CRISPR-Cpf1 assisted genome editing of Corynebacterium glutamicum

Yu Jiang1,2, Fenghui Qian1,2, Junjie Yang1,2, Yingmiao Liu1,2,3,4, Feng Dong1,2, Chongmao Xu1,2, Bingbing Sun1,2,5, Biao Chen1,2, Xiaoshu Xu1, Yan Li6, Renxiao Wang6 & Sheng Yang1,2,3

Corynebacterium glutamicum is an important industrial metabolite producer that is difficult to genetically engineer. Although the Streptococcus pyogenes (Sp) CRISPR-Cas9 system has been adapted for genome editing of multiple bacteria, it cannot be introduced into C. glutamicum. Here we report a Francisella novicida (Fn) CRISPR-Cpf1-based genome-editing method for C. glutamicum. CRISPR-Cpf1, combined with single-stranded DNA (ssDNA) recombineering, precisely introduces small changes into the bacterial genome at efficiencies of 86–100%. Large gene deletions and insertions are also obtained using an all-in-one plasmid consisting of FnCpf1, CRISPR RNA, and homologous arms. The two CRISPR-Cpf1-assisted systems enable N iterative rounds of genome editing in 3N + 4 or 3N + 2 days. A proof-of-concept, codon saturation mutagenesis at G149 of γ-glutamyl kinase relieves L-proline inhibition using Cpf1-assisted ssDNA recombineering. Thus, CRISPR-Cpf1-based genome editing provides a highly efficient tool for genetic engineering of Corynebacterium and other bacteria that cannot utilize the Sp CRISPR-Cas9 system.
Corynebacterium glutamicum, a high-GC content, Gram-positive soil bacterium, is an important organism for the industrial production of amino acids1–3, and it has been engineered to produce a variety of compounds, including polymer subunits and biofuels4,5. Many approaches have been developed for introducing site-directed mutations into C. glutamicum6. For example, allelic exchange methods involve either introducing a selectable marker into the edited locus or a two-step process that includes a counter-selection system6–9. More recently, phage recombination proteins have been used for recombineering, a technique that promotes homologous recombination of linear DNA or oligonucleotides10,11. However, the absence of positive selection for mutations and a low recombineering efficiency often require screening of many colonies. Due to the industrial importance of C. glutamicum, finding a way to take advantage of a highly efficient Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system for robust genome editing is of high priority.

The Class 2 CRISPR-Cas prokaryote immunity system has been developed successfully as a tool for genome editing and transcription regulation, most often based on the Cas9 nuclease from Streptococcus pyogenes1–15. Deactivated Cas9 (dCas9)-based CRISPR interference technology has been successfully applied to C. glutamicum for target-specific knock-down of the expression of pgi, pck and pyk with an efficiency up to 97% (ref. 16). However, in our hands, genome editing of C. glutamicum cannot be achieved by simply replacing dCas9 with Cas9 or Cas9 nickase, possibly due to toxicity in the recipient cells. Cpf1, a single-strand RNA-guided endonuclease of the class 2 CRISPR-Cas system that cleaves target DNA with features distinct from those of Cas9 (ref. 17), such as utilizing a T-rich protospacer-adjacent motif (PAM) and cutting in staggered ends, may serve as an alternative to Cas9 for genome editing.

In the present work, we find that CRISPR-Cpf1 systems, when applied to C. glutamicum and related species, generate nucleotide substitutions, insertions and gene deletions with high efficiency (Fig. 1). When combine with single-stranded (ss) DNA recombineering technology, the CRISPR-Cpf1 system enables us to achieve genomic in situ codon saturation mutagenesis without relying on laborious pre-construction of libraries. CRISPR-Cpf1-assisted genome editing tools for C. glutamicum allow us to complete each round of iterative metabolic engineering in as little as 3 days, less than half the time required for sacB-based allelic exchange protocols8.

Results
Selection and functionality of Cas effector proteins. To develop a Cas9-based genome editing method, three C. glutamicum-Escherichia coli shuttle plasmids were constructed: pXJM19ts-Pncas9 carried SpCas9 with its native promoter and a kanamycin resistance cassette, pXJM19ts-Pcas9 carried SpCas9 under the control of the constitutive promoter PlacM6, and pXJM19ts-Pcas9n carried SpCas9 nickase and PlacM. Unexpectedly, none of these plasmids could generate C. glutamicum (ATCC13032) transformants. In contrast, pXJM19ts-Plpf1, carrying Cpf1 from Francisella novicida (FnCpf1), transformed C. glutamicum at a high efficiency, close to that achieved by the pXJM19ts control (Fig. 2a). This surprising difference in the transformation efficiency between SpCas9 and FnCpf1 suggests that the expression of SpCas9 in C. glutamicum is toxic. However, a Basic Local Alignment Search Tool (BLAST) search indicated that the C. glutamicum genome lacks a typical CRISPR RNA (crRNA) sequence of SpCas9.

To confirm that FnCpf1 is functional in C. glutamicum, we constructed plasmid pXJM19ts-Plpf1-crRNAcrYf, which carried FnCpf1, crRNA-targeting crtYf (cg0718), and a spectinomycin resistant, pXJM19ts compatible plasmid, pyS2_crtYf, which carried only the crtYf crRNA. Plasmid pXJM19ts-Plpf1-crRNAcrYf was introduced into C. glutamicum by bacterial transformation, and pyS2_crtYf was introduced into C. glutamicum harbouring pXJM19ts-Plpf1. The transformation efficiencies were several orders of magnitude lower than those of their respective control plasmids (Fig. 2a). Regardless of whether the FnCpf1 and crRNA modules were combined into a single plasmid or placed separately in two compatible plasmids, they efficiently cleaved the C. glutamicum genome at the crtYf locus, thereby leading to reduced transformant survival, which was a prerequisite for positive selection of genetically edited recombinants that resist FnCpf1-mediated cleavage at the targeted site(s).

Optimization of CRISPR-Cpf1-assisted ssDNA recombineering. To develop a CRISPR-Cpf1-assisted recombineering method, three RecT-expressing plasmids based on pXJM19ts-Plpf1 were constructed: pYJS1Ptac, pYJS1Ppet, and pYJS1Pefu, each which express E. coli recT under the control of an inducible promoter, Ptac or Ppet11, or a constitutive promoter, Pefu18 (Fig. 1a; Supplementary Fig. 1). To measure editing efficiency, we designed mutations of crtY that would alter its protospeccer seed sequence with a two-nucleotide mismatch. The mismatch was targeted by pyS2_crtYf with the help of editing template oligonucleotide recombineering (GC to TT, and introduction of an HpaI restriction site, Fig. 2b).

Plasmid pYJS1Ptac, pYJS1Ppetu, or pYJS1Ppet was introduced into C. glutamicum, and the resulting strains were transformed with 500 ng of pyS2_crtYf plus various amounts of editing template (59 nucleotides or 75 oligonucleotides that target the lagging strand, O_crtYf59-1 or O_crtYf75-, respectively). Neither the length and amount of editing template nor the presence of RecT affected the recovery of transformant colony forming units (c.f.u.; Fig. 2c, Supplementary Data 1). Although the editing efficiency with recT transcribed from an induced Ppet was less than 20%, the fraction of edited cells increased to more than 80% if recT was transcribed from Pefu or induced Ptac, using 1 μg of O_crtYf59-1. Increasing the amount of O_crtYf59-1 from 1 μg to 10 μg did not significantly increase the editing efficiency when recT was transcribed from either Pefu or Ptac. Increasing the oligonucleotide length from 59 to 75 nucleotides did not result in any significant difference in the editing efficiency (80% to 90 ± 10%, Fig. 2c, Supplementary Data 1, Supplementary Fig. 2) using Pefu-transcribed recT. However, when Ptac-transcribed recT was used, the editing efficiency decreased by 4-fold for unknown reasons (Fig. 2c, Supplementary Data 1, Supplementary Fig. 2). There was a strong strand bias for editing efficiency: the oligonucleotide that targeted the lagging strand (O_crtYf59-) recombined four-fold more efficiently than its complementary oligonucleotide (O_crtYf59 +) (Fig. 2d,e), a finding that is consistent with data obtained with Lactobacillus reuteri19,20. Nucleotide sequence determination of 10 HpaI-digested, correct mutants confirmed not only the presence of the introduced target mutations, but also the absence of unintended mutations in the target region (Fig. 2f). These data indicate that the CRISPR-Cpf1-assisted ssDNA recombineering system can serve as a robust and precise tool for introducing nucleotide substitutions in C. glutamicum.

By using the optimized CRISPR-Cpf1-assisted ssDNA recombineering system with co-transformation of the pyS2 plasmid series using 1 μg of the 59-nucleotide lagging-strand oligonucleotide into C. glutamicum harbouring pYJS1Ptac, we succeeded in obtaining a 50 bp deletion of crtYf; the editing efficiency was about

---

2 NATURE COMMUNICATIONS | DOI: 10.1038/ncomms15179 | www.nature.com/naturecommunications
15% (Fig. 3a, Supplementary Data 1, Supplementary Fig. 3). However, the method failed to produce a 500 bp deletion (Fig. 3a, Supplementary Data 1, and Supplementary Fig. 3). The editing efficiency increased to 40% for the 17 bp deletion and 100% for the 2-nucleotide substitution in argR (cg1585) for the corresponding double plasmids and oligonucleotides, respectively (Supplementary Fig. 4, Supplementary Data 1). Therefore, CRISPR-Cpf1-assisted ssDNA recombineering is more suitable for small alterations than for introducing large deletions or insertions into the C. glutamicum genome.

CRISPR-Cpf1-assisted gene deletion and insertion. To improve the efficiency of the CRISPR-Cpf1-assisted genome editing...
Figure 2 | Optimization of double-plasmid-based CRISPR-Cpf1-ssDNA recombineering in crtYf of C. glutamicum. (a) Growth of C. glutamicum cells expressing the nuclease, with or without the combined expression of crRNA targeting crtYf. Asterisks indicate c.f.u. were not detected. NA, not applicable. (b) Overview of ssDNA-mediated nucleotide substitution in crtYf. (c) C. glutamicum cell growth and the mutation efficiency of crtYf after editing by 59 nt (O_crtYf59-1) or 75 bp (O_crtYf75-) template oligonucleotide targeting the lagging strand mediated by the pJYS1Ptac/pJYS2_crtYf and pJYS1 series carrying RecT under the control of various promoters. The mutation efficiency was determined by colony PCR, followed by HpaI digestion, as indicated in Supplementary Fig. 3. When a 1 kb oligonucleotide targeted the lagging strand, a 1.6 kb, 1.1 kb, and 0.2 kb fragment was observed in WT, 1.6 kb and 0.2 kb fragment, indicating wild-type genotype, whereas the presence of 1.1 kb and 0.2 kb fragments indicated recombinant genotypes. (n/N): n, number of correctly edited, positive transformants; N, number of transformants tested. DNA ladder mix (GeneRuler, Thermo Scientific) was used as a marker. (f) Ten representative crtYf recombinants identified by HpaI digestion were further sequenced, which revealed the substitution of GC by TT for all samples, as expected. Details in transformants recording and editing efficiency calculation are listed in Supplementary Data 1. Experiments were performed in duplicates. Bar represents mean ± s.d.

Method for gene deletions and insertions, a set of all-in-one CRISPR-Cpf1 plasmids was constructed. Upstream and downstream homologous arms (1 kb each) were inserted into plasmid pXM19ts-Ptcp1-crRNAcrYf to create pJYS3_AcrtYf, which aimed to delete 705 bp within crtYf (Fig. 1b). Transformation with 1 µg of pJYS3_AcrtYf into C. glutamicum competent cells produced more than 500 c.f.u., among which ~15% were correctly edited (Fig. 3a, Supplementary Data 1, and Supplementary Fig. 3). Using homologous arms flanking a 7.5 kb region from crtYf (cg0718) to cg0723, pJYS3_AcrtYf to cg0718 edited C. glutamicum with an efficiency of 10% (Fig. 3a, Supplementary Data 1, and Supplementary Fig. 3). When a 1 kb tdcB gene was inserted between the homologous arms of pJYS3_AcrtYf, 1 µg of the resulting plasmid, pJYS_AcrtYf:tdcB, when transformed...
cleavage activity was similar among these constructs (Fig. 3b).

Since reducing the number of CRISPR escapers (bacteria that are not cut by the Cpf1 nuclease despite of the existence of the Cpf1 cleavable target) may help improve editing efficiency, we tried to enhance Cpf1 cleavage by replacing the PlacM promoter driving FnCpf1 expression in the pJYS3 plasmid series with two strong constitutive promoters, Psod and Pefu, or with three inducible promoters, Ptrc, PtcR, and Ptet.11. This construction led to the pJYS3Psod, pJYS3Pefu, pJYS3Ptac, pJYS3Ptrc and pJYS3Ptet plasmid series. Transformants were recovered with these plasmids at similar frequencies, indicating that Cpf1 cleavage activity was similar among these constructs (Fig. 3b).

The efficiencies of editing with these plasmids were also similar to those of the corresponding pJYS3 series in which FnCpf1 expression was under the control of PlacM (Fig. 3a; Supplementary Data 1). The in vivo activity of FnCpf1 and the predicted FnCpf1 nickase R1218A were compared. Details in transformants recording and editing efficiency calculation are listed in Supplementary Data 1. (c) CRISPR–Cpf1 plasmid-curing efficiency for iterative genome manipulations. Cells carrying double plasmids pYS1Pac/pYS2_crtYf or pJYS3_crtYf were cultured in liquid BHISG with or without the indicated antibiotics at the indicated temperatures. Cultures were diluted and plated on BHISG plates with or without the indicated antibiotics and incubated for 48 h at 30 °C. The relative plasmid-curing efficiency was calculated by dividing the number of c.f.u. obtained from antibiotic-containing plates by the c.f.u. obtained from plates lacking the corresponding antibiotics. Experiments were performed in duplicates. Bar represents mean ± s.d.

Figure 3 | Double and all-in-one plasmid-based deletion and insertion in the crtYf locus of C. glutamicum. (a) Overview of crtYf editing by the ssDNA-mediated double-plasmid-based CRISPR–Cpf1 system (pJYS1Ptac/pJYS2_crtYf) or the all-in-one CRISPR–Cpf1 plasmid (pJYS3 series). The editing efficiency and numbers of c.f.u. of the corresponding transformants were determined by colony PCR and calculated as indicated in Supplementary Data 1 and Supplementary Fig. 3. (b) C. glutamicum growth after transformation with CRISPR–Cpf1 plasmids (pJYS3_crtYf, pJYS3Pefu_crtYf, pJYS3Psod_crtYf, pJYS3Ptac_crtYf, pJYS3Ptrc_crtYf, or pJYS3Ptet_crtYf) containing FnCpf1 under the control of various promoters (PlacM, Pefu, Psod, PtcR or Ptet) and the crRNA targeting crtYf locus. CRISPR–Cpf1 plasmid pJYS3 lacking crRNA was used as a control. The in vivo activity of FnCpf1 and the predicted FnCpf1 nickase R1218A were compared. Details in transformants recording and editing efficiency calculation are listed in Supplementary Data 1. (c) CRISPR–Cpf1 plasmid-curing efficiency for iterative genome manipulations. Cells carrying double plasmids pYS1Pac/pYS2_crtYf or pJYS3_crtYf were cultured in liquid BHISG with or without the indicated antibiotics at the indicated temperatures. Cultures were diluted and plated on BHISG plates with or without the indicated antibiotics and incubated for 48 h at 30 °C. The relative plasmid-curing efficiency was calculated by dividing the number of c.f.u. obtained from antibiotic-containing plates by the c.f.u. obtained from plates lacking the corresponding antibiotics. Experiments were performed in duplicates. Bar represents mean ± s.d.

Figure 3 | Double and all-in-one plasmid-based deletion and insertion in the crtYf locus of C. glutamicum. (a) Overview of crtYf editing by the ssDNA-mediated double-plasmid-based CRISPR–Cpf1 system (pJYS1Ptac/pJYS2_crtYf) or the all-in-one CRISPR–Cpf1 plasmid (pJYS3 series). The editing efficiency and numbers of c.f.u. of the corresponding transformants were determined by colony PCR and calculated as indicated in Supplementary Data 1 and Supplementary Fig. 3. (b) C. glutamicum growth after transformation with CRISPR–Cpf1 plasmids (pJYS3_crtYf, pJYS3Pefu_crtYf, pJYS3Psod_crtYf, pJYS3Ptac_crtYf, pJYS3Ptrc_crtYf, or pJYS3Ptet_crtYf) containing FnCpf1 under the control of various promoters (PlacM, Pefu, Psod, PtcR or Ptet) and the crRNA targeting crtYf locus. CRISPR–Cpf1 plasmid pJYS3 lacking crRNA was used as a control. The in vivo activity of FnCpf1 and the predicted FnCpf1 nickase R1218A were compared. Details in transformants recording and editing efficiency calculation are listed in Supplementary Data 1. (c) CRISPR–Cpf1 plasmid-curing efficiency for iterative genome manipulations. Cells carrying double plasmids pYS1Pac/pYS2_crtYf or pJYS3_crtYf were cultured in liquid BHISG with or without the indicated antibiotics at the indicated temperatures. Cultures were diluted and plated on BHISG plates with or without the indicated antibiotics and incubated for 48 h at 30 °C. The relative plasmid-curing efficiency was calculated by dividing the number of c.f.u. obtained from antibiotic-containing plates by the c.f.u. obtained from plates lacking the corresponding antibiotics. Experiments were performed in duplicates. Bar represents mean ± s.d.
Compared with the control plasmid lacking FnCpf1 (Fig. 3b). Thus, the FnCpf1 nickase cannot be used for positive selection of edited targets that tolerate FnCpf1 cleavage.

Plasmid curing for iterative genome editing. We next cultivated edited C. glutamicum strains in medium supplemented with kanamycin at 30°C to maintain the pJYS1-derivative plasmids but cure cells of the pJYS2 derivatives (they are segregationally unstable due to using a pGA1 replicon that possesses the replicative protein Rep, but not the distribution protein PeP24). To obtain plasmid-free strains, edited C. glutamicum strains were then cultivated in antibiotic-free medium at 34°C because the pJYS1 and pJYS3 derivatives are all based on the pBL15 replicon, which is temperature sensitive23. Loss of the corresponding plasmids was nearly 100% after an overnight incubation (Fig. 3c). This feature facilitated iterative genome editing. The operation time was 3N + 2 or 3N + 4 days for N targets with the pJYS1/pJYS2- or pJYS3-derived double or the all-in-one plasmid systems, respectively (Fig. 1).

Applicability in other Corynebacterium species. To determine whether CRISPR-Cpf1-assisted recombineering is applicable to C. glutamicum-related bacteria other than strain C. glutamicum ATCC13032, we transformed pJYS1Ptac into six industrial Corynebacterium strains using ATCC13032 as a control. Because electroporation-competent cells were prepared from cultures growing in tubes rather than in flasks, the overall efficiency of ATCC13032 transformation with the control plasmid pJYS1Ptac was several orders of magnitudes lower (3.56 ± 0.05 × 10^2 c.f.u.) (Fig. 4a; Supplementary Data 1) than observed with competent cells prepared with flask cultures (Supplementary Data 1). Nevertheless, we did not observe any obvious toxicity with the FnCpf1-expressing plasmid for the Corynebacterium strains tested. The resulting strains were then transformed with pJYS2_crtYf singly or together with O_crtYf59-1. The editing efficiency for a 2-nucleotide substitution was 100% for all seven Corynebacterium strains (Fig. 4b); the total transformant numbers varied from 34 ± 22 to 350 ± 40 among the various strains (Fig. 4b; Supplementary Data 1). Interestingly, except for ATCC13032, no other Corynebacterium strain escaped cleavage after the introduction of pJYS2_crtYf alone (Fig. 4a).

Application in codon saturation mutagenesis. The ~100% editing efficiency of CRISPR-assisted ssDNA recombineering raised the possibility of performing codon saturation mutagenesis (mutagenesis that causes a change from a wild-type amino acid to all other amino acids) in chromosomal DNA of C. glutamicum. To demonstrate feasibility, we edited the γ-glutamyl kinase gene (proB) of C. glutamicum ATCC13032 to relieve feedback inhibition by l-proline (a useful chemical that is produced commercially by C. glutamicum26), thereby facilitating l-proline production (Fig. 5a).

The first and rate-limiting step in the synthesis of l-proline from l-glutamate is catalysed by γ-glutamyl kinase. With E. coli, this enzyme is inhibited by l-proline, and an D107N or a E143A substitution results in a loss of allosteric regulation27. No l-proline feedback-insensitive γ-glutamyl kinase variant from C. glutamicum has been reported.

A sequence alignment of ProB from Burkholderia thailandensis (btProB), E. coli (ecProB), and C. glutamicum (cgProB) indicated that D154 and N155 of cgProB resemble D148 and N149 of ecProB previously reported as the location for the binding motif of l-proline28 (Fig. 5b). Assuming that D154 and N155 of cgProB are part of the active site, 3D protein structure modelling indicated that five amino-acid residues (G149, G153, D154, N155, and D156) participate in a complex hydrogen-bonding network with l-proline (Fig. 5b). Four residues (other than G149) are highly conserved among btProB, ecProB, and cgProB (Fig. 5b), and they are predicted to bind the substrate glutamate as determined by a previous study of D148, N149, and D150 of ecProB28,29. In addition, G149 of cgProB matches E143 of ecProB, a site involved in l-proline inhibition27.
Therefore, G149 was chosen as the target site for saturation mutagenesis.

