Title: Phage DNA polymerase prevents on-target damage and enhances precision of CRISPR editing

Authors: Qiaoyan Yang[1], Jonathan S. Abebe[1], Michelle Mai[1], Gabriella Rudy[1], Orrin Devinsky[2], Chengzu Long[1, *]

Affiliations:

[1] NYU Cardiovascular Research Center, Leon H. Charney Division of Cardiology, Department of Medicine, NYU Grossman School of Medicine, New York, NY, USA.

[2]New York University Langone Health Comprehensive Epilepsy Center, New York, NY, USA.

*Corresponding author. Email: Chengzu.Long@nyulangone.org

Abstract:

CRISPR/Cas9-mediated genome editing in mammalian cells can generate undesired chromosomal alterations, including deleterious on-target large deletions and chromosomal translocations. Currently few approaches effectively prevent these on-target DNA damage. Here we show that the association of engineered phage DNA polymerases with Cas9 protein substantially inhibits the production of undesired on-target chromosomal alterations. Our "CasPlus" editing platform significantly increases the proportion of precise 1- to 3-base-pair insertions at target sites. CasPlus efficiently corrected common mutations in the *DMD* (del exon 52) and *CFTR* (F508del) genes by generating precise insertions without an exogenous template with fewer deleterious edits than Cas9 alone. CasPlus offers safer and more efficient targeting strategies optimized for human applications.

One-Sentence Summary:

Engineered phage DNA polymerases impact CRISPR/Cas9 gene editing outcomes by influencing DNA repair pathway choice.

Main Text:

The engineered clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR/Cas9) system is a powerful tool for sequence-specific gene editing (*1-4*). However, it can also generate undesired on-target large deletions(*5, 6*), chromosomal translocations(*7*), chromothripsis(*8*), and other complex chromosomal rearrangements(*9*) as well as off-target effect. Although numerous strategies have been developed to minimize CRISPR/Cas9-mediated off-target effects (*10*), few approaches can mitigate collateral on-target DNA damage. Cas9 cleaves target DNA producing blunt ends or staggered ends with 5′ overhangs (*11*). Repair of these ends without exogenous templates typically occurs through canonical non-homologous end joining (c-NHEJ) or microhomology-mediated end joining (MMEJ)(*12*). The choice of repair pathway determines Cas9 editing outcomes. MMEJ repair often results in deletions, particularly large deletions (*13, 14*). Systematic analyses of template-free Cas9 editing have revealed that small insertions arising from the c-NHEJ pathway are precise and predictable (*15-17*). The frequency and pattern of Cas9-generated small insertions depend highly on the local sequence context and staggered ends created (also known as templated insertions) (*10*). In budding yeast, DNA polymerase 4 (Pol 4) is essential for Cas9-induced 1- to 3-base-pair (bp) insertions (*18*). Hence, systematic screening various DNA polymerases could identify enzymes that promote c-NHEJ over MMEJ pathway, biasing for insertions over deletions during Cas9 editing in mammalian cells.

Duchenne muscular dystrophy (DMD) is characterized by degeneration of cardiac and skeletal muscles (*19*) and results from inactivating mutations in the X-linked dystrophin gene (*DMD*)(*20*). Dilated cardiomyopathy (DCM) is a common and lethal feature of DMD (*21*). Previously, we used Cas9 to rectify *DMD* mutations in cultured human cells and *mdx* mice(*22-25*); however, undesired

on-target DNA damage at edited *DMD* sites, a safety concern in human therapy, were not evaluated. Here we aim to identify DNA polymerases that can inhibit the large deletions associated with regular Cas9 editing and evaluate the safety and efficacy of the gene-editing platforms including these polymerases in correcting DMD and other diseases.

Results

Identification of T4 and RB69 DNA polymerases as favoring insertions over deletions during Cas9 editing.

To identify DNA polymerases that can enhance the fill-in step, leading to a ligation process favoring insertions over deletions during Cas9 editing (**Fig. 1A**), we first established a stable HEK293T cell line expressing a mutant tdTomato gene with a 1-bp deletion of adenine (A) at position 151 (tdTomato-del151A) (**Fig. 1B**). Next, we investigated the effects on Cas9 editing of DNA polymerases that can fill in 5′ overhangs, including the X-family polymerases Pol λ, Pol μ, and Pol β of human origin, and Pol 4 from yeast (*18, 26*), the family A polymerases Pol I and Klenow fragment (KF)(*27*) from bacteria and the B-family polymerases T4 DNA pol from bacteriophage T4 (*28*). We constructed MS2 (MS2 bacteriophage coat protein)-tagged vectors expressing these DNA polymerases with optimized human codons (**Fig. 1C**). We transfected tdTomato-d151A reporter cells with vectors containing Cas9, Green fluorescent protein (GFP), and a tdTomato-targeted single guide RNA (tdTomato-sgRNA) alone or combined with one of the DNA polymerases, sorted the transfected cells into populations expressing either only GFP ($tdTomato^{-}GFP^{+}$) or both tdTomato and GFP ($tdTomato^{+}GFP^{+}$). We then amplified the targeted DNA for high-throughput sequencing (HTS) (**Fig. 1D**). HTS results for the $tdTomato^{-}GFP^{+}$ populations indicated that overexpression of T4 DNA polymerase along with Cas9 increased the

frequency of 2-bp insertions—by ~6-fold, mostly at the expense of deletions—versus the Cas9-only treatment control (CTR) (**Fig. 1E)**. We obtained similar results from overexpressing RB69 DNA polymerase(*29*), a B-family DNA polymerase with 61% amino acid homology to T4 DNA polymerase, but not with the A-family T7 DNA polymerase (*30*) (**Fig. 1F**). All treatments had a similar ratio of templated to non-templated 2-bp insertions (**Fig. 1G**); No substantial changes were observed in tdTomato^{+}GFP^{+} populations (**fig. S1, A to D**). Western blot analysis confirmed the expression of all these DNA polymerases (**fig. S1E**). Subsequent experiments focused on T4 DNA polymerase because it strongly favored insertions over deletions at the tdTomato-d151A site. Unlike T4 DNA polymerase overexpression, overexpression of Cas9 and T4 DNA polymerase fusion proteins (Cas9-L-T4 and T4-L-Cas9) strongly induced deletions at the protospacer-adjacent motif (PAM) distal region in both the tdTomato^{-}GFP^{+} and tdTomato^{+}GFP^{+} populations (**fig. S1, F and G**).

We next tested the polymerase activity with guide RNAs (gRNAs) targeting an additional 26 target sites (TS1–TS26; **fig. S1H**). Compared to CTR, T4 DNA polymerase selectively increased 1–3-bp insertions or 1–2-bp deletions, depending on the gRNA context (**fig. S1, I to L**); Additionally, all increases in insertions across different Cas9 target sites occurred at the expense of deletions. Collectively, these results revealed that T4 DNA polymerase favors insertions over deletions at Cas9 target sites. This effect of T4 DNA polymerase on Cas9 editing outcomes is "CasPlus editing."

Engineering T4 DNA polymerase mutants capable of improving CasPlus editing efficiency.

T4 DNA polymerase possesses a polymerase activity that extends the polynucleotide chain, and a 3′-to-5′ proofreading exonuclease activity that excises erroneously incorporated nucleotides (*31*).

We hypothesized that T4 DNA polymerase mutations in the exonuclease or DNA polymerase domains might enhance CasPlus editing efficiency (**Fig. 2A**). We constructed seven T4 DNA polymerase mutants that had a decreased or an increased DNA mutation frequency relative to wild-type T4 DNA polymerase (T4-WT), and one N-terminal truncation mutant lacking the exonuclease domain (Δ1-377) (*32-35*) (**Fig. 2B**). Compared to T4-WT, one mutant, T4-D219A, showed an additional 2.4-fold increased 1-bp insertions at the expense of deletions at TS11 (**Fig. 2C**). Both T4-WT and the T4 DNA polymerase mutants exhibited a similar pattern of 1-bp insertions at TS11 (**fig. S2A**). T4-D219A led to 1-to-6-fold increases in insertions at the expense of deletions across 17 genomic sites tested (**fig. S2, B to D**). Consistent with these observations, mutation of the corresponding Asp-222 residue to Ala in RB69 DNA polymerase (RB69-D222A) (*29*) also resulted in a 2-to-4-fold higher frequency of insertions at the TS2, TS11, and TS12 than RB69-WT (**fig. S2E**). Together, these results support that in CasPlus editing, T4-D219A outperforms T4-WT and RB69-D222A outperforms RB69-WT.

