttcctgggctgcgtc	UHA
*svibla*4-UR	tgggtgttggtggtgttggcgaggccgtgctcccgttccatgtcc	UHA
*svibla*4-DF	ggacatggaacgggagcacggcctcgccaacaccaccaacaccca	DHA
*svibla*4-DR	tcgcgcgcggccgcggatccTCTAGAttgtacccgcccctgtggctgat	DHA
g-DNA		
***E. coli* JM109 (DE3)**		
g-*lacZ* (222 bp) T7 promoter-19 bp rrnB_T1 terminator-44 bp	ggctgcaggtcgactctagaGGATCCtaatacgactcactatagggaatatt**gtcctcatcgccccttcgaggggtcgcaac***cccgcccggtgcagtatgaaggcggcggagccgacaccacg***gtcctcatcgccccttcgaggggtcgcaac**ataaaacgaaaggctcagtcgaaagactgggcctttcgttttatGAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
***C. glutamicum* ATCC 13032**		
g-*ldh* (366 bp) ptuf promoter-179 bp rrnB T1 terminator-87 bp	ccacagggtagctggtagtttgaaaatcaacgccgttgcccttaggattcagtaactggcacattttgtaatgcgctagatctgtgtgctcagtcttccaggctgcttatcacagtgaaagcaaaaccaattcgtggctgcgaaagtcgtagccaccacgaagttccaggaggacataca**gtcctcatcgccccttcgaggggtcgcaac***cattccaagccggagaatttccacactgcgtaggtcagga***gtcctcatcgccccttcgaggggtcgcaac**caaataaaacgaaaggctcagtcgaaagactgggccttttcgttttatctgttgtttgtcggtgaacgctctcctgagtaggacaaat	PAM: ttc R: 30+30 bp S: 40 bp
HEK293T		
g-*DROSHA* (389 bp) snR52 promoter-269 bp	tctttgaaaagataatgtatgattatgctttcactcatatttatacagaaacttgatgttttctttcgagtatatacaaggtgattacatgtacgtttgaagtacaactctagattt	PAM: ttc R: 30+30 bp

SUP4 terminator-20 bp	tgtagtgccctcttgggctagcggtaaaggtgcgcatttttcacaccctacaat gttctgttcaaaagattttggtcaaacgctgtagaagtgaaagttggtgcgcatg tttcggcgttcgaaacttctccgcagtgaaagataaatgatc**gtcctcatcgcc ccttcgaggggtcgcaac***tttcaacagtgggcacaatacaatagagggcc atgcccag***gtcctcatcgcccccttcgaggggtcgcaac**ttttttttgttttttatgt ct	S: 40 bp (outside of the exon3)
g-*DROSHA*-9 (352 bp) U6 promoter-241+8 bp *Sp* terminator-41 bp	gagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttag agagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgt gacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaat ggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatct tgtggaaaggac**gaaacacc***gtacaaagtctggtcgtgga***gttttagagct agaaatagcaagttaaaataaggctagtccg**ttatcaacttgaaaaagtgg caccgagtcggtgctttttt	PAM: ggg R: 42 bp S: 20 bp
g-*CAMKMT* (359 bp) snR52 promoter-269 bp SUP4 terminator-20 bp	tctttgaaaagataatgtatgattatgctttcactcatatttatacagaaacttgatgt tttctttcgagtatatacaaggtgattacatgtacgtttgaagtacaactctagattt tgtagtgccctcttgggctagcggtaaaggtgcgcatttttcacaccctacaat gttctgttcaaaagattttggtcaaacgctgtagaagtgaaagttggtgcgcatg tttcggcgttcgaaacttctccgcagtgaaagataaatgatc**tcgcaac***ttgaat gttgaagatgtccttaccagctttgacaatacag***gtcctcatcgcccccttcga gggg**ttttttgttttttatgtct	PAM: tcc R: 7+23 bp S: 40 bp (Exon3=64+1 nt)
NIH-3T3		
g-*Lepr* (359 bp) snR52 promoter-269 bp SUP4 terminator-20 bp	tctttgaaaagataatgtatgattatgctttcactcatatttatacagaaacttgatgt tttctttcgagtatatacaaggtgattacatgtacgtttgaagtacaactctagattt tgtagtgccctcttgggctagcggtaaaggtgcgcatttttcacaccctacaat gttctgttcaaaagattttggtcaaacgctgtagaagtgaaagttggtgcgcatg tttcggcgttcgaaacttctccgcagtgaaagataaatgatc**tcgcaac***gttcct gagttatccaaaacagtcttccactgttgctttg***gtcctcatcgcccccttcgag ggg**ttttttgttttttatgtct	PAM: tac R: 7+23 bp S: 40 bp (Exon2=328+2 nt)
g-*Vegfa*-9 (352 bp) U6 promoter-241+8 bp *Sp* terminator-41 bp	gagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttag agagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgt gacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaat	PAM: ggg R: 42 bp S: 20 bp

	ggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatct tgtggaaaggac**gaaacacc***cgcttaccttggcatggtgg***gttttagagcta gaaatagcaagttaaaataaggctagtccg**ttatcaacttgaaaaagtggc accgagtcggtgctttttt	
***S. cerevisiae* LYC4**		
g-*crtE* (389+17＝406 bp) snR52 promoter-269 bp SUP4 terminator-20 bp	cgGAATTCtctttgaaaagataatgtatgattatgctttcactcatatttatacag aaacttgatgttttctttcgagtatatacaaggtgattacatgtacgtttgaagta caactctagattttgtagtgccctcttgggctagcggtaaaggtgcgcatttttcc acaccctacaatgttctgttcaaaagattttggtcaaacgctgtagaagtgaaag ttggtgcgcatgtttcggcgttcgaaacttctccgcagtgaaagataaatgatc**g tcctcatcgccccttcgaggggtcgcaac***tagtgcctcatcttcttcttcagca tcctccgaaaatggt***gtcctcatcgccccttcgaggggtcgcaac**ttttttgtt ttttatgtctCTCGAGcgg	PAM: ttc R: 30+30 nt S: 40 nt (*crtE*=1131 nt)
***S. virginiae* IBL14**		
g-*sviipe* (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtataatactagt**gtcctcatcgccccttcgaggggtcgcaac***gggaac ctgcacaaatgggggtacgccccctccggcagcg***gtcctcatcgccccttc gaggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttt gGAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
g-*svipam*1 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtataatactagt**gtcctcatcgccccttcgaggggtcgcaac***cgagctc ggcccccgaccgcggcaccggctacgccgtcgc***gtcctcatcgccccttc gaggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttt gGAATTCgtaatcatgtcatagctgtt	PAM: tcc R: 30+30 bp S: 40 bp
g-*svipam*2 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtataatactagt**gtcctcatcgccccttcgaggggtcgcaac***ggtgtac aggtcctggtagtcggccatggcgttggtgatg***gtcctcatcgccccttcga ggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttttgG AATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp

g-*svibla*1 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtatatactagt**gtcctcatcgccccttcgaggggtcgcaac***taccag* *gagtccacccgtgaggggcagttgaggctgctgt***gtcctcatcgccccttcg** **aggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttttg GAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
g-*svibla*2 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtatatactagt**gtcctcatcgccccttcgaggggtcgcaac***aagccc* *gccgggatgacgcacaccagcaccgagttctccg***gtcctcatcgccccttc** **gaggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttttt gGAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
g-*svibla*3 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtatatactagt**gtcctcatcgccccttcgaggggtcgcaac***gactac* *atgagcgagaaactgaccgccgccggagcccccg***gtcctcatcgcccctt** **cgaggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttt tgGAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
g-*svibla*4 (182+52=234 bp) pJ23119 promoter: 40 bp *S. pyogenes* terminator: 42 bp	ggctgcaggtcgactctagaGGATCCgatccttgacagctagctcagtcc taggtatatactagt**gtcctcatcgccccttcgaggggtcgcaac***ctggccc* *gccgcgtccactacaccgcggccgaggtcgacg***gtcctcatcgcccctcg** **aggggtcgcaac**ttatcaacttgaaaaagtggcaccgagtcggtgctttttttg GAATTCgtaatcatgtcatagctgtt	PAM: ttc R: 30+30 bp S: 40 bp
t-DNA		
***E. coli* JM109 (DE3)**		
t-Δ*lacZ*-1144 bp (*lacZ*=**3075** nt) (*Escherichia coli* str. K-12 substr. MG1655, ~4.6 Mbp)	Atgagcgtggtggttatgccgatcgcgtcacactacgtctgaacgtcgaaaac ccgaaactgtggagcgccgaaatcccgaatctctatcgtgcggtggttgaact gcacaccgccgacggcacgctgattgaagcagaagcctgcgatgtcggtttc cgcgaggtgcggattgaaaatggtctgctgctgctgaacggcaagccgttgct gattcgaggcgttaaccgtcacgagcatcatcctctgcatggtcaggtcatgga tgagcagacgatggtgcaggatatcctgctgatgaagcagaacaactttaacg ccgtgcgctgttcgcattatccgaaccatccgctgtggtacacgctgtgcgacc gctacggcctgtatgtggtggatgaagccaatattgaaacccacggcatgg	UHA: 423 bp DHA: 721 bp Del (**3075**-2655 nt): 420 bp