To obtain a high mutation efficiency at codon G149, we first assessed the functionality of three PAMs (PAM1, TTTG; PAM2, ATTC; and PAM3, GTTG) near G149 of proB (Fig. 5e). Then, three corresponding pJYS2 plasmids (pJYS2-proB1, pJYS2-proB2, and pJYS2-proB3), were constructed. While pJYS2-proB2 transformed C. glutamicum ATCC13032 harbouring pJYS1Peftu at high efficiency (similar to the control plasmid pTrcmob_spc), the number of transformants was lower by about three orders of magnitude for pJYS2-proB1 or pJYS2-proB3 plasmids (Fig. 5d). These results are consistent with previous statistical data identifying functional PAM diversity across CRISPR-Cas systems30.

We next determined how many adjacent mismatches (all silent mutations) must be introduced to evade the mismatch repair system. Two oligonucleotides, F152FG149D and A146AG149D (Supplementary Data 2), were designed to construct the G149D substitution along with a silent mutation at the +1 position of PAM1 or PAM3 to inactivate PAM. No mutation was detected in 10 transformants from each of the two types. After the addition of two more silent mutations at the +1 and +3 protospacer positions, 0, 1, 2, 3, and 4 oligo recombination frequencies were determined. Three PAM sequences near G149 of ProB are highlighted in yellow, and the wild-type codon 149 is underlined in bold. Double site-directed mutagenesis of F152FG149D and A146AG149D, and multiple mutageneses of A146AT147TT148TG149D, A146AT147TT148TG149A, A146AT147TT148TG149C, and A146AT147TT148TG149F were performed. (e) Forty-two proB recombinants with substitutions at codon 149 (table) screened by 96-well fermentation exhibited different degrees of L-proline formation (blue bars). Experiments were performed in duplicates. Bar represents mean ± s.d.
position of PAM3 A146AT147TT148TG149D, at least 50% of the transformant colonies were identified as mutants (Fig. 5d). These results are consistent with a previous work with E. coli in which changing the wobble (third) position in several nearby codons resulted in high levels of ssDNA recombination, even though the wild-type amino-acid sequence is maintained at these nearby codons. Multiple substitutions of A146AT147TT148TG149A, A146AT147TT148TG149C and A146AT147TT148TG149F, which cover all substitution types at G149 of ProB, resulted in mutation efficiencies of 50–90% (Fig. 5d).

Finally, a mixture of oligonucleotides was designed for codon saturation mutagenesis of G149. We designed 20 59-oligonucleotides, each with 1 to 3 nucleotide substitutions in ProB G149, to introduce 20 codons covering 19 non-glycine amino acids and a stop codon (Supplementary Table 1). One microgram of each of the 20 oligonucleotides was mixed and co-transformed with pYS2_proB3 into C. glutamicum harbouring pYS1Peftu (Fig. 5a). Thirty transformants were analysed; 25 incorporated the desired changes, which covered 13 amino acids (Supplementary Table 2). Of the 190 colonies that were picked, 115 exhibited an l-proline titre higher than 2 g l$^{-1}$, as compared to <1 g l$^{-1}$ observed with wild-type ATCC13032. For 42 colonies, the l-proline titre was higher than 4 g l$^{-1}$ (Supplementary Table 3). These mutants contained 15 amino-acid substitutions at ProB G149 (Supplementary Table 4); their effect on l-proline titres from high to low, was in the order of K→T→D→V→R→Q>N→H→W→L>S>A→Y>C→G (Fig. 5e). The highest l-proline titre (6.6 ± 1.0 g l$^{-1}$) occurred with a G149K variant.

Discussion

CRISPR-Cas systems are now revolutionizing biotechnology, biology and medicine. So far, nearly all applications have been based on Cas9, the Cas effector protein of the type II CRISPR-Cas system. Cpf1 is a different Cas effector protein that is derived based on Cas9, the Cas effector protein of the type II CRISPR-Cas system. CRISPR-Cas systems are now revolutionizing biotechnology, even though the wild-type amino-acid sequence is maintained at these nearby codons. Multiple substitutions of A146AT147TT148TG149A, A146AT147TT148TG149C and A146AT147TT148TG149F, which cover all substitution types at G149 of ProB, resulted in mutation efficiencies of 50–90% (Fig. 5d).

To our knowledge, complete failure of Cas9 for genetic editing has not been reported previously. Because our SpCas9- or nickase-expressing plasmids, pXM191ts-Pncas9, pXM191ts-Pcas9 and pXM191ts-Pcas9n, lack a customized crRNA cassette, their toxicity cannot be explained by off-target cleavage. The toxicity is also unlikely to be caused by an endogenous crRNA, because we could not find any typical SpCas9 crRNA sequence in the C. glutamicum genome. Thus we speculate that SpCas9 binds tightly to PAMs, even without a crRNA, thereby explaining why an inactivated SpCas9, expressed from a de-repressed Ptac promoter, produced no transformant.

Given the diversity of species in which SpCas9 has been used, it is surprising that SpCas9 is toxic to C. glutamicum while FnCpf1 is not, although both enzymes are expressed from the same plasmid vector. One could argue that C. glutamicum, which belongs to Actinobacteria, a high-GC content bacterial phylum, possesses more PAM sequences for SpCas9 (NGG) than for FnCpf1 (TGTN). Interestingly, only Cas9-CRISPR interference, not Cas9-CRISPR editing, has been established in Mycobacterium tuberculosis, another GC-rich bacterium of the Actinobacteria phylum. Thus, M. tuberculosis may also be intolerant to SpCas9. However, SpCas9 works well in Streptomyces species that are also GC-rich. CRISPR-Cas is exclusively derived from prokaryotes; therefore, the introduction of heterologous CRISPR-Cas components may interfere with the native CRISPR system. Clarifying differences between Cas9 and Cpf1 requires more data from other Class 2 CRISPR-Cas systems, including type V FnCpf1 and type II SpCas9 orthologs.

While addressing the problem of SpCas9 toxicity is of scientific interest, from a practical perspective, our finding that FnCpf1 behaves very differently from SpCas9 suggests that when problems of CRISPR-Cas setup are encountered, assessing different CRISPR-Cas systems, rather than fine-tuning the expression of Cas effector proteins and/or the transcrRNA, may be an effective solution. CRISPR nucleases having a longer PAM and different base compositions may also help reduce off-target effects.

Methods

Strains. All plasmids were introduced by electroporation into competent cells of E. coli DH5α grown in Luria–Bertani broth (LB) with ampicillin (100 mg l$^{-1}$) at 37°C. C. glutamicum strains were generally cultured in brain–heart infusion (BHI; Bacto) supplemented with 50 mM sorbitol (BHIS) unless otherwise indicated. Antibiotic concentrations, when necessary, were 25 mg l$^{-1}$ kanamycin and 50 mg l$^{-1}$ spectinomycin at 30°C. Agar was added at 1.5 g l$^{-1}$ for plates. All recombinant strains were transformed by electroporation. All strains were kept as glycerol stocks prepared in LB or BHI broth containing 20% glycerol at −80°C. Strains and plasmids are listed in Table 1 and Supplementary Table 5.

Plasmid constructions and oligonucleotides. All constructs used in the study are listed in Table 1 and Supplementary Table 5. Sequences of the primers, crRNAs, and oligonucleotides used in the study are listed in Supplementary Table 5. Plasmids and chromosomal DNA were extracted using AxyPrep kits (Corning). PCR primers used Taq (Thermo Scientific) or KOD-plus-neo polymerases (Toyobo). Cloning used restriction endonucleases, T4 DNA ligase (Thermo Scientific), and the isothermal assembly method.

homologous recombination by improving the efficiency of large gene deletions and insertions. However, dsDNA recombination activities associated with the Rac prophage recombinease and the exonuclease RecE/RecT from E. coli are rather low in C. glutamicum. Other Cpf1 orthologs, such as those from Lachnospiraceae bacterium and an Acidaminococcus sp, have been reported to be more efficient endonucleases in human embryonic kidney 293 cells, which will allow further improvement in deletion/insertion efficiency.
Plasmid pXM19ts-Pncas9 was cloned by the isothermal assembly using four DNA fragments: a temperature-sensitive replicon fragment, pBLts, a kanamycin resistance gene (Kan), E. coli replicon pSC101, and cas9 with its native promoter. Plasmids pXM19ts-PcMcas9, pXM19ts-PcMcas9a, pXM19ts-PcMcrRNAcrtYf were generated using pXM19ts-Pncas9 as a template to amplify the pSC101-PBLts-Kan fragment, and they were assembled with a fragment containing various nucleic genes with a given promoter (PlacM-cas9, PlacM-cas9D10A, PlacM-cpf1 and PlacM-cpf1n, respectively) using the isothermal assembly method. To generate pXM19ts-Pncas9, pXM19ts was used as a template to amplify a ~1.5 kb fragment using primers P1/P2 and a 1.3 kb fragment using primers P3/P4. A temperature-sensitive replicon fragment, pBLts, was obtained by overlap-extension PCR of these two fragments. A ~1.2 kb fragment of the kanamycin resistance gene (Kan) was amplified using pTRCmob as a template and primers P5/P6. A ~1.8 kb fragment of E. coli replicon pSC101 was amplified using pMWJ19 as a template and primers P7/P8. A ~4.6 kb pncas9 fragment was amplified using the pCas plasmids as a template and primers P9/P10. These four fragments were purified and recovered for isothermal assembly. To generate pXM19ts-PcMcas9a, pXM19ts-PcMcas9 was used as a template to amplify a ~5.7 kb fragment containing the Cas9 gene. The pCas fragment was amplified by two steps using the primers P11/P12, as well as a ~4.2 kb fragment of PlacM-cas9 using primers P13/P14. These two fragments were purified and recovered for isothermal assembly. To generate pXM19ts-PcMcas9a, pXM19ts-PcMcas9 was used as a template to amplify 5.7 and 4.2 kb fragments with primers P15/P16 and P17/P18, respectively.

All-in-one CRISPR-Cpf1 plasmid

pJYS2_crtYf: rep oriVcep, Sp1 pMB1 oriVE, rep P23191-cRNA targeting crtYf

To generate pXM19ts-Plcas9, pXM19ts-Plcas9 was used as a template to amplify a 1.2 kb fragment of the kanamycin resistance gene (Kan) using primers P26/P27. Primers P28/P29 were used to amplify a 93 bp fragment of the kanamycin resistant gene using primers P30/P31. These fragments were purified and recovered for isothermal assembly. To generate pXM19ts-Plcas9n, pXM19ts-Plcas9 was used as a template to amplify a 1.2 kb fragment of the kanamycin resistance gene (Kan) using primers P32/P33. These fragments were purified and recovered for isothermal assembly. To generate pXM19ts-Plcas9n, pXM19ts-Pncas9 was used as a template to amplify a ~5.7 kb fragment containing the Cas9 gene. The pCas fragment was amplified by two steps using the primers P34/P35, as well as a ~1.2 kb fragment of PlacM-cas9 using primers P36/P37. These two fragments were purified and recovered for isothermal assembly. To generate pXM19ts-PcMcas9a, pXM19ts-PcMcas9 was used as a template to amplify a ~1.5 kb fragment using primers P1/P2 and a ~1.3 kb fragment using primers P3/P4. A temperature-sensitive replicon fragment, pBLts, was obtained by overlap-extension PCR of these two fragments. A ~1.2 kb fragment of the kanamycin resistance gene (Kan) was amplified using pTRCmob as a template and primers P5/P6. A ~1.8 kb fragment of E. coli replicon pSC101 was amplified using pMWJ19 as a template and primers P7/P8. A ~4.6 kb pncas9 fragment was amplified using the pCas plasmids as a template and primers P9/P10. These four fragments were purified and recovered for isothermal assembly. To generate pXM19ts-PcMcas9a, pXM19ts-PcMcas9 was used as a template to amplify a ~5.7 kb fragment containing the Cas9 gene. The pCas fragment was amplified by two steps using the primers P11/P12, as well as a ~4.2 kb fragment of PlacM-cas9 using primers P13/P14. These two fragments were purified and recovered for isothermal assembly.

To generate pXM19ts-PcMcas9a, pXM19ts-PcMcas9 was used as a template to amplify 5.7 and 4.2 kb fragments with primers P15/P16 and P17/P18, respectively.
pJYS3Ptac_A0716/0723 was generated by assembling fragments 11, 12, and 14 via the isothermal assembly method. pJYS3Ptet_A0716/0723 was generated by assembling fragments 11, 12, and 14 via the isothermal assembly method.

To generate pJYS3P tetAcrtYf/scdb, pJYS3AcrtYf/scdb was linearized with KpnI/Xhol to generate a 1.3 kb fragment (fragment 13), which was assembled with fragments 9 and 10 via the isothermal assembly method. pJYS3P tetAcrtYf/scdb was generated by spheroplast transformation at 34 °C in BHISG and BHISG+plates to obtain bacteria in which the pJYS1 and pJYS2 series were both lost.

When using the all-in-one CRISPR-Cpf1 plasmid series for continuous genome manipulation, curing of the pJYS3 series was performed in the same manner as curing the pJYS1 series described above.

Sequence alignment and molecular modelling. To determine whether pJYS3_0723 has a single CRISPR-Cpf1 sequence, the genome sequence of ATCC13032 (NC_006958) was uploaded to CRISPRs finder online server (http://crispr.i2bc.paris-saclay.fr/Server/). An alignment was conducted using the BLAST at the National Center for Biotechnology Information.

The three-dimensional structural model of cproB was generated through homology modelling using the SWISS-MODEL server (https://swissmodel.expasy.org/). The available structure of cproB (PDB code: 4q1t, which shares 38% sequence similarity to cproB) was used as the template. Then, the binding mode of proline to cproB was predicted by molecular docking using Schrodinger software (Schrodinger LLC). The cproB protein structure was prepared using the Protein Preparation module to set the protonation state of the protein and optimize its structure using Maestro, on which the secondary amine group was protonated to bear a positive charge. A cubic grid required for docking was generated by centering at the centroid of 48 and 115 with a length of 1 nm for each dimension. Then, the proline molecule was docked into the binding site using the GLIDE module with GlideScore (v 5.0)35–37 as the scoring function. The standard precision mode of the scoring function was used to rank the docking poses. The all-atom optimized potentials for liquid simulations force field was applied to molecular docking. Other parameters were set to the default values.

L-Proline producer construction and cultivation assay. Electrocopentoe cells of C. glutamicum ATCC13032 containing pJYS3Peuf were prepared as described above. For cultivation, the strain was inoculated with a loopful of ProB G149, a mixture of 20 oligonucleotides (1 µg each of 59 bp oligonucleotides with 20 µg total per mixture) was used for electroporation. Transformants were inoculated into 96-well plates for l-proline fermentation tests. Fermenation volume was 600 µl, incubation temperature was 30 °C, rotation speed was 290 r.p.m. (Innova 43, Eppendorf, Germany), and incubation time was 72 h. Proline fermentation medium was composed of glucose 100 g l−1, corn steep liquor (Meihua Group) 20 g l−1, (NH₄)₂SO₄ 30 g l−1, MgSO₄·7H₂O 0.4 g l−1, KH₂PO₄ 1.2 g l−1, urea 2.9 g l−1, CaCO₃ 30.1 g l−1, pH 7.2. Strains producing l-proline were chosen for nucotide sequence determination.

To quantify l-proline, high-pressure liquid chromatography was performed using an Agilent 1200 series equipped with an Eclipse Plus C18, 3.5 µm, 150 mm, and a diode array detector (DAD G1321A) via a 6-fluoroaldehyde and 9-fluorenylmethyl chloroformate derivative reaction according to the high-speed amino-acid analysis on 1.8 µm reversed-phase columns instruction provided by Agilent. The mobile phase was composed of a gradient of 4 different eluents, and the flow rate was set at 1.0 ml min−1 at 40 °C. Glucose was measured by high pressure liquid chromatography (HPLC) and oxygen consumption using a combination for use and care of Aminex Resin-Based Columns (Bio-Rad) using an Agilent 1200 series equipped with an Aminex HXP-87H, 7.8 × 300 mm (Bio-Rad), and a refractive index detector (RID G1362A) at 60 °C with mobile phase composed of 5 mM H₂SO₄ and Flow rate of 1.0 ml min−1.

Data availability. The authors declare that all the data supporting the findings of this study are available within the paper and its Supplementary Information files or are available from the corresponding author on request.

References
1. Wendisch, V. F., Jorge, J. M., Perez-Garcia, F. & Sgobba, E. Updates on Corynebacterium glutamicum. World J. Microbiol. Biotechnol. 32, 105 (2016).
2. Lee, J. Y., Na, A. Y., Kim, E., Lee, H. S. & Kim, P. The Actinobacterium Corynebacterium glutamicum, an Industrial Workhorse. J. Microbiol. Biotechnol. 26, 807–822 (2016).
3. Eggeling, L. & Bott, M. A giant market and a powerful metabolism: l-lysine provided by Corynebacterium glutamicum. Appl. Microbiol. Biotechnol. 99, 3387–3394 (2015).
4. Heider, S. A. & Wendisch, V. F. Engineering microbial cell factories: Metabolic engineering of Corynebacterium glutamicum with a focus on non-natural products. Biotechnol. J. 10, 1641–1654 (2015).
5. Nesvöra, J. & Patek, M. Tools for genetic manipulations in Corynebacterium glutamicum and their applications. Appl. Microbiol. Biotechnol. 90, 1641–1654 (2011).
6. Tan, Y., Xu, D., Li, Y. & Wang, X. Construction of a novel sacB-based system for marker-free gene deletion in Corynebacterium glutamicum. *Plasmid* **67**, 44–52 (2012).

7. Okibe, N., Suzuki, N., Inui, M. & Yukawa, H. Efficient markerless gene replacement in *Corynebacterium glutamicum* using a new temperature-sensitive plasmid. *J. Microbiol. Methods* **85**, 155–163 (2011).

8. Schafer, A. et al. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene** **145**, 69–73 (1994).

9. Schwarzer, A. & Pühler, A. Manipulation of *Corynebacterium glutamicum* by gene disruption and replacement. *Biotechnol. (N Y)** **9**, 84–87 (1991).

10. Ma, W. et al. Development of a markerless gene replacement system in *Corynebacterium glutamicum* using up as a counter-selection marker. *Biotechnol. Lett.* **37**, 609–617 (2015).

11. Binder, S., Siddler, S., Marienhagen, J., Bott, M. & Eggeling, L. Recombining in *Corynebacterium glutamicum* combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res.* **41**, 6360–6369 (2013).

12. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science** **339**, 823–826 (2013).

13. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).

14. Qi, L. & Doudna, J. A. A programmable system for specific interference in gene expression. *Cell** **152**, 1153–1163 (2013).

15. Cleto, S., Jensen, J. V. K., Wendisch, V. F. & Lu, T. K. From zero to hero--design-based systems metabolic engineering of *Corynebacterium glutamicum* using upp as a counter-selection marker. *Biotechnol. J.* **3**, 1431–1446 (2007).

16. Becker, J., Zelder, O., Hafner, S., Schroder, H. & Wittmann, C. From zero to hero--design-based systems metabolic engineering of *Corynebacterium glutamicum* using upp as a counter-selection marker. *Biotechnol. J.* **3**, 1431–1446 (2007).

17. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell** **163**, 793–807 (2015).

18. Becker, J., Zelder, O., Hafner, S., Schroder, H. & Wittmann, C. From zero to hero--design-based systems metabolic engineering of *Corynebacterium glutamicum* using upp as a counter-selection marker. *Biotechnol. J.* **3**, 1431–1446 (2007).

19. Halgren, T. A. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell** **163**, 793–807 (2015).

20. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell** **163**, 793–807 (2015).

21. Nakanishi, T., Yonekura, H., Hattori, K., Hirao, T. & Azuma, T. Processes for producing L-proline by fermentation. *EP Patent 0,098,122B1* (1989).

22. Okibe, N., Suzuki, N., Inui, M. & Yukawa, H. Efficient markerless gene replacement in *Corynebacterium glutamicum* using a new temperature-sensitive plasmid. *J. Microbiol. Methods* **85**, 155–163 (2011).

23. Yamano, T. & Nakamura, J. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* **81**, 2506–2514 (2015).

24. Chen, T., Zhu, N. & Xia, H. Aerobic production of succinate from arabinose by metabolically engineered *Corynebacterium glutamicum*. *Biotechnol. Biofuels* **15**, 141–144 (2011).

25. Grissa, I., Vergnaud, G. & Pourcel, C. CRISPRFinder: a web tool to identify clustered regularly interspersed short palindromic repeats. *Nucleic Acids Res.* **32**, 293–296 (2004).

26. Benkert, P., Biasini, M. & Schwede, T. The SWISS-MODEL workbench: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).

27. Qi, L. & Doudna, J. A. A programmable system for specific interference in gene expression. *Cell** **152**, 1153–1163 (2013).

28. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimou, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).

29. Maestro, version 10.5 https://www.schrodinger.com/citations. (Schrodinger, New York, 2016).

30. Friesner, R. A. et al. Precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196 (2006).

31. Friesner, R. A. et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

32. Halgren, T. A. et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759 (2004).

Acknowledgements

We thank Dr. Xin Zhao and Karl Drlica for critical comments on the manuscript, and Dr. Liuyang Diao and Ms. Xiaoxue Chen for technical support. This work was supported in part by the National Basic research program of China (973: 2014CB745101) and Shanghai scientific research project (14XD1424900).

Author contributions

S.Y. and Y.I. conceived the project and designed the experiments. F.Q., Y.I., F.D., C.X., B.S. and B.C. performed the experiments. J.Y. analysed the sequence data. X.X., Y.L. and X.X. contributed to the writing of the manuscript. This work was supported in part by the National Basic research program of China (973: 2014CB745101) and Shanghai scientific research project (14XD1424900).

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Jiang, Y. et al. CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum* lactic acid. *Nat. Commun.* **8**, 15179 doi: 10.1038/ncomms15179 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.