# Expansion of the Prime editing Modality with Cas9 from *Francisella novicida*

**Seung Hwan Lee** ( lsh080390@kribb.re.kr )
KRIBB: Korea Research Institute of Bioscience and Biotechnology https://orcid.org/0000-0003-0856-8023

**Yeounsun Oh**
KRIBB: Korea Research Institute of Bioscience and Biotechnology https://orcid.org/0000-0001-9042-9326

**Wi-jae Lee**
KRIBB: Korea Research Institute of Bioscience and Biotechnology https://orcid.org/0000-0002-8231-8658

**Hanseop Kim**
KRIBB: Korea Research Institute of Bioscience and Biotechnology https://orcid.org/0000-0001-6186-7570

**Lee Wha Gwon**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Young-Hyun Kim**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Young-Ho Park**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Chan Hyoung Kim**
KRIBB: Korea Research Institute of Bioscience and Biotechnology https://orcid.org/0000-0001-8230-9164

**Kyung-Seob Lim**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Bong-Seok Song**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Jae-Won Huh**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Sun-Uk Kim**
KRIBB: Korea Research Institute of Bioscience and Biotechnology

**Bong-Hyun Jun**
Konkuk University - Seoul Campus: Konkuk University

**Cheulhee Jung** ( damo363@korea.ac.kr )
Korea University - Seoul Campus: Korea University https://orcid.org/0000-0001-7392-5862

**Research Article**

**Keywords:** Prime editing, Target expansion, CRISPR-Cas9, Ortholog, Francisella novicida

**Posted Date:** June 21st, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-628506/v1

# Abstract

Prime editing can induce a desired base substitution, insertion, or deletion in a target gene using reverse transcriptase (RT) after nick formation by CRISPR nickase. In this study, we developed a technology that can be used to insert or replace external bases in the target DNA sequence by linking reverse transcriptase to the Francisella novicida Cas9, which is a CRISPR-Cas9 ortholog. Using FnCas9(H969A) nickase, the targeting limitation of existing Streptococcus pyogenes Cas9 nickase [SpCas9(H840A)]-based prime editing was dramatically extended, and accurate prime editing was induced specifically for the target genes in human cell line.


# Background

The recent development of reverse transcriptase (RT)-based target DNA editing technology (i.e., prime editing) is based on the SpCas9(H840A) module [1-3]. Since prime editing technology can induce various types of mutations [2, 4-11] compared to base editing [12, 13] that only induces deamination-mediated base substitutions (A to G or C to T), most of the pathogenic mutations reported to date will be corrected without double-stranded DNA cleavage [2]. However, since SpCas9 module-based prime editing technology can insert, delete, and replace bases within the upstream of 3 bp from the protospacer adjacent motif (PAM) sequence (NGG) on the target DNA, it shows a PAM limitation [14, 15]. In order to apply prime editing to various biological systems, it is important to improve the fundamental limitations generated by the CRISPR module. To this end, we tried to accurately induce the mutation in target sequence using a different type of CRISPR-Cas module and extended the function of the existing prime editor. Compared to SpCas9 module [16], FnCas9 (*F. novicida*)[17-20] shows an overhang pattern of double strand break for the protospacer sequence. Therefore, the FnCas9(H969A) nickase module, which shows a different nicking property on the non-target strand of protospacer, has the advantage of expanding the region recognized as the reverse transcription template (RTT) following the primer binding site (PBS) sequence for prime editing. In this study, we improved the effectiveness limitations of SpCas9(H840A) based prime editor with a new approach using *F. novicida* Cas9, a CRISPR-Cas9 ortholog[19], in human-derived cell lines. Using FnCas9(H969A) linked RT (i.e., FnCas9(H969A)-RT, FnCas9 prime editor) developed in this study, precise genome editing in a range that was not previously applied with SpCas9 is possible. As the number of regions that can be precisely edited using this technology increases, it is expected that FnCas9(H969A)-RT can be applied to various biological systems.

# Results And Discussion

For the application of prime editing with a new ortholog, the characteristics of FnCas9 were compared and analyzed with that of generally used SpCas9. First, wild-type FnCas9 and SpCas9 were purified in the form of recombinant protein, and cleavage experiments were performed on PCR amplicons including various target nucleotide sequences (*c-Myc, NRAS, AAVS1, EMX1, HEK3*) in vitro **(Additional file 1: Figure S1)**. At this time, FnCas9 and SpCas9 recognized the same PAM (NGG) for each target nucleotide

sequence, and directly compared the cleavage results using a protospacer of the same length. In both FnCas9 and SpCas9, it was confirmed that cleavage occurred at 3bp in front of PAM (NGG) in common for TS (target strand) **(Additional file 1: Figure S1)**. On the other hand, for NTS (non-target strand), FnCas9 cleavage occurred at 6-8 bp upstream of PAM (NGG) and SpCas9 cleavage occurred at 3-4 bp upstream PAM (NGG), respectively. Based on the cleavage property, we were sought to test the target-specific prime editing with FnCas9(H969A)-RT for various target sites **(Additional file 1: Figure S2)** in human-derived cell lines **(Fig. 1a, b)**. First, an RT enzyme was connected to the FnCas9(H969A) nickase module to optimize the performance of the prime editor on the target gene. To optimize the efficiency of the newly applied FnCas9 prime editor, FnCas9(H969A)-RT or prime editing guide RNA (pegRNA) with various linker lengths were prepared **(Fig. 1c,d top, Additional file 1: Table S1)**. The pegRNA was constructed to allow for insertion of a "TT" di-nucleotide sequence at the expected cleavage point [6 and 3 nucleotides upstream from the PAM (NGG) sequence for Fn and Sp, respectively] for the endogenous locus by considering the nicking point on each target sequence based on SpCas9 or FnCas9 nickase modules. The HEK293FT cell line was then co-transfected with pegRNA/FnCas9(H969A)-RT and nicking-guide RNA (ngRNA) expression plasmids **(Additional file 1: Figure S3)**. We found that the bi-nucleotide TT was accurately inserted in the genomic DNA (*c-Myc*, *NRAS*) at a site of 6 nucleotides upstream from the PAM in the target sequence when using FnCas9(H969A)-RT compared to SpCas9(H840A)-RT as expected **(Fig. 1b)**. During the optimization, pegRNA without a linker showed the highest efficiency in *HEK3* locus **(Fig. 1c)**, and optimized TT insertion was achieved when using the FnCas9 prime editor with a 1× linker length for *HEK3* and *NRAS* locus **(Fig. 1d)**.

To confirm the effect of the length of the PBS and RTT in pegRNA on prime editing efficiency, several candidate pegRNAs with different lengths of PBS and RTT were designed and produced **(Additional file 1: Table S1)**, and a TT insertion was tested at various genes (*HEK3, c-Myc, NRAS*) in human cell lines **(Fig. 1e, Additional file 1: Figure S4a, c, e)**. The external TT insertion efficiency at the *HEK3* site increased as the length of the PBS increased **(Fig. 1e)**. FnCas9(H969A)-RT resulted in 1.65% (*HEK3*), 1.69% (*c-Myc*), and 65.77% (*NRAS*) of precise TT insertion [6 bp upstream from PAM (NGG)] compared to SpCas9(H840A)-RT on average **(Fig. 1e, Additional file 1: Figure S4b, d, f)**, and the efficiency slightly varied according to the length of RTT. Notably, the TT insertion efficiency was significantly different for each targeted gene (*HEK3, EMX1, AAVS1*) according to the location of the nicking guide RNA **(Fig. 1f, Additional file 1: Table S1)** on the target-strand side when applying the prime editing with PE3 manner **(Additional file 1: Figure S3)**. To directly compare the efficiency of FnCas9(H969A)-RT with that of SpCas9(H840A)-RT for various genes, TT insertion experiment was performed in parallel for the same target sequence in various genes (*c-Myc*, *NRAS*) by using both prime editing systems **(Additional file 1: Figure S5)**. Next-generation sequencing **(Additional file 1: Table S3)** showed that compared to SpCas9-RT (29.34-95.18%), FnCas9-RT exhibited greater variation in TT insertion efficiency (3.82-87.56%), which depends on the target sequences **(Additional file 1: Figure S5a)**. These results may reflect the weaker nicking property of FnCas9(H969A) compared to SpCas9(H840A) which is shown in double nicking experiment **(Additional file 1: Figure S6)**. As reported in many previous papers[2, 7, 9], although TT insertion was accurately induced for a given target sequence, unintended indel formation also caused by

the PE3 method occurred at a frequency of 4.79% (*NRAS*) and 1.13% (*c-Myc*) when using SpCas9(H840A)-RT, and at a frequency of 1.32% (*NRAS*) and 1.84% (*c-Myc*) in the case of FnCas9(H969A)-RT **(Additional file 1: Figure S5a)**. Next, based on the sensitivity to the target sequence of FnCas9[17], it was predicted that the off-target editing effect would be small, the editing efficiency (%) was analyzed for off-target candidates which predicted from in-silico analysis **(Additional file 1: Table S4)**. Indeed, when analyzing the off-target effect of the FnCas9-based prime editor for each target gene, we found that no significant off-target mutation was observed by FnCas9(H969A)-RT **(Additional file 1: Figure S5b, c)**.

