Fusion of SpCas9 to E.coli Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells

Lin, L., Petersen, T.S., Jensen, K.T., Bolund, L., Kühn, R., Luo, Y.

Journal of Biotechnology
2017 APR 10 ; 247: 42-49
2017 MAR 01 (first published online: final publication)
doi: 10.1016/j.jbiotec.2017.02.024

Publisher: Elsevier

Title: Fusion of SpCas9 to *E.coli* Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells

Author: Lin Lin Trine Skov Petersen Kristopher Torp Jensen Lars Bolund Ralf K*ü*hn Yonglun Luo

PII: S0168-1656(17)30087-1
DOI: http://dx.doi.org/doi:10.1016/j.jbiotec.2017.02.024
Reference: BIOTEC 7805

To appear in: *Journal of Biotechnology*

Received date: 9-11-2016
Revised date: 3-2-2017
Accepted date: 25-2-2017

Please cite this article as: <doi>http://dx.doi.org/10.1016/j.jbiotec.2017.02.024</doi>

Fusion of SpCas9 to *E.coli* Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells

Lin Lin 1, Trine Skov Petersen 1, Kristopher Torp Jensen 1, 2, Lars Bolund 1, Ralf Kühn 3, Yonglun Luo 1, *

1. Department of Biomedicine, Aarhus, Aarhus University
2. University of Cambridge, UK
3. Max-Delbrück-Center for Molecular Medicine, 13125 Berlin, Germany
Berlin Institute of Health, 10117 Berlin, Germany

* All correspondings should be addressed to Y.L. (alun@biomed.au.dk)

Abstract:

Mammalian cells repair double-strand DNA breaks (DSB) by a range of different pathways following DSB induction by the engineered clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein Cas9. While CRISPR-Cas9 thus enables predesigned modifications of the genome, applications of CRISPR-Cas9-mediated genome-editing are frequently hampered by the unpredictable and varying pathways for DSB repair in mammalian cells. Here we present a strategy of fusing Cas9 to recombinant proteins for fine-tuning of the DSB repair preferences in mammalian cells. By fusing *Streptococcus Pyogenes* Cas9 (SpCas9) to the recombinant protein A (Rec A, NP_417179.1) from *E. coli*, we create a recombinant Cas9 protein (rSpCas9) which enhances the generation of indel mutations at DSB sites in mammalian cells, increases the frequency of DSB repair by homology-directed single-strand annealing (SSA), and represses homology-directed gene conversion by approximately 33%. Our study thus proves for the first time that fusing SpCas9 to recombinant proteins can influence the balance between DSB repair pathways in mammalian cells. This approach may form the basis for further investigations of the applications of recombinant Cas9 proteins to fine-tuning DSB repair pathways in eukaryotic cells.

Keywords:

CRISPR; SpCas9; Rec A; SpCas9(1.1); DSB; SSA

1. Introduction

The bacterial clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9 (Cas9) has revolutionized genome-editing in various cells and organisms (Cong et al., 2013; Jinek et al., 2012; Jinek et al., 2013; Mali et al., 2013). The most widely used Cas9 is from *Streptococcus pyogenes* (SpCas9) and can be programmed by a single small guide RNA (gRNA) to target any desired genomic locus that is followed by a protospacer-adjacent motif (PAM, 5'NGG). The Cas9 protein then catalyzes the formation of a double-strand DNA break (DSB) 3 nucleotides upstream of the PAM (Cong et al., 2013; Jinek et al., 2012). In mammalian cells, unrepaired DSBs are frequently lethal, and cells have evolved several different mechanisms to repair DSBs (Huertas, 2010). These include simply rejoining the two DNA fragments with little or no further processing (non-homologous end joining, NHEJ) (Lieber, 2008) and a set of pathways using homologous sequences for DSB repair, collectively known as homologous recombination (HR) (Krogh and Symington, 2004). Single-strand annealing (SSA) is a DSB repair pathway that anneals and ligates two homologous regions flanking the DSB site resulting in a precise deletion of the intervening sequences (Huertas, 2010; Krogh and Symington, 2004). It has been discovered that microhomology-mediated end joining (MMEJ), which shares molecular signatures with both NHEJ and SSA (McVey and Lee, 2008), is also frequently used along with NHEJ/SSA for DSB repair by mammalian cells. These error-prone repair mechanisms have greatly facilitated the use of Cas9 for gene knockout applications as they commonly result in non-sense mutations. However, a challenge in CRISPR-Cas9 based genome editing is the lack of tools for controlling the choice of DSB repair pathways in cells. This has great increased the workload necessary to generate genetically modified cells with desired mutations due to the necessity for extensive screening and analysis.

It has previously been shown that fusion of a catalytically inactive Cas9 (dCas9) to functional proteins or catalytic domains can greatly broaden the CRISPR-Cas9-based genome and epigenome editing toolbox. Prominent examples include the fusion of dCas9 to the C terminus of the transcriptional activation domain VP64 (dCas9-VP64) for gene activation (Gilbert et al., 2013) or the Krüppel associated box

(KRAB) transcriptional repressor domain (dCas9-KRAB) for gene inhibition (Qi et al., 2013). It has also previously been shown that catalytically active Cas9 remains bound to both ends of the cleaved DNA following generation of a DSB (Sternberg et al., 2014). We therefore hypothesized that a similar fusion approach can be used to influence the preferences of DSB repair pathways by fusing a catalytically active wild-type Cas9 protein to the N terminus of a protein or domain of interest to yield multifunctional recombinant fusion proteins. Here we show that this strategy can be used to generate fusion proteins which retain the Cas9 endonuclease activity while simultaneously influencing the balance of DSB repair pathways in mammalian cells.

2. Materials and Methods

2.1. CRISPR gRNA design

CRISPR gRNAs were designed using the online prediction tool (http://crispr.mit.edu/). *EMX1* and *HEK293 site 2* gRNA target sites were selected from a previous validated study (Tsai et al., 2015). All gRNA target sequences are listed in **Supplementary Table 1**.

2.2. DNA oligonucleotide synthesis and Sanger sequencing

All DNA oligonucleotides in this study were synthesized from Sigma Aldrich (Denmark, Europe). Sanger sequencing was conducted using the Mix2seq kit from Eurofin Genomics (Germany).

2.3. Generation of fusion eSpCas9 expression vectors, gRNA expression vectors, and C-Check vectors

The SpCas9 (Addgene plasmid # 48139) and eSpCas9(1.1) (Addgene plasmid # 71814) plasmids were a gift from Feng Zhang. The *E.Coli* RecA coding region (NP_417179.1) was human codon optimized and synthesized by Invitrogen GeneArt Gene Synthesis (ThermoFisher Scientific). All fusion plasmids (SpCas9-RecA and eSpCas9-RecA) were validated by restriction enzyme digestion and Sanger sequencing to ensure correct fusion with no unintended mutations introduced by PCR and will be available through Addgene (https://www.addgene.org/Yonglun_Luo/). Addgene plasmid ID: 87263 and 87264.

To generate the gRNA expression vectors, complementary gRNA oligonucleotides were annealed in 1X NEB buffer 2, cloned into a single gRNA expression plasmid with a human-U6 promoter, and subsequently validated by Sanger sequencing. To generate the C-Check reporter vectors, complementary DNA oligonucleotides that contain a gRNA target site were annealed and cloned into the C-Check vector (Addgene plasmid #66817) by golden gate assembly, following our protocol described previously (Zhou et al., 2016). All C-Check vectors were validated by Sanger sequencing

2.4. Cell culture

HEK293T, HeLa and U2OS cells (ATCC) were cultivated in Dulbecco's modified Eagle's medium (DMEM) (LONZA) supplied with 10% FBS (Gibco), 1% penicillin-streptomycin (Sigma), 1% GlutaMAX (Gibco) in a 37°C incubator with 5% CO2 and maximum humidity. The TLR cell line was a gift from Chu *et al.* 2015. At approximately 80% confluence, cells were detached by 0.05% Trypsin-EDTA and passaged at a ratio of 1:8.

2.5. Transfection

In this study, all transfections were performed using the X-tremeGENE9 DNA transfection reagent (Roche) following the manufacturer's instructions. Briefly, cells were seeded into 24-well plates one day before transfection with a cell density of $5X10^4$ cells per well for HEK293T cells and $1X10^4$ cells per well for HeLa and U2OS cells. For each transfection, a total amount of 250 ng of plasmids was used (3:1 ratio of Cas9 to gRNA). A pUC19 plasmid was used for transfection of control groups in the same amount as the combined plasmids of non-control groups.

2.6. Flow cytometry (FCM)

Flow cytometry analysis was performed using a four-laser FACSAria III cell sorter at the FACS CORE facility of Aarhus University. At least 10,000 events were analyzed per biological replicate (independent transfection).

2.7. C-Check analysis

For each transfection, Cas9, gRNA and C-Check plasmid was used in a 3:1:4 ratio. 48 hours after transfection, cells were harvested using 0.025% trypsin-PBS-EDTA (phenol red free) and washed twice with PBS-5%FBS, fixed with 4% Formaldehyde-PBS solution for 10 min at R.T and washed with PBS twice before analysis. The percentage of GFP+ cells out of the AsRED+ cells was quantified by FCM.

2.8. TLR analysis

24 hours before transfection, 50,000 TLR cells were seeded per well in a 24-well plate. For each transfection, Cas9 (90 ng), gRNA (30 ng) and donor plasmid (30 ng) were used.

For the TLR-like plasmid based reporter system, 50,000 HEK293T cells were seeded per well in a 24-well plate 24 hours before transfection. For each transfection, the following plasmids were used: Cas9 (90 ng), gRNA (30 ng), donor plasmid (70 ng) and TLR-like plasmid (10 ng).

Transfections were performed using the X-tremeGENE9 DNA transfection reagent (Roche) according to the manufacturer's instructions. 24 hours after transfection, fresh cell culture medium was added to the cells, and transfected cells were harvested for FCM analysis 72 hours after transfection. At least 10,000 events were analyzed per biological replicate (independent transfection).

2.9 Detection of gene deletion mediated by pairs of gRNAs.

To evaluate gene deletion facilitated by pairs of gRNAs, HEK293T cells were transfected with Cas9 (SpCas9 or rSpCas9) and a pair of gRNA expression plasmids. 48 hours after transfection, genomic DNA was purified from the transfected cells (DNeasy Blood & Tissue Kit, Cat No. 69504) and 50 ng genomic DNA was used for PCR using primers flanking the deleted regions. All PCR primers are listed in supplementary table 1. Quantification of deletion (KO) and wildtype (wt) band intensity was performed with Image J.

2.10. Western blot analysis

HEK293T cells were transfected with Cas9 expression plasmids in a 6-well plate. Two days after transfection, cells were harvested by trypsinization and washed with PBS three times. The cells were then resuspended in lysis buffer (10mM Tris-HCl (pH 7.4)), 137mM NaCl, 10% Glycerol, 1% NP40 and protease inhibitors) by vortexing. Following resuspension, the cells were placed on dry ice for 3 min, transferred to 37°C, and vortexed. The freeze-thaw cycles were repeated 3 times and the lysates were placed on ice for 30 min. After centrifugation at 10,000 rpm at 6°C for 10min, the supernatant was transferred into a new tube and protein concentrations were determined with Qubit Fluorometer (ThermoFisher Scientific). 2µg lysate was boiled at 100 °C for 5 min and loaded on gradient SDS-PAGE gels (4-15%). Blots were probed with anti-FLAG M2 (F1804, Sigma) and anti-

beta-tubulin (ab6046, abcam) antibodies. Blots were developed with secondary goat anti-rabbit IgG HRP (sc-2004, Santa Cruz Biotechnology) or anti-mouse IgG HRP (sc-2005, Santa Cruz Biotechnology) and bands visualized using the Immobilon Western Chemiluminescent HRP Substrate (WBKLS0100). Images were taken with ImageQuant™ TL (GE Healthcare Life Sciences) and quantified with Image J.

2.11. TIDE assay

Quantifications of indel frequencies were performed using the TIDE assay as described previously (Brinkman et al., 2014). Briefly, two days after transfection, cells were harvested in Lysis buffer CS, 0.5% Tween20, 0.5% NP40, proteinase K followed by heating at 65°C for 30 min and 95 °C for 10 min (Luo et al., 2011). 1µl of lysate was used as template for PCR with the high fidelity polymerase Platinum® Pfx DNA Polymerase. PCR products were purified with the NucleoSpin Gel and PCR Clean Up kit and directly sequenced by Sanger sequencing (Mix2seq, EurofinGenomics). % indels was quantified by TIDE software (https://tide.nki.nl/). A p-value cutoff of $p<0.001$ and indel size range from -10 to +10 was used for all analyses, and decomposition windows were optimized for each gRNA according to TIDE guidelines. For analysis of TIDE data, control values from groups transfected with a plasmid expressing an empty gRNA backbone were subtracted as these indicate the non-specific error at each site from the TIDE calculation.

2.12. Statistical analyses

All data are presented as mean ± standard deviation. Student's T-test and one-way analysis of variance (ANOVA) with Bonferroni correction were used for statistical analyses of multiple comparisons in the study. All statistical analyses were conducted using Stata (version 10). P values less than 0.05 were considered statistically significant.

3. Results

3.1. Fusing SpCas9 to bacterial RecA (rSpCas9) enhances gene knockout and deletion

The bacterial RecA protein is a single strand DNA (ssDNA) binding protein and ATP-dependent recombinase involved in bacterial homologous recombination (Chen et al., 2008). It has previously been demonstrated that overexpression of RecA promotes homology-directed gene conversion in mammalian cells (Shcherbakova et al., 2000). Since RecA is also involved in several HR-related functions ranging from homology-search to strand-interactions, we hypothesised that it could be an alternative to mammalian protein domains such as Rad51 for promoting homologous recombination. We therefore hypothesized that fusing active SpCas9 to the N-terminus of RecA would yield a fusion protein that promotes homology-directed DSB repair.

A plasmid expressing a recombinant SpCas9 fused to the N terminus of the human codon-optimized RecA protein (NP_417179.1) was created (**Figure 1A**). To prevent potential interference with correct protein folding, a flexible peptide linker (Gly – Gly – Gly – Gly – Ser, G4S) was used to join SpCas9 and RecA. We first confirmed the correct size of the SpCas9-RecA fusion protein (hereafter rSpCas9) and its expression by Western blot. Our results showed that rSpCas9 expression levels were slightly lower but not significantly different from the wild-type SpCas9 (**Figure 1A**). We next investigated whether the rSpCas9 fusion endonuclease could still be programmed by gRNAs to cleave chromosomal DNA in mammalian cells. Five genomic loci, including two previously validated gRNAs targeting *EMX1* and *HEK293 site 2* (Tsai et al., 2015), were selected for this analysis (**Figure 1B**). Two days after co-transfecting HEK293T cells with plasmids expressing either SpCas9 or rSpCas9 together with a plasmid expressing the gRNA in question, the percentage of indel mutations was quantified by the TIDE assay (Brinkman et al., 2014). Surprisingly, we observed that in four out of five target sites, rSpCas9 was significantly more efficient at generating indel mutations than SpCas9 (**Figure 1B.** *EMX1*: 51% increase, *HEK293 site 2*: 39%, Chr1: 57%, and Chr18: 93%, $p < 0.05$ Student's T-test). To prove that this effect is not cell type-specific, the increased indel mutation rate by rSpCas9 was further verified in HeLa and U2OS cells (**Supplementary Figure 1**).

We further confirmed the applicability of rSpcas9 to gene knockout by comparing gene deletion efficiencies for SpCas9 and rSpCas9 targeting *uPA, IGF1R* and *CRP* in HEK293T cells, each with two gRNAs, where a general increase in gene deletion was observed for rSpCas9 (**Figure 1C**, 32-84% increase, $p < 0.01$, Student's T-test).

3.2. rSpCas9 promotes DSB repair by SSA

To investigate whether rSpCas9 enhances DSB repair by SSA, we used our previously developed C-Check dual fluorescent reporter system (Zhou et al., 2016). The C-Check system relies on SSA-mediated repair of a DSB at a target site introduced between two sections of GFP (500 nt homology) in a plasmid constitutively expressing AsRED. A DSB repaired by SSA will generate a full-length GFP which can be quantified by flow cytometry (**Figure 2A**). HEK293T cells were co-transfected with plasmids expressing either SpCas9 or rSpCas9, and a C-Check plasmid together with one of two gRNAs targeting the C-Check plasmid in question. The SSA efficiency was quantified by normalizing the number of GFP^+ cells to the number of $AsRed^+$ cells 48 hours after transfection. As shown in **Figure 2B, C**, rSpCas9 exhibited significantly higher activities in the C-Check assay than SpCas9 (2.1- and 2.5-fold increase, $p<0.001$), indicating an increased preference for SSA-mediated DSB repair. Our results thus suggest that fusion of SpCas9 to *E.coli* Rec A facilitates DSB repair by homology-directed SSA in mammalian cells.

3.3. rSpCas9 inhibits DSB repair by homology-directed gene conversion in traffic light reporter cells

We next investigated the effect of rSpCas9 on DSB repair by homology-directed gene conversion using a 'traffic light' reporter (TLR) cell line (Chu et al., 2015). Following generation of a DSB at the TLR locus, repair by homology-directed gene conversion results in expression of a green fluorescent protein (Venus), whereas a fraction of the DSB repair events by NHEJ or MMEJ lead to +2 reading frame mutations resulting in red fluorescent protein (RFP) expression (**Figure 3A**). We co-transfected TLR cells with either SpCas9 or rSpCas9, a TLR gRNA targeting the integrated mutated *Venus* gene, and a Venus donor plasmid. 72 hours after transfection, % Venus positive and RFP positive cells were quantified by flow cytometry. We observed that rSpCas9 exhibited a lower frequency of homology-

directed gene conversion than SpCas9 as quantified by the fraction of Venus positive cells (**Figure 3B, C**, decreased by 33 %, $p < 0.01$), indicating that the fusion protein inhibits DSB repair by homology-directed gene conversion. Significantly higher RFP+ frequencies were observed in the TLR cells expressing rSpCas9 compared to SpCas9 (**Figure 3B, C**, 29% increase, $p<0.001$), which again suggests a preference for DSB repair by NHEJ/MMEJ when using rSpCas9. Finally, the TLR system allowed us to compare the preference for DSB repair by homology-directed gene conversion relative to NHEJ/MMEJ for each of the two proteins by comparing the ratio of $Venus^+$ to RFP^+ cells, where a 48% decrease is seen for rSpCas9 compared to SpCas9 (**Figure 3C**, $p < 0.001$). This suggests that rSpCas9 can be a useful tool for gene KO experiments by facilitating error-prone repair over HDR.

To investigate whether rSpCas9 also affects DSB repair by HDR in plasmid-based systems, we generated a TLR-like reporter plasmid (Figure 3D). This TLR-like plasmid contains the same mutated Venus coding sequence and TLR gRNA target site. In addition, it contains the same CMV-AsRED expression cassette as the C-Check system for normalization. However, using this plasmid-based HDR reporter system, no significant difference in plasmid-based HDR repair was observed between SpCas9 and rSpCas9 (**Figure 3E, F**).

3.4. Fusion of RecA to eSpCas9(1.1) does not affect its mismatch tolerance

A novel, enhanced specificity, SpCas9 variant (eSpCas9(1.1)) which can only tolerate one mismatch between gRNA and the target locus has recently been generated (Slaymaker et al., 2016). To investigate whether fusing RecA to eSpCas9(1.1), hereafter referred to as eSpCas9-RecA, can enhance gene knockout (on-target activity) without interfering with its mismatch tolerance, we transfected HEK293T cells with a C-Check reporter vector, one of the Cas9 variants (eSpCas9(1.1.) and eSpCas9-RecA), and one gRNA carrying 0-2 mismatches at varying positions (**Figure 4A**). Control cells were transfected with a scrambled gRNA and one of the Cas9 variants, or transfected with a control plasmid pUC19. 48 hours after transfection, the relative gene editing efficiencies of the Cas9 proteins were quantified by flow cytometry by normalizing the number of GFP positive cells to the number of AsRED positive cells. This assay was repeated in triplicates three times for each of two target sites (target site 1 and target site 2, **Figure 4**). We observed

that eSpCas9-RecA generally showed significantly higher efficiencies in the assay than eSpCas9(1.1) but the relative increase was less than that observed for rSpCas9 compared to wt SpCas9. Similar increases in targeting efficiency by fusion of RecA compared to eSpCas9(1.1) were observed for both on-target sites and the off-target sites with one mismatch. Importantly, there was no detectable off-target cleavage for either protein when targeting the C-check vectors with gRNAs containing 2 mismatches.

Since we have previously observed a difference between plasmid-based and genomic assays (**Figure 3**), we next investigated whether the observed activities of eSpCas9-RecA are replicated when targeting genomic loci. We first targeted three genomic loci in HEK293T cells (Chr1, Ch3, and Chr8). Unlike for rSpCas9 and the reporter system in **Figure 4B**, despite a consistent increase in the frequency of indels, the difference was not significant for eSpCas9-RecA compared to eSpCas9(1.1). Furthermore, to investigate the mismatch tolerance in the genomic context, we generated six gRNAs containing 1-2 mismatches to the EMX1 target site. Using TIDE assay, no difference was observed in either on-target activity or mismatch tolerance (except OT-1, significantly increased in eSpCas9-RecA) between eSpCas9(1.1) and eSpCas9-RecA. Both eSpCas9(1.1) and eSpCas9-RecA had a significantly lower mismatch tolerance than SpCas9 (**Figure 4C**). Taken together, our results show that fusing RecA to eSpCas9(1.1) only results in significantly increased cleavage efficiency ($p<0.05$) in a plasmid based assay but not in a genomic context, and that fusing RecA to eSpCas9(1.1) did not affect its mismatch tolerance.

4. Discussion

In this study, we have created a recombinant Cas9 protein (SpCas9-RecA; rSpCas9) with modified preferences for DNA DSB repair in mammalian cells. Previous studies have found that the majority of indel mutations at target sites resulted from repair of Cas9/gRNA-induced DBSs by mutation-prone end-joining pathways such as classical nonhomologous end-joining (c-NHEJ) and microhomology mediated end-joining (MMEJ) (Canver et al., 2015; Mandal et al., 2014; van Overbeek et al., 2016). In this study, we observed that rSpCas9 mediates indel formation 30% more

efficiently than wild type SpCas9 with a decreased frequency of HDR events, suggesting that fusion of RecA to SpCas9 enhances DSB-repair by c-NHEJ/MMEJ at the Cas9/gRNA-induced DSBs relative to error-free repair pathways. In addition, we and others have shown that cells also utilized single strand annealing (SSA) pathways to repair DSBs created by programmable nucleases such as TALENs and Cas9/gRNA (Kuhar et al., 2014; Yasuda et al., 2016; Zhou et al., 2016). In this study, we show that fusing RecA to SpCas9 can also increase the frequency of DSB repair by SSA.

Furthermore, we have investigated the effect of fusing RecA to SpCas9 on DSB repair by HDR. Our data suggests that the fusion has a negative effect on DSB repair by HDR in a genomic context although this effect was not observed in a plasmid-based HDR assay. Additionally, the effects of fusing Cas9 to RecA were generally less pronounced for eSpCas9(1.1) than for native SpCas9. The discrepancy between plasmid-based and genomic results might be due to the copy number of TLR cassette present in each cell, as well as other factors such as chromatin accessibility and modification, interactions of the RecA protein with chromatin components etc. We also obtained a relatively low basal HDR rate in our study compared to a study by Liang et al. (Liang et al., 2017). Although similar cells (HEK293) have been used between Liang et al. and our studies, many other factors such Cas9 format (RNP in Liang's study, and plasmid in ours), transfection methods (electroporation in Liang's study, and X-tremeGENE 9 in ours, donor template (ssDNA in Liang's study, and circular plasmid in ours) etc. might contribute to the difference in basal HDR rate. Thus, future studies are still required to further decipher the effect of SpCas9 and recombinant variants such as rSpCas9 on DSB repair by HDR in mammalian cells.

An early study by Shcherbakova *et al.* reported that overexpression of *E. Coli* Rec A increases homology-directed gene conversion (Shcherbakova et al., 2000). However, our study reveals that fusion to RecA promotes DSB repair by SSA/MMEJ rather than homology-directed repair. The discrepancy between the previous study by Shcherbakova *et al.* and ours is likely due to: (1) we introduced a DSB whereas no DSB is introduced in the study by Shcherbakova *et al.*, and (2) the physical

interaction of Cas9-gRNA with the DSB site may affect repair preferences. Since the Cas9-gRNA duplex still remains bound to the DSB site after cutting (Sternberg et al., 2014) and RecA can form RecA-ssDNA filaments following 5' to 3' exonuclease processing of the broken DNA ends (Chen et al., 2008), RecA may facilitate the search for homologous sequences between the two adjacent DNA fragments. This could also explain our observation that RecA facilitates SSA/MMEJ-mediated repair of DSBs generated by CRISPR-Cas9. An increased preference for DSB repair by SSA was also achieved by overexpression of RecA and SpCas9 (**Supplementary Figure 2**), but to a lower extent compared to the SpCas9-RecA fusion. This suggests that recruitment of RecA to the DSB site is crucial for the enhanced DSB repair by cNHEJ/MMEJ/SSA.

Recently, it has been shown that transfecting cells with single-stranded non-homologous DNA oligonucleotides may alter DNA repair outcomes and facilitate NHEJ (Richardson et al., 2016). It was hypothesized that cells undergo cycles of error-free repair and cutting upon treatment with Cas9 until the occurrence of error-prone repair terminates the cycle, and that the inclusion of non-homologous oligonucleotides increases the frequency of such error-prone repair. Fusion of RecA to SpCas9 also increases the frequency of error-prone repair of DSBs but without the need for co-treatment with other external factors such as non-homologous DNA molecules or small chemical inhibitors. This renders rSpCas9 a useful tool for gene knockout applications. Furthermore, it also suggests that fine-tuning the preference of DSB repair by using recombinant Cas9 proteins could serve as an alternative strategy to enhance gene editing by CRISPR/Cas9.

In summary, we show that fusion of SpCas9 to RecA increases the efficiency of gene knockout and deletion. As such, we provide a new element to the Cas9 toolbox, but also demonstrate for the first time that mammalian repair pathway preferences can be altered by fusing active Cas9 to other catalytically active domains. This is an area with future potential in the ongoing search for methods to enhance the frequency of homology-directed repair in order to improve the utility of CRISPR/Cas9 for efficient genome editing in mammalian cells.

Acknowledgements

We thank Andrew Bassett (Head of Genome Engineering Oxford University of Oxford) for critical comments and suggestions to the study, the FACS Core Facility of Aarhus University for technical help with all FCM analyses. This work was funded by grants from the Danish Research Council for Independent Research (DFF-1337-00128, to Y.L.), the Sapere Aude Young Research Talent Prize (DFF-1335-00763, to Y.L.), the Innovation Fund Denmark (BrainStem, to Y.L.), the Lundbeck Foundation (R173-2014-1105, to Y.L.; R151-2013-14439, to L.B.), and the German Ministry of Education and Research within the VIP program (TAL-CUT 03V0261, to R.K.).

Author contributions

L.L. and Y.L. conceived the idea.

L.L., L.B., R.K. and Y.L. planned and oversaw the study

L.L., T.S.P, K.T.J and Y.L. performed experiments and analyzed the data.

L.L., K.T.J, and Y.L. prepared the manuscript and figures, and all authors revised the manuscript.

Competing Interests Statement

The authors declare no competing financial interests.

Figure Captions

Figure 1. Fusing SpCas9 to RecA increases the efficiency of gene knockout and deletion

A) Top: Schematic illustration of the SpCas9 and rSpCas9 fusion protein structures; Middle: Western blot analysis of Cas9 expression (anti FLAG tag, FLAG) and loading control (anti beta III Tubulin, b-Tub); and Bottom: Quantification of SpCas9 and rSpCas9 expression. Error bars represent one s.d. (n=3, independent transfections).

B) Quantification of indel mutations at five target sites by TIDE assay. The percentage of Indel mutations at five genomic loci (*EMX1, HEK293site2, Chr1, Chr3, and Chr18*) was quantified by TIDE assay for SpCas9 and rSpCas9. Values represent Mean and one s.d. (n=3, independent transfections). Asterisks represent $p < 0.05$ (*), and $p < 0.01$ (**), Student's T-test.

C) Schematic representation of the gRNA binding sites (top), gel electrophoresis pictures (middle), and quantification of Cas9-mediated deletion of gene segments (bottom) in *uPA* (left), *IGF1R* (middle) and *CRP* (right) by SpCas9 and rSpCas9. Error bars represent one s.d. Controls were transfected with the Cas9 expression plasmids only (n=3, independent transfections). Asterisks represent $p < 0.01$ (**), Student's T-test.

Figure 2. rSpCas9 promotes the frequency of single strand annealing-mediated repair of DSBs

A) Schematic illustration of the C-Check dual fluorescent reporter system.

B) Representative flow cytometry plots of untransfected cells, cells transfected with the C-Check vector only, and cells transfected with the C-Check vector together with either a scrambled gRNA (Ctrl), T1 gRNA or T2 gRNA, and SpCas9 or rSpCas9.

C) Quantification of the percentage of GFP+/AsRED+ cells as mean ± one s.d. (n=3, independent transfections). Asterisks (***) represents a p value < 0.001 between the two groups.

Figure 3. rSpCas9 inhibits the frequency of homology-directed repair of DSBs in TLR cells

A) Schematic illustration of the traffic light reporter system (TLR).

B) Representative flow cytometry plots of TLR cells co-transfected with either SpCas9 (left) or rSpCas9 (right) and either a scrambled gRNA (Ctrl, top) or a TLR gRNA (bottom).

C) % Venus$^+$ (HDR, left), % RFP$^+$ (NHEJ/MMEJ, middle) and the ratio of % Venus$^+$ to % RFP$^+$ (right) for TLR cells transfected with either SpCas9 or rSpCas9 and either a scrambled gRNA or the TLR gRNA. Quantified by FCM, error bars represent one s.d. (n=3, independent transfections). Asterisks (***) represents a p value < 0.001, ANOVA.

D) Schematic illustration of the TLR-like HDR plasmid reporter system.

E) Representative flow cytometry plots of HEK293T cells co-transfected with TLR gRNA, either SpCas9 (left) or rSpCas9 (right), and without (top) or with (bottom) TLR donor plasmid.

F) % of Venus$^+$/AsRED$^+$ cells. Mean and one s.d. (n=3, independent transfections).

Figure 4. Comparison of on-target activity and mismatch tolerance for eSpCas9(1.1) and eSpCas9-RecA.

A) Percentage of GFP+ out of AsRED positive cells in HEK293T cells transfected with a C-Check reporter vector, one of the Cas9 variants (eSpCas9(1.1) and eSpCas9-RecA), and one gRNA containing 0-2 mismatches to the target sites (target site 1 or target site 2) in the C-Check vector or a scrambled gRNA. Negative controls were transfected with C-Check and a pUC19 plasmid. Error bars represent one SD (n=3, independent transfections). Asterisks indicate a P value < 0.05, ANOVA.

B) The percentage of Indel mutations by eSpCas9(1.1) and eSpCas9-RecA at three genomic loci (*Chr1, Chr3, and Chr18*) in HEK293T cells was quantified by TIDE assay as mean +- one s.d. (n=3, independent transfections).

C) Percentage of Indel mutations by eSpCas9(1.1) and eSpCas9-RecA at the EMX1 target site in HEK293T cells was quantified by TIDE assay. ON, OT1, OT2, OT3, OT4, OT5,

OT6 represents on-target gRNA and gRNAs with 1-2 mismatches as indicated at the bottom. Red boxes indicate mismatch between gRNA and the target site. Background value from a group transfected with a control plasmid (pUC19) is replotted for comparison. Presented as mean +- one s.d. (n=3, independent transfections).

D)

Supplementary

Figure S1. Percentage of Indel mutations by SpCas9 and rSpCas9 at the EMX1, Chr1, Chr3, and Chr18 loci in U2OS (**A**) and EMX1 locus in HeLa cells (**B**), quantified by TIDE assay. Presented as mean and one s.d. (n=3, independent transfections). Asterisks represents $p < 0.05$ (*), $p < 0.01$ (*), Student's T-test.

Figure S2. C-Check quantification of SSA efficiency of SpCas9, SpCas9 in combination with RecA overexpression, and rSpCas9. The percentage of GFP+ out of the AsRED+ cells was quantified by flow cytometry. Data represents mean ± one s.d. (n=3, independent transfections).

Reference:

Brinkman, E.K., Chen, T., Amendola, M., van Steensel, B., (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 42, e168.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., Luc, S., Kurita, R., Nakamura, Y., Fujiwara, Y., Maeda, T., Yuan, G.C., Zhang, F., Orkin, S.H., Bauer, D.E., (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 527, 192-197.

Chen, Z., Yang, H., Pavletich, N.P., (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature 453, 489-484.

Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., Kuhn, R., (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33, 543-548.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819-823.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442-451.

Huertas, P., (2010) DNA resection in eukaryotes: deciding how to fix the break. Nat Struct Mol Biol 17, 11-16.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., (2013) RNA-programmed genome editing in human cells. eLife 2, e00471.

Krogh, B.O., Symington, L.S., (2004) Recombination proteins in yeast. Annu Rev Genet 38, 233-271.

Kuhar, R., Gwiazda, K.S., Humbert, O., Mandt, T., Pangallo, J., Brault, M., Khan, I., Maizels, N., Rawlings, D.J., Scharenberg, A.M., Certo, M.T., (2014) Novel fluorescent genome editing reporters for monitoring DNA repair pathway utilization at endonuclease-induced breaks. Nucleic Acids Res 42, e4.

Liang, X., Potter, J., Kumar, S., Ravinder, N., Chesnut, J.D., (2017) Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. J Biotechnol 241, 136-146.

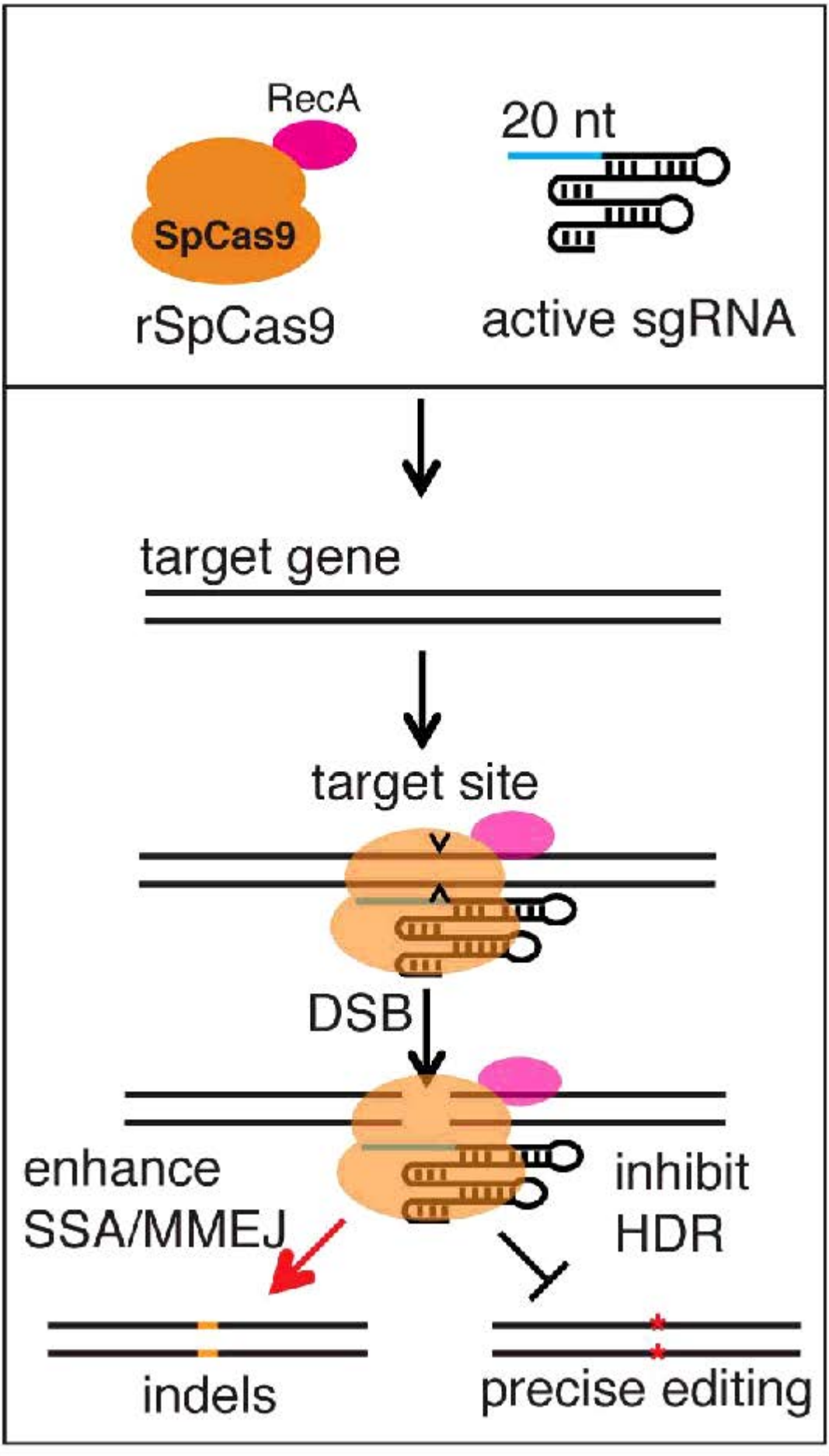
Lieber, M.R., (2008) The mechanism of human nonhomologous DNA end joining. J Biol Chem 283, 1-5.

Luo, Y., Li, J., Liu, Y., Lin, L., Du, Y., Li, S., Yang, H., Vajta, G., Callesen, H., Bolund, L., Sorensen, C.B., (2011) High efficiency of BRCA1 knockout using rAAV-mediated gene targeting: developing a pig model for breast cancer. Transgenic Res 20, 975-988.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., Dicarlo, J.E., Norville, J.E., Church, G.M., (2013) RNA-Guided Human Genome Engineering via Cas9. Science.

Mandal, P.K., Ferreira, L.M., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanac, V., Garrison, B.S., Stortchevoi, A., Bryder, D., Musunuru, K., Brand, H., Tager, A.M., Allen,

T.M., Talkowski, M.E., Rossi, D.J., Cowan, C.A., (2014) Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. Cell Stem Cell 15, 643-652.
McVey, M., Lee, S.E., (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet 24, 529-538.
Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152, 1173-1183.
Richardson, C.D., Ray, G.J., Bray, N.L., Corn, J.E., (2016) Non-homologous DNA increases gene disruption efficiency by altering DNA repair outcomes. Nature communications 7, 12463.
Shcherbakova, O.G., Lanzov, V.A., Ogawa, H., Filatov, M.V., (2000) Overexpression of bacterial RecA protein stimulates homologous recombination in somatic mammalian cells. Mutat Res 459, 65-71.
Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., Zhang, F., (2016) Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84-88.
Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., Doudna, J.A., (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507, 62-67.
Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., Aryee, M.J., Joung, J.K., (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33, 187-197.
van Overbeek, M., Capurso, D., Carter, M.M., Thompson, M.S., Frias, E., Russ, C., Reece-Hoyes, J.S., Nye, C., Gradia, S., Vidal, B., Zheng, J., Hoffman, G.R., Fuller, C.K., May, A.P., (2016) DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. Mol Cell 63, 633-646.
Yasuda, H., Kim, E., Reza, A.M., Kim, J.H., (2016) A highly efficient method for enriching TALEN or CRISPR/Cas9-edited mutant cells. J Genet Genomics 43, 705-708.
Zhou, Y., Liu, Y., Hussmann, D., Brogger, P., Al-Saaidi, R.A., Tan, S., Lin, L., Petersen, T.S., Zhou, G.Q., Bross, P., Aagaard, L., Klein, T., Ronn, S.G., Pedersen, H.D., Bolund, L., Nielsen, A.L., Sorensen, C.B., Luo, Y., (2016) Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. Cell Mol Life Sci 73, 2543-2563.


RecA
SpCas9
rSpCas9
20 nt
active sgRNA
target gene
target site
DSB
enhance
SSA/MMEJ
inhibit
HDR
indels
precise editing

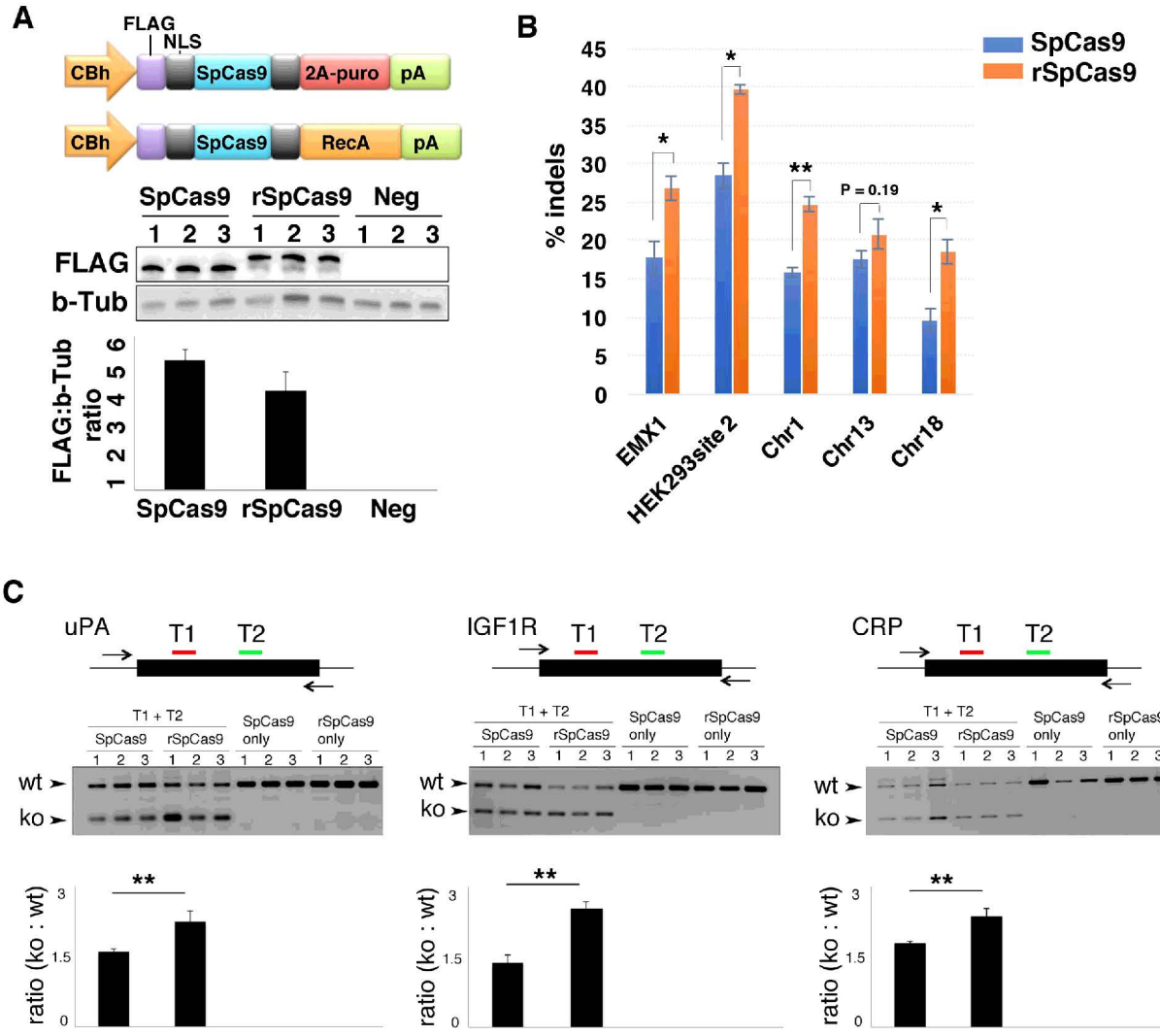

A
FLAG
NLS
CBh
SpCas9
2A-puro
pA
CBh
SpCas9
RecA
pA
SpCas9
rSpCas9
Neg
1 2 3 1 2 3 1 2 3
FLAG
b-Tub
FLAG:b-Tub ratio
1 2 3 4 5 6
SpCas9
rSpCas9
Neg
B
SpCas9
rSpCas9
% indels
45
40
35
30
25
20
15
10
5
0
*
*
**
P = 0.19
*
EMX1
HEK293site 2
Chr1
Chr13
Chr18
C
uPA
T1
T2
IGF1R
T1
T2
CRP
T1
T2
T1 + T2
SpCas9
rSpCas9
SpCas9 only
rSpCas9 only
1 2 3 1 2 3 1 2 3 1 2 3
wt
ko
ratio (ko : wt)
3
1.5
0
**
**
**

Figure 1

Figure 2

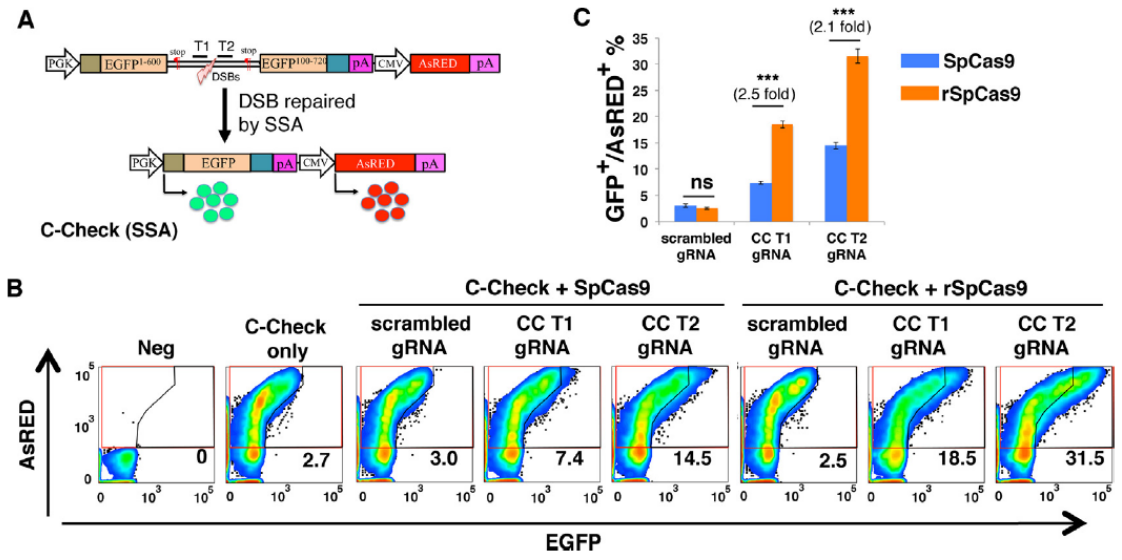


Figure 2: rSpCas9 promotes the frequency of single strand annealing-mediated repair of DSBs. (A) Schematic illustration of the C-Check dual fluorescent reporter system. (B)Representative flow cytometry plots of untransfected cells, cells transfected with the C-Check vector only, and cells transfected with the C-Check vector together with either a scrambled gRNA (Ctrl), T1 gRNA or T2 gRNA, and SpCas9 or rSpCas9. (C) Quantification of the percentage of GFP+/AsRED+ cells as mean ± one s.d. (n = 3, independent transfections). Asterisks (***) represents a p value < 0.001 between the two groups.

Figure 3

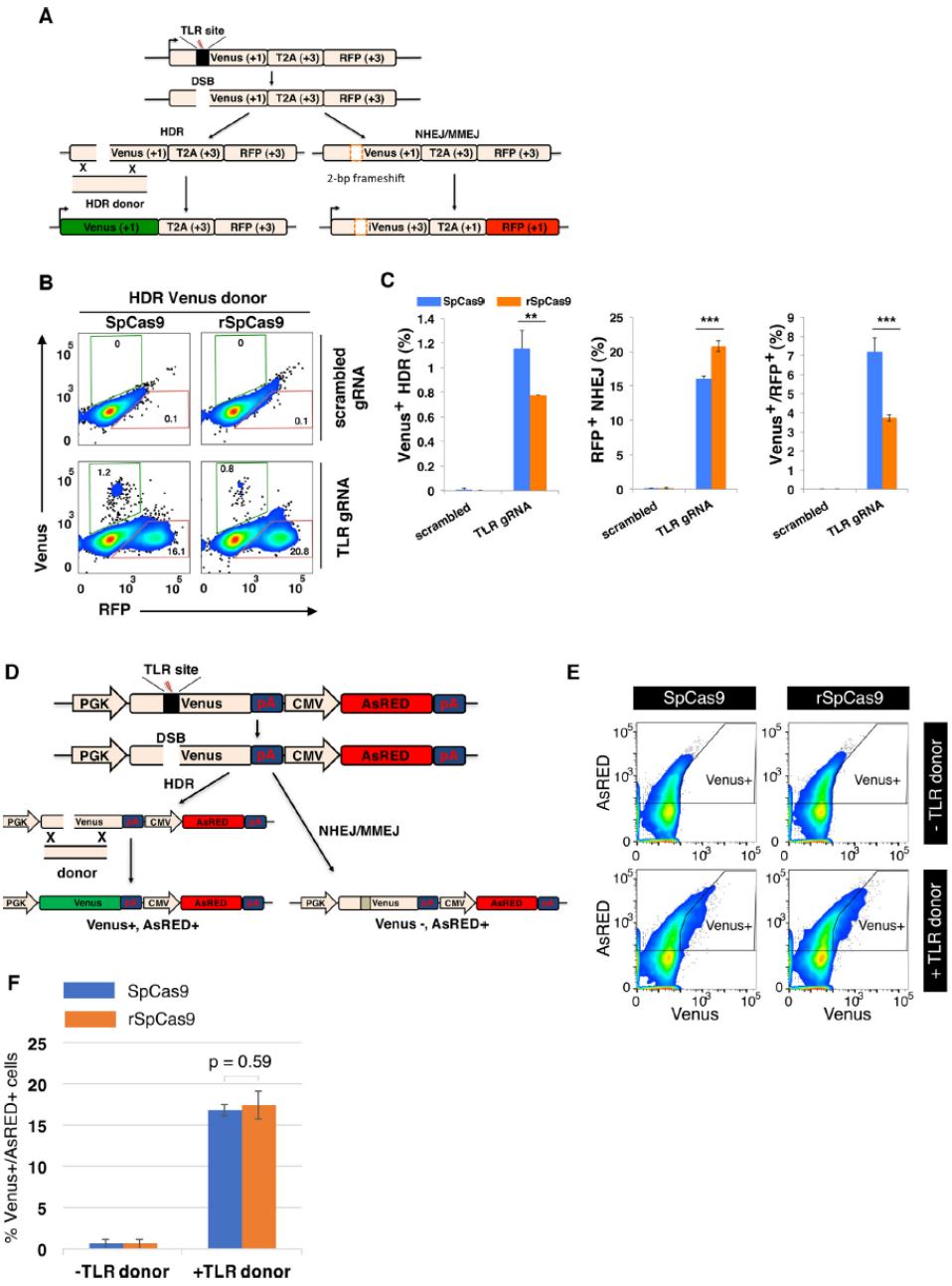
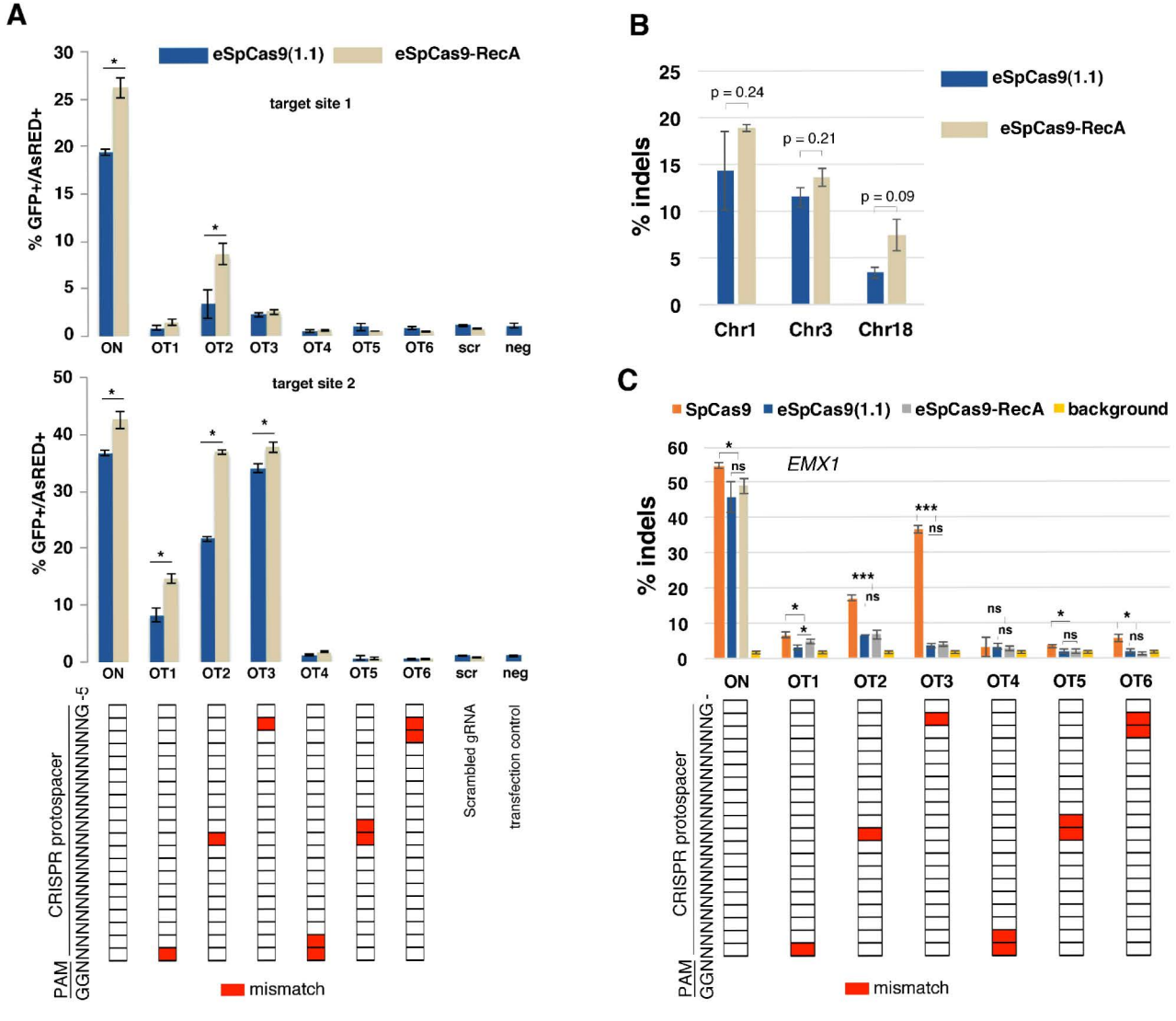


Figure 3: rSpCas9 inhibits the frequency of homology-directed repair of DSBs in TLR cells. (A) Schematic illustration of the traffic light reporter system (TLR). (B) Representative flow cytometry plots of TLR cells co-transfected with either SpCas9 (left) or rSpCas9 (right) and either a scrambled gRNA (Ctrl, top) or a TLR gRNA (bottom). (C) % Venus+ (HDR, left), % RFP+ (NHEJ/MMEJ, middle) and the ratio of % Venus+ to % RFP+ (right) for TLR cells transfected with either SpCas9 or rSpCas9 and either a scrambled gRNA or the TLR gRNA. Quantified by FCM, error bars represent one s.d. (n = 3, independent transfections). Asterisks (***) represents a p value < 0.001, ANOVA. (D) Schematic illustration of the TLR-like HDR plasmid reporter system. (E) Representative flow cytometry plots of HEK293T cells co-transfected with TLR gRNA, either SpCas9 (left) or rSpCas9 (right), and without (top) or with (bottom) TLR donor plasmid. (F) % of Venus+/AsRED+ cells. Mean and one s.d. (n = 3, independent transfections).

Figure 4


A
eSpCas9(1.1)
eSpCas9-RecA
target site 1
% GFP+/AsRED+
ON
OT1
OT2
OT3
OT4
OT5
OT6
scr
neg
target site 2
CRISPR protospacer
PAM
GGNNNNNNNNNNNNNNNNNNG -5
Scrambled gRNA
transfection control
mismatch
B
% indels
p = 0.24
p = 0.21
p = 0.09
Chr1
Chr3
Chr18
C
SpCas9
eSpCas9(1.1)
eSpCas9-RecA
background
EMX1
ns
GGNNNNNNNNNNNNNNNNNNG -