	Gtttacagggcggcttcgtctgggactgggtggatcagtcgctgattaaatatg atgaaaacggcaacccgtggtcggcttacggcggtgattttggcgatacgccg aacgatcgccagttctgtatgaacggtctggtctttgccgaccgcacgccgcat ccagcgctgacggaagcaaaacaccagcagcagtttttccagttccgtttatcc gggcaaaccatcgaagtgaccagcgaatacctgttccgtcatagcgataacg agctcctgcactggatggtggcgctggatggtaagccgctggcaagcggtga agtgcctctggatgtcgctccacaaggtaaacagttgattgaactgcctgaact accgcagccggagagcgccgggcaactctggctcacagtacgcgtagtgca accgaacgcgaccgcatggtcagaagccgggcacatcagcgcctggcagc agtggcgtctggcggaaaacctcagtgtgacgctccccgccgcgtcccacgc catcccgcatctgaccaccagcgaaatggatttttgcatcgagctgggtaataa gcgttggcaatttaaccgccagtcaggctttctttcacagatgtggattggcgat aaaaaacaactgctgacgccgctgcgcgatcagttcacccgtgcaccgctgg ataacgacattggcgtaagtgaagcgacccgcat	
t-Δ*lacZ*::*cat*-1400 bp (*lacZ*=**<u>3075</u>** nt) (*Escherichia coli* str. K-12 substr. MG1655, ~4.6 Mbp)	Cgatgtcggtttccgcgaggtgcggattgaaaatggtctgctgctgctgaacg gcaagccgttgctgattcgaggcgttaaccgtcacgagcatcatcctctgcatg gtcaggtcatggatgagcagacgatggtgcaggatatcctgctgatgaagca gaacaactttaacgccgtgcgctgttcgcattatccgaaccatccgctgtggta cacgctgtgcgaccgctacggcctgtatgtggtggatgaagccaatattgaaa cccacggcatgg Agacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatc actaccgggcgtatttttgagttgtcgagattttcaggagctaaggaagctaaa atggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgt aaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccg ttcagctggatattacggccttttaaagaccgtaaagaaaaataagcacaagtt ttatccggcctttattcacattcttgcccgcctgatgaatgctcatccggaattccg tatggcaatgaaagacggtgagctggtgatatgggatagtgttcacccttgttac accgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacg acgatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtga aaacctggcctatttccctaaagggtttattgagaatatgtttttcgtctcagccaa tccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttctt	UHA: 278 bp *cat*: 807 bp DHA: 315 bp Del (**<u>3075</u>**-2797 nt): 441 bp

	cgcccccgttttcaccatgggcaaatattatacgcaaggcgacaaggtgctgat gccgctggcgattcaggttcatcatgccgtctgtgatggcttccatgtcggcag aatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaag gcgcgccatttaaatgaagttcctattccgaagttcc Gggactgggtggatcagtcgctgattaaatatgatgaaaacggcaacccgtg gtcggcttacggcggtgattttggcgatacgccgaacgatcgccagttctgtat gaacggtctggtctttgccgaccgcacgccgcatccagcgctgacggaagca aaacaccagcagcagtttttccagttccgtttatccgggcaaaccatcgaagtg accagcgaatacctgttccgtcatagcgataacgagctcctgcactggatggt ggcgctggatggtaagccgctggcaagcggtgaagtgcctctggatgtcg	
***C. glutamicum* ATCC 13032**		
t-Δ*ldh*::*egfp*-2402 bp (*ldh*=**<u>945</u>** nt) (*Corynebacterium glutamicum* ATCC 13032, ~3.3 Mbp)	Ttaatactgtgctgggttaattcgccggtgatcagcagcgcgccgtaccccaa ggtgccgacactaatgcccgcgatcgtctccttcggtccaaaattcttctgccca atcagccggatttgggtgcgatgcctgatcaatcccacaaccgtggtggtcaa cgtgatggcaccagttgcgatgtgggtggcgttgtaaattttcctggatacccg ccggttggttctggggaggatcgagtggattcccgtcgctgccgcatgcccca ccgcttgtaaaacagccaggttagcagccgtaacccaccacggtttcggcaac aatgacggcgagagagcccaccacattgcgatttccgctccgataaagccag cgcccatatttgcagggaggattcgcctgcggtttggcgacattcggatcccc ggaactagctctgcaatgacctgcgcgccgagggaggcgaggtgggtggca ggttttagtgcgggtttaagcgttgccaggcgagtggtgagcagagacgctag tctggggagcgaaaccatattgagtcatcttggcagagcatgcacaattctgca gggcataggttggttttgctcgatttacaatgtgatttttttcaacaaaaataacactt ggtctgaccacattttcggacataatcgggcataattaaaggtgtaacaaagga atccgggcacaagctcttgctgattttctgagctgctttgtgggttgtccggttag ggaaatcaggaagtgggatcgaaa Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag	UHA: 776 bp *egfp*: 717+9 bp DHA: 900 bp Del (**<u>945</u>**+1053 nt): 1998 bp

	tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Ccctcctggaacacaactacgaccgctcccgggtctacggcatccccgccgt agttcagcgcatcaacctcaaagtcggcgaccgcctcatccttaccgacgaag aactcacctacgatccatccctcggatccggccgcacaccacgcatcagctgc acccttccacaagcagtcgatgcaattaaagtcgggcaccgcgtgcttttcgac gacggagccatcgccgcagtctgcatcgacaagacctccactgccgacggc cacaacgacgtagaattggaagtcacccacgcccgcccacaaggcgtaaac ctggccgcatacaagggaatcaacctcccagactccgaacttccactcccaag cctcactgaagaagacctccaacacctgcgctttgtcgtcaaatacgccgacat cgcagccatctccttcatccgaaacgtcgccgacgtggaatacctcctccaag cactcgccgacatcggagatccagtagccgtcgaacgccttggcctcgtcctt aaaatcgagaccatcccaggctacgaaggcctcgcccaaatcctcctgaccg gcatgcgccacgaaaacttcggcatcatgatcgcccgcggagacctcgccgt cgaactcggcttcgaccgcatggcagaagtcccccaactgatcatggcccttg ccgaagccgcccacgtcccaaccatcttggccacccaagtcctggaaaacat ggccaaaaacggactcccatctcgcgcagaaatcaccgacgcagcaatggc acttcgcgctgaatgcgtcatgctgaacaagggaccacacatcaacgacgcc atcaaggtcctcaccgaaatgagccgcaaacttggtgcatcccaacgaaaga gtaggctgc	
t-Δ*ldh*::*egfp* (2001 bp) (*ldh*=**945** nt)	Atgaaagaaaccgtcggtaacaagattgtcctcattggcgcaggagatgttgg agttgcatacgcatacgcactgatcaaccagggcatggcagatcaccttgcga tcatcgacatcgatgaaaagaaactcgaaggcaacgtcatggacttaaaccat	UHA: 375 bp *egfp*: 717+9 bp DHA: 900 bp

(*Corynebacterium glutamicum* ATCC 13032)	ggtgttgtgtgggccgattcccgcacccgcgtcaccaagggcacctacgctg actgcgaagacgcagccatggttgtcatttgtgccggcgcagcccaaaagcc aggcgagacccgcctccagctggtggacaaaaacgtcaagattatgaaatcc atcgtcggcgatgtcatggacagcggattcgacggcatcttcctcgtggcgtc caaccca Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Ccctcctggaacacaactacgaccgctcccgggtctacggcatccccgccgt agttcagcgcatcaacctcaaagtcggcgaccgcctcatccttaccgacgaag aactcacctacgatccatccctcggatccggccgcacaccacgcatcagctgc acccttccacaagcagtcgatgcaattaaagtcgggcaccgcgtgcttttcgac gacggagccatcgccgcagtctgcatcgacaagacctccactgccgacggc cacaacgacgtagaattggaagtcacccacgcccgcccacaaggcgtaaac ctggccgcatacaagggaatcaacctcccagactccgaacttccactcccaag cctcactgaagaagacctccaacacctgcgctttgtcgtcaaatacgccgacat cgcagccatctccttcatccgaaacgtcgccgacgtggaatacctcctccaag cactcgccgacatcggagatccagtagccgtcgaacgccttggcctcgtcctt aaaatcgagaccatcccaggctacgaaggcctcgcccaaatcctcctgaccg	Del (**945**-375+1053 nt): 1623 bp

	gcatgcgccacgaaaacttcggcatcatgatcgcccgcggagacctcgccgt cgaactcggcttcgaccgcatggcagaagtcccccaactgatcatggcccttg ccgaagccgcccacgtcccaaccatcttggccacccaagtcctggaaaacat ggccaaaaacggactcccatctcgcgcagaaatcaccgacgcagcaatggc acttcgcgctgaatgcgtcatgctgaacaagggaccacacatcaacgacgcc atcaaggtcctcaccgaaatgagccgcaaacttggtgcatcccaacgaaaga gtaggctgc	
HEK293T		
t-Δ*DROSHA*::*egfp* (1390 bp) (Exon3=**<u>66</u>** nt) (Homo sapiens, chrV~181.54 Mbp)	Atgcctataagtctttttagctgctgatagttttgagttggcaaaactagtaataat gacattttaaaatttaaggacattaaaaatttagttggaaactttaacacagttcaa catgcttttaattttgtttgtaggctaagtgtattggaaactacaaaattagtgttatg gtaattgcagaaatattttgtagcattggcagtttgggaaacatgtttatgtctgttt tattctttctgtttttttagagcttatattctctgtggaagatgtgacatatccaggcg gaacatcatg Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Attaaaggtggaattaggtccttagaaggtggattgtaattattgtaggacctgc taagaagcagcacagagtgttgtgagagcaaggaacctaacaaagccctag	UHA: 299 bp *egfp*: 717+9 bp DHA: 365 bp Del (**<u>17</u>**+247 nt): 264 bp

	ggtcagtgacagtggttgtagaaaggccagcaaggatattattataatagccat gtaagaggtgatgatgtgaagtagatgaattcaggaaatatttatgaagcagaa acaattggatttttgtttatttttattaagggatagttgccatataagtgacagcaa ttttaaagtaaatttttagctaaggatgggggggtgagggaggacaagtcagatg attccctgttacatgtcttaaggactattaatggtgtc	
t-Δ*DROSHA*::*egfp*-9 (1498 bp) (Exon4=**834** nt) (Homo sapiens, chrV~181.54 Mbp)	Gaagtggagcagctttaaggaatggtcggtccagggactaaacatggcagtg ggctgggtggccaagaaaggggaagaggagtttattgagcatgagggtttca gatgttagcaggaacccagtattaaatgggtggcctattccagtggcttgacta gggggtctttgagttgctcatcaagatggtcaggcatttatgaaaccctgtttaca tagtaagaattatttttttaaaaaaactttttccctttttctttttctgccatgaagtcacag aatgtcgttccacccgggacgagggtgtccccgaggacgaggaggacatgg agccagaccctcagcaccatcctttaggccccaaaatctgaggctgcttcacc ctcagcagcctcctgtgcaatatcaatatgaacctccaa Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Cagtcatgccgcagcaggttaattatcagtaccctccgggctattctcaccaca acttcccacctcccagttttaatagtttccagaacaaccctagttctttcctgccca gtgctaataacagcagtagtcctcatttcagacatctccctccatacccactccc	UHA: 417 bp *egfp*: 717+9 bp DHA: 355 bp Del (**834**-159-382): 303 bp

	aaaggctcccagtgagagaaggtccccagaaaggctgaaacactatgatgac cacaggcaccgagatcacagtcatgggcgaggtgagaggcatcggtccctg gatcggcgggagcgaggccgcagtcccgacaggagaagacaagacagcc ggtacagatctgattatgaccgagggagaacaccatc	
t_{g1}-Δ*CAMKMT*::*egfp* (1302 bp) (Exon1=**<u>423</u>** nt) (Homo sapiens, chr II~242.19 Mbp)	Ggctctggacacaaaccagtgaaggcaggccggagaacctgaagctctca gaggggagcaggtgtcaccgcaggcaagtccagccgaagtctgcgttccgc agcccacagaacgacaacttacccagagccgcctagagctggtttgcagcac gccaatctacgtaacctcaatctagcacgagcaacaggcagatttcgccatctt tgttgtggtcaaggagtcttcttgggttcctgggttctttagtctcgaaatatataa gacttttcgttcttcgctagtcttctgagctcgagatgaagcacagaaggctaaa aca Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtaa Gtaagggagaacctgctcgcctcacctttgcctctggtcactcctctctcacgt accgggggagcgactgttcgcagttctttcctcttggcagctctgggagaccct tctcagtttgccgggagccacctgcgaagctccccctcgcttcacctaagcgtc ccatccaagaggaggaatcatgtggggttgggtggggagcgggacagccaa ggctgagctcctgggaaactcagatcgaagtagaagaacccccttcta	$t_{g1:}$ gene edit template for exon1 UHA: 322 bp *egfp*: 717+3 bp DHA: 260 bp Del=Exon1=**<u>423</u>** bp

t_{b2}-Δ*CAMKMT*::*egfp* (191 bp) (Exon2=**<u>173</u>** nt) (Homo sapiens, chr II~242.19 Mbp)	ATTTCAAAGgt**A**ctgaa**A**caaaaacacctggatgattgcctgcgaca tgtatctgtaagaagatttgaatcatttaatctgttttcagtaacagaaggcaaag aaagggaaactgaagaggaggttggtgcatgggtccaatatacaagcatcttc tgtcctgaatacagtatctc**A**t**C**aagGTAACCATT	t_{b2}: base edit template for exon2 Capital letters with dash-line denote complementary regions with the outer sides of the exon2 Bold capital letters indicate mutant bases (T→A, G→A,C→A and T→C, in proper order)
t-Δ*CAMKMT*::*egfp* (1506 bp) (Exon3=**<u>64</u>**+1 nt) (Homo sapiens, chr II~242.19 Mbp)	Tcagcttgattctctaagtatactgtgtgggtttttggctttaaaaaggttaagaac ttccctttacataatgaaaacccaacaagtaaatattatttgttagtactttaagatg tagattaatcatttaaaaagtcattatgttaaatattctgaggaccatactgttatttc tttccttctccctctagtgggcagaaaatattcttaaaatatagcatggaataattta ttgctattgggtataccataatatacagtctcagtctcagctttttgaatactaaaa atgttaacatttcagaatcattaaagcaaacgctttactccttctaggcataatag tgga Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Gagatggagagagtaggtcagaaagaacactggacactgaagtactgtccct tttatttttgtttccattttgacctcagttaattcagtagtggggctgtttcttccttcttt	UHA: 351 bp *egfp*: 717+9 bp DHA: 429 bp Del (**<u>64</u>**-12+254 nt): 306 bp

	gctttgtttcgggtgctttctttgccattccctgcttctttactttcactgtagaataaa cccagaacatatttttcttgatgctactatacattaataccagagtgtttttgtttgat gcagtattacaatattacagaatatattgccttttcatcaaagcttctgattaattgg tgatagatttgtccactgagaaaatcacttaaatatttcttataattcttgtaattaatt tcatgtgatgaaagacatcagccaagaataccttttatgtcactgagttaaacagc ataccatcataggagctatggac	
NIH-3T3		
t-Δ*Lepr*::*egfp* (1230 bp) (Exon2=**328**+2 nt) (Mus musculus, chr IV ~156.51 Mbp)	Gaaggtagacgctcagggttgtttattagactccgatccagcactctgaaggtt aaatatttctacttgaaaatgctctaccatttaatctcaagtacttcctagttaatctt cgtgacaatgatggcttgttttctctgtctttttgctgtcctagcagaatttctttatgt gatagctgcacttaacctggcatatccaatctctccctggaaatttaagttgttttg tggaccaccgaacacaaccgatgactcctttctctcacctgctggagccccaa acaatgcctcggctttgaaggggcttctgaagcaattgttgaagctaaatttaa ttcaagtggtatc Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa	UHA: 351 bp *egfp*: 717+9 bp DHA: 153 bp Del (**328**-244 nt): 84 bp

	Actgaagggaagacactggcttcagtagtgaaggcttcagtttttcgccagct aggtaagtacttttatgccccaagtatatatagtgttgtgttctgccttatgcaggtt aagaatctctctgaaacaatcaatacatgttatgctgagtga	
t-Δ*Vegfa*::*egfp*-9 (1444 bp) (Exon1=**<u>600</u>** nt) (Mus musculus, chr XVII ~94.99 Mbp)	Acttttcgtccaacttctgggctcttctcgctccgtagtagccgtggtctgcgcc gcaggagacaaaccgatcggagctgggagaagtgctagctcgggcctgga gaagccggggcccgagaagagaggggaggaagagaaggaagaggagag ggggccgcagtgggcgctcggctctcaggagccgagctcatggacgggtg aggcggccgtgtgcgcagacagtgctccagccgcgcgcgcgccccaggcc ccggcccgggcctcggttccagaagggagaggagcccgccaaggcgcgc aagagagcgggctgcctcgcagtccgagccggagagggagcgcgagccg cgccggccccggacgggcctccgaaaccatgaac Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtc gagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggg cgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcg tgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacg acggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatc ggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtcc gccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggag ttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaaggaatt ctaa Ccacccaccagagtcaccgcacgtacgatctgggccgagcagcggagggc gggagccagaggaggaggctgaggggctgggcttgtgccgaggctggc ggcagaagtttgctccgggtcgcgggtccccggagaactggaagtccgggc aaaggggcgggagtccggagcccagcgggcatgcctgggggtgctcgg accttggacccggggagggcagagatcgtggaggggcagggcgcgggc	UHA: 385 bp *egfp*: 717+9 bp DHA: 333 bp Del (**<u>600</u>**-540+223): 283 bp

	gaccgaggggctttgctgtcactgccgtttgggtctctgaggcccttgcagtg agtttggggaaagttttaggatggattgctg	
***S. cerevisiae* LYC4**		
t-Δ*crtE* (807 bp) (*crtE*=**<u>1131</u>** nt) (*Saccharomyces cerevisiae* S288C, chr IV~1.53 Mbp)	Cagccattccattggagtttactccacaagatgatattgttttgttggaaccatat cattatttgggaaaaaatcctggtaaagaaattcgttcacaattgattgaagccttt aattattggttggatgttaaaaaagaggaccttgaagttattcaaaatgttgttggt atgttgcatactgcttctttgctcatggatgatgttgaagatagttccgttttacgtc ggggttcaccagttgctcatttgatctacggtattccacaaactattaatacagcc aattatgtttacttttttagcctatcaagaaatttttaaattgcgtccaacacctattcc tatgccagtt Ggtggtttgtttcgtattgccgttcgtttgatgatggccaaatctgaatgtgatatt gattttgttcaattagttaatttgattagtatctattttcaaattcgtgatgattatatga atcttcaatctagtgaatatgcacataacaaaaattttgccgaggaccttactgaa ggtaaattttcctttccaactattcatagtattcatactaatccaagtagtcgtttagt tattaatacattacagaaaaaatctacatccccagaaattttgcatcattgtgttaat tatatgcgtactgaaacacattcctttgaatatactcgtgaagttttgaatactttgt ccggtgccttggaacgtgaattgggtcgtttgcaagaagaatttgccgaagcta atagtcgtatggatttgggcgacgttgaatccgaaggtcgtacagggaaaa	UHA: 353 bp DHA: 454 bp Del (**<u>1131</u>**-879 nt): 252 bp
***S. virginiae* IBL14**		
t-Δ*sviipe*::*cat*-1399 bp (*sviipe*=**<u>1143</u>** nt) (*sviipe* / KY243071)	Gcacctcaaccacggctcctacggggcggtcccgctgcccgtccaggaagc ccaggaggccctgcgcttccaggcccacgccgaccccgacgccttcttcaac ggggtctgcgaacggctcaccgccgcccggggccggatcgcggcgcgcct gggagcggacccggacggcctcgcgttcatctccaacgccaccgaggg <u>Agacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatc actaccgggcgtatttttgagttgtcgagattttcaggagctaaggaagctaaa</u> atggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgt aaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccg ttcagctggatattacggccttttaaagaccgtaaagaaaaataagcacaagtt ttatccggcctttattcacattcttgcccgcctgatgaatgctcatccggaattccg tatggcaatgaaagacggtgagctggtgatatgggatagtgttcacccttgttac accgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacg	UHA: 201 bp *cat*: 807 bp DHA: 391 bp Del (**<u>1141</u>**-692 nt): 449 bp

	acgatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtga aaacctggcctatttccctaaagggtttattgagaatatgtttttcgtctcagccaa tccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttctt cgcccccgttttcaccatgggcaaatattatacgcaaggcgacaaggtgctgat gccgctggcgattcaggttcatcatgccgtctgtgatggcttccatgtcggcag aatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaag gcgcgccatttaaatgaagttcctattccgaagttcc Catgggaggaccaccagggcttcccgctctccgtcgagtaccgcgccaccttt cgactacaccggctggctggccgcccccgagggactcgacctcctggagcg gctcggcgccgaccgggtccgggagcacaacagcgcgctggccggctacg gcgccgggctgctcgcccggatccccggcctcaccccgctgccgcacaccc cgggcctggccctgcgctccctgttgctgccgcccggcgtcgccgagaccc gcgaggcggccaccgccctgcgcgaggaactcgccgcgaagctccgcatc cgggtcctggtctggccccgcgagggcggcggcgggctgcgggtctgcgg gcaggtctacaaccggcccgaggagtacgaacgcct	
t-Δ*svipam*1-2047 bp (*svipam*1=**2901** nt) (*svipam*1 / KY243072)	Cgagaccgacgccttcctgcgcaccctcggctggcgtcaggtcgcgcagga ggagtacgacaagaagctcacgcccgagacgaagaagtacctccaggccta cgccgacggggtcaacgcgtggctgaaggggaagtccggcaaggacctct ccgtcgaacacgccgccctgaagatcaccgacggctacaagcccgagaagt ggacgccggtcgactcggtggcctggctcaaggcgatggcctgggacctgc gcggcaacatgcaggacgagatcgaccgctcgctgatgtcgaacaagctca cgcaggcgcagatcgacgagctctacccggcgtacccccttcgaccgcaaca agccgatcgtcgagggcggcaccgtcgcgggcggcaagtacgccccccag ggcaccgcgggcgccggcaccggctcctccccgacgggcgccaccggca cgggcacgggcaccggctccaccacgggcgccgccgccggtacgggtac gggcgccggtacgggcaccggctcgcaggccacccccaacggcggcacc aaccccgccaccggcctcgccgacaacgcgagcgcccagggcgcgaccgt cggcctgcgcacccagctgaccgccctgtccaagaccctggacgacatccc ggccatcctcggccccaacggcagcggcatcggctcgaactcctgggtcatc tcgggcaagtacacgaccaccggcaagccgctgctcgcgaacgacccgca cctctccccgcagctgccctcggtctggtaccagatgggcctgcactgccgcg	UHA:945 bp DHA:1102 bp Del (**2901**-2574): 327 bp

	ccgtctcggcccggtgccagtacgacgtcgccggcttcaccttctccggcatg cccggcgtggtcatcggccacaacgccgacatcgcctggggcatgaccaac ctcggcgccgacgtgaccgacctctacctgga Ggtcccgtcgcagaacctcatctacgccgacaacaagggccccaacggcaa catcggctaccaggccccgggccgcatcccggtccgcggccagggcgacg gccgcatgccggcccccggctgggacccgaagtacgcctggaagggcggc cgcgacggcaacgccggctacatcccgcagaacgagctgccctgggagctc aacccgcagcgcggctacatcgtcaccgccaaccaggccgtcaccgagtcc ggcaccggcgcgggcaagtacccgtacgtgctgaccaccgactggggctac ggcgcccgcagccagcggatcaacgacctcatcgaggcgaagatcaagga cagcggcaagatctccaccgacgacatgcgcaccatgcagatggacaacag cagcgagatcgccgcgctgctgaccccgatgctggccaagatcccggtctcc gacccggacgtgcgctcggcgcagaagctgctggagggctggaactacacc caggagaacgactcggcggcggccgcgtacttcaacgcggtctggcgcaac atcctcaagctggcgttcggcgacaagatgccgaaggagctgcgcgtcgag ggcagctgcatgaacgtcgtcgaggagagcaccggccccaacgacgacct ggccaagaccgtccgcgagtgcggcacgcgtggtcccgactccgcgcagc cggacggcggcgaccgctggttcgaggtggtccgccgcctggtcaaggacg agaagtcgccgtggtggaccgccaccccgaagaccgtccacgacaagccc atcaagacccgcgacgagctcttcgcccgggcgatggaggacgcgcgctgg gagctgaccgccaagctcggcaaggaccagtcgacctggagctggggccg gctgcaccagctgacgctgaagaaccagaccatcggcaaggagggcccgg gcttcatgcagtggctcctcaaccgcggcccgtggaacctgggcggcggcg aggccaccgtcaacgccaccggctggaacgcctccagcgggtacgacgtca cctgggtgccgtcgatgcggatggtcgtcaacctcaac	
t-Δ*svipam*2-1030 bp (*svipam*2=**2049** nt) (*svipam*2 / KY243073)	Accaccgaggcctacgaggtctaccgcgacagctggggcatcccgcacct gcgcgcgtccgacccgctccgcctctcctacgcccagggccggaccaccgc ccgggaccgggcctggcagctggaggtggagcggcaccgcgcccagggc tccagcgcctccttcctcggggccgactgcgtgccctgggaccgcttcgccc gccaggcccgcctcgacgacaccgcccgccggtgcttcgaggccctggac gcggacaccgccgcctgggtcaccgcctacgtcgacggggtcaacgcggg	UHA: 414 bp DHA: 616 bp Del (**2049**-1602 nt):447 bp

	cctggccgagggcgcggcccgcgacgaccgcttcgcggccgcgggccac acccccgccccctggcagccgtgggtcccgctctccatctggatcggcaccc acatcctcttc Gtcgaggtcgagatcctggagaccgcccggggcccggtgatcgtcaacga cgggtccgccaccctctccctgcgctacccgccccgcgtccgctccgacctc ggcttcgccgccctgcccgccctgctgcgcgcccgcaccgtcgccgacgtc gaccgcgccctcgacggctgggccgagcccgtcaacgtcgtgcacgccgc cgactccgagggcggcctgctgcaccgtgtcgcgggcgccgtcccgctgcg cgaggccgcgaaccggctgcgtcccgtgcccgcctgggacgcccgccacg cctggcggggctgggccgagacccccgccgagcccgtgcggggcgtcgc cgtgatggcgaacgcgcgcgggatcgcgtccccctcggcgtggagttcgc gccgccacaccgcgccgaccggatccgtgccctgctcggcggttcggccga ctggtccccgaaggccatggccgaggtccaccgcgacacccacctggcctct gccgcgccgctgctcgccctgctgcccgggctcgacgccctctccccgcg gccgccgcgctccgcgaccggctgctggcctgggaccggcacatggaggc cgacagca	
t-Δ*svibla*1-541 bp (*svibla*1=**2082** nt) (*svibla*1 / KY243074)	Ctcccccgaacacccctactacgccggggccgtggtcctcgccgggcgcg gccgcacggtcgccctgcaccgggcgatgggggacgccgtccggtacgcg gactacgacgggcgcaccgaccggcccccgggagttcccggccgccgaccg gatcgccatggccgaggacaccgtcttcgacctggcctccctcagcaagctgt tcacctccatcatcgc Gcctccgtcgaactgctcttcaccgactacaacaccgccttcccgggcgacg accacggcctcggcttcgagctctaccagcactggtacatgggcgccctggc caccccccactcggccggccacaccggcttcaccgggacctccctggtcctc gacccctccaccgactccttcctcgtcctcctcggcaactccgtgcaccccgtc cgcacctggcgcgccggcagcgcaccccgtgtggccaccggcaaccgcttc gcccgggccgtacccgtccgtacgaagcacggcggcccggcctggttctcc gggatggaca	UHA: 219 bp DHA: 322 bp Del (**2082**-1505 nt): 577 bp
t-Δ*svibla*2-1066 bp (*svibla*2=**1542** nt) (*svibla*2 /KY243075)	Atcctgctcacagccgggacggccgcggcggagccgaccccgcccccgc cccccagaccccgcatcagcgacgccgccgtcgccacggcggtgagccgc ctcgacgcggacgtcgcggacgtcatgcgccgcaccggcatccccggtgtg	UHA: 418 bp DHA: 648 bp Del (**1542**-1140 nt): 402 bp

	gcggtggccgtggtccaccgcgacaaggtgctctacctcaagggcttcggcc tccgccgcacgggcgagagcgcgaaggtggacccggacaccgtcttccag ctggcctccgtctcgaagccgctgtcctccaccgtcgtcgccggcgccctgac cgaccccgccgcctgggacgagcccctcggctcctccctgcccggcttcgc gctgaaggacccctgggtgacctcccacgtcacccccgccgacctgttctccc accgcagcg Gggtcctcgacgggaaacggatcatccccgccgaccacctgaaccgcacc cgcctccccgagatcgtctcgcaggcccagaccttccgcggcgtcccgcagt tctacgggctcggctggaacgtcagctacgacgacgacggccggctgcgcc tgggccactcgggggcttcgaactcggcgccaacaccaacgtcaccatgct cccgctcgaacagctcggcatcgtcgtcctgaccaacgcggcccggtcgg ccaggccgacgcgatcgccctggacttcttcgacaccgccgaacacggcaa gcggaccgccgactggctggccctgaccggcgcgctctacgcacaggagct caacgagggccaccgcccggccaccgactacgcccacccgccggccggc gccaagcccgcccgggggccggcacctacaccggtacctacgacaaccc cttctacggccccctcaccgtcaccgccgacgccaagggcgccctcaccctc tccctcggccccaagccccagcgcttccccctcacccactacgacggggaca cgttcagcttcgtgaccgccggcgagaacgccgtcggccgcaccggggtga ccttctccgacggcaaggtccgcatcgagtacctgg	
t-Δ*svibla*3-838 bp (*svibla*3=**1245** nt) (*svibla*3 /KY243076)	Ctctcaccgatgtcgcagaactgataccagttctagttcgaggagtccacggt ccggacgagggtgccggaggatccgatggacggtcaggactgttgggggga cggtcgacttgggctactactggaacgcgttctcgtcagagcggatcgagtcg gccagggagcccccatgtcagaggtcatgtcacaggtcaccgaccacggcg ggggtgtgtggggcatcaaggtccccatccccgacaacccccctgggacaca ccctcgtccacgtcctcgacacggacgcaggccccgtcctcgtcgacaccgg ctgggacgaccccgcgtcctgggacgccctcaccggcggtctcggcgcact cggcatcgccgtggacgacgtgcacggcgtggtcatcacccaccaccaccc cg Gccacgtctgcctgcacctggaggaggcccaccccgcgaacctgcccggc aacggccggctcttctccggcgaccacctgctgcccgggatctccccgcaca tcggcctgtacgaggccccggacgacacccggggtcaccgacccctgggc	UHA: 414 bp DHA: 424 bp Del (**1245**-927 nt): 318 bp

	gactacctcgactccctcgaacgcgtcggccgcctggagcccgccgaggtg ctcccggcccaccagcacgccttcaccgacgcccccgcccgcgtacgggag ctcctcgcgcaccacgaggagcggctggccgggctgtggcggctgctgctg gccgagccgctgaccccgtgggggctcgcggagcggatggagtggaaccg gccctgggagcagatcccgtacggctcccggaacatcgccgtctcggaagc ggaagcccatctgcggcg	
t-Δ*svibla*4-1559 bp (*svibla*4=**921** nt) (*svibla*4 / KY243077)	Gctggtgttcctgggctgcgtcctccacgacctggggttgtcggaacagggc aacggccaccagcggttcgaggtggacggtgcggacctggcggcccggttc ctgcgcgaacagggcgccggcgaggaggccgtgacggtggtgtgggacg ccgtcgccctgcacacgtcggagggcatcgcctcccgcaagggcccggag atcgcactggcccacgccggtatctcggtggacgtcctggggcgcggcaag gagctcctgcccgaggggttcgccgaccgggtgcacgccgcgttcccccgc gccgacctcggctacgccatcacggacgtgatcgtccggcagatgctggaca accccgccaaggccggcccgctcagcttccccgggcagctgctgcgcgccc acgtgccggccggcacgctgcccgactggtccgacctgatcgacgcggccg gctggggagaccggccggcgggcgggccggtgcaaccgtcctgagacggt tcgcgtgtttcctcccgtagtgatgatcaacgatgaggagacgatgtgaagcg cgtggcaagccgtccgtcccggcgtacggtcctgaccctcgcgaccggggg cctggccgccctcggcctcggcgccgccgctccggcgtccgcgtccgcgga cacgcggccgggcccgggccggggcgccccgcgccggctgcgggacatg gaacgggagcacggc Ctcgccaacaccaccaacacccatcgcttcaaggagggcctgcccggctgg accctggccgacaagaccggcggaggcgacgactacggcgtcgccaacga cgtgggcgtggtgtggagcccggccggtgtgccggtggtgatggcggtcct gacgaccaggccgcaggccccgcacggcccgaacgacagcgagctcgtc gccgagacggcccgcctgctgggcgaaaccgtcctggactgacccgcacc ggcccgcaccggccccatcccgctcgcaccgacccgcacgggccggctcc tcatggcgctcgccgcgcgggcgtacggcggtccgtgggcgggaggggcg cgatgccggtgaaggggcaccgggcgctcccgggtgcgggggtaccggcc gaagcggccggtgctgcgacgggacgggggcggaggtcgtgcggatggtc ggtgatgtcgagggccagcttgtggggttggaggatggtgcgggtgctgaag	UHA: 727 bp DHA: 832 bp Del (**921**-408 nt): 513 bp

	acgggcgtggtgcccggggtgcggttgatctggtagctgctggtgagggtgg cggtggtcacggtcatcgcgacggtggcgaagtgggttcccaggcaggtgc gcggtccgccgccgaagggcatgaacgcgtacttgggcagcggatcggcg ggggagcggccggtccagcggtcggggaggaaggcgtccggctggtggta gtggcgggggtcgcgctggaggacgaaggggctgacggcgacgcgctgc ccggggcgcagggggaagccgccgagttcggtgggttcttcgacggtccttt cgatcagccacaggggcgggtacaa	

*Capital letters represent the site for restriction enzyme digestion; letters with dash-line represent complementary region with plasmid; letters with single underline represent promoter; letters with double underline represent terminator; boldface letters represent direct repeat; italic letters represent spacer; UHA: upstream homologous arm; DHA: downstream homologous arm; *cat*: a gene acronym for chloramphenicol acetyl transferase; Fsc / Rsc: primers for the construction of Cas expression plasmids in genome editing of *S. cerevisiae* LYC4; UF / UR: primers used for amplification of UHA by PCR; DF / DR: primers used for amplification of DHA by PCR; vF / vR: primers used for verification of edited sequences (two fragment sequences in a genome normally locate at both sides of an edited sequence but not in t-DNA); -9: represents the Cas9-mediated genome editing; NCBI database: KY243071- KY243077.

Extended Data Table 3. Genome editing efficiency in the template-based genome editing directed by the *Svi*Cas3 or *Sp*Cas9*

Conditions / Cells	HCN	MCN	NCGEC	TE (%)	HRE (%)	GEE (%)
RNA-guided genome editing by the *Svi*Cas3						
***S. virginiae* IBL14**				**1.1**	**92.0**	**1.0**
Svi^{p} (×10^5/ml)	4.13±0.24	NT	NT			
Svi-M (×10^3/ml)	NT	4.43±0.89	NT			
S. virginiae IBL14-Δ*gene*	NT	NT	46			
***E. coli* JM109 (DE3) (*lacZ*)**				**3.28×10^{-4}**	**27.3**	**9.0×10^{-5}**
Ec^{c} (×10^6/ml)	7.5±1.12	NT	NT			
Ec-BW (×10^3/ml)	NT	2.46±1.60	NT			
Ec-W (×10^3/ml)	NT	1.29±0.95	NT			
E. coli JM109 (DE3)-Δ*lacZ*	NT	NT	26			
***C. glutamicum* ATCC 13032 (*ldh*)**				**1.8×10^{-6}**	**48.0**	**8.6×10^{-7}**
Cg-pEK^{c} (×10^7/ml)	2.59±0.25	NT	NT			
Cg^{m} (×10/ml)	NT	4.7±0.19	NT			
C. glutamicum ATCC 13032-Δ*ldh::egfp*	NT	NT	24			
***S. cerevisiae* LYC4 (*crtE*)**				**6.8**	**94.0**	**6.4**
Sc^{c} (×10^8/ml)	3.77±0.61	NT	NT			
Sc-M (×10^7/ml)	NT	2.57±1.20	NT			
S. cerevisiae LYC4-Δ*crtE*	NT	NT	47			
HEK293T (*DROSHA* and *CAMKMT*)				**12.2**	**36.0**	**4.4**
Cell number in Bright filed	584.2±67.5	NT	NT			
Cell number with mCherry	NT	71.3±33.6	NT			
Cell number with EGFP	NT	NT	25.7±7.3			
NIH-3T3 *(Lepr)*				**11.9**	**61.3**	**7.3**
Cell number in Bright filed	337.3±98.3	NT	NT			
Cell number with mCherry	NT	40.3±24.1	NT			
Cell number with EGFP	NT	NT	24.7±12.5			
RNA-guided genome editing by the *Sp*Cas9						

HEK293T (*DROSHA*)				**44.8**	**1.5**	**0.7**
Cell number in Bright filed	578±125.9	NT	NT			
Cell number with mCherry	NT	259±32.5	NT			
Cell number with EGFP	NT	NT	4±4.2			
NIH-3T3 (*Vegfa*)						**0.5**
Cell number in Bright filed	592.7±28.7	NT	NT			
Cell number with mCherry	NT	NT	NT			
Cell number with EGFP	NT	NT	3.0±4.4			
DNA-guided genome editing by Cas3						
***S. cerevisiae* LYC4 (*crtE*)**				**6.6**	**98.0**	**6.5**
Sc^{c} (×10^{8}/ml)	1.89±1.37	NT	NT			
Sc-M (×10^{7}/ml)	NT	1.25±0.28	NT			
S. cerevisiae LYC4-Δ*crtE*	NT	NT	49			
HEK293T (*DROSHA* and *CAMKMT*)				**11.8**	**65.8**	**7.8**
Cell number in Bright filed	278.7±249.4	NT	NT			
Cell number with mCherry	NT	33.0±25.4	NT			
Cell number with EGFP	NT	NT	21.7±22.9			
NIH-3T3 *(Lepr)*				**14.8**	**37.3**	**5.5**
Cell number in Bright filed	398.3±207.2	NT	NT			
Cell number with mCherry	NT	59.0±55.2	NT			
Cell number with EGFP	NT	NT	22.0±22.1			

*Concentration and number of cells are expressed as mean ± SEM of n ≥ 3; NT: not tested.

Svi^{p}: *S. virginiae* IBL14 protoplasts growing on MR2YE plates without antibiotics; *Svi*-M: mutants of *S. virginiae* IBL14 transformed by pKC1139-t/g-*gene abbreviation* growing on MR2YE plates with apramycin; *EC*^{c}: *E. coli* JM109 (DE3) competent cells growing on LBPET plates without antibiotics; *Ec*-BW: blue, white and blue-white colonies of *E. coli* JM109 (DE3) transformed by pCas-*cas*3 and pKC1139-t/g-Δ*lacZ* growing on LBPET plates with kanamycin and apramycin; *Ec*-W: white and blue-white colonies of *E. coli* JM109 (DE3) transformed by pCas-*cas*3 and pKC1139-t/g-Δ*lacZ* growing on LBPET plates with kanamycin and apramycin (about one half of the tested *Ec*-W colonies are correctly gene-edited mutants); *Cg*-pEK^{c}: competent cells of *C. glutamicum* ATCC 13032-pEC-XK99E-*cas*3 growing on LBHIS plates with kanamycin; *Cg*^{m}: mutants of *C. glutamicum* ATCC 13032 transformed by pXMJ19-t/g-Δ*ldh::egfp*

959 and pEC-XK99E-*cas*3 growing on LBHIS plates with chloramphenicol and kanamycin; *Sc*^{c}: *S. cerevisiae* LYC4 competent cells growing
960 on YPD plates (yeast extract peptone dextrose medium); *Sc*-M: mutants of *S. cerevisiae* LYC4 transformed by pRS415-*cas*3 and pYES2-
961 NTA-t/g-Δ*crtE* growing on SD-Leu/Ura plates (synthetic dropout medium-leucine/uracil).
962 HCN: host cell number; MCN: mutant cell number; NCGEC: number of correctly gene-edited cells identified from 50 potential gene-
963 edited cells by PCR and DNA sequencing; TE: transformation / transfection efficiency; HRE: homologous recombination efficiency;
964 GEE: genome editing efficiency.
965

Extended Data Table 4. Three confirmed CRISPRs in *S. virginiae* IBL14 genome*

***Svi*-CRISPR**	**Direct repeat (R)**	**Spacer (S)**
1	***30 bp, 67%GC, Tm 71ºC***	***35-37 bp, 51-86%GC, Tm 66-84ºC***
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ggtccgcaccctggtggagaggctcccgtacgtgct
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	aggtcgccgaggcggagcgggacaccgccgtaaaa
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	cccgtctacggcgctcgcgccgacatcctgcacgccgc
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ctggtacctggcccgctgggccgagtggatcgcctg
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ttcaccatcaacgccaccgcatgcgggaccctcgc
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	cgagccgggcggtggctgggatgtccaccgtcgcgc
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	*cccgccgtccggcgggccg*ggtgccttccgtcagcggg-*gvgl*005738
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ccaccgccgcccgcacgagcgggcggagtcgagtct
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ctgacccttagatcgccaacacgcaaccccaatccca
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	tcgggtgacgccctgggcggcggctgcgtcaagcatcgcgt
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	ccttgatcgaccattcctgactacaacaggaggac
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	cgggccgcgcagcggccggcctcggcaggcactgc
	gttgcga**ccccct**cgt**aggggg**cgatgaggac	The 13 R sequences in the CRISPR 1 are the same.
2	***30 bp, 57-70%GC, Tm 67-71ºC***	***35-37 bp, 66-83%GC, Tm 75-83ºC***

	gtcctcatcg**ccccct**gcg**aggggg**tcgcaac	gtctccggc*tcggccagctcgatcagcc*cgtgaccgg-*gvgl*005245
	gtcttcatcg**ccccct**gcg**aggggg**tcgcaac	ccgaggcg*accggcggccaggccgatc*cgctcgccg-*gvgl*002878
	gtcttcatcg**ccccctgcaggggg**gtcacaac	atggatccggtacggctgatcagcgcggcggtctg
	gtcctcatc**gacccct**gga**agggtc**cgcaac	aggcgaagggcaagtttgaggaggcccaggccgccg
	gttctaatcg**ccccct**ttg**aggggg**tcgcaac	The first R sequence in the CRISPR 2 is the same as the ninth and eleventh R sequences in CRISPR 3.
3	***30 bp, 57-70%GC, Tm 67-74ºC***	***36-44 bp, 62-81%GC, Tm 75-84ºC***
	gtcctcatcg**ccccct**cga**ggggg**tcgcaac	tgcggagcg*gggaggccggccagctcggc*ccagtcggc-*gvgl*001077
	gtcctcatcg**ccccct**cga**ggggg**tcgcaac	ctcccagacgaggacggcgtactgcccgacgaccag
	gtcctcatcg**ccccct**cga**ggggg**tcgcaac	aggaggcgggtatggatg*gtgccgccgccctcgacc-gvgl*006237
	gtcctcatcg**ccccct**cga**ggggg**tcgcaac	gcctcgggcacctcgtcctccaggcgcagccccgca
	gtgctcatcg**ccccct**gcg**aggggg**tcgcaac	accagatcgcgttcgacgacaccgtcacccccggca
	gtcctcatcg**ccccct**gtg**aggggg**tcgcaat	ctgttgcgctcatggcgcgcagccttgcgggctccgg
	gtcctcat**cgaccc**tatgaag**gggtcg**caa	cgtggcgaggctggcagcggggcagcggggcagcgggtgacgtt
	gtcctcat**cga**t**cct**ctg**agg**ggg**tcg**caac	atgaacaggctgcggccctcgccgtcctggttggtc
	gtcctcatcg**ccccct**gcg**aggggg**tcgcaac	gtctggacgcccgtgccccagcggccgatctcgcca
	gtcttcatcg**ccccct**gac**aggggg**tcgcaac	ataaccagcagcgctgtcgatcgcgggtccaacccgt

	gtcctcatcg**ccccc**tgcg**aggggg**tcgcaac	The sequences of the first four Rs in the CRISPR 3 are the same and were utilized in this study

967 * Purple, orange and red letters indicating the same base sequences in the repeats, respectively; Boldface letters with blue color indicating
968 complementary sequences; Italic letters with green color indicating that the fragment sequences in the spacers can also be found in other
969 sites of the *S. virginiae* IBL14 genome; *gvgl*00xxxx representing the temporary gene number in the *S. virginiae* IBL14 genome.
970

Extended Data Table 5. Main biochemical properties of *Svi*Cas proteins in the subtype I-B-*Svi* Cas system* (http://web.expasy.org/protparam) [7]

Name	ID in NCBI	Size (AA)	MW(Da)	C/M	pI	I-ind	AI	GRAVY	Functions
*Svi*Cas1	KY176008	326	36722.99	2/**6**	9.59	48.45	**99.69**	-0.196	A metal-dependent DNAase
*Svi*Cas2	KY176009	***87***	***9976.31***	1/***0***	6.19	46.59	99.66	-0.164	A small endo-RNAase
*Svi*Cas3	KY176010	**771**	**84352.40**	2/4	***5.78***	***35.49***	95.91	**-0.153**	Both Hel-D (cas3') and HD (cas3'')
*Svi*Cas4	KY176011	165	18758.08	5/***0***	8.65	**58.89**	***77.58***	***-0.557***	A RecB-like nuclease
*Svi*Cas5	KY176012	220	23514.74	1/2	**9.72**	56.70	86.00	-0.165	A RNAase generating crRNAs
*Svi*Cas7	KY176013	332	35946.55	***0***/2	6.18	40.91	90.78	-0.315	A subunit of Cascade positioning the RNA guide for DNA binding
GVGL007770 / *Svi*Cas6	KY243078	469	51276.86	**6**/4	6.18	43.09	88.83	-0.272	Potential Cas6 generating crRNAs and forming a hairpin structure on the 3'flank from pre-crRNA

*:C/M: cysteine/methionine; pI: theoretical isoelectric point; I-ind: instability index (smaller than 40 is predicted as stable); AI: aliphatic index (regarded as a positive factor for the increase of thermostability of globular proteins); GRAVY: grand average of hydropathicity (+: hydrophobic, -: hydrophilic); bold letters represent the highest value; bold italic letters represent the lowest value.

Extended Data Table 6. Main biochemical characteristics of the typical signature proteins in various type CRISPR-Cas systems * (http://web.expasy.org/protparam; http://www.ncbi.nlm.nih.gov; http://www.addgene.org/) [3,46,47]

Protein / AA	*Svi*Cas3 type I	*Sp*Cas9 type II	*Cj*Cas9 type II	*Ll*Cas10 type III	*As*Cas12a type V	*Ls*Cas13a type VI
Ala (A)	**14.5%**	5.3%	6.5%	7.5%	6.0%	***2.0%***
Arg (R)	**9.2%**	5.6%	4.3%	5.3%	4.4%	***3.9%***
Asn (N)	***1.2%***	5.1%	6.8%	6.7%	5.7%	**10.7%**
Asp (D)	6.9%	7.2%	6.0%	**7.5%**	***5.5%***	7.3%
Cys (C)	0.3%(2)	***0.1%(2)***	0.7%(7)	**0.8%(6)**	0.6%(8)	0.5%(7)
Gln (Q)	3.2%	3.8%	3.0%	3.4%	**4.5%**	***1.7%***
Glu (E)	***6.4%***	7.9%	8.7%	6.7%	7.6%	**10.4%**
Gly (G)	**7.7%**	5.0%	3.7%	4.6%	4.4%	***2.7%***
His (H)	**3.8%**	2.3%	1.7%	1.6%	3.0%	***0.8%***
Ile (I)	***3.2%***	6.8%	6.1%	7.4%	7.0%	**12.7%**
Leu (L)	**12.3%**	10.8%	10.5%	10.0%	10.6%	***8.4%***
Lys (K)	***1.3%***	11.0%	14.7%	7.5%	9.0%	**15.3%**
Met (M)	***0.5%***	1.6%	1.0%	**2.9%**	1.4%	1.2%
Phe (F)	***3.4%***	4.6%	**6.2%**	4.8%	5.8%	5.2%
Pro (P)	**5.7%**	2.6%	1.8%	1.2%	3.7%	***0.7%***
Ser (S)	***4.0%***	5.6%	5.9%	**6.7%**	5.7%	4.3%
Thr (T)	5.4%	4.8%	***3.3%***	5.4%	**6.0%**	3.4%
Trp (W)	**1.7%**	0.5%	***0.4%***	0.8%	0.8%	***0.4%***
Tyr (Y)	***2.2%***	4.0%	4.4%	**5.3%**	4.3%	5.0%
Val (V)	**7.1%**	5.3%	4.3%	3.7%	3.9%	***3.5%***
Total number of AA	771	1368	984	757	1307	1389
MW (Da)	84352.40	158441.41	114896.12	87181.92	151206.55	166277.01
pI	5.78	8.98	9.37	5.75	8.01	8.76
I-ind	35.49	37.57	30.50	33.83	35.53	40.26
AI	95.91	89.44	83.49	86.26	85.99	94.23
GRAVY	-0.153	-0.594	-0.678	-0.429	-0.492	-0.749

*AA: amino acid; *Svi*Cas3: from *Streptomyces virginiae* IBL14; *Sp*Cas9: from *Streptococcus pyogenes*; *Cj*Cas9: from *Campylobacter jejuni*; *Ll*Cas10: from *Lactococcus lactis* subsp. Lactis; *As*Cas12a / Cpf1: from *Acidaminococcus sp*. BV3L6; *Ls*Cas13a / C2c2: from *Leptotrichia shahii*; pI: theoretical isoelectric point; I-ind: instability index (smaller than 40 is predicted as stable); AI: aliphatic index (regarded as a positive factor for the increase of thermostability of globular proteins); GRAVY: grand average of hydropathicity (+: hydrophobic, -: hydrophilic); bold letters represent the highest value; bold italic letters represent the lowest value.

Extended Data Table 7. The main features of typical gene editing techniques [37,62,63,74]

Item*	ZFN	TALEN	*Sp*Cas9	*Svi*Cas3
Reco-Mode	Protein	Protein	RNA	DNA or RNA
Size-TDS	~24 bp	~30 bp	~20 bp	≥40 bp?
Reco-element	ZFA	TALEA	sgRNA	D-/R-loop?
Endonuclease	Dimer of FokI	Dimer of FokI	*Sp*Cas9	*Svi*Cas3
MW (kDa) / AA	65.4 / 587	65.4 / 587	158.4 / 1368	84.4 / 771
Effector complex	ZFA::FokI	TALEA::FokI	sgRNA::Cas9	D/R-loop::*Svi*Cas3
Signature-TD	3 bp / ZF array	T-5’	PAM	unlimited
Resulted DNA	DSB	DSB	DSB	SSB or DSB?
Cytotoxicity	high	mediate	mediate	Low ?
Off-target	low	low	high	not detected
NoTS	single	single	multiple	multiple
Efficiency	low	low	high	high
Operation	difficult	difficult	mediate	facile

Cast	high	high	mediate	low

* Reco-Mode: Recognition mode of engineered-nucleases to recognize target DNA, Size-TDS: the size of target DNA sequence recognized by an endonuclease, Reco-element: recognition element, ZFA: zine finger array protein, TALEA: transcription activator like effector array protein, sgRNA: single guide RNA, MW: molecular weight, AA: the number of amino acid residues, Signature-TD: the signature of target DNA, NoTS: number of target site per round operation, DSB: double strand break; SSB: single strand break, ?: to be further determined.

Supplementary references

1 Maier, L. K. *et al.* Essential requirements for the detection and degradation of invaders by the *Haloferax volcanii* CRISPR/Cas system I-B. *RNA Biol* **10**, 865-874, doi:10.4161/rna.24282 (2013).

2 Goren, M. G. *et al.* Repeat Size Determination by Two Molecular Rulers in the Type I-E CRISPR Array. *Cell Rep* **16**, 2811-2818, doi:10.1016/j.celrep.2016.08.043 (2016).

3 Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science (New York, N.Y.)* **353**, aaf5573, doi:10.1126/science.aaf5573 (2016).

4 Stern, A., Keren, L., Wurtzel, O., Amitai, G. & Sorek, R. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet* **26**, 335-340, doi:10.1016/j.tig.2010.05.008 (2010).

5 Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Current opinion in microbiology* **37**, 67-78, doi:10.1016/j.mib.2017.05.008 (2017).

6 Hidalgo-Cantabrana, C., Goh, Y. J., Pan, M., Sanozky-Dawes, R. & Barrangou, R. Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 15774-15783, doi:10.1073/pnas.1905421116 (2019).

7 Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477, doi:10.1038/nrmicro2577 (2011).

8 Mohanraju, P. *et al.* Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science (New York, N.Y.)* **353**, aad5147, doi:10.1126/science.aad5147 (2016).

9 Majumdar, S., Ligon, M., Skinner, W. C., Terns, R. M. & Terns, M. P. Target DNA recognition and cleavage by a reconstituted Type I-G CRISPR-Cas immune effector complex. *Extremophiles*, doi:10.1007/s00792-016-0871-5 (2016).

10 Gasiunas, G., Sinkunas, T. & Siksnys, V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell Mol Life Sci* **71**, 449-465, doi:10.1007/s00018-013-1438-6 (2014).

11 Sinkunas, T. *et al.* Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* **30**, 1335-1342, doi:10.1038/emboj.2011.41 (2011).

12 He, L., St John James, M., Radovcic, M., Ivancic-Bace, I. & Bolt, E. L. Cas3 Protein-A Review of a Multi-Tasking Machine. *Genes* **11**, doi:10.3390/genes11020208 (2020).

13 Sinkunas, T., Gasiunas, G. & Siksnys, V. Cas3 nuclease-helicase activity assays. *Methods in molecular biology (Clifton, N.J.)* **1311**, 277-291, doi:10.1007/978-1-4939-2687-9_18 (2015).

14 Chylinski, K., Makarova, K. S., Charpentier, E. & Koonin, E. V. SURVEY AND SUMMARY Classification and evolution of type II CRISPR-Cas systems. *Nucleic acids research* **42**, 6091-6105 (2014).

15 Rani, R. *et al.* CRISPR/Cas9: a promising way to exploit genetic variation in plants. *Biotechnol Lett* **38**, 1991-2006, doi:10.1007/s10529-016-2195-z (2016).

16 Mulepati, S. & Bailey, S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *The Journal of biological chemistry* **286**, 31896-31903, doi:10.1074/jbc.M111.270017 (2011).

17 Dolan, A. E. *et al.* Introducing a Spectrum of Long-Range Genomic Deletions in Human Embryonic Stem Cells Using Type I CRISPR-Cas. *Molecular cell* **74**, 936-950 e935, doi:10.1016/j.molcel.2019.03.014 (2019).

18 Xiao, Y., Luo, M., Dolan, A. E., Liao, M. & Ke, A. Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science (New York, N.Y.)* **361**, doi:10.1126/science.aat0839 (2018).

19 Nimkar, S. & Anand, B. Cas3/I-C mediated target DNA recognition and cleavage during CRISPR interference are independent of the composition and architecture of Cascade surveillance complex. *Nucleic acids research* **48**, 2486-2501, doi:10.1093/nar/gkz1218 (2020).

20 Li, X. *A construction method of recombinant Streptomyces viginiae IBL-14 used for penicillin producation,* MS thesis, Anhui University, (2016).

21 Yong, D.-X. *The CRISPR-Cas system found in Streptomyces viginiae IBL-14 and its gene editing procedures,* MS thesis, Anhui University, (2016).

22 Zhang, Y. *Study on the Physiological Function and Steroid Transformation Mechanism of CYP105 Family Genes in Streptomyces virginiae IBL14*, PhD thesis, Anhui University, (2017).

23 Qiu, C.-H. *Primary studies on both the Type I-B-svi CRISPR-Cas system in Streptomyces virginiae IBL14 and the self-targeting gene editing to the genes involved in penicillin metabolism,* MS thesis, Anhui University, (2017).

24 Xu, X. *Functional Analyses of the sviscsn Gene Cluster in Streptomyces virginiae IBL14,* MS thesis, Anhui university, (2017).

25 Cao, S.-L. *Self-targeting gene editing of O-methyltransferase genes in Streptomyces virginiae IBL14,* MS thesis, Anhui University, (2018).

26 Huang, H., Zheng, G., Jiang, W., Hu, H. & Lu, Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim Biophys Sin (Shanghai)* **47**, 231-243, doi:10.1093/abbs/gmv007 (2015).

27 Li, Z.-Z. *et al.* Identification and functional analysis of cytochrome P450 complement in *Streptomyces virginiae* IBL14. *BMC Genomics* **14**, 130 (2013).

28 Jervis, A. J. *et al.* A plasmid toolset for CRISPR-mediated genome editing and CRISPRi gene regulation in *Escherichia coli*. *Microb Biotechnol*, doi:10.1111/1751-7915.13780 (2021).

29 Pyne, M. E., Moo-Young, M., Chung, D. A. & Chou, C. P. Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in *Escherichia coli*. *Applied and environmental microbiology* **81**, 5103-5114, doi:10.1128/aem.01248-15 (2015).

30 Xue, C. & Sashital, D. G. Mechanisms of Type I-E and I-F CRISPR-Cas Systems in Enterobacteriaceae. *EcoSal Plus* **8**, doi:10.1128/ecosalplus.ESP-0008-2018 (2019).

31 Hochstrasser, M. L., Taylor, D. W., Kornfeld, J. E., Nogales, E. & Doudna, J. A. DNA Targeting by a Minimal CRISPR RNA-Guided Cascade. *Molecular cell* **63**, 840-851, doi:10.1016/j.molcel.2016.07.027 (2016).

32 Xiao, Y. *et al.* Structure Basis for Directional R-loop Formation and Substrate Handover Mechanisms in Type I CRISPR-Cas System. *Cell* **170**, 48-60 e11, doi:10.1016/j.cell.2017.06.012 (2017).

33 Wilkinson, M. E. *et al.* Structural plasticity and in vivo activity of Cas1 from the type I-F CRISPR-Cas system. *Biochem J* **473**, 1063-1072, doi:10.1042/bcj20160078 (2016).

34 Morisaka, H. *et al.* CRISPR-Cas3 induces broad and unidirectional genome editing in human cells. *Nature communications* **10**, 5302, doi:10.1038/s41467-019-13226-x (2019).

35 Csörgő, B. *et al.* A compact Cascade-Cas3 system for targeted genome engineering. *Nature methods*, doi:10.1038/s41592-020-00980-w (2020).

36 Liu, J. *et al.* Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microbial cell factories* **16**, 205, doi:10.1186/s12934-017-0815-5 (2017).

37 Nerys-Junior, A., Braga-Dias, L. P., Pezzuto, P., Cotta-de-Almeida, V. & Tanuri, A. Comparison of the editing patterns and editing efficiencies of TALEN and CRISPR-Cas9 when targeting the human CCR5 gene. *Genetics and molecular biology* **41**, 167-179, doi:10.1590/1678-4685-gmb-2017-0065 (2018).

38 Ferrari, S. *et al.* Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nature biotechnology* **38**, 1298-1308, doi:10.1038/s41587-020-0551-y (2020).

39 Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. & Yang, S. H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular therapy. Nucleic acids* **4**, e264, doi:10.1038/mtna.2015.37 (2015).

40 Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research* **24**, 132-141, doi:10.1101/gr.162339.113 (2014).

41 Kim, D. *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nature methods* **12**, 237-243, 231 p following 243, doi:10.1038/nmeth.3284 (2015).

42 Wang, G., Du, M., Wang, J. & Zhu, T. F. Genetic variation may confound analysis of CRISPR-Cas9 off-target mutations. *Cell discovery* **4**, 18, doi:10.1038/s41421-018-0025-2 (2018).

43 Rutkauskas, M. *et al.* Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep* **10**, 1534-1543, doi:10.1016/j.celrep.2015.01.067 (2015).

44 Pallarès-Masmitjà, M. *et al.* Find and cut-and-transfer (FiCAT) mammalian genome engineering. *Nature communications* **12**, 7071, doi:10.1038/s41467-021-27183-x (2021).

45 Qu, J. *et al.* Identification of a Novel Cleavage Site and Confirmation of the Effectiveness of NgAgo Gene Editing on RNA Targets. *Mol Biotechnol* **63**, 1183-1191, doi:10.1007/s12033-021-00372-1 (2021).

46 Kim, E. *et al.* In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nature communications* **8**, 14500, doi:10.1038/ncomms14500 (2017).

47 Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771, doi:10.1016/j.cell.2015.09.038 (2015).

48 Kuhstoss, S. & Rao, R. N. Analysis of the integration function of the streptomycete bacteriophage phi C31. *J Mol Biol* **222**, 897-908, doi:10.1016/0022-2836(91)90584-s (1991).

49 Rausch, H. & Lehmann, M. Structural analysis of the actinophage phi C31 attachment site. *Nucleic acids research* **19**, 5187-5189, doi:10.1093/nar/19.19.5187 (1991).

50 Feng, X., Bednarz, A. L. & Colloms, S. D. Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic acids research* **38**, 1204-1216, doi:10.1093/nar/gkp1068 (2010).

51 Liang, F., Han, M., Romanienko, P. J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5172-5177, doi:10.1073/pnas.95.9.5172 (1998).

52 Carroll, D. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther* **15**, 1463-1468, doi:10.1038/gt.2008.145 (2008).

53 Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**, 181-211, doi:10.1146/annurev.biochem.052308.093131 (2010).

54 Hinnen, A., Hicks, J. B. & Fink, G. R. Transformation of yeast. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 1929-1933, doi:10.1073/pnas.75.4.1929 (1978).

55 Capecchi, M. Gene targeting. How efficient can you get? *Nature* **348**, 109, doi:10.1038/348109a0 (1990).

56 Silva, G. *et al.* Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Current gene therapy* **11**, 11-27, doi:10.2174/156652311794520111 (2011).

57 Pingoud, A. & Silva, G. H. Precision genome surgery. *Nature biotechnology* **25**, 743-744, doi:10.1038/nbt0707-743 (2007).

58 Paques, F. & Duchateau, P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Current gene therapy* **7**, 49-66 (2007).

59 Fernandez-Martinez, L. T. & Bibb, M. J. Use of the meganuclease I-SceI of *Saccharomyces cerevisiae* to select for gene deletions in actinomycetes. *Scientific reports* **4**, 7100, doi:10.1038/srep07100 (2014).

60 Schleifman, E. B., Chin, J. Y. & Glazer, P. M. Triplex-mediated gene modification. *Methods in molecular biology (Clifton, N.J.)* **435**, 175-190, doi:10.1007/978-1-59745-232-8_13 (2008).

61 Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 1156-1160, doi:10.1073/pnas.93.3.1156 (1996).

62 Mani, M., Smith, J., Kandavelou, K., Berg, J. M. & Chandrasegaran, S. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochemical and biophysical research communications* **334**, 1191-1197, doi:10.1016/j.bbrc.2005.07.021 (2005).

63 Gutierrez-Guerrero, A. *et al.* Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus. *Human gene therapy* **29**, 366-380, doi:10.1089/hum.2017.047 (2018).

64 Persikov, A. V., Osada, R. & Singh, M. Predicting DNA recognition by Cys2His2 zinc finger proteins. *Bioinformatics* **25**, 22-29 (2009).

65 Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169-1175 (2002).

66 Li, T. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic acids research* **39**, 359-372, doi:10.1093/nar/gkq704 (2011).

67 Zhao, J., Sun, W., Liang, J., Jiang, J. & Wu, Z. A One-Step System for Convenient and Flexible Assembly of Transcription Activator-Like Effector Nucleases (TALENs). *Molecules and cells* **39**, 687-691, doi:10.14348/molcells.2016.0140 (2016).

68 Shmakov, S. *et al.* Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol* **15**, 169-182, doi:10.1038/nrmicro.2016.184 (2017).

69 Pausch, P. *et al.* CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science (New York, N.Y.)* **369**, 333-337, doi:10.1126/science.abb1400 (2020).

70 Csörgő, B. *et al.* A compact Cascade-Cas3 system for targeted genome engineering. *Nature methods* **17**, 1183-1190, doi:10.1038/s41592-020-00980-w (2020).

71 Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* **14**, 8096-8106, doi:10.1128/mcb.14.12.8096 (1994).

72 Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nature biotechnology* **34**, 933-941, doi:10.1038/nbt.3659 (2016).

73 Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, doi:10.1038/s41586-019-1711-4 (2019).

74 Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.)* **346**, doi:10.1126/science.1258096 (2014).