To further investigate the capabilities of T4 DNA polymerase in CRISPR/Cas editing, we turned to Cas12a, an RNA-guided endonuclease that creates 5′ staggered ends with 5–8-bp overhangs(*36*). We reasoned that T4 DNA polymerase could fill in those overhangs, creating 5–8 nucleotides insertions (**fig. S2F**). We tested this by overexpressing T4 DNA polymerase along with Cas12a. As expected, without T4 DNA polymerase, Cas12a produced <2% insertions, whereas insertions were produced in 17% with T4-WT and Cas12a and 39% with T4-D219A and Cas12a (**fig. S2, G and H**).

Expanding the target range for CasPlus editing by combining Cas9 variants with T4 DNA polymerase.

CasPlus editing requires DNA ends with 5′ overhangs. At target sites where wild-type Cas9 (Cas9-WT) often generates blunt ends, some Cas9 variants can preferentially produce staggered ends with 1-bp overhangs(*37*). We hypothesized that combining these Cas9 variants with T4 DNA polymerase would favor 1-bp insertions at these target sites (**Fig. 2, D and E**). At TS11, the Cas9 variants F916P, F916del, R919P, and Q920P produced an average of 15% 1-bp insertions, whereas Cas9-WT produced <5% 1-bp insertions (**Fig. 2F**). Combining these Cas9 variants with T4-WT resulted in an average of ~50% 1-bp insertions and ~4.5% deletions whereas combining Cas9-WT and T4-WT yielded ~15% 1-bp insertions and ~59% deletions (**Fig. 2G**). We further tested five target sites that rarely induced 1-bp insertions when using Cas9-WT alone or with T4-WT. Cas9-F916P led to an average 4.3-fold and Cas9-F916del led to an average 5.1-fold increase in 1-bp insertions at the expense of deletions across the five target sites with T4-D219A compared to Cas9-WT or Cas9 variants alone (**fig. S2I**). Thus, this strategy expanded the range of targetable sites for CasPlus editing. We further hypothesized that Cas9 variants with T4 DNA polymerase could produce longer insertions (2–4-bp) at target sites where Cas9-WT and T4-WT produce 1–3-bp insertions (**fig. S2J**). With T4-WT, Cas9 variants F916P, F916del, or Q920P substantially increased 3-bp insertions at tdTomato-d151A site where Cas9-WT produced 2-bp insertions (**fig. S2, K and L**). With T4-WT or T4-D219A, Cas9-F916P and Cas9-F916del promoted 2-bp insertions at TS5, 3-bp insertions at TS17, and 4-bp insertions at TS18 whereas Cas9-WT produced 1-bp insertions at TS5, 2-bp insertions at TS17 and 3-bp insertions at TS18 (**fig. S2M**).

The four versions of the CasPlus system (**Fig. 2H**) described here generate distinct editing outcomes. At target sites where CasPlus-V1 predominantly favors 1–3-bp insertions, CasPlus-V2 produces similar effects to V1 but offers even higher efficiency. CasPlus-V3 and -V4 can increase

the length of insertions by 1 bp. At target sites where CasPlus-V1 mainly generates 1- or 2-bp deletions, CasPlus-V2, -V3, and -V4 favor 1-bp insertions (**fig. S2N**).

Enhanced correction of *DMD* exon 52 deletion by CasPlus editing

CasPlus editing enhances 1–2-bp insertions, which can efficiently correct disease-causing frameshift mutations as in DMD (**Fig. 3A**). Precise insertion of 1 bp at the 3′ end of exon 51 or the 5′ end of exon 53 could reframe mutated *DMD* genes with exon 52 deleted (**Fig. 3B**). We screened all possible gRNAs that can reframe *DMD* del exon 52 and determined their editing efficiencies in HEK293T cells (**Fig.3C**). The gRNAs G10 and Ex51-G2 targeting exon 51 and G9 and Ex53-G3 targeting exon 53 had the highest editing efficiencies (**fig. S3A**). Like G10 and G9 (**fig. S2B**), Ex51-G2 and Ex53-G3 with CasPlus-V1 or -V2 produced a 1.6-to-2-fold increase in correction efficiency compared to Cas9 due to increased 1-bp insertions (**fig. S3B**). Thus, CasPlus editing is more flexible and efficient in correcting *DMD* mutations than Cas9 editing. CasPlus-V1 and -V2 led to a 4.3-to-6.5-fold increase in correction efficiency in exon 51 using G10 and a 2.1-to-2.5-fold increase in correction efficiency in exon 53 using G9 versus Cas9 in a human induced pluripotent stem cell (iPSC) line with *DMD* exon 52 deletion (DMD-del52) (**Fig. 3D and fig. S3C**). We differentiated the pool of edited iPSCs and a single clone (SC) with 1-bp insertions into induced cardiomyocytes (iCMs). Compared to Cas9 editing, frequencies of in-frame mRNAs ($3n+1$) in iCMs resulting from CasPlus-V1 and -V2 editing were ~2.5- fold higher in exon 51 and ~1.5-fold higher in exon 53 (**Fig. 3E**). Western blot analysis confirmed that CasPlus-V1 and -V2 editing produced 1.5-to-3-fold greater dystrophin expression than Cas9 editing (**Fig. 3F**).

To investigate whether co-expression of T4 DNA polymerase and Cas9 increases genome-wide off-target effects, we performed whole-genome sequencing (WGS) in HEK293T cells with or

without Cas9 or CasPlus editing using G10 (**table S1**). We observed no apparent differences in insertion/deletion (indel) or single-nucleotide polymorphism (SNP) profiles when comparing CasPlus to Cas9 editing (**fig. S3, D to G**, and **table S2**). We did not detect gene editing at predicted off-target sites in Cas9- or CasPlus-edited cells (**fig. S3, H and I,** and **table S3**).

Exogenous template-independent correction of CFTR-F508del mutation by sequential CasPlus editing

CasPlus editing that induces a single, dominant outcome without exogeneous templates can sequentially correct genetic diseases caused by small deletions, as in cystic fibrosis (CF). CF results from mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene (*38*). A 3-bp deletion (c.1521-1523delCTT) resulting in the deletion of Phe-508 (F508del) accounts for ~75% of all pathogenic *CFTR* mutations (**fig. S3J**). CRISPR/Cas mediated gene editing via homology directed repair (HDR) with exogenous templates is inefficient in the non-dividing airway epithelial cells which are the most therapeutically target cells for CF patients (*39, 40*). Hence, exogenous template-independent correction of CFTR-F508del mutation via CasPlus editing offers a potential therapy. Based on the sequences flanking CFTR-F508del mutation site, we predicted three potential outcomes of restoring CFTR expression via sequential CasPlus editing: 1) two-step editing in which AT is inserted in the first step and T in the second step, 2) two-step editing in which T is inserted in the first step and TT in the second step, and 3) three-step incorporation of TTT with one T inserted at each step (**Fig. 3G**). We designed and tested gRNAs targeting the CFTR-F508del allele (Guide A) or intermediates containing an insertion of either T (Guide B), AT (Guide D), or TT (Guide C and E) in an HEK293T cell line homozygous for the CFTR-F508del mutation (**Fig. 3H and fig. S3K**). CasPlus editing produced <4% 3-bp insertions with Guide A alone (**fig. S3L**). In sequential CasPlus editing with two guides, combined A and D

produced ~8% to 15% 3-bp insertions while A and B generated ~ 3% to 6% 3-bp insertions (**fig. S3, M and N**). In sequential CasPlus editing with triple guides, combined A, B, and D induced ~16% to 19% 3-bp insertions. The highest rate of 3-bp insertions was 12.8% when combining A, B, and E (**Fig. 3I**). We analyzed the 3-bp insertion patterns produced by sequential CasPlus editing and observed that >90% of edits resulted in a silent mutation, producing the same protein sequence as found in WT CFTR (**Fig. 3G** and **fig. S3, O and P**). Thus, sequential CasPlus editing can efficiently and precisely correct the CFTR-F508del mutation without exogenous templates.

Repression of on-target large deletions by CasPlus editing.

On-target large deletions can arise from the long-range end resection that occurs during Cas9 editing (*41*) (**Fig. 4A**). In our corrected DMD-del52 iCMs, at the mRNA level, we detected the unexpected deletion of the whole exon 51 using G10 and exon 53 using G9 in Cas9 editing (**Fig. 4, B** and **C**). We speculated that these exons loss resulted from on-target large deletions. To investigate this, we PCR-amplified a 1–2-kilobase (kb) region in pools of edited cells (**fig. S4A**). We observed several lower bands representing deletions of ~0.5 kb distal to G10 and ~1.2 kb distal to G9 in Cas9- but not CasPlus-edited iPSCs and iCMs (**Fig. 4, D and E, and fig. S4, B to E**). We next amplified a ~5-kb region around the *DMD* exon 51 and 53 target sites from pools of edited iPSCs and sequenced the PCR amplicons using PacBio sequencing technology. Up to 23.0% of the PacBio reads contained deletions of 0.2–3 kb around the cut site of exon 51 in Cas9-edited cells (**Fig. 4F, fig. S4F**, and **table S4)**. We did not observe this effect in untreated cells (~2.0%) or cells edited with CasPlus-V1 (~3.2%) or -V2 (~3.5%). In untreated cells, we detected ~3-kb deletions around *DMD* exon 53 in 13.2% of the PacBio reads. This result likely resulted from artifact due to PCR amplification process, as 3-kb deletions of similar scale were observed in all tested cells [Cas9 (11.1%); CasPlus-V1 (9.4%); CasPlus-V2 (14.8%)]. On *DMD* exon 53, Reads

with deletions of 0.2–3.5 kb around the cut site were greatly decreased in cells edited with CasPlus-V1 (9.5%) and -V2 (17.4%) versus Cas9 (48.9%) (**Fig. 4G, fig. S4F and table S4**). We genotyped single clones sorted from edited iPSC pools to confirm our results. Three of 33 clones (9.1%) using Cas9 with G10 and three of 13 clones (23.1%) using Cas9 with G9 contained deletions ranging from 300 bp to 10 kb. No clones from CasPlus-V1- or CasPlus-V2 edited cells bore deletions >150 bp (**fig. S4, G to I**). Hence, CasPlus-V1- and -V2 editing efficiently repressed on-target large deletions while enhanced the correction efficiency of DMD.

Repression of on-target chromosomal translocations between two chromosomes by CasPlus editing.

Chromosomal translocations can occur when two simultaneous DSBs are present on two chromosomes (**Fig. 5A**). To investigate whether CasPlus editing can reduce chromosomal translocations, we recapitulated previously described translocation events between the genes *CD74* and *ROS1* in HEK293T cells (*42*) (**Fig. 5B**). We PCR-amplified the breakpoint junction regions on the fused chromosomes and assessed translocation efficiencies. We detected and verified both *ROS1-CD74* and *CD74-ROS1* translocations induced by Cas9 and CasPlus editing (**Fig. 5C and fig. S5A**). The translocation frequencies were ~5-fold lower with CasPlus-V1 and ~2-fold lower with CasPlus-V2 versus Cas9 editing (**Fig. 5, C and D)**. The frequencies of insertions at *ROS1* and *CD74* individual sites were higher with CasPlus-V1 and -V2 editing compared to Cas9 editing (**Fig. 5E**). We observed similar trends of repression of chromosomal translocations in iPSCs (**Fig. 5, F to H**). Genotyping revealed that three of 95 clones (3.2%) from the Cas9 edited iPSC pool but no clones from the CasPlus edited iPSC pools contained *ROS1-CD74* and *CD74-ROS1* translocations (**fig. S5, B to E**).

Repression of on-target chromosomal translocations among multiple chromosomes by CasPlus editing.

Multiple translocations were observed between cleaved chromosomes in Cas9-engineered T cells when delivered with three gRNAs targeting genes *PDCD1*, *TRBC1/2* and *TRAC*, on chromosome 2, 7 and 14, respectively, in a clinical trial (NCT03399448) (*6*). We, therefore, used the same three gRNAs combined with Cas9 or CasPlus to investigate the translocations in these three chromosomes in HEK293T cells (**Fig. 6A and fig. S6A**). The PCR results demonstrated that CasPlus-V1 decreased all translocations by a 2.5-to-4.5-fold in these three chromosomes (**Fig. 6, B and C and fig. S6, B and C**). CasPlus-V1 editing induced a comparable knockout efficiency at these four individual genes when compared to Cas9 editing (**Fig. 6D**). CasPlus-V2 induced a similar knockout efficiency but was less efficient in repressing chromosomal translocations compared to CasPlus-V1. The TaqMan assays revealed that CasPlus-V1 resulted in an average 5.8-fold and CasPlus-V2 resulted in an average 3.9-fold decreased translocations compared to Cas9 editing (**Fig. 6E**). These results indicate that CasPlus editing significantly represses Cas9-mediated on-target chromosomal translocations and may be a safer editing strategy for engineered T cell and other *ex vivo* cell therapies.

Discussion

Despite successes of CRISPR/Cas9-mediated genome editing in cultured cells and animal models, its clinical potential is limited by the safety and efficacy of gene editing in the human body. Unexpected on-target large deletions and chromosomal translocations from regular Cas9 editing are underappreciated risk factors for *in vitro*, *in vivo* and *ex vivo* applications. Our results show that in cultured human cells, the CasPlus platform can substantially reduce deleterious on-target

large deletions and chromosomal translocations while maintaining equal or higher efficiency of desired editing outcomes (**Fig. 6F**). We will next evaluate the safety and efficacy of CasPlus in animal models using adeno-associated virus (AAV) and lipid nanoparticles delivery systems.

Our CasPlus system predominantly favors templated 1–3-bp insertions, which can precisely correct genetic disease caused by small deletions, such as CFTR-F508del. The Cas9-mediated precise gene editing also relies on HDR. However, this approach is limited to the dividing cells (*40*). Because our CasPlus system depends on c-NHEJ pathway to elicit improved Cas9 edits, it is applicable to various cell types including non-dividing cells.

Our CasPlus system can enhance safety and thereby expand ex vivo applications, such as generating chimeric antigen receptor T (CAR-T) cells for cancer or autoimmune disease immunotherapies. The US Food and Drug Administration recently placed a clinical hold on a gene-editing-based investigational new drug application (IND) related to an allogeneic CAR-T development candidate and requested additional control data about genomic rearrangement assessments (*43*). As eukaryotic DNA repair mechanisms are conserved, our data from HEK293T and iPS cells suggest that CasPlus may generate engineered cells with fewer chromosomal abnormalities for cell therapy applications.

References and Notes

1. M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
2. M. Jinek *et al.*, RNA-programmed genome editing in human cells. *Elife* **2**, e00471 (2013).

3. L. Cong *et al.*, Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).

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

5. M. Kosicki, K. Tomberg, A. Bradley, Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* **36**, 765-771 (2018).

6. A. D. Nahmad *et al.*, Frequent aneuploidy in primary human T cells after CRISPR-Cas9 cleavage. *Nat Biotechnol*, (2022).

7. E. A. Stadtmauer *et al.*, CRISPR-engineered T cells in patients with refractory cancer. *Science* **367**, (2020).

8. M. L. Leibowitz *et al.*, Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat Genet* **53**, 895-905 (2021).

9. J. Boutin *et al.*, ON-Target Adverse Events of CRISPR-Cas9 Nuclease: More Chaotic than Expected. *CRISPR J* **5**, 19-30 (2022).

10. F. Uddin, C. M. Rudin, T. Sen, CRISPR Gene Therapy: Applications, Limitations, and Implications for the Future. *Front Oncol* **10**, 1387 (2020).

11. X. Shi *et al.*, Cas9 has no exonuclease activity resulting in staggered cleavage with overhangs and predictable di- and tri-nucleotide CRISPR insertions without template donor. *Cell Discov* **5**, 53 (2019).

12. H. H. Y. Chang, N. R. Pannunzio, N. Adachi, M. R. Lieber, Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506 (2017).

13. D. D. G. Owens *et al.*, Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res* **47**, 7402-7417 (2019).

14. M. Kosicki *et al.*, Cas9-induced large deletions and small indels are controlled in a convergent fashion. *Nat Commun* **13**, 3422 (2022).

15. M. W. Shen *et al.*, Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646-651 (2018).

16. F. Allen *et al.*, Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol*, (2018).

17. R. T. Leenay *et al.*, Large dataset enables prediction of repair after CRISPR-Cas9 editing in primary T cells. *Nat Biotechnol* **37**, 1034-1037 (2019).

18. B. R. Lemos *et al.*, CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc Natl Acad Sci U S A* **115**, E2040-E2047 (2018).

19. K. F. O'Brien, L. M. Kunkel, Dystrophin and muscular dystrophy: past, present, and future. *Mol Genet Metab* **74**, 75-88 (2001).

20. F. Muntoni, S. Torelli, A. Ferlini, Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol* **2**, 731-740 (2003).

21. R. Adorisio *et al.*, Duchenne Dilated Cardiomyopathy: Cardiac Management from Prevention to Advanced Cardiovascular Therapies. *J Clin Med* **9**, (2020).

22. C. Long *et al.*, Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. *Sci Adv* **4**, eaap9004 (2018).

23. C. Long *et al.*, Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400-403 (2016).

24. C. Long *et al.*, Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* **345**, 1184-1188 (2014).

25. E. N. Olson, Toward the correction of muscular dystrophy by gene editing. *Proc Natl Acad Sci U S A* **118**, (2021).

26. D. A. Ramsden, K. Asagoshi, DNA polymerases in nonhomologous end joining: are there any benefits to standing out from the crowd? *Environ Mol Mutagen* **53**, 741-751 (2012).

27. K. Bebenek, C. M. Joyce, M. P. Fitzgerald, T. A. Kunkel, The fidelity of DNA synthesis catalyzed by derivatives of Escherichia coli DNA polymerase I. *J Biol Chem* **265**, 13878-13887 (1990).

28. Y. Liu *et al.*, In Vitro CRISPR/Cas9 System for Efficient Targeted DNA Editing. *mBio* **6**, e01714-01715 (2015).

29. M. Hogg, W. Cooper, L. Reha-Krantz, S. S. Wallace, Kinetics of error generation in homologous B-family DNA polymerases. *Nucleic Acids Res* **34**, 2528-2535 (2006).

30. K. Hori, D. F. Mark, C. C. Richardson, Deoxyribonucleic acid polymerase of bacteriophage T7. Characterization of the exonuclease activities of the gene 5 protein and the reconstituted polymerase. *J Biol Chem* **254**, 11598-11604 (1979).

31. T. L. Capson *et al.*, Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry* **31**, 10984-10994 (1992).

32. L. J. Reha-Krantz, Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. Relating structure to function. *J Mol Biol* **202**, 711-724 (1988).

33. L. J. Reha-Krantz, Regulation of DNA polymerase exonucleolytic proofreading activity: studies of bacteriophage T4 "antimutator" DNA polymerases. *Genetics* **148**, 1551-1557 (1998).

34. A. K. Abdus Sattar, T. C. Lin, C. Jones, W. H. Konigsberg, Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase. *Biochemistry* **35**, 16621-16629 (1996).

35. L. J. Reha-Krantz *et al.*, DNA polymerization in the absence of exonucleolytic proofreading: in vivo and in vitro studies. *Proc Natl Acad Sci U S A* **88**, 2417-2421 (1991).

36. B. Zetsche *et al.*, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771 (2015).

37. J. Shou, J. Li, Y. Liu, Q. Wu, Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Mol Cell* **71**, 498-509 e494 (2018).

38. B. P. O'Sullivan, S. D. Freedman, Cystic fibrosis. *Lancet* **373**, 1891-1904 (2009).

39. S. D. Patel, T. R. Bono, S. M. Rowe, G. M. Solomon, CFTR targeted therapies: recent advances in cystic fibrosis and possibilities in other diseases of the airways. *Eur Respir Rev* **29**, (2020).

40. W. Sun *et al.*, Strategies for Enhancing the Homology-Directed Repair Efficiency of CRISPR-Cas Systems. *CRISPR J* **5**, 7-18 (2022).

41. K. W. Yoo, M. K. Yadav, Q. Song, A. Atala, B. Lu, Targeting DNA polymerase to DNA double-strand breaks reduces DNA deletion size and increases templated insertions generated by CRISPR/Cas9. *Nucleic Acids Res* **50**, 3944-3957 (2022).

42. P. S. Choi, M. Meyerson, Targeted genomic rearrangements using CRISPR/Cas technology. *Nat Commun* **5**, 3728 (2014).

43. "Beam Therapeutics announces FDA clinical hold on BEAM-201 IND application" (Tech. Rep., Beam Therapeutics., 2022)

44. F. Jiang *et al.*, Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867-871 (2016).

45. F. A. Ran et al., Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281-2308 (2013).

46. L. Pinello *et al.*, Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol* **34**, 695-697 (2016).

47. Statistical Genomics. Methods and Protocols. *Anticancer Res* **36**, 3224 (2016).

48. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).

49. A. McKenna *et al.*, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).

50. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).

51. S. Bae, J. Park, J. S. Kim, Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473-1475 (2014).

Acknowledgements: We thank C. Zhao, M. Khurram, and Y. Xiao for the cloning and sequencing of guide RNAs and plasmids; X. Wang, L. Dow, P. Hsu, and L. Bu for subcloning or lentivirus

packaging vector plasmids; H. Stower and R. Barr for manuscript editing; A. Namboodiripad, R. Iyer, M. Huang and U. Andréo for comments and suggestions.

Funding: This work was supported by grants from the Departmental Start-Up Grant, NYU Langone Health, and Kids Connect Charitable Fund.

Author contributions:

Conceptualization: QY, CL

Data curation: QY, JA

Formal analysis: JA

Funding acquisition: CL

Investigation: QY, MM

Methodology: QY

Project administration: QY, CL

Software: JA

Supervision: CL

Visualization: QY, CL

Writing – original draft: QY, OD, CL

Writing – review & editing: QY, GR, OD, CL

Competing interests: C.L. and O.D. are co-founders of Script Biosciences. C.L and Q.Y. are listed on two patents related to this work (U.S. Application No. 63/335,625 and No. 63/109,909). All other authors declare that they have no competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Supplementary Materials

Materials and Methods

Figs. S1 to S6

Tables S1 to S7

References (*45–51*)

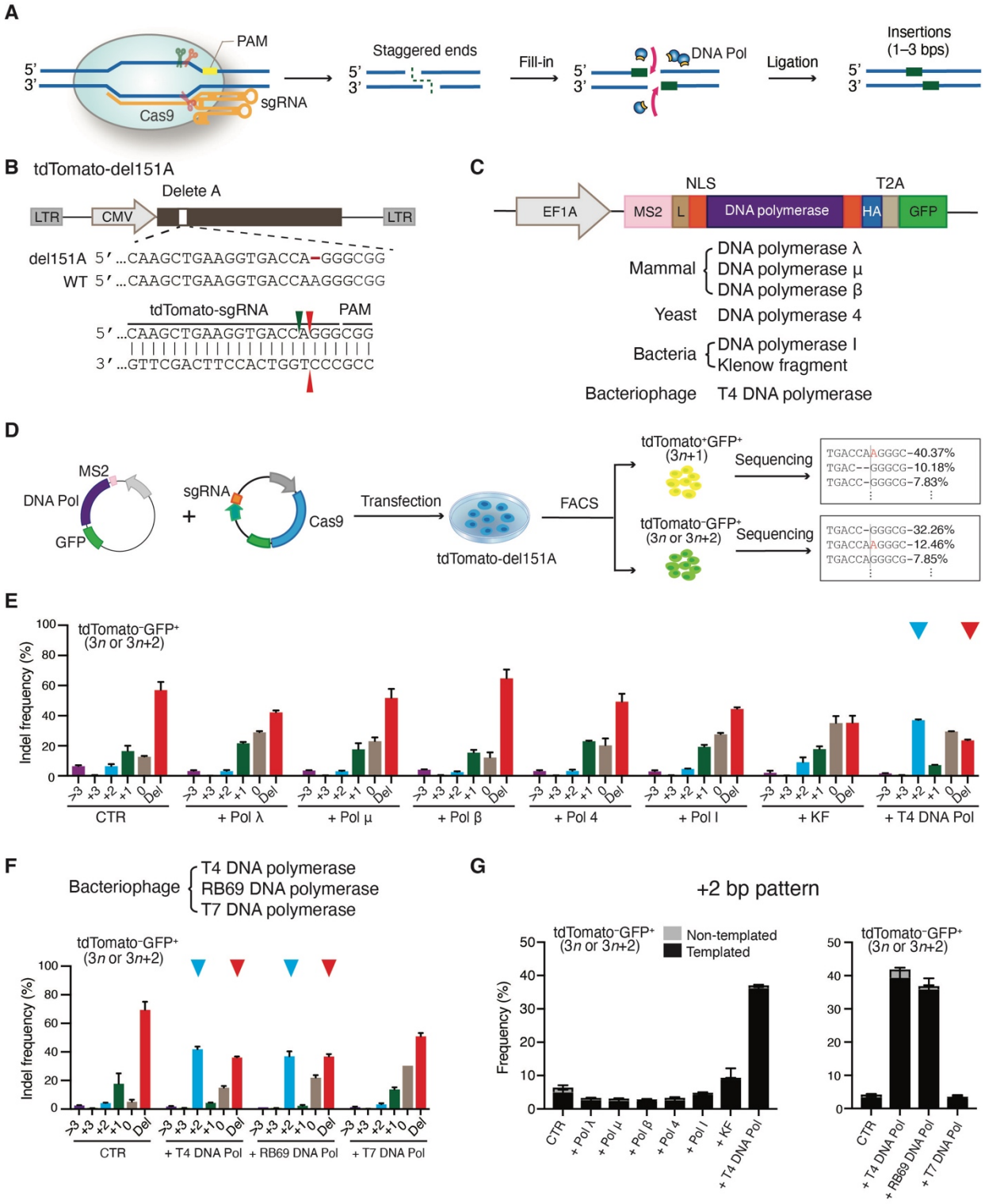

Fig. 1. Identification of T4 and RB69 DNA polymerases as favoring insertions over deletions during Cas9 editing. (**A**) Schematic showing the filling-in of Cas9-induced staggered ends by

exogenous DNA polymerase. (**B**) Schematic of lentiviral vector containing tdTomato-d151A (top) and the gRNA used to target it (bottom). Arrowheads, cut sites; red dash, missing nucleotide; LTR, long terminal repeat. (**C**) Architecture of vectors expressing distinct DNA polymerases (see method). (**D**) Workflow of the DNA polymerase selection process in tdTomato-d151A reporter cells. (**E** and **F**) Frequencies of Cas9-induced indels in tdTomato^{-}GFP^{+} populations with or without overexpression of a distinct DNA polymerase. (**G**) Patterns of 2-bp insertions in **E** and **F**.

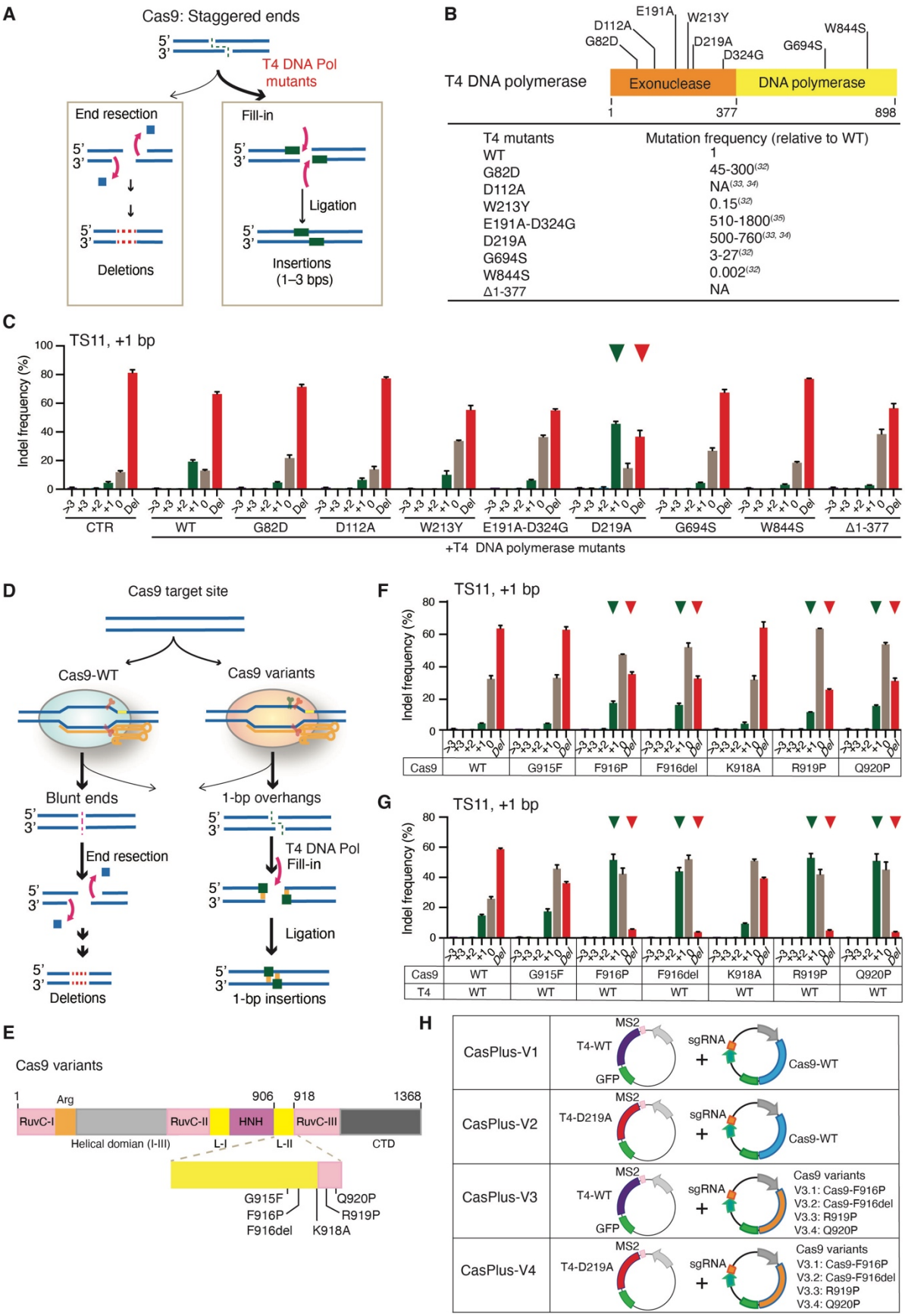
A
Cas9: Staggered ends
5'
3'
T4 DNA Pol mutants
End resection
Fill-in
Ligation
Deletions
Insertions (1–3 bps)
B
T4 DNA polymerase
Exonuclease
DNA polymerase
1
377
898
E191A
W213Y
D112A
G82D
D219A
D324G
G694S
W844S
T4 mutants | Mutation frequency (relative to WT)
WT | 1
G82D | 45-300[32]
D112A | NA[33, 34]
W213Y | 0.15[32]
E191A-D324G | 510-1800[35]
D219A | 500-760[33, 34]
G694S | 3-27[32]
W844S | 0.002[32]
Δ1-377 | NA
C
TS11, +1 bp
Indel frequency (%)
CTR
WT
G82D
D112A
W213Y
E191A-D324G
D219A
G694S
W844S
Δ1-377
+T4 DNA polymerase mutants
D
Cas9 target site
Cas9-WT
Cas9 variants
Blunt ends
1-bp overhangs
End resection
T4 DNA Pol Fill-in
Ligation
Deletions
1-bp insertions
E
Cas9 variants
1
Arg
906
918
1368
RuvC-I
RuvC-II
HNH
RuvC-III
Helical domian (I-III)
L-I
L-II
CTD
G915F
F916P
F916del
K918A
Q920P
R919P
F
TS11, +1 bp
Indel frequency (%)
Cas9
WT
G915F
F916P
F916del
K918A
R919P
Q920P
G
TS11, +1 bp
Indel frequency (%)
Cas9
T4
H
CasPlus-V1
CasPlus-V2
CasPlus-V3
CasPlus-V4
MS2
T4-WT
T4-D219A
GFP
sgRNA
Cas9-WT
Cas9 variants
V3.1: Cas9-F916P
V3.2: Cas9-F916del
V3.3: R919P
V3.4: Q920P

Fig. 2. Engineering of T4 DNA polymerase mutants and Cas9 variants capable of improving CasPlus editing efficiency. (**A**) Cas9-induced staggered ends can be processed for either end resection that generates deletions, or fill-in that produces insertions. T4 DNA polymerase mutants can enhance the fill-in process, thereby increasing 1–3-bp insertions. (**B**) Schematic showing the amino acid alterations introduced into T4 DNA polymerase; the corresponding relative mutation frequencies are below. NA, non-available. (**C**) Frequency of Cas9-induced indels at TS11 with or without T4 DNA polymerase mutants. (**D**) At target sites where Cas9-WT often generates blunt ends, leading to deletions (left), engineered Cas9 variants can preferentially generate staggered ends with 1-bp overhangs, enabling T4 DNA polymerase to produce 1-bp insertions (right). (**E**) Amino acid alterations introduced within Cas9 (*37*, *44*). (**F** and **G**) Frequency of Cas9-induced indels at TS11 in cells transfected with Cas9-WT or distinct Cas9 variants individually (**F**) or along with T4-WT (**G**). (**H**) Four distinct versions of the T4 DNA polymerase mediated CasPlus system.

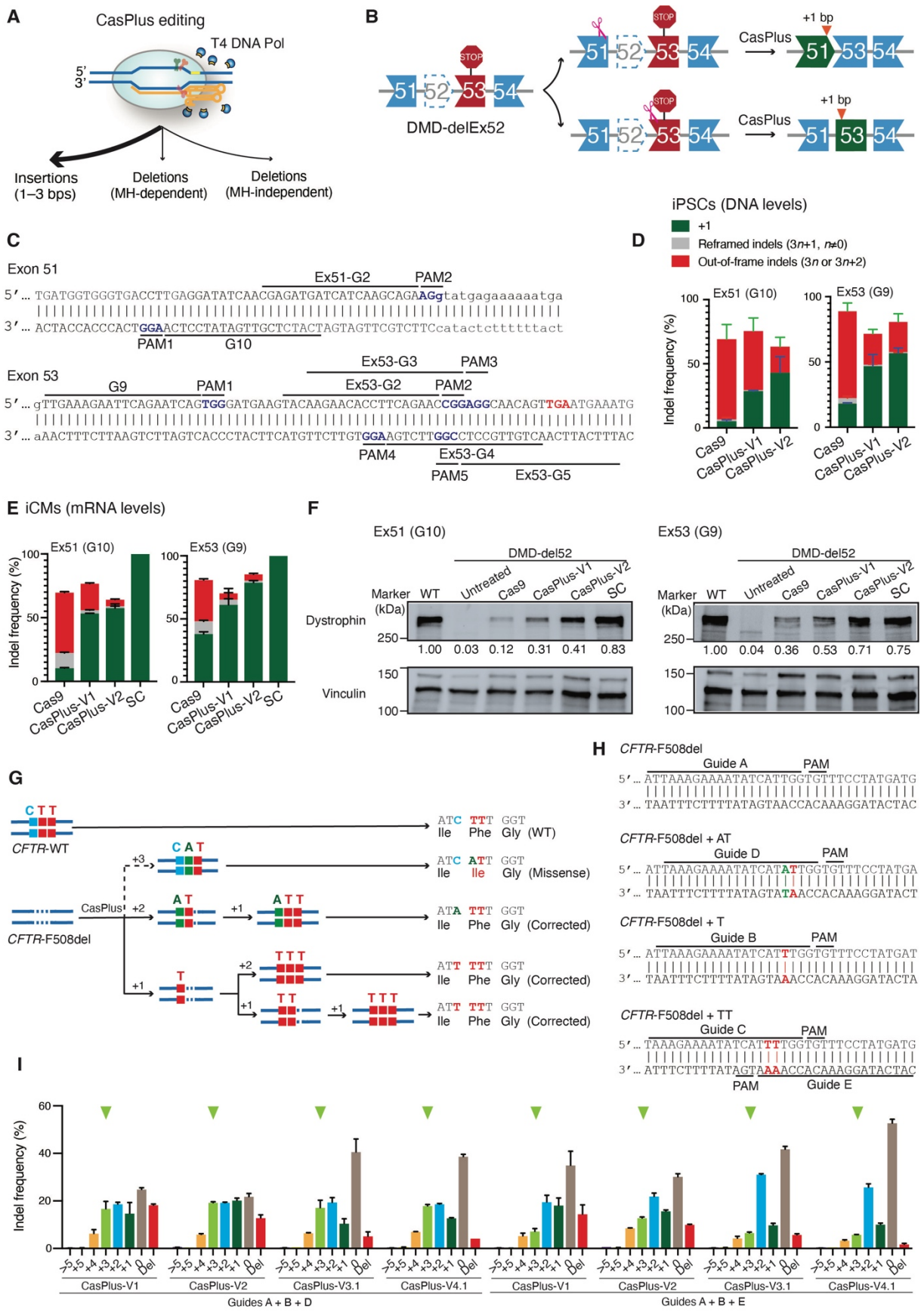
A
CasPlus editing
T4 DNA Pol
5'
3'
Insertions (1–3 bps)
Deletions (MH-dependent)
Deletions (MH-independent)
B
STOP
51 52 53 54
DMD-delEx52
CasPlus
+1 bp
C
Exon 51
Ex51-G2
PAM2
PAM1
G10
Exon 53
Ex53-G3
PAM3
G9
Ex53-G2
Ex53-G4
PAM4
PAM5
Ex53-G5
D
iPSCs (DNA levels)
+1
Reframed indels (3n+1, n≠0)
Out-of-frame indels (3n or 3n+2)
Ex51 (G10)
Ex53 (G9)
Indel frequency (%)
Cas9
CasPlus-V1
CasPlus-V2
E iCMs (mRNA levels)
SC
F
Ex51 (G10)
Ex53 (G9)
DMD-del52
Untreated
Marker (kDa)
WT
Dystrophin
Vinculin
1.00 0.03 0.12 0.31 0.41 0.83
1.00 0.04 0.36 0.53 0.71 0.75
G
CFTR-WT
CFTR-F508del
CasPlus
ATC TTT GGT
Ile Phe Gly (WT)
ATC ATT GGT
Ile Ile Gly (Missense)
ATA TTT GGT
Ile Phe Gly (Corrected)
ATT TTT GGT
Ile Phe Gly (Corrected)
H
CFTR-F508del
Guide A
PAM
CFTR-F508del + AT
Guide D
CFTR-F508del + T
Guide B
CFTR-F508del + TT
Guide C
Guide E
I
Indel frequency (%)
CasPlus-V1
CasPlus-V2
CasPlus-V3.1
CasPlus-V4.1
Guides A + B + D
Guides A + B + E

Fig. 3. Enhanced correction of *DMD* exon 52 deletion and exogenous template-independent correction of CFTR-F508del by CasPlus editing. (**A**) CasPlus editing enhances 1–3-bp insertions at the expense of MH-dependent or -independent deletions. (**B**) Deletion of *DMD* exon 52 creates a premature stop codon in exon 53, preventing dystrophin expression. Two CasPlus editing strategies can restore dystrophin expression by generating 1-bp insertions. (**C**) All potential gRNA sequences with an NGG PAM on *DMD* exons 51 and 53 available for restoring the *DMD* exon 52 deletion mutation. Red, premature stop codon. (**D** and **E**) Frequencies of in-frame ($3n$+1) and out-of-frame ($3n$ or $3n$+2) indels induced by Cas9, CasPlus-V1, and CasPlus-V2 with G10 or G9 in DMD-del52 iPSCs (**D**) and iCMs (**E**). (**F**) Western blot analysis of the expression of dystrophin and vinculin in DMD-WT iCMs, untreated DMD-del52 iCMs, DMD-del52 iCMs with edits at *DMD* exon 51 (left) or 53 (right), or DMD-del52 single clone containing 1-bp insertions. The normalized dystrophin quantitation is shown below the gel image. (**G**) Schematic showing four potential CasPlus editing outcomes for correction of the CFTR-F508del mutation. The one-step insertion of 3 bp produces a missense mutation. The two- and three-step strategies that insert 3 bp correct the original F508del mutation. (**H**) Guide RNAs used for sequential correction of CFTR-F508del. (**I**) Frequency of indels produced by sequential CasPlus editing when combining the Guides A, B, and D or A, B, and E.

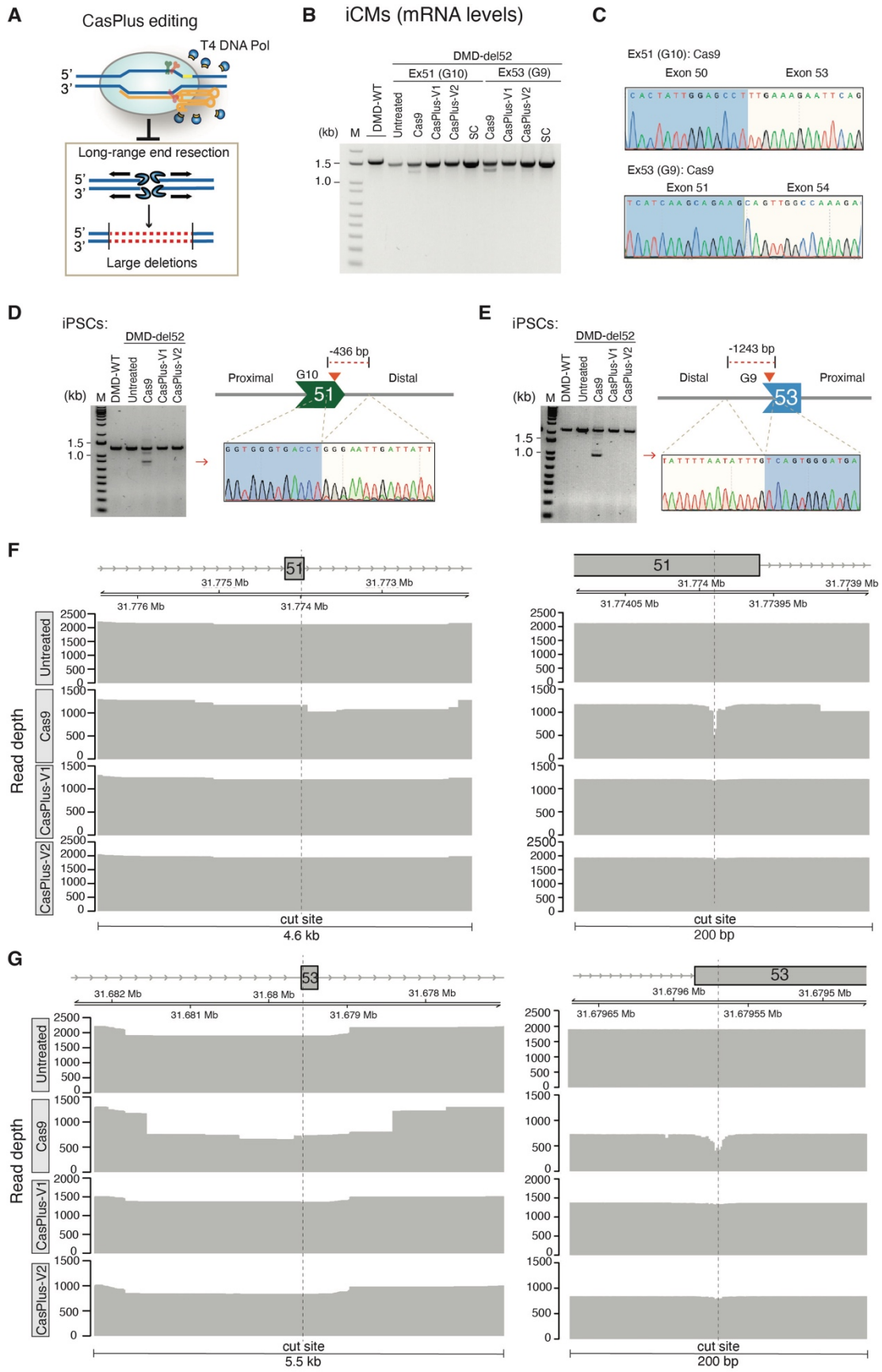
A
CasPlus editing
T4 DNA Pol
5'
3'
Long-range end resection
Large deletions
B
iCMs (mRNA levels)
DMD-del52
Ex51 (G10)
Ex53 (G9)
DMD-WT
Untreated
Cas9
CasPlus-V1
CasPlus-V2
SC
(kb)
M
1.5
1.0
C
Ex51 (G10): Cas9
Exon 50
Exon 53
Ex53 (G9): Cas9
Exon 51
Exon 54
D
iPSCs:
-436 bp
Proximal
Distal
G10
51
E
-1243 bp
G9
53
F
31.775 Mb
31.773 Mb
31.776 Mb
31.774 Mb
31.774 Mb
31.7739 Mb
31.77405 Mb
31.77395 Mb
Read depth
cut site
4.6 kb
200 bp
G
31.682 Mb
31.68 Mb
31.678 Mb
31.681 Mb
31.679 Mb
31.6796 Mb
31.6795 Mb
31.67965 Mb
31.67955 Mb
5.5 kb

Fig. 4. Repression of on-target large deletions by CasPlus editing. (**A**) CasPlus editing represses the generation of large deletions. (**B**) Representative gel image showing the RT-PCR products amplified from DMD-WT iCMs, untreated or edited DMD-del52 iCMs pools, or DMD-del52 single clone containing 1-bp insertions. Lower bands were purified and sequenced. (**C**) Sanger sequencing results illustrating the skipping of the whole exon 51 using G10 and exon 53 using G9 in Cas9-edited DMD-del52 iCMs. (**D** and **E**) Representative gel images showing the PCR products amplified from DMD-WT iPSCs, untreated DMD-del52 iPSCs or DMD-del52 iPSCs with edits at *DMD* exon 51 (**D**) or 53 (**E**). Lower bands (red arrowheads) were purified and sequenced. (**F** and **G**) Depth of PacBio reads at *DMD* exon 51 (**F**) or 53 (**G**) in untreated, Cas9-, CasPlus-V1-, and CasPlus-V2-edited DMD-del52 iPSCs.

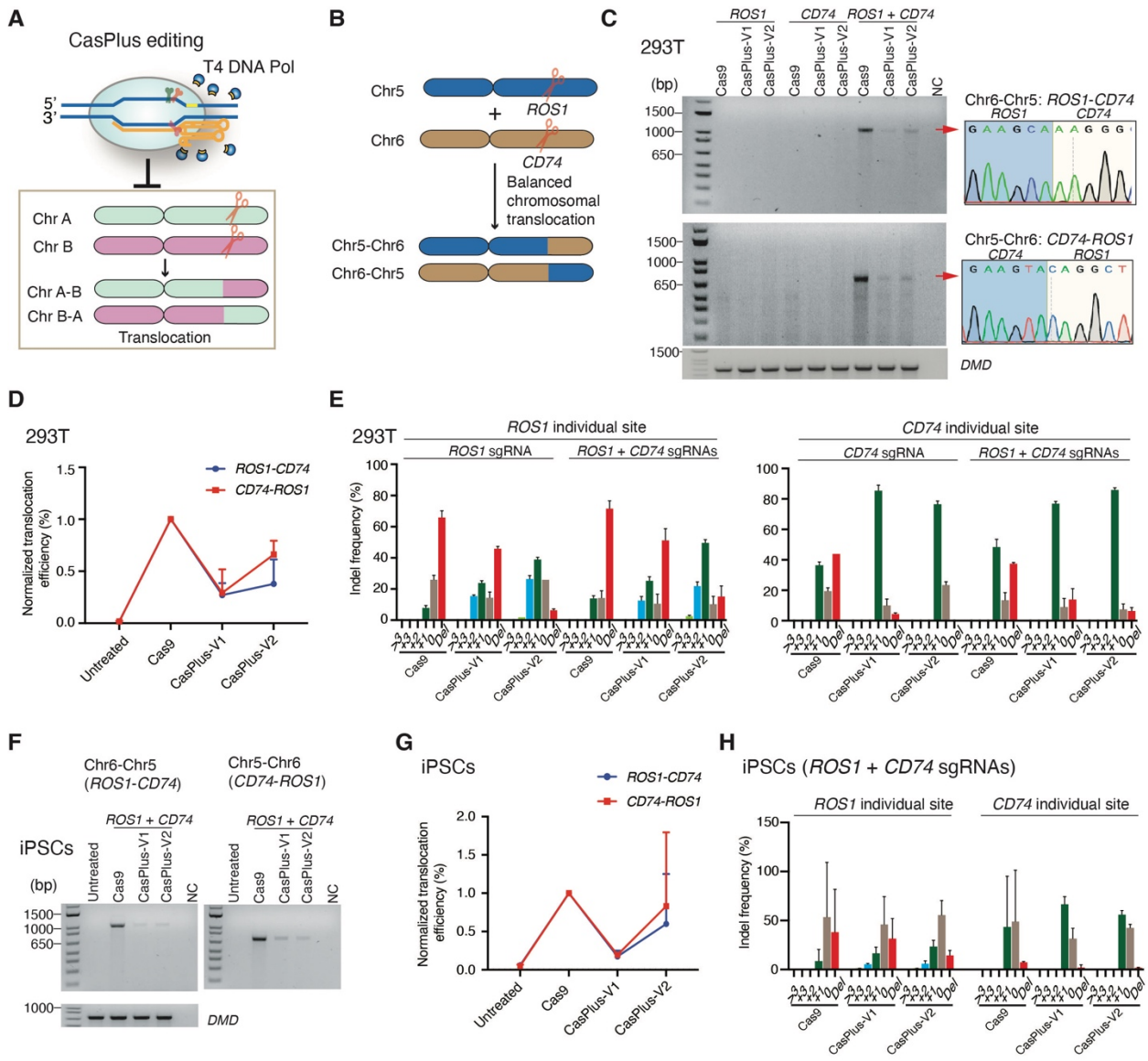

Fig. 5. Repression of on-target chromosomal translocations between two chromosomes by CasPlus editing. (**A**) CasPlus editing represses Cas9-mediated chromosomal translocations. (**B**) Schematic illustrating the generation of *ROS1-CD74* and *CD74-ROS1* fused chromosomes. (**C**) Representative gel images showing *ROS1-CD74* and *CD74-ROS1* translocations in HEK293T cells during Cas9, CasPlus-V1, or CasPlus-V2 editing (see methods). Bands with expected size (red arrowheads) were purified, TA-cloned and sequenced. *DMD* is a control for intensity normalization. NC, negative control. (**D**) Normalized quantification of data in **C**. Band intensity

obtained from Cas9-edited cells is set as 1. Values and error bars reflect mean ± SEM of n=3 replicates. (**E**) Frequencies of indels at *ROS1* and *CD74* individual sites in HEK293T cells during Cas9, CasPlus-V1, or CasPlus-V2 editing. Frequencies were calculated based on ICE (https://ice.synthego.com/#/) analysis on Sanger sequencing files. Values and error bars reflect mean ± SEM of n=3 replicates. (**F**) Representative gel images demonstrating the *ROS1-CD74* and *CD74-ROS1* translocations in iPSCs during Cas9, CasPlus-V1, or CasPlus-V2 editing. (**G**) Normalized quantification of data in **F**. (**H**) Frequency of indels at *ROS1* and *CD74* individual sites in iPSCs during Cas9, CasPlus-V1, or CasPlus-V2 editing. Frequency was calculated based on ICE analysis on Sanger sequencing files.

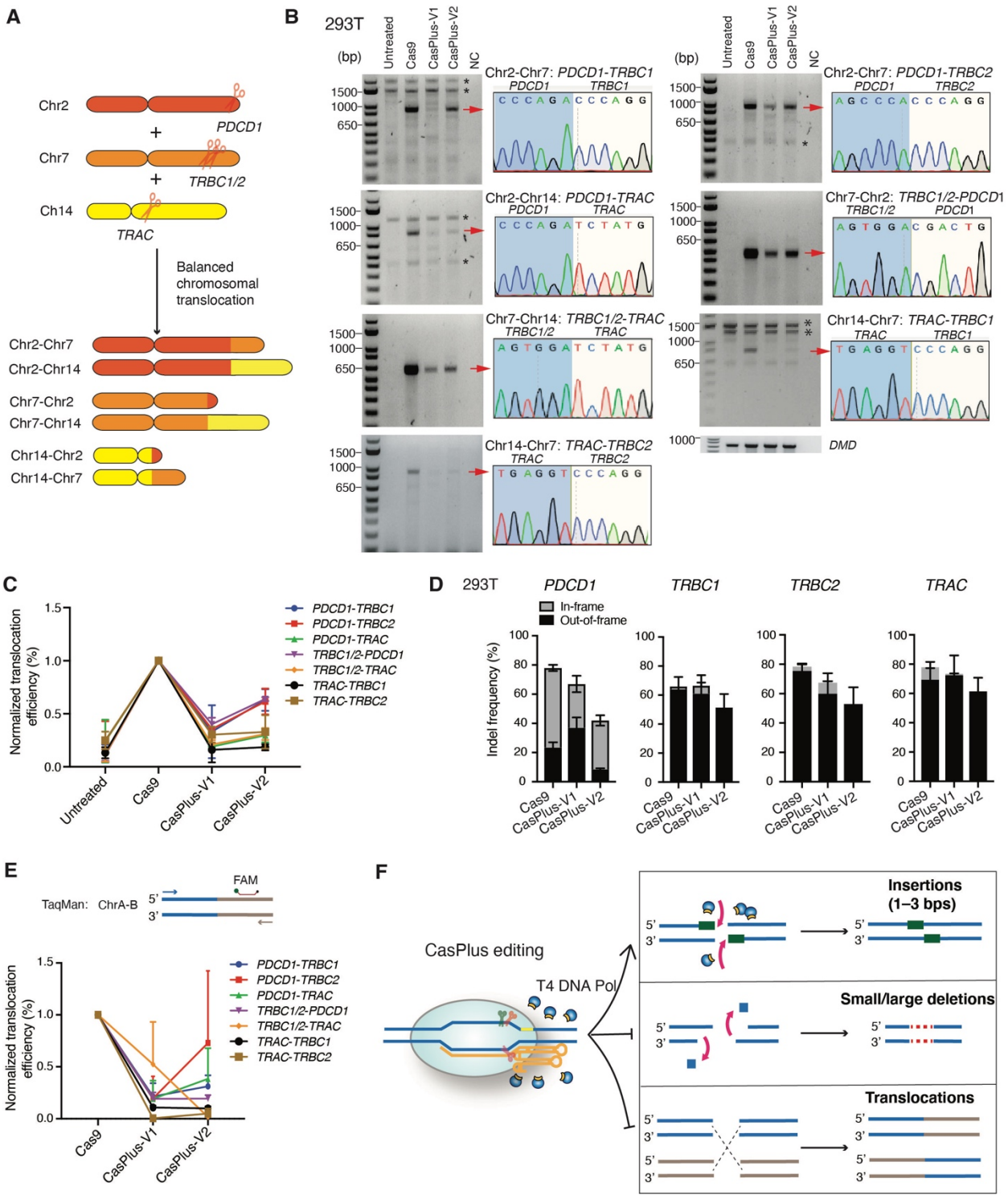

Fig. 6. Repression of on-target chromosomal translocations among multiple chromosomes

by CasPlus editing. (**A**) Schematic illustrating the balanced translocations among genes *PDCD1*, *TRBC1/2*, and *TRAC*. (**B**) Representative gel images demonstrating the balanced translocations detected in HEK293T cells during Cas9, CasPlus-V1, or CasPlus-V2 editing (see method). Balanced translocation of Chr14-Chr2: *TRAC-PDCD1* was undetectable by PCR. Bands with expected size (red arrowheads) were purified, TA-cloned and sequenced. Asterisk, non-specific bands. (**C**) Normalized quantification of data in **B**. (**D**) Frequencies of out-of-frame and in-frame indels at four individual sites in HEK293T cells during Cas9, CasPlus-V1, or CasPlus-V2 editing. Frequencies were calculated based on ICE analysis on Sanger sequencing files. (**E**) Evaluation of balanced chromosomal translocations detected in HEK293T cells during Cas9, CasPlus-V1, or CasPlus-V2 editing by TaqMan assay (see method). For each assay, the translocation efficiency in Cas9 edited cells is set as 1. Top panel, schematic showing the location of primers and probes. (**F**) Features of CasPlus editing. CasPlus editing utilizes T4 DNA polymerase to fill in the Cas9-created overhangs, thereby biasing insertions over small or large deletions. CasPlus editing can also repress chromosomal translocations that potentially occur between either on-target and off-target site during Cas9-mediated single site editing or different on-target sites during multiplex gene editing.