Because the FnCas9 nickase (H969A) module induces far distance of nicking point from PAM (NGG) on the non-target strand side of the target gene in contrast to the SpCas9 nickase (H840A) module **(Additional file 1: Figure S1)**, reverse transcription can be initiated by RT from a position farther away from the nucleotide sequence recognized as the PAM (NGG). According to this difference, we directly compared the range that can induce prime editing based on FnCas9-RT and SpCas9-RT with respect to the same target sequence **(Fig. 2)**. In HEK293FT cells, the efficiency of prime editing was compared for the same target sequence in the *c-Myc* and *EMX1* genes. The point at which the base is inserted (AA for *c-Myc*, TT for *EMX1*) was indicated according to the point at which nicks were formed (from -6 to +3) by the both nickase module **(Fig. 2a, Additional file 1: Figure S7)**. SpCas9 nickase(H840A)-RT showed an editing efficiency of 27–39% for the AA insertion at *c-Myc* from the −1 to +3 position in the sequence based on the nick position **(Fig. 2b,c, top)**. By contrast, FnCas9 nickase(H969A)-RT showed a 2–8% insertion efficiency from position −6 to +3 based on the nick position **(Fig. 2b,c, bottom)**. These results showed that based on the shared PAM sequence (NGG) in the same target, FnCas9(H969A)-RT greatly expanded the prime editing range. In particular, FnCas9(H969A)-RT can effectively induce prime editing of various point mutations or insertions in regions (-3 to -2 position) that cannot be edited using the SpCas9 system **(Fig. 3)**. It was found that various base substitutions and insertions were possible from -3 to +3 positions for each of *c-Myc* **(Fig. 2d, f)** and *NRAS* **(Fig. 2e, g)** genes by using FnCas9(H969A)-RT. Similar results were found for the TT insertion efficiency for the same target sequence in the *EMX1* gene **(Additional file 1: Figure S7)**, in which the range of genome editing by FnCas9(H969A)-RT was significantly expanded compared to that of SpCas9(H840A)-RT editing.

Finally, we examined whether the optimized FnCas9(H969A)-RT can be used in combination with SpCas9(H840A)-RT to induce prime editing simultaneously on the target gene **(Fig. 2h, i)**. The TT insertion was respectively induced by FnCas9(H969A)-RT and SpCas9(H840A)-RT in two adjacent nucleotide sequences on the *NRAS* gene **(Fig. 2i)**. To increase the efficiency of simultaneous editing of the two sites, nickases on the target-strand side were treated at various positions (**Additional file 1: Figure S2c**), resulting in different editing efficiencies **(Fig. 2h)**. In samples treated with both prime editors simultaneously, the TT insertion was generally induced with high efficiency (25.81–47.06%) by SpCas9(H840A)-RT, whereas that induced by FnCas9(H969A)-RT showed more varied efficiency (0.12–44%) according to the nickase position on the target-strand side. The overall rate of simultaneous TT insertion by the two prime editors was relatively lower (0.02–1.3%) than single TT insertion.

# Conclusions

In this study, we confirmed that the FnCas9 nickase formed a nick 6-8 bp from the PAM on the non-target strand side, and extended the editable region by prime editor. This effect could be also applicable to other endonucleases that induce the nick formation in different fashion [21], and in particular, the farther the nick is generated upstream from the PAM, the wider the range of target expansion will be. In addition to this, the FnCas9(H969A)-RT based editable region could be further extended by engineering the PAM recognition domain within the FnCas9 nickase as previous study [20]. In summary, we developed a new prime editing module that can dramatically expand the capabilities of SpCas9-based prime editing technology. Compared to the SpCas9(H840A)-RT prime editor, FnCas9(H969A)-RT shows precise, expanded range of prime editing and versatile editing profiles. In addition, prime editing can be achieved with both SpCas9 and FnCas9 modules simultaneously. These type of prime editor with expanded target range and diversity is expected to have a big ripple effect on the bio/medical field by enabling target-specific and various types of gene editing in the future. In particular, FnCas9(H969A)-RT is possibly applied to most of the existing human disease-causing mutations, and it is thought that it will increase the possibility of application as a therapeutic agent by effectively restoring the original form of information at the DNA level.

# Methods

### Purification of SpCas9 and FnCas9 proteins

Subcloned pET28a-SpCas9 and pET28a-FnCas9 bacterial expression vectors were transformed into *Escherichia coli* BL21(DE3) cells, respectively, and colonies were grown at 37°C. After growing to an optical density of 0.6 in a 500 ml culture flask, induction was performed at 18°C for 48 h by isopropyl β-d-1-thiogalactopyranoside treatment. The cells were then precipitated by centrifugation and resuspended in lysis buffer [20 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM β-mercaptoethanol, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride(PMSF)]. The cells were disrupted by sonication on ice for 3 min, and the cell lysates were separated by centrifugation at 20,000 ×*g* for 10 min. The further purification process was carried out as previously described [22]. Each purified SpCas9 and FnCas9 protein was replaced with a Centricon filter (Amicon Ultra) as a storage buffer [200 mM NaCl, 50 mM HEPES (pH 7.5), 1 mM dithiothreitol (DTT), and 40% glycerol] for long-term storage at −80°C. The purity of the purified proteins was confirmed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (8–10%), and protein activity was tested using an *in vitro* polymerase chain reaction (PCR) amplicon cleavage assay.

### *In vitro* transcription for guide RNA synthesis

For *in vitro* transcription, DNA oligos containing an sgRNA sequence **(Additional file 1: Table S2)** corresponding to each target sequence were purchased from Cosmo Genetech. After extension PCR (denaturation at 98°C for 30 s, primer annealing at 62°C for 10 s, and elongation at 72°C for 10 s, 35 cycles) using the DNA oligos, DNA was purified using GENECLEAN® Turbo Kit (MP Biomedicals). The

purified template DNA was mixed with an *in vitro* transcription mixture [T7 RNA polymerase (NEB), 50 mM $MgCl_2$, 100 mM rNTPs (rATP, rGTP, rUTP, rCTP), 10× RNA polymerase reaction buffer, murine RNase inhibitor (NEB),100 mM DTT, and DEPC) and incubated at 37°C for 8 h. The DNA template was completely removed by incubation with DNase I (NEB) at 37°C for 1 h, and the RNA was further purified using GENECLEAN® Turbo Kit (MP Biomedicals) for later use. The purified RNA was concentrated by lyophilization (20,000 ×*g* at −55 °C for 1 h) and stored at −80°C.

### *In vitro* cleavage assay and DNA sequencing for nicking point analysis

Using the purified FnCas9 or SpCas9 recombinant protein, an *in vitro* cleavage assay was performed to determine the location of nicks in the non-target strand of the DNA sequence to be prime-edited. Each target site was cloned into a T-vector, and then T-vectors were purified and incubated in cleavage buffer (NEB3, 10 μl) with FnCas9 at 37°C for 1 h. The cleavage reaction was stopped by adding a stop buffer (100 mM Ethylenediaminetetraacetic (EDTA) acid, 1.2% SDS), and only the plasmid DNA was separated through the column (Qiagen, QIAquick® PCR purification Kit). The cleavage point of FnCas9 or SpCas9 was confirmed by run-off sequencing analysis of the cut fragments. The last nucleotide sequence was confirmed by the A-tailing of polymerase and the cleavage pattern was analyzed by comparative analysis with reference sequences.

### Design and cloning of the prime editor and pegRNA expression vectors

The cytomegalovirus promoter-based SpCas9(H840A)-RT and FnCas9(H969A)-RT expression vectors were constructed to induce genome editing in human cell lines. To optimize the efficiency of the newly developed FnCas9 prime editor, FnCas9 prime editors with various linker lengths were prepared or the linker length of pegRNA was diversified **(Additional file 1: Table S1)**. In addition, pegRNA containing the corresponding nucleotide sequence (TT or AA) was prepared so that the nucleotide could be inserted into the target nucleotide sequence (*HEK3, AAVS1, c-Myc, NRAS, EMX1*). The PBS and RTT regions in pegRNA were designed and produced according to the non-target strand nick site generated by the FnCas9 nickase module **(Additional file 1: Figure S1)**. As a positive control, the pegRNA of SpCas9-RT was constructed so that a TT base was inserted 3 bp in front of the PAM in consideration of the nicking point. By contrast, in FnCas9-RT, pegRNA was constructed so that the TT base was inserted at the expected cleavage point (6 bp upstream from the PAM). The pegRNA expression vector is driven by the *U6* promoter, and was designed and manufactured so that only the protospacer, PBS, and RTT can be replaced with restriction enzymes according to the target nucleotide sequence.

### Cell culture and transfection

Human-derived cell lines (HEK293FT and HeLa) were purchased from Invitrogen (R70007) and American Type Culture Collection (CCL-2), respectively. The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (both from Gibco) at 37°C in the presence of 5% $CO_2$. Cells were subcultured every 48 h to maintain 70% confluency. For target sequence editing, $2 \times 10^5$ HEK293FT or HeLa cells were transfected with plasmids expressing pegRNAs (240 pmol), the prime editor

expression plasmid [SpCas9(H840A)-RT (Addgene no. 132775), FnCas9(H969A)-RT (developed in this study)], and nicking sgRNA (60 pmol) via electroporation using an Amaxa electroporation kit (V4XC-2032; program: CM-130 for HEK293FT cells, CN-114 for HeLa cells). In parallel with the SpCas9 prime editor, the FnCas9 prime editor was transfected targeting the same nucleotide sequence in HEK293FT cells. Transfected cells were transferred to a 24-well plate containing DMEM (500 μl/well), pre-incubated at 37°C in the presence of 5% $CO_2$ for 30 min, and incubated under the same conditions for subculture.

### Purification of genomic DNA and construction of target site amplicons

Genomic DNA (gDNA) was extracted from the cultured cells 72 h after genome editing. The gDNA was isolated using DNeasy Blood and Tissue Kit (Qiagen). Target amplicons were obtained through PCR (denaturation at 98°C for 30 s, primer annealing at 58°C for 30 s, elongation at 72°C for 30 s, 35 cycles**; Additional file 1: Table S3**) from gDNA extracted from cells treated with each SpCas9 or FnCas9 prime editor. Targeted amplicon next-generation sequencing was performed using nested PCR (denaturation: 98°C for 30 s, primer annealing: 58°C for 30 s, elongation: 72°C for 30 s, 35 cycles) to analyze the efficiency of base (TT or AA) insertion.

### Targeted amplicon sequencing and data analysis

To prepare the targeted amplicon library, gDNA was extracted from the cells and further amplified using DNA primers **(Additional file 1: Table S3)**. Nested PCR (denaturation at 98°C for 30 s, primer annealing at 62°C for 15 s, and elongation at 72°C for 15 s, 35 cycles) was performed to conjugate adapter and index sequences to the amplicons. All targeted amplicon sequencing and data analysis were performed as suggested in a previous study [22].

# Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

Targeted deep sequencing data are available at NCBI Sequence Read Archive (SRA) under accession number PRJNA736246.

### Competing interests

The authors declare no competing interests.

### Funding

This research was supported by grants from the National Research Foundation funded by the Korean Ministry of Education, Science and Technology (NRF-2019R1C1C1006603); the Technology Innovation Program funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) (20009707); the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 9991006929, KMDF_PR_20200901_ 0264); and KRIBB Research Initiative Program (KGM1051911, KGM4252122, KGM5382113, KGM4562121, KGM5282113).

### Authors' contributions

Conceptualization: Y.O., W.L., and S.H.L.; Methodology, Y.O., W.L., H.K., L.W.G., and S.H.L.; Software, Y.H.K., Y.H.P., and S.H.L.; Validation, Y.O. and C.H.K.; Formal Analysis, Y.O., W.L. and S.H.L.; Investigation, Y.O., W.L., K.S.L., S.B.S., J.W.H., S.U.K., and S.H.L.; Resources, Y.H.K., Y.H.P., K.S.L., and B.H.J.; Data Curation, Y.O., W.L., and S.H.L.; Writing-Original Draft, Y.O., W.L., C.J., and S.H.L.; Writing-Original Draft, Y.O., W.L., C.J., and S.H.L.; Visualization, Y.O., W.L., and S.H.L.; Supervision, C.J. and S.H.L.

### Acknowledgements

Authors thank to members of the National Primate Research Center (NPRC) for helpful discussions.

# References

1. Anzalone AV, Koblan LW, Liu DR: **Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors.** *Nat Biotechnol* 2020, **38:**824-844.

2. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR: **Search-and-replace genome editing without double-strand breaks or donor DNA.** *Nature* 2019, **576:**149-157.

3. Marzec M, Braszewska-Zalewska A, Hensel G: **Prime Editing: A New Way for Genome Editing.** *Trends Cell Biol* 2020, **30:**257-259.

4. Chemello F, Chai AC, Li H, Rodriguez-Caycedo C, Sanchez-Ortiz E, Atmanli A, Mireault AA, Liu N, Bassel-Duby R, Olson EN: **Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing.** *Sci Adv* 2021, **7**.

5. Gao P, Lyu Q, Ghanam AR, Lazzarotto CR, Newby GA, Zhang W, Choi M, Slivano OJ, Holden K, Walker JA, 2nd, et al: **Prime editing in mice reveals the essentiality of a single base in driving tissue-specific gene expression.** *Genome Biol* 2021, **22:**83.

6. Lin Q, Zong Y, Xue C, Wang S, Jin S, Zhu Z, Wang Y, Anzalone AV, Raguram A, Doman JL, et al: **Prime genome editing in rice and wheat.** *Nat Biotechnol* 2020, **38:**582-585.

7. Liu P, Liang SQ, Zheng C, Mintzer E, Zhao YG, Ponnienselvan K, Mir A, Sontheimer EJ, Gao G, Flotte TR, et al: **Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice.** *Nat Commun* 2021, **12:**2121.

8. Petri K, Zhang W, Ma J, Schmidts A, Lee H, Horng JE, Kim DY, Kurt IC, Clement K, Hsu JY, et al: **CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells.** *Nat Biotechnol* 2021.

9. Bosch JA, Birchak G, Perrimon N: **Precise genome engineering in Drosophila using prime editing.** *Proc Natl Acad Sci U S A* 2021, **118**.

10. Xu R, Li J, Liu X, Shan T, Qin R, Wei P: **Development of Plant Prime-Editing Systems for Precise Genome Editing.** *Plant Commun* 2020, **1:**100043.

11. Kim Y, Hong SA, Yu J, Eom J, Jang K, Yoon S, Hong DH, Seo D, Lee SN, Woo JS, et al: **Adenine base editing and prime editing of chemically derived hepatic progenitors rescue genetic liver disease.** *Cell Stem Cell* 2021.

12. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR: **Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.** *Nature* 2017, **551:**464-471.

13. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR: **Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.** *Nature* 2016, **533:**420-424.

14. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, Noda T, Abudayyeh OO, Gootenberg JS, Mori H, et al: **Engineered CRISPR-Cas9 nuclease with expanded targeting space.** *Science* 2018, **361:**1259-1262.

15. Walton RT, Christie KA, Whittaker MN, Kleinstiver BP: **Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants.** *Science* 2020, **368:**290-296.

16. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity.** *Science* 2012, **337:**816-821.

17. Acharya S, Mishra A, Paul D, Ansari AH, Azhar M, Kumar M, Rauthan R, Sharma N, Aich M, Sinha D, et al: **Francisella novicida Cas9 interrogates genomic DNA with very high specificity and can be used for mammalian genome editing.** *Proc Natl Acad Sci U S A* 2019, **116:**20959-20968.

18. Chen F, Ding X, Feng Y, Seebeck T, Jiang Y, Davis GD: **Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting.** *Nat Commun* 2017, **8:**14958.

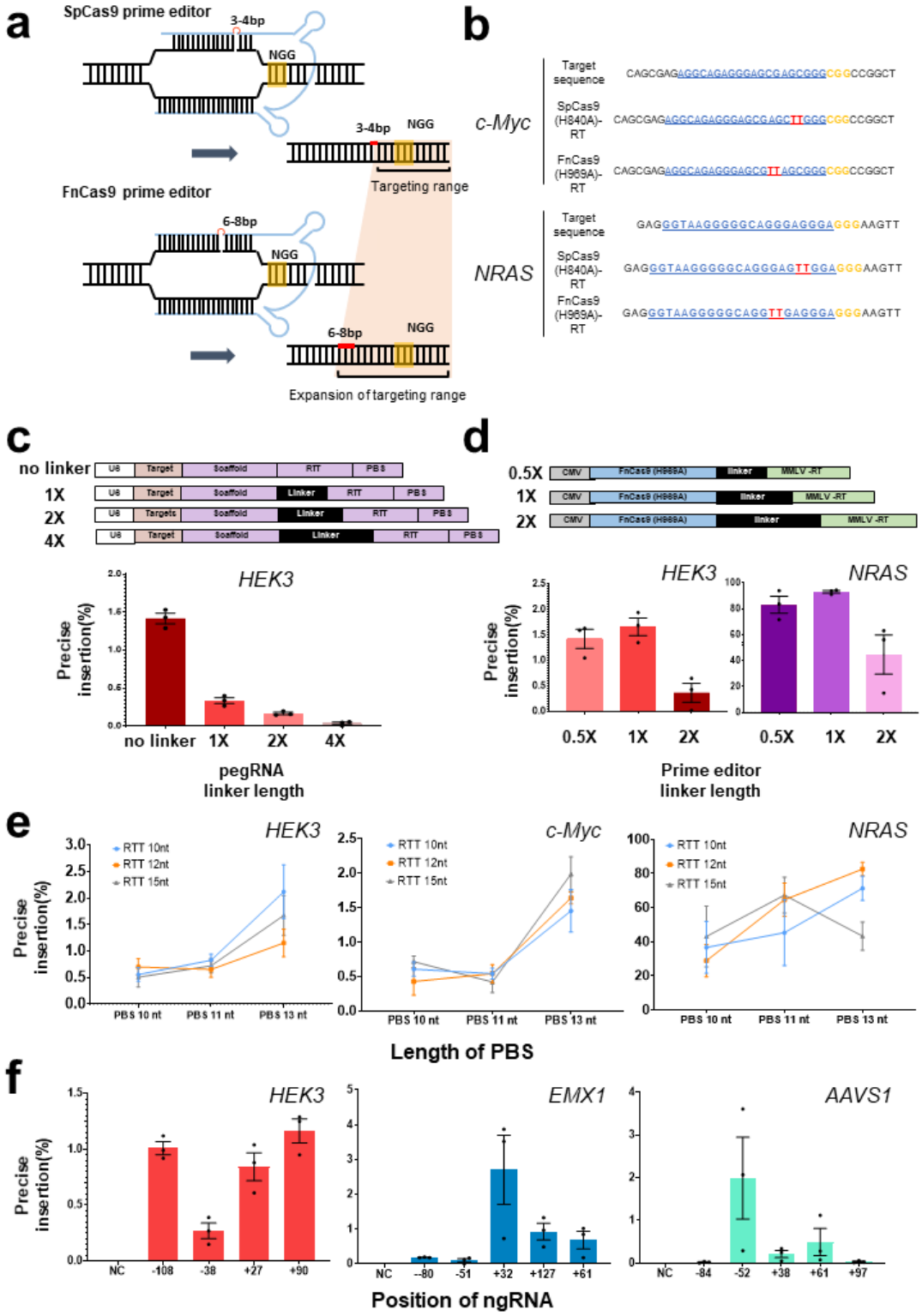
19. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lecrivain AL, Bzdrenga J, Koonin EV, Charpentier E: **Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems.** *Nucleic Acids Res* 2014, **42:**2577-2590.

20. Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, Hsu PD, Nakane T, Ishitani R, Hatada I, Zhang F, et al: **Structure and Engineering of Francisella novicida Cas9.** *Cell* 2016, **164:**950-961.

21. Gasiunas G, Young JK, Karvelis T, Kazlauskas D, Urbaitis T, Jasnauskaite M, Grusyte MM, Paulraj S, Wang PH, Hou Z, et al: **A catalogue of biochemically diverse CRISPR-Cas9 orthologs.** *Nat Commun* 2020, **11:**5512.

22. Kang SH, Lee WJ, An JH, Lee JH, Kim YH, Kim H, Oh Y, Park YH, Jin YB, Jun BH, et al: **Prediction-based highly sensitive CRISPR off-target validation using target-specific DNA enrichment.** *Nat Commun* 2020, **11:**3596.

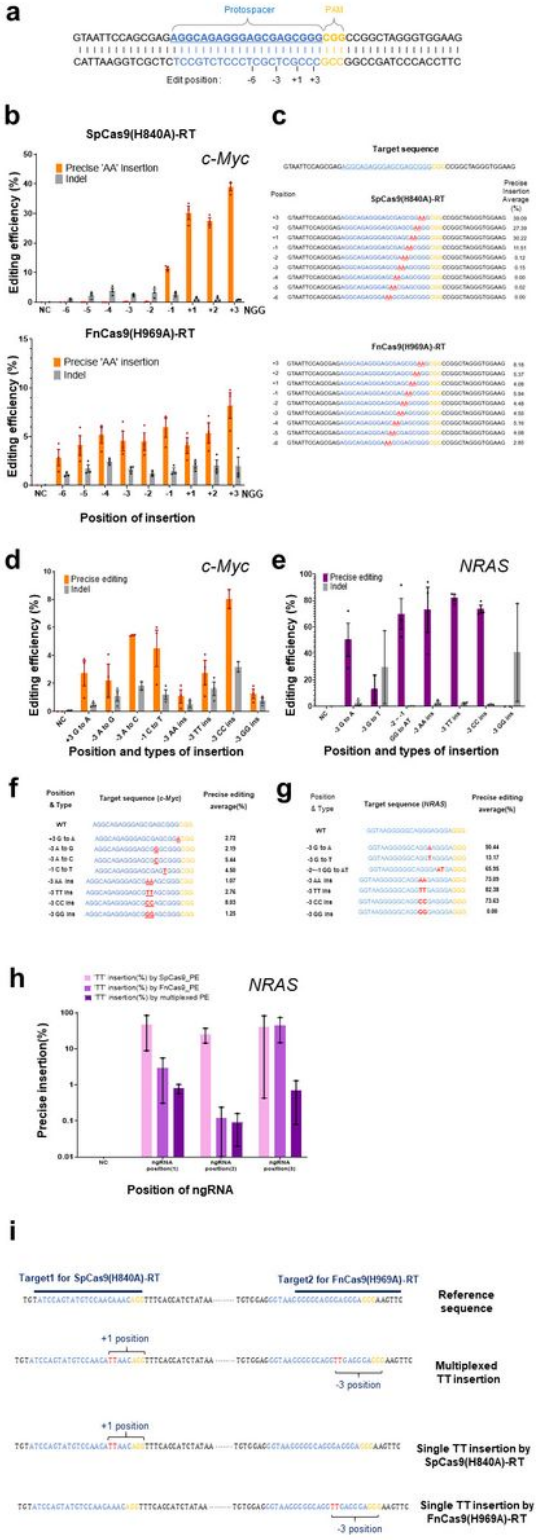
# Figures




## Figure 1

Targeted prime editing and optimization with FnCas9(H969A)-RT. a. Comparison between SpCas9(H840A) and FnCas9(H969A) nickase-based prime editing. SpCas9(H840A) forms a nick 3–4 bp upstream from the PAM (NGG) and FnCas9 (H969A) forms a nick 6–8 bp upstream from the PAM (NGG). Only pegRNAs are shown in blue for clarity, each PAM (NGG) sequence is shown in yellow, and the targeted base insertion is shown in red. b. Comparison of target-specific nucleotide insertion into the c-

Myc and NRAS genes according to the different nicking positions of SpCas9(H840A)-RT and FnCas9(H969A)-RT. PAM (NGG) sequences are shown in yellow, protospacers are shown in blue, and inserted base sequences are shown in red. c. Preparation of various pegRNAs for optimizing the prime editing efficiency and comparison of nucleotide insertion efficiency according to linker length (no linker, 1×, 2×, 4× linkers) for HEK3. d. Construction of FnCas9(H969A)-RT with linkers of various lengths (0.5×, 1×, 2×) and comparison of nucleotide insertion efficiency for HEK3 and NRAS. e. Comparison of gene editing efficiency of FnCas9(H969A)-RT according to the PBS and RTT length in pegRNA on various genes (HEK3, c-Myc, NRAS). f. Comparison of base insertion efficiency according to ngRNA targeting at various positions (HEK3, c-Myc, NRAS) with the PE3 method of FnCas9 (H969A)-RT. Each histogram was plotted by applying standard error of the mean values to repeated experimental values (n = 3). PBS: primer binding site; RTT: reverse transcription template; pegRNA: prime editing guideRNA; ngRNA: nicking guideRNA; MMLV-RT: reverse transcriptase domain from Moloney Murine Leukemia Virus.




## Figure 2

Expansion of the range of prime editing by FnCas9(H969A)-RT. a. Schematics of the protospacer (blue), PAM (yellow), and base insertion site (indicated numbers below the target sequence) on the target gene (c-Myc) recognized by SpCas9(H840A)-RT and FnCas9(H969A)-RT. b. Direct comparison of the efficiency of base insertion by SpCas9(H840A)-RT and FnCas9(H969A)-RT. Top: Comparison of efficiency according to the base insertion site caused by SpCas9(H840A)-RT. Bottom: Comparison of efficiency according to

the base insertion site caused by FnCas9(H969A)-RT. c. Next-generation sequencing result of site-specific base insertion induced by SpCas9(H840A)-RT and FnCas9(H969A)-RT. Protospacers are shown in blue, PAM (NGG) sequences are shown in yellow, and inserted sequences (AA) are shown in red. Each position and efficiency in which the AA bases are inserted are indicated to the left and right of the target sequence, respectively. Each histogram was plotted by applying standard error of the mean values to repeated experimental values (n = 3). d, e, Next-generation sequencing result of site-specific (c-Myc, NRAS) base insertion induced by FnCas9(H969A)-RT. The editing efficiency (%) according to each position and type of prime editing is plotted in the histogram. Each histogram was plotted by applying standard error of the mean values to repeated experimental values (n = 3). f, g, NGS results and average values of prime editing using FnCas9 obtained from (d, e). Protospacers are shown in blue, PAM (NGG) sequences are shown in yellow, and edited sequences are shown in red. Each position and average efficiency (%) of prime editing by FnCas9(H969A)-RT is indicated to the left and right of the target sequence, respectively. h. Comparison of base insertion efficiency of multiplexed prime editing at NRAS locus by SpCas9(H840A)-RT and FnCas9(H969A)-RT according to ngRNA targeting at various positions (1-3, Additional file 1: Figure S2c) using the PE3 method. Each histogram was plotted by applying standard error of the mean values to repeated experimental values (n = 3). i. Next-generation sequencing results of multiplexed TT insertion using SpCas9(H840A)-RT and FnCas9(H969A)-RT. Each target sequence [protospacer shown in blue and PAM (NGG) shown in yellow] for SpCas9(H840A)-RT and FnCas9(H969A)-RT is indicated. The targeted TT insertion is shown in red and each position is shown above the base. The dotted line indicates the omission of the base. PE3: prime editing with target-strand nicking; ngRNA: nicking guide RNA.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterial.pdf