Boosting activity of high-fidelity CRISPR/Cas9 variants using a tRNA\textsuperscript{Gln}-processing system in human cells

Received for publication, January 30, 2019, and in revised form, April 11, 2019. Published Papers in Press, April 22, 2019.

Xiexie Liu
100049, China
Biology, Chinese Academy of Sciences, Beijing 100101, China, and the **University of Chinese Academy of Sciences, Beijing

Laboratory of Plant Cell and Chromosome Engineering and Center for Genome Editing, Institute of Genetics and Developmental

Edited by Ronald C. Wek

CRISPR/Cas9 nucleases are widely used for genome editing but can induce unwanted off-target mutations. High-fidelity Cas9 variants have been identified; however, they often have reduced activity, constraining their utility, which presents a major challenge for their use in research applications and therapeutics. Here we developed a tRNA\textsuperscript{Gln}-processing system to restore the activity of multiple high-fidelity Cas9 variants in human cells, including SpCas9-HF1, eSpCas9, and xCas9. Specifically, acting on previous observations that small guide RNAs (sgRNAs) harboring an extra A or G (A/G) in the first 5’ nucleotide greatly affect the activity of high-fidelity Cas9 variants and that tRNA–sgRNA fusions improve Cas9 activity, we investigated whether a GN20 sgRNA fused to different tRNAs (G-tRNA-N\textsubscript{20}) could restore the activity of SpCas9 variants in human cells. Using flow cytometry, a T7E1 assay, deep sequencing–based DNA cleavage activity assays, and HEK-293 cells, we observed that a tRNA\textsuperscript{Gln}–sgRNA fusion system enhanced the activity of Cas9 variants, which can be harnessed for efficient correction of a pathogenic mutation in the retinoschisin 1 (RS1) gene, resulting in 6- to 8-fold improved Cas9 activity. We propose that the tRNA-processing system developed here specifically for human cells could facilitate high-fidelity Cas9-mediated human genome-editing applications.

CRISPR/Cas9, a type II system of CRISPR derived from the prokaryotic adaptive immune system, has great potential for genome editing and is under intense investigation at the present time (1, 2). The nuclease activity of \textit{Streptococcus pyogenes} Cas9 (SpCas9), the most widely used nuclease at present, can be triggered by guide RNA targeting imperfectly matched off-target genomic sites. These off-target effects not only confound interpretation of results in the laboratory but also severely undermine the safety and reliability of clinical applications of the technology (3, 4). To address this issue, various strategies and efforts have been employed to minimize off-target activity, such as direct delivery (ribonucleoprotein complex) (5, 6), tunable systems (intein-inactivated Cas9, light-activated and small-molecule induction of Cas9) (7–9), separate Cas9 binding strategies (paired Cas9 nickases) (10), and truncated sgRNA (small guide RNA)\textsuperscript{2} (11). With the delineation and optimization of Cas9 structure, high-fidelity Cas9 variants have been identified, including SpCas9-HF1, eSpCas9(1.1), HypaCas9, evoCas9, xCas9(3.7), Sniper-Cas9, and SpCas9-NG (12–18).

Recent studies revealed that sgRNA transcribed from the U3 or U6 promoter harboring an extra A or G/A/G in the first nucleotide of the sgRNA may affect the activity of high-fidelity Cas9 variants (19, 20). Thus, selection of endogenous A/G in the first nucleotide position of the 20-nt target sequence is potentially useful. Additionally, different strategies have been adopted to address the extra nucleotide. The tRNA\textsuperscript{Gly} from rice, expressed as a fusion with the guide sequence that would be processed by RNaseP and RNaseZ, boosted the activity of SpCas9-HF1 and eSpCas9(1.1) (19). Meanwhile, self-processing Hammerhead (HH) ribozymes, which self-cleave at their 3’ terminus, have also been used in combination with sgRNA to remove the extra G in mammalian cells (20). Here we investigated whether the perfectly matched GN\textsubscript{19} could restore the activity of SpCas9 variants in human cells and sought to identify optimized RNA to achieve high-activity and high-fidelity genome editing mediated by Cas9 variants.

Results

Low relative activity of high-fidelity Cas9 variants

According to previous studies, the activity of high-fidelity Cas9 variants is greatly affected by the presence of an extra 5’ terminal nucleotide (G) in the sgRNA after transcription from

\textsuperscript{2}The abbreviations used are: sgRNA, small guide RNA; HH, Hammerhead; EGF, enhanced GFP; PAM, protospacer adjacent motif; indel, insertion or deletion.

This is an open access article under the CC BY license.

© 2019 He et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.
the U6 promoter. Their performance would be degraded in the presence of the mismatched G, unlike WT SpCas9, which retains its activity (19, 20) (Fig. 1A). Therefore, to confirm whether this is the case, we selected the two most commonly used SpCas9 variants, SpCas9-HF1 (HF1) and eSpCas9(1.1) (eCas9), and used an EGFP reporter cell line (293-SC1), with which it is easy to measure the cleavage activity via flow cytometry (21). We chose 16 target sites in the EGFP gene, grouped by the identity of the first nucleotide (AX19, TX19, CX19, and GX19) and constructed sgRNA expression plasmids of 19 nt (GN19) and 20 nt (GN20) (Fig. S1 and Table S1).

For further comparison, coding sequences of WT, HF1, or eCas9 were inserted into the same backbone of pX330. sgRNAs and Cas9 expression plasmids were co-transfected into 293-SC1 cells (Fig. 1B). The results illustrated that HF1 with GN20 possesses no activity at any of the tested sites and that eCas9 with GN20 shows similar results, with the exception of two sites (A2 and G2), which have an activity of 15% and 32%, respectively (Fig. S1B and Fig. 1C). Moreover, compared with the
Boosting activity of Cas9 variants by tRNA processing

The efficiency of Cas9 variants can be improved by using tRNA processing. For example, GN20, a high-fidelity Cas9 variant, has lower activity than WT Cas9. However, by using a human tRNA processing system, we were able to restore the activity of GN20 to almost the same levels as WT Cas9. This was achieved by processing tRNA with the GN20 variant of Cas9. The results showed that tRNA processing could greatly boost the activity of high-fidelity Cas9 variants, making them more efficient in genome editing.

Effect of a human tRNA-processing system on the activity of xCas9

Recently, xCas9, a high-fidelity Cas9 variant with a flexible PAM, has been identified. We therefore sought to determine whether its activity is the same as WT SpCas9 and whether the effect of a human tRNA processing system would be beneficial. We initially tested xCas9 with EGFP target sites harboring the NGG PAM. The results show that xCas9 has lower activity at all tested sites (Fig. 2F). Similar to the above results, xCas9 with tRNA
tg shows increased activity (Fig. 2G). Specifically, the average efficiency was changed from about 10% (xCas9-GN20) to 21% (xCas9-tRNA
tg-N20) (Fig. 2H). Because it has been reported that xCas9 has flexible PAM selectivity, we tested its activity against 20 EGFP target sites that contain NGT, NGC, NGA, GAN, and NAA PAMs (Fig. S3). Surprisingly, xCas9 showed low activity at all tested sites, even with identical sequences (PAM of TGC, GGA, and GAT) reported in the literature, which claimed higher activity compared with WT Cas9 (Fig. S3B). We could only detect activity at sites with a GGTT PAM (Fig. S3B). We then sought to determine whether tRNA
tg is functional combined with xCas9 in noncanonical PAMs in EGFP. No boosting effects were observed, with the exception of the GGTT site (35% (tRNA
tg-N20) versus 24% (GN20)) (Fig. 2J).

Restoration of the activity of high-fidelity Cas9 variants using a human tRNA-processing system

We reasoned that, because the target cells are of human origin, human tRNA may work better than rice tRNA in improving the activity of high-fidelity Cas9 variants. Also, human tRNA has been utilized to boost the activity of Cpf1 by enhancing the stability of crRNA (CRISPR RNA) for Cpf1 (22). Therefore, we investigated whether human tRNA
tg and tRNA
tg, two of the best-performing tRNAs in Cpf1-mediated genome editing in human cells (22), could be exploited to restore the activity of Cas9 variants. To test this, initially we chose and investigated three sites, sgEGFP-A1 ~A3. The results revealed that tRNA
tg substantially rescued Cas9 activity at these three sites (16% (tRNA
tg-N20) versus 5% (GN20) at site sgEGFP-A1, 26% versus 7% at site sgEGFP-A2, and 30% versus 3% at site sgEGFP-A3). Our data showed that tRNA
tg possesses lower efficiency than tRNA
tg (Fig. 2A and B). We then investigated whether tRNA length would affect its efficiency. We tested different lengths (0 nt, 5 nt, 10 nt, 15 nt, and 20 nt) of the upstream sequence of mature tRNA
tg with HF1 at the sgEGFP-A3 site and found that mature tRNA
tg and 5nt-tRNA
tg had the best performance, about 5% higher than others (Fig. 2A and C). To confirm this, we tested additional target sites with HF1. These results established that different lengths of tRNA (mature tRNA
tg and 5nt-tRNA
tg) were rescued with a comparable efficiency and substantially outperformed tRNA
tg (Fig. S2). As
tion (stop codon) is removed, generating a novel ORF (when 3n bp indels are generated) of \( RS1 \) (Fig. 3A). We chose two sites (sites 1 and 2) that contain an NGG PAM sequence to test whether human tRNAGln can boost the activity of WT Cas9 and high-fidelity variants. The results showed that the tRNA had better performance at site 2, which is reasonable for the TAG stop codon to be cleaved at a distance of 3–4 bp from the PAM (Fig. 3, A–C). We observed 6- to 8-fold improved activity of variants compared with GN20, with a 2-fold increase over WT Cas9 at site 2 (Fig. 3C). We also observed the incremental EGFP signal because of tRNA-mediated restoration at site 2 (Fig. 3B). However, nonhomologous end joining cannot perform accurate repair. Therefore, we hypothesized that oligo template-mediated homologous recombination could increase positive EGFP expression. We designed an 86-nt oligo template and transfected it with Cas9-sgRNA (tRNAGln-N20) (Fig. 3D). As a result, we detected higher EGFP expression of WT (1.6-fold), HF1 (9.3-fold), eCas9 (1.3-fold), and xCas9 (1.5-fold) with templates than with Cas9-sgRNA alone (Fig. 3E).

Fidelity of WT Cas9 and high-fidelity variants with human tRNA

To investigate whether tRNA would increase or decrease the off-target effects of Cas9, we compared WT Cas9 and its variants with tRNAGln. We systematically mutated the guide sequence to introduce single mismatches at positions 1 to 20 at site sgEGFP-A3 (Fig. 4A). As before, we co-transfected matched or mismatched tRNA–sgRNAs with Cas9, and EGFP disruption was analyzed by flow cytometry. We transformed the ratio of EGFP disruption of mismatched tRNA–sgRNAs with Cas9 with tRNA\(^{\text{Gln-N20}}\) grouped by the first nucleotide of the 20-nt target sequence. Figure 2. Rescue of Cas9 variant activity with human tRNA. A, schematic of different tRNAs (rice tRNA\(^{\text{Gly}}\), human tRNA\(^{\text{Gln}}\), and human tRNA\(^{\text{Arg}}\)). B, comparison of three tRNAs at sgEGFP-A1, -A2, and -A3 sites. C, comparison of different-length (0 nt, 5 nt, 10 nt, 15 nt, and 20 nt) tRNAGln upstream sequences tested at the A3 site. D, EGFP disruption by WT SpCas9 and variants with GN\(_{20}\) and tRNA\(^{\text{Gln-N20}}\) grouped by the first nucleotide of the 20-nt target sequence. E, comparison of the WT and variants GN\(_{20}\) and tRNA\(^{\text{Gln-N20}}\) in EGFP disruption. F, EGFP disruption by WT and xCas9 with GN\(_{20}\). G, rescue of xCas9 with tRNA\(^{\text{Gln-N20}}\) grouped by the first nucleotide of the 20-nt target sequence. H, comparison of WT and xCas9, GN\(_{20}\) and tRNA\(^{\text{Gln-N20}}\) in EGFP disruption. I, EGFP disruption of the non-NGG PAM (GGT, TGC, GGA, GAT) with GN\(_{20}\) and tRNA\(^{\text{Gln-N20}}\). Values represent the average ratio of tRNA\(^{\text{Gln-N20}}\)/GN\(_{20}\) at each site. Error bars, S.D.; n = 3; NC, negative control; ns, not significant; *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.0001.
Figure 3. tRNAGln-processing system–mediated genome editing in 293-RS1 cells. A, schematic of 293-RS1 cells and target sites. mRS1, mutant RS1 gene with the stop codon TAG (bold font); PURO, puromycin resistance gene. The target sequence is shown in red, and the PAM is shown in blue. B, images of eCas9-mediated genome editing with GN20 and tRNAGln-N20 at site 2. Scale bars = 100 μm. C, percentage of positive EGFP edited by the WT and variants with GN20 and tRNAGln-N20 at site 1 and site 2. Values represent the average ratio of tRNAGln-N20/GN20. D, schematic of Cas9-sgRNA and oligo template-mediated homologous recombination at site 2 in 293-RS1 cells. E, percentage of positive EGFP edited by WT and variants with tRNAGln-N20 and oligo templates at site 2 in 293-RS1 cells. Error bars, S.D.; n = 3. NC, negative control; *, p < 0.05; **, p < 0.01; ***, p < 0.001.
WT/HF1/eCas9/xCas9 and matched tRNA–sgRNAs with WT/HF1/eCas9/xCas9 to heatmaps (Fig. 4A). The results revealed that HF1 induced the lowest indel levels among three variants, with notably improved specificity at both the 5′ and 3′ termini (Fig. 4A and Fig. S5). The other two variants, eCas9 and xCas9, showed slightly reduced activity at off-target sites only at the 5′ terminus (Fig. 4A). To further characterize their performance at genomic loci, we chose two endogenous sites in EMX1 and VEGFA. As expected, there was a substantial increase in the on-target efficiency of WT, HF1 and eCas9, which was consistent with previous results with the EGFP reporter system. At the EMX1 site, with tRNA^{Gln}, the cleavage efficiency was increased from 20% (without tRNA^{Gln}) to 27% (with tRNA^{Gln}, WT), 0% to 16% (HF1), and 4% to 27% (eCas9) (Fig. 4, B and C). At the VEGFA site, the cleavage efficiency was increased from 20% to 35% (WT), 10% to 41% (HF1), and 6% to 56% (eCas9) (Fig. 4, B and E). In the case of xCas9, with tRNA^{Gln}, relative increased activity or detectable cleavage was observed with EMX1 and VEGFA (Fig. 4, B, C, and E). With tRNA^{Gln}, high-fidelity Cas9 variants had higher activity, and no off-target effects were observed compared with its parentals (Fig. 4, D and F, and Fig. S6). Not surprisingly, we observed notable off-target effects with WT Cas9. To further evaluate the performance of tRNA, we utilized deep sequencing to test on-target sites and off-target sites examined with T7E1. The results revealed a similar outcome that tRNA restored the efficiency of WT, HF1, eCas9, and xCas9. Also, no increasing off-target effects of high-fidelity variants were observed (Fig. 4, G–J). We noticed that the efficiency detected by deep sequencing was almost two times the efficiency observed using T7E1, which is consistent with a previous study showing that T7E1 is not sensitive enough to detect all indels (23) but is still an simple and quick method to test activity and off-target effects.

Discussion

In this study, we demonstrated that the first G nucleotide transcribed from the U6 promoter is extremely important in the activity of Cas9 variants because of their sensitivity to spacer length and mismatch in human cells. Our results also provide direct evidence that there is no apparent regularity of the first nucleotide in the 20-nt target sequence. Curiously, even with a perfectly matched G, there was no benefit for Cas9 or HF1 may be involved in RNA–DNA heteroduplex recognition (12), which showed improved performance compared with tRNAGly from rice, which has been reported to boost Cas9 activity in plants (19). However, this approach is inefficient in human cells and even somewhat attenuated the activity of WT Cas9. Currently, the reasons underlying this observation are not clear, and further studies will dissect the molecular mechanism involved. To rescue the activity of high-fidelity Cas9 variants, we utilized tRNA^{Gln} from rice, which has been reported to boost Cas9 activity in plants (19).

The results illustrated that, in the presence of mature tRNA^{Gln}, Cas9 variant activity was restored to different extents. At EGFP target sites, HF1 only recovered an average of 54% activity relative to WT Cas9 compared with almost complete recovery with eCas9 (Fig. S4D). This is consistent with the results observed in endogenous genes (Figs. 3C and 4, B–J). To account for this, we speculate that some mutated residues of HF1 may be involved in RNA–DNA heteroduplex recognition (14), inducing the diminished catalytic activity of Cas9, or per-
haps there is an interaction between tRNA–sgRNA and RNase that may hinder tRNA processing. Theoretically, after processing, tRNA-N20 of site G1~G4 is perfectly matched with GN19. However, increased Cas9 activity with tRNA-N20 at site G1~G4 has been observed (Fig. 2D). We suspect that tRNA may play another role in structural interactions that enhance Cas9 activity.

Notably, we found that xCas9 is sensitive to the extra G but not to the same extent as HF1. With tRNA processing, xCas9 only recovered about 62% activity of WT Cas9 at NGG PAM sites in EGFP, moderately higher than HF1 (Fig. S4D). The activity of xCas9 with tRNA was maintained at endogenous sites compared with the other two variants. For non-NGG PAM sites, we found that xCas9 possesses rather low activity with sgRNA-GN19, even at the same sites (PAM of TGC, GGA, PAM sites, we found that xCas9 possesses rather low activity compared with the other two variants. For non-NGG high-fidelity variants.

An applicable approach to genome editing with CRISPR/Cas9 processing system enabled development of a promising and applicable approach to additional Cas9 variants, such as evoCas9, variants in human cells, which broadens the editing regions and subjects harbored insertions or deletions (indels) were amplified by PCR using the primer sets listed in Table S2. For the T7E1 assay, 300 ng of purified PCR products was mixed with 1 μl of T7E1 and subjected to a reassambling process to enable heteroduplex formation (24). After reannealing, products were treated with T7 Endonuclease I (New England Biolabs). Quantification was based on relative band intensities. The indel percentage was determined by the formula 100 × (1 − sqrt (b + c)/(a + b + c)), where a is the integrated intensity of the undigested PCR product and b and c are the integrated intensities of the cleavage product. Primers and target sequence are summarized in Table S3.

**Experimental procedures**

**Plasmid constructions**

Plasmids pX330 (harboring the WT SpCas9 coding sequence), eSpCas9 (1.1), SpCas9-HF1, and xCas9(3.7)-ABE(7.10) were obtained from Addgene (plasmids 42230, 71814, 72247, and 108382, respectively). To make them comparable, we generated all constructs with the same backbone. Specifically, the coding sequences for SpCas9 from SpCas9-HF1 and xCas9(3.7)-ABE(7.10) were amplified and inserted into the backbone of pX330 (without the SpCas9 coding sequence). The coding sequence of SpCas9 from xCas9(3.7)-ABE(7.10) was in the dead form, which was mutated to the active form with site-directed mutagenesis for this study. sgRNA oligos were annealed and inserted into the backbone of pX330 using a standard protocol. DNA sequencing confirmed the desired specific sequences in the constructs. The oligonucleotide, target, and primer sequences in the present study are summarized in Tables S1–S15.

**Cell culture and cell transfection**

HEK-293 cells (catalog no. CRL-1573) expressing EGFP were generated by lentiviral transduction as described previously (21). Drug-resistant single colonies of transduced HEK-293 cells were isolated and named 293-SC1. To maintain EGFP expression, the medium for 293-SC1 culture included puromycin. Cells were cultured in advanced DMEM (Life Technologies) supplemented with 10% FBS and 1% penicillin–streptomycin at 37 °C with 5% CO2. 1.8 × 10^5 293-SC1 cells were seeded per well of a 12-well plate on day 1. The 293-SC1 cells were transfected with 750 ng of plasmids (250 ng of sgRNA and 500 ng of Cas9 expression plasmids) using 3.0 μl of TurboFect (Thermo Fisher Scientific) on day 2. Fresh medium was added to the transfected 293-SC1 cells on day 3. The cells were harvested for flow cytometry, T7 Endonuclease I (T7E1) assay, or deep sequencing on day 4.

**DNA cleavage activity assay**

The flow cytometry experiments were performed as described previously (21). The background of EGFP disruption was gated at in a range of 1% to 5% for all experiments. Fragments harboring insertions or deletions (indels) were amplified by PCR using the primer sets listed in Table S2. For the T7E1 assay, 300 ng of purified PCR products was mixed with 1 μl of NEBuffer 2 and ultrapure water to a final volume of 14.5 μl and subjected to a reannealing process to enable heteroduplex formation (24). After reannealing, products were treated with T7 Endonuclease I (New England Biolabs). Quantification was based on relative band intensities. The indel percentage was determined by the formula 100 × (1 − sqrt (b + c)/(a + b + c)), where a is the integrated intensity of the undigested PCR product and b and c are the integrated intensities of the cleavage product. Primers and target sequence are summarized in Table S3.

**Targeted deep sequencing assay**

Genomic regions of interest were amplified by PCR with flanking high-throughput sequencing (HTS) primer pairs that are listed in Table S15. PCR reactions were carried out under the following conditions: 95 °C for 5 min, then 34 cycles of 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 15 s, followed by a final 72 °C extension for 5 min. PCR products were verified on a 2% agarose gel and purified using a DNA extraction kit (Vazyme). Purified products were sequenced on an Illumina HiSeq X Ten instrument at Novogene (Tianjin, China). Sequencing data were obtained using standard protocols.

**Statistics**

All data replicated three times were expressed as mean ± S.D. Differences between two groups were determined by two-tailed Student’s t test or Mann–Whitney test. The criteria for statistical significance were as follows: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.


**Acknowledgments**—We thank Xianglian Ge, Xiujuan Lv, Mengjun Tu, Fanfan Li, Huihui Sun, and Jie Liu for technical assistance.
Boosting activity of Cas9 variants by tRNA processing

References


