Cas9-sgRNA target residence alters DNA double strand break repair pathway choices in CRISPR/Cas9 genome editing

An-Yong Xie (anyongxie@zju.edu.cn)
Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine https://orcid.org/0000-0002-6608-2550

Si-Cheng Liu
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine

Yi-Li Feng
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine

Xiu-Na Sun
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine

Ruo-Dan Chen
Department of Biochemistry and Molecular Biology, Zhejiang University School of Medicine

Qian Liu
Institute of Translational Medicine, Zhejiang University School of Medicine

Jing-Jing Xiao
Institute of Translational Medicine, Zhejiang University School of Medicine

Ji-Feng Xiang
Institute of Translational Medicine, Zhejiang University School of Medicine

Guo-Qiao Chen
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine

Yi Yang
Institute of Translational Medicine, Zhejiang University School of Medicine

Chao Lou
Shurui Tech Ltd

Hao-Dan Li
Shurui Tech Ltd

Shi-Ming Xu
Institute of Translational Medicine, Zhejiang University School of Medicine

Hui Lin
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine

Article

Keywords:

Posted Date: February 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-266312/v1

Abstract

Given its different target-binding affinity and varying residence time after DNA cleavage, the Cas9-sgRNA complex that remains bound to cleaved DNA may influence DNA double strand break (DSB) repair pathway choices, contributing to highly heterogeneous mutations in genome editing. Here, we found that DSB repair pathway choices vary significantly at different sites for Cas9-induced DSBs. Reduced target-binding affinity of Cas9-sgRNA promotes a bias toward classical NHEJ (c-NHEJ), but inactivation of c-NHEJ aggravates off-target activities. By dislodging Cas9-sgRNA from the cleaved target, local DNA replication suppresses c-NHEJ in favor of other DSB repair pathways by altering DSB configuration and repair timing. It also results in sister chromatid fusion, which may promote chromosomal breakage-fusion-bridge cycles. This suggests that varying Cas9-sgRNA target-binding affinity or residence duration modulate DSB repair pathway choices for Cas9-induced DSBs and widen the mutational spectra in CRISPR/Cas9 genome editing.

Introduction

CRISPR/Cas9 genome editing relies on the binding of the Cas9 nuclease, in complex with a single guide RNA (sgRNA), to a DNA target to induce a site-specific DNA double strand break (DSB) and its subsequent repair[1,2]. Upon DSB induction by Cas9, different repair pathways compete for DSB repair, generating the desired DNA edits including substitutions, insertions, deletions or translocations among varieties of repair products[3]. The two major DSB repair mechanisms in mammalian cells include non-homologous end joining (NHEJ) and homology-directed repair (HDR). While classical NHEJ (c-NHEJ) is the primary NHEJ pathway, alternative NHEJ (a-NHEJ) could also be employed to re-ligate the ends of DSBs if either of the core NHEJ factors including DNA-PKcs, Ku70/Ku80, XRCC4/DNA ligase 4 is deficient or not engaged[4]. If the ends of DSBs are readily ligatable, such as Cas9-induced blunt ends and I-SceI-induced 3'-overhanging ends, c-NHEJ generates largely accurate end-joining products whereas a-NHEJ remains mostly mutagenic[5,6]. Additionally, using homologous sequences as a template, HDR is the preferred pathway for accurate substitutions and insertions in CRISPR/Cas9 genome editing.

The DSB repair pathway choice is governed by a host of factors, including cell cycle stage, DNA end configurations, surrounding chromatin context and local DNA metabolism[7]. Uniqueness in DSB induction by CRISPR/Cas9 may also participate in this regulation[8]. In CRISPR/Cas9 genome editing, targeting Cas9 to a given site is mediated by several interactions, including the contacts between Cas9 and the protospacer adjacent motif (PAM) of the target, the base pairing of the sgRNA spacer with target strand and non-specific interactions between Cas9 and target DNA[2]. *In vitro* and *in vivo* studies have indicated that these interactions entail strong and persistent binding of the Cas9-sgRNA complex to its target and help maintain its target residence for hours (h) even after Cas9-induced DNA cleavage[9–13]. Repair kinetics reveals that repair of Cas9-induced DSBs is generally slow and often lasts for more than 20 h in mammalian cells; this is likely due to the concealing of DSBs by the Cas9-sgRNA complex retained at the cleaved DNA[13,14]. Owing to intrinsic disparity in the interactions that mediate the binding of Cas9-sgRNA

to its target, the binding affinity of Cas9-sgRNA varies at different sites along with altered target residence. It is likely that Cas9-sgRNA could be spontaneously released from its target, or may encounter local DNA replication, transcription or chromatin remodeling, leading to the release of Cas9-sgRNA from cleaved DNA and the exposure of Cas9-induced DSBs[8,15–18]. However, it is unclear what effect either form of Cas9-sgRNA dissociation from its cleaved targets has on repair of Cas9-induced DSBs.

Nevertheless, the interplay between these regulation factors in DSB repair, in combination with competing choices of repair pathway, lead to varieties of insertions, deletions and/or chromosomal rearrangements in repair products and make it difficult to efficiently generate the desired products or accurately predict mutations in CRISPR/Cas9 genome editing. We hypothesize that the target-binding affinity of Cas9-sgRNA or its post-cleavage residence at the target may influence DSB repair pathway choice, contributing to mutational heterogeneity in CRISPR/Cas9 genome editing. Here, we demonstrated that the target-binding affinity or target residence of Cas9-sgRNA regulates the repair pathway choice for Cas9-induced DSBs by altering repair timing and DSB configuration.

Results

Repair of Cas9-induced DSBs involves c-NHEJ to varying degrees at different targets

If persistent binding and residence of Cas9-sgRNA at its target after DNA cleavage influence DSB repair pathway choice, the involvement of c-NHEJ in repair of Cas9-induced DSBs would change between targets with different binding affinities and residence duration for Cas9-sgRNA. Inactivation of c-NHEJ would thus lead to varying degrees of mutagenic NHEJ (m-NHEJ) stimulation at these sites. To test this hypothesis, we used Cas9-sgRNA to induce site-specific DSBs in an NHEJ reporter integrated in the genome of mouse embryonic stem cells (mESC) as done before[19], and analyzed the extent of the c-NHEJ involvement in repair of Cas9-induced DSBs at different sites (Fig. 1a). In this NHEJ reporter, only m-NHEJ can lead to *GFP*$^{+}$ cells, and the frequency of *GFP*$^{+}$ cells thus represents the relative efficiency of m-NHEJ as part of both c-NHEJ and a-NHEJ (Fig. 1a). We found that neither DNA-PKcs inhibition by NU7441 nor *XRCC4* deletion changed the frequencies of Cas9-induced m-NHEJ at the two sites targeted by the sgRNA gEJ_W3 or gEJ_W7, while the inactivation of c-NHEJ inhibited Cas9-induced m-NHEJ at the four sites targeted by gEJ_C5, gEJ_W4, gEJ_W5 and gEJ_W6 to different extents (Fig. 1b). This indicates that the participation of c-NHEJ varies in repair of Cas9-induced DSBs at different targets.

Surprisingly, no stimulation of m-NHEJ was observed in repair of Cas9-induced DSBs by inactivation of c-NHEJ at some targets. To determine whether this was unique to repair of Cas9-induced DSBs, we used the same NHEJ reporter cells but with the first I-SceI recognition site being destroyed by CRISPR/Cas9 genome editing to imitate Cas9-induced single DNA breakage in the reporter so that *GFP*$^{+}$ cells can only be generated by m-NHEJ in repair of I-SceI-induced DSBs (**Fig. S1a**). Inhibition of c-NHEJ with NU7441 increases I-SceI-induced m-NHEJ by more than 2-fold (**Fig. S1b,c**). We then wondered why inactivation of

c-NHEJ stimulates I-SceI-induced m-NHEJ, but not Cas9-induced m-NHEJ. Given that the repair of Cas9-induced DSBs is mostly accurate[5,6], it is possible that Cas9-sgRNA in the cells is abundant and could recleave target DNA until the targeted sites are mutated by NHEJ, thus accumulating m-NHEJ events (Fig. 1c). Inactivation of c-NHEJ would then reduce m-NHEJ events (Fig. 1b). By reducing the amount of Cas9 or sgRNA transfected, we predicted that the target recleavage would be reduced and Cas9-induced m-NHEJ would be stimulated by DNA-PKcs inhibition (Fig. 1c). We found that overall Cas9-induced m-NHEJ was reduced with a low amount of Cas9-gEJ_W6 as expected (**Fig. S1d**). While NU7441 suppressed m-NHEJ induced by a high amount of Cas9 and gEJ_W6 typically at 0.25μg each, the inhibitor started to stimulate m-NHEJ when the amount of Cas9 and gEJ_W6 was both reduced to 0.001μg (**Fig. S1d,e**). In contrast, DNA-PKcs inhibition did not alter m-NHEJ induced by Cas9 and gEJ_W7 at an amount ranging from 0.25μg to 0.0001μg, indicating no involvement of c-NHEJ in repair of Cas9-induced DSB at the gEJ_W7 target (**Fig. S1f,g**). At the two sites targeted by gEJ_C5 or gEJ_W5 with the transfection amount of Cas9-sgRNA at 0.001μg, DNA-PKcs inhibition and *XRCC4* deletion still inhibited Cas9-induced m-NHEJ but this inhibition was reduced (Fig. 1d). However, this c-NHEJ inactivation abolished or even reversed the suppression of Cas9-induced m-NHEJ to stimulation at the targets with gEJ_W4 and gEJ_W6 and remained to exert no effect on Cas9-induced m-NHEJ at the two sites with gEJ_W3 or gEJ_W7 (Fig. 1d). This further confirms that the involvement of c-NHEJ varies at different targets in repair of Cas9-induced DSBs.

We also used a single-copy *GFP* reporter gene and two natural genome loci to further assess variable involvement of c-NHEJ in repair of Cas9-induced DSBs (**Fig. S2a**). Among 7 sgRNAs designed for *GFP* gene knock-out (KO), all converted over 60% GFP^+ cells into GFP^- cells (**Fig. S2b**). DNA-PKcs inhibition by NU7441 had no significant effect on GFP KO at all 7 sites (**Fig. S2b**). To limit recleavage by Cas9, we transfected mESC with much lower amount of Cas9-sgRNA at 0.001μg. While limited recleavage lowered CRISPR/Cas9-mediated *GFP* KO editing, DNA-PKcs inhibition slightly increased *GFP* KO at 1 site targeted by gG_C7, but made little change on *GFP* KO at the other 6 sites (**Fig. S2c**). Using targeted PCR amplicon deep sequencing, we additionally measured Cas9-induced m-NHEJ at two natural genome loci *Cola1* and *Rosa26* in mESC. We found that DNA-PKcs inhibition reduced the editing efficiency at the sites targeted by *Cola1* gC2 and *Rosa26* gR3, stimulated at the sites by *Cola1* gC3, and had minimal effect at the rest of the sites including gC1 and gC4 for *Cola1* and gR1, gR2 and gR4 for *Rosa26* (**Fig. S2d,e**). These results together indicated variable involvement of c-NHEJ in CRISPR/Cas9 genome editing at different sites or even no involvement of c-NHEJ at some sites.

Inactivation of c-NHEJ induces varying stimulation of Cas9-induced HDR among targets

Inactivation of c-NHEJ also channels DSBs for repair from c-NHEJ towards HDR with a certain probability in each repair round, causing the increased use of HDR regardless of DNA recleavage by Cas9. Thus, using a single-copy HDR reporter in mESC[20], we analyzed the impact of c-NHEJ inactivation on HDR induced by Cas9-sgRNA targeting at 5 different sites surrounding the I-SceI site in the reporter (Fig. 1e). Like I-SceI-induced HDR, Cas9-induced HDR was increased by NU7441 at the sites targeted by gHR_C1,

gHR$_C$2 and gHR$_C$3 and the extent of this stimulation was different among these three targets (Fig. 1f). Surprisingly, DNA-PKcs inhibition did not elevate HDR induced by Cas9-gHR$_C$4 and Cas9-gHR$_C$5 (Fig. 1f). We also used CRISPR/Cas9 gene editing to generate isogenic wild-type, *DNA-PKcs*$^{-/-}$ and *Ku80*$^{-/-}$ mESC clones containing the HDR reporter (**Fig. S3a,b**). Using one of these clones, along with isogenic *XRCC4*$^{+/+}$ and *XRCC4*$^{-/-}$ HDR reporter mESC previously established[21], we found that deletion of *DNA-PKcs*, *Ku80 or XRCC4* significantly enhanced HDR induced by gHR$_C$1, gHR$_C$2 or gHR$_C$3 in complex with SpCas9, as well as HDR induced by I-SceI (Figs. 1g and S3c). The extents of the HDR stimulation varied among these three different targets (Fig. 1g). However, deletion of *DNA-PKcs* or *Ku80*, like DNA-PKcs inhibition, caused little increase in HDR induced by Cas9-gHR$_C$4 and Cas9-gHR$_C$5 whereas *XRCC4* deletion had modest stimulation (Fig. 1g). This suggests that inactivation of c-NHEJ stimulates Cas9-induced HDR to varying extents among targets due to different involvement of c-NHEJ at these targets. At the targets where Cas9-induced HDR is not stimulated by either inhibition of DNA-PKcs or deletion of *DNA-PKcs* or *Ku80*, it is possible that repair of Cas9-induced DSBs by NHEJ does not involve DNA-PKcs or Ku70/Ku80.

Reduced target-binding affinity of Cas9-sgRNA shifts repair of Cas9-induced DSBs towards c-NHEJ

To further determine whether c-NHEJ repair of Cas9-induced DSBs is influenced by target-binding affinity or post-cleavage target residency of Cas9-sgRNA, we mutated either sgRNA or SpCas9 for two sites targeted by gEJ$_C$5 and gEJ$_W$7 in the NHEJ reporter to reduce the target-binding affinity of the Cas9-sgRNA complex. Consistent with previous observation that weakened Cas9-sgRNA target binding generally reduces the efficiency of genome editing[22–25], Cas9-induced m-NHEJ was less efficient with mismatched or truncated sgRNA variants (i.e. C2A, T15A and 16nt for gEJ$_C$5, and A1T, A4C and T15A for gEJ$_W$7) and with SpCas9 variants eSpCas9 and SpCas9-HF1, both of which were engineered to have lower binding affinity and higher specificity to target DNA (Fig. 2a). As in Fig. 1b, *XRCC4* deletion and DNA-PKcs inhibition reduced Cas9-induced m-NHEJ at the site targeted with gEJ$_C$5, again suggesting significant DNA recleavage by Cas9 (Fig. 2a). In contrast, the gEJ$_C$5 variants (C2A, T15A and 16nt), eSpCas9 and SpCas9-HF1 alleviated or even reversed this reduction (Fig. 2a). This suggests that in repair of Cas9-induced DSBs, the weaker the Cas9-sgRNA target binding is, the less Cas9-induced DNA recleavage occurs and the more preferentially c-NHEJ is engaged. At the site targeted by gEJ$_W$7, neither DNA-PKcs inhibition nor *XRCC4* deletion had effect on Cas9-induced m-NHEJ as shown in Fig. 1b (Fig. 2a), indicating no engagement of c-NHEJ at this site. However, gEJ$_W$7 mismatch variants (i.e. A1T, A4C and T15A), eSpCas9 and SpCas9-HF1 allowed significant stimulation of Cas9-induced m-NHEJ up to 3.5-fold by DNA-PKcs inhibition or *XRCC4* deletion (Fig. 2a), suggesting that weakened target binding of Cas9-sgRNA would induce strong involvement of c-NHEJ.

Using the integrated single-copy GFP gene (**Fig. S2a**), we also analyzed whether weakened Cas9-sgRNA target binding promoted c-NHEJ, which is largely accurate in repair of Cas9-induced DSBs, and suppressed GFP KO. As expected, Cas9-mediated GFP KO at the gG$_W$5 and gG$_C$15 sites was less efficient with mismatched sgRNA mutants (i.e. C1T, T15A and C16G for gG$_W$5, and A2C, A15T and G16C for

gG_C15) or SpCas9 variants eSpCas9 and xCas9 (**Fig. S4a,b**). Although DNA-PKcs inhibition had no effect on GFP KO mediated by wild-type SpCas9-sgRNA at either site (**Fig. S2bc**), it stimulated GFP KO mediated by these Cas9-sgRNA variants (**Fig. S4a,b**). Using endogenous genomic loci, we also found that the editing efficiency with the mismatch variants of *Cola1* gC4 (i.e. C1T and G16C) and *Rosa26* gR4 (i.e. A16T) was reduced due to weaker target binding, but elevated by DNA-PKcs inhibition (Fig. 2b,c), although DNA-PKcs inhibition had minimal effect on the editing efficiency at the sites targeted by *Cola1* gC4 and *Rosa26* gR4 (**Fig. S2d,e**). This indicates that reduced Cas9-sgRNA target binding promotes c-NHEJ. Taken together, these results suggest that the variable target-binding affinity of Cas9-sgRNA may determine the extent of c-NHEJ involvement in repair of Cas9-induced DSBs.

Reduced Cas9-sgRNA target-binding affinity enhances stimulatory effect of c-NHEJ inactivation on Cas9-induced HDR

Consistent with previous studies[21,26], inactivation of c-NHEJ stimulates HDR induced by CRISPR nucleases as well as I-SceI (Figs. 1f,g and S3c). We expected that this stimulatory effect would be further enhanced if HDR were induced by Cas9-sgRNA variants with reduced target-binding affinity because reducing Cas9-sgRNA target binding promotes c-NHEJ. Therefore, we compared HDR induced by mutated Cas9-sgRNA between cells proficient and deficient in c-NHEJ. Due to reduced efficiency of DNA cutting, Cas9-induced HDR was generally less efficient with mismatched or truncated sgRNA variants (i.e. G1C, G2C and 17nt for gHR_C4, and A1T, C2A and 17nt for gHR_C2) and SpCas9 variants eSpCas9, SpCas9-HF1 and xCas9 (Fig. 2d,e). At the site targeted by gHR_C4, where Cas9-induced HDR was not affected by either DNA-PKcs inhibition, *DNA-PKcs* deletion or *Ku80* deletion, and modestly stimulated by deletion of *XRCC4*, Cas9-induced HDR with the sgRNA variants or the SpCas9 variants was significantly elevated by inactivating c-NHEJ (Fig. 2d). At the site targeted by gHR_C2, where Cas9-induced HDR was increased by DNA-PKcs inhibition or deletion of *DNA-PKcs*, *Ku80* or *XRCC4*, respectively, stimulation of Cas9-induced HDR by inactivation of c-NHEJ was further enhanced with the Cas9-sgRNA variants (Fig. 2e). This confirms that reducing target-binding affinity of Cas9-sgRNA promotes c-NHEJ, providing the basis for the enhanced stimulatory effect of c-NHEJ inactivation on Cas9-induced HDR.

c-NHEJ deficiency increases off-target activity of CRISPR/Cas9

As mismatches in base pairing between sgRNA and off-target sites reduce the binding affinity of Cas9-sgRNA to off-target sites and lower the occurrence of DNA recleavage at these sites, it is anticipated that c-NHEJ would be engaged proportionally more at off-target sites than at on-target sites. Thus, inactivation of c-NHEJ would also increase the use of m-NHEJ at off-target sites, exacerbating off-target effects in CRISPR/Cas9 genome editing. To test this hypothesis, we analyzed the effects of DNA-PKcs inhibition and *XRCC4* deletion on off-target activities of Cas9 at 7 potential off-target sites for gPnpla3 and 6 potential off-target sites for gMertk. We found that both NU7441 and *XRCC4* deletion slightly reduced on-target editing by Cas9-gPnpla3 and Cas9-gMertk, indicating significant on-target DNA

recleavage. In contrast, off-target editing efficiencies were not reduced by either DNA-PKcs inhibition or *XRCC4* deletion, but even increased at some of these off-target sites, e.g. OT5 and OT7 for gPnpla3 and OT1 for gMertk (Fig. 3a,b). This suggests that inactivation of c-NHEJ exerts more stimulatory effect on off-target editing than on on-target editing, aggravating off-target effect in CRISPR/Cas9 genome editing.

Chemical inhibition and genetic inactivation of c-NHEJ are often used to increase the efficiency of Cas9-induced HDR-mediated gene knock-in or replacement[27–31]. Given that NU7441 stimulated Cas9-induced HDR in the HDR reporter at the targets by gHR_C1 and gHR_C2 (Fig. 1f), we also performed off-target analysis through amplicon deep sequencing of 6 potential off-target sites for Cas9-gHR_C1 and Cas9-gHR_C2 respectively. The on-target editing efficiencies of Cas9-gHR_C1 and Cas9-gHR_C2 were slightly lowered by DNA-PKcs inhibition, again indicating significant on-target DNA recleavage. Unlike on-target editing, the editing efficiencies at the 6 off-target sites were not reduced by NU7441. Instead, relative off-target effects represented by the ratio of the off-target activity to the on-target activity were elevated (Fig. 3c,d). We also found that relative off-target effects were significantly enhanced at the off-target sites of gHR_C1 and gHR_C2 in *XRCC4*$^{-/-}$ cells (Fig. 3c,d). This again suggests that inactivation of c-NHEJ could exacerbate off-target effects in CRISPR/Cas9 genome editing.

Local transcription does not affect involvement of c-NHEJ in repair of Cas9-induced DSBs

After DNA cleavage by Cas9-sgRNA, some ends can be exposed by spontaneous dissociation of Cas9-sgRNA from the cleaved target DNA and engage c-NHEJ, and the others may require local transcription machinery or DNA replication fork to dislodge the target-bound Cas9-sgRNA complex. If the dislodgement of Cas9-sgRNA by transcription underlies the disengagement of c-NHEJ, we reasoned that the gene silencing activity (i.e. the transcription-blocking capability) of catalytically dead Cas9 (dCas9)-sgRNA at a target would be negatively correlated with the extent of c-NHEJ participation in repair of Cas9-induced DSBs at the same site. Thus, using the integrated single-copy *GFP* in mESC, we examined any potential correlation between the gene silencing activities of dSpCas9 at various sites and c-NHEJ involvement in SpCas9-induced *GFP* KO editing at the same sites. While dSpCas9-sgRNA exhibited variable gene silencing activities at many of these sites (Fig. 4a), the effect of DNA-PKcs inhibition on *GFP* KO varied from no effect to 4-fold stimulation among targets (Figs. 4b and S4c). No apparent bias towards either template strand of transcription or non-template strand was detected in both transcriptional silencing and DNA-PKcs involvement (Fig. 4c). Importantly, no correlation was observed between the gene silencing activities by dSpCas9-sgRNA and the DNA-PKcs participation in SpCas9-induced m-NHEJ (Fig. 4d; P = 0.78).

To further determine the effect of the collision between local transcription and Cas9-sgRNA on the engagement of c-NHEJ for Cas9-induced m-NHEJ, we used catalytically dead *Staphylococcus aureus* Cas9 (dSaCas9)-sgRNA to block the translocating RNA polymerase (RNAP), preventing its collision with downstream site-specific DSBs induced by SpCas9 (Fig. 4e). Among 6 sgRNAs tested for transcriptional blockage, only $gSaG_W1$ and $gSaG_W2$, in complex with dSaCas9, mediated efficient gene silencing,

indicating a strong capability of blocking RNAP (Fig. 4e). The *GFP* KO frequency induced by SpCas9-gG_C4 and SpCas9-gG_W5 at a transfection amount of 0.25μg or 0.001μg was not altered by DNA-PKcs inhibition, and this non-effect was not changed by co-transfection with either dSaCas9-$gSaG_W1$ or dSaCas9-$gSaG_W2$ (Fig. 4f). This suggests that transcription blockage by dSaCas9-sgRNA would not affect the extent of c-NHEJ engagement in repair of SpCas9-induced DSBs and further confirms that a collision with local transcription do not control c-NHEJ engagement in repair of Cas9-induced DSBs.

Local replication abolishes c-NHEJ engagement at Cas9-induced DSBs

To address whether a replication fork could locally dissociate the Cas9-sgRNA complex from cleaved target DNA, disfavoring c-NHEJ, we transfected HEK293 cells with a plasmid containing an SV40 origin-ATG-GFP-P2A-FLuc NHEJ reporter cassette, together with expression plasmids for SV40 large T antigen (LT), I-SceI or the SpCas9-gEJ_W10 complex, and the *Renilla luciferase* (*RLuc*) gene as internal control. The expression of SV40 *LT* drives bidirectional DNA replication *via* the SV40 origin, and the expression of I-SceI or SpCas9-gEJ_W10 induces a site-specific DSB between the “Koz-ATG” and the “ATG-GFP-P2A-FLuc” (Fig. 5a). Repair of I-SceI- or Cas9-induced DSBs by m-NHEJ can proportionally reframe the originally out-of-frame *firefly luciferase* (*FLuc*) gene in the NHEJ reporter plasmids to in-frame in the cells and induce synthesis of active firefly luciferase. The frequency of this NHEJ repair can thus be measured as a relative ratio of FLuc to RLuc by luminescence assays. Treatment with NU7441 reduced I-SceI-induced NHEJ in this assay, but the level of this reduction was similar with or without the expression of SV40 *LT* (Fig. 5a), suggesting little effect of SV40 LT on I-SceI-induced NHEJ. However, while Cas9-induced NHEJ was also suppressed by NU7441, DNA replication initiated by SV40 LT significantly attenuated this repressive effect (Fig. 5a). This suggests that local DNA replication driven by SV40 LT might inhibit the involvement of c-NHEJ in repair of Cas9-induced DSBs.

We also wondered whether a collision with local DNA replication would prevent potential competition between c-NHEJ and HDR in repair of Cas9-induced DSBs by blocking c-NHEJ engagement, thus removing the stimulatory effect of DNA-PKcs inhibition on Cas9-induced HDR. Using U2OS cells containing an integrated single-copy HDR reporter (Fig. 5b), in which an SV40 origin is located between TrGFP and I-SceI-GFP, we analyzed the effect of DNA-PKcs inhibition on HDR induced by I-SceI and SpCas9. DNA-PKcs inhibition stimulated HDR induced by SpCas9 in complex with gHR_C1, gHR_C2, gHR_C3, gHR_C4 and gHR_C5 to different degrees, as well as by I-SceI (Figs. 5b and S5a), indicating variable but detectable engagement of the competing c-NHEJ pathway in repair of these I-SceI- or Cas9-induced DSBs. After expression of SV40 *LT*, both I-SceI- and Cas9-induced HDR were repressed in a dose-dependent manner (Figs. 5b and S5b,c). DNA-PKcs inhibition stimulated I-SceI- or Cas9-induced HDR, and the expression of SV40 *LT* attenuated this stimulation or even abolished it at a transfection amount of 0.16 μg (1/5 of total DNA transfected) (Figs. 5b and S5b,c). This suggests that local DNA replication driven by SV40 LT could collide with both I-SceI and SpCas9-sgRNA after DNA cleavage to dislodge I-SceI

and Cas9-sgRNA from its cleaved target and restrict the engagement of c-NHEJ in repair of exposed DSBs.

By restricting c-NHEJ due to a collision with replication fork, the DSB repair pathway choice would be biased toward HDR. To test this possibility, we used the HDR reporter to measure the bias between HDR and NHEJ in repair of the same DSB induced by SpCas9-sgRNA that tightly bound with its target and by SpCas9-sgRNA variants with weakened target-binding affinities. In the HDR reporter, repair of the same Cas9-induced DSBs around the I-SceI site of I-SceI-GFP by HDR generates the "WT GFP", whereas NHEJ generates "mutant GFP" due to disruption of the I-SceI site (Fig. 5c). We can separate these two repair outcomes in mESC by nested PCR and evaluate the HDR bias (i.e. the ratio of HDR to total edited) by deep sequence analysis. After HDR and NHEJ induced by SpCas9-gHR_C4 in mESC, we found the HDR bias was nearly 3-fold lower with gHR_C4 variants (G1C and 17nt) than with gHR_C4, indicating a reduced HDR preference when the SpCas9-sgRNA binding affinity to its target is weakened (Fig. 5d). At the site targeted by gHR_C2, where the HDR stimulation by DNA-PKcs inhibition was nearly abolished by the expression of SV40 *LT* in U2OS cells (Fig. 5b), SV40 *LT* expression increased the HDR bias, indicating a repair pathway shift from NHEJ to HDR (Fig. 5d). Therefore, for Cas9-sgRNA tightly bound to its cleaved target in cells, it is likely that its release from cleaved DNA requires local DNA replication, thus promoting HDR over NHEJ in repair of exposed DSBs.

While spontaneous dissociation of Cas9-sgRNA from cleaved DNA results in a conventional two-ended DSB, DNA replication that releases Cas9-sgRNA from its cleaved target may generate a three-ended DSB, with the leading strand likely forming a blunt end on one sister chromatid and the lagging strand a 3'-overhanging end with long ssDNA on the other sister chromatid (Fig. 5e). These two ends each can rejoin with the other blunt end of the DSB, or have a potential to directly ligate with each other, generating a palindromic chromosome from sister chromatid fusion (SCF) and potentially promoting chromatid breakage-fusion-bridge (BFB) cycles[32–34] (Fig. 5e). Because neither DNA-PKcs nor Ku80 is engaged at Cas9-induced DSBs at the gHR_C4 target site for repair in the HDR reporter in mESC (Fig. 1f,g), it is likely that Cas9-gHR_C4 at this site may collide with a replication fork after DNA cleavage, generating a three-ended DSB and allowing subsequent fusion of two sister chromatids and production of a palindromic chromosome. Because the product contains the palindromic DNA sequence surrounding the junctions, a single primer could in theory be annealed to both the leading strand template and the newly synthesized lagging strand in the repair product for PCR amplification. However, no PCR products were detected from repair of Cas9-induced DSBs at the gHR_C4 target site in the HDR reporter in mESC and U2OS cells with a single primer e.g. TF1, TF2 or TF3 (data not shown), likely due to the interference in PCR amplification by palindromic DNA sequences[35]. Given potentially asymmetric sequence deletion at the junction of SCF, we thus paired a distal primer to the break (TF2 or TF3) with the most proximal primer TF1 to minimize the length of palindromic DNA sequence in PCR amplification of repair products induced by SpCas9-gHR_C4 in the HDR reporter and detected PCR bands over 250bp in mESC (Figs. 5f and S6a). In U2OS cells, these PCR bands were detected only after expression of SV40 *LT* (Figs. 5f and S6a). This is consistent with the

observations that DNA-PKcs inhibition stimulates HDR induced by Cas9-gHR$_C$4 in U2OS cells, but neither in mESC nor in U2OS cells highly expressing SV40 *LT*.

To further confirm that the PCR bands for these repair products were indeed fusions of sister chromatids via end ligation of Cas9-induced DSBs, we first cloned PCR products into a plasmid for Sanger sequencing. Among 40 clones for PCR bands with TF1 and TF2, 17 were from mESC and the rest from U2OS cells. Among 31 clones for PCR bands with TF1 and TF3, 29 were from mESC and the rest from U2OS cells. Sanger sequencing revealed only two sequence variations in each PCR band: DL251R6 and DL268R1 for the PCR band with TF1 and TF2 and DL231R5 and DL386R45 for the PCR band with TF1 and TF3 (Figs. 5g and S6b). They all contained some GFP sequences inverted around the break site but no palindromic GFP sequences, indicating that SCF may occur but palindromic sequences may be lost during repair or may not be amplified by PCR (**Table S1**). In addition, the deletion length in each sequence was distinctly asymmetric surrounding the break point, long at 231bp, 251bp, 268bp or 386bp at one direction and short at 1bp, 5bp, 6bp or 45bp at the other direction (Figs. 5g and S6b). It is likely that the collision between DNA replication and Cas9-sgRNA could generate long ssDNA at the lagging strand end and little or no ssDNA overhang at the leading strand end. Long ssDNA could be easily degraded, generating long deletion. PCR targeted amplicon sequencing also confirmed inverted GFP sequences with no palindromic fragments around the sites of Cas9-induced DSBs, but with more junction sequence variations (**Fig. S7a,b**). Taken together, these results indicate that Cas9-sgRNA could be released from its cleaved DNA target upon encountering local DNA replication, generating three-ended DSBs and subsequently inverted duplication *via* end-joining of sister chromatids.

Discussions

As Cas9-sgRNA binds DNA targets with varying affinities and remains bound for variable time even after DNA cleavage, one key issue often ignored in the development and application of CRISPR/Cas9 genome editing is possible effects of target-binding affinity and post-cleavage target-residence duration of Cas9-sgRNA on DSB repair pathway choice in repair of Cas9-induced DSBs. These effects could hamper our efforts in predicting and improving the efficiency and specificity of CRISPR/Cas9 genome editing. In this study, we demonstrate that the binding affinity and residence duration of Cas9-sgRNA to its target modulates c-NHEJ involvement in repair of Cas9-induced DSBs, shaping the choices of repair pathway that differ among targets with different Cas9-sgRNA binding affinities or residence durations (Fig. 5h). Even at a same target, due to different duration of Cas9-sgRNA post-cleavage target residence, Cas9-sgRNA can be dissociated from cleaved DNA either spontaneously or by local transcription or DNA replication (Fig. 5h), exposing Cas9-induced DSBs with different end configurations for specific repair pathways. Weaker binding and shorter residence permit more frequent, spontaneous dissociation or transcription-mediated dissociation of Cas9-sgRNA from its cleaved targets. DSBs exposed in this way can readily engage c-NHEJ, which are largely accurate. In contrast, stronger binding and longer residence delay DSB exposure and increase the probability of a collision between Cas9-sgRNA and local replication fork, generating DSB ends unsuitable for c-NHEJ and inducing inverted ligation of sister chromatids (Fig.

5h). This may lead to extensive structural abnormalities in chromosomes. Therefore, this regulation of DSB repair pathway choice is potentially a new source for the heterogeneity of mutation profiles in CRISPR genome editing (Fig. 5h).

It has been shown that dSpCas9-sgRNA could block translocating RNAP at some targets, thus repressing gene expression[36]. Further, sgRNAs targeting the coding DNA strand of transcription generally demonstrate better gene silencing than sgRNAs targeting the template strand[17,36]. This raises a possibility that Cas9-sgRNA bound to its cleaved target could encounter translocating RNAP and be removed from cleaved DNA by a collision with transcription in a strand-biased manner. While the collision with transcription on template strand might facilitate genome editing more efficiently than on the coding strand[21], our data indicate that c-NHEJ involvement is not altered by the collision of Cas9-sgRNA with transcription no matter which strand the sgRNA is paired with. It is likely that throughout the cell cycle, transcriptional collision, like spontaneous dissociation, may expose DSBs with cleans ends that can be recognized and bound easily by Ku70/Ku80 and rejoined by c-NHEJ using DNA ligase 4/XRCC4. However, replication-coupled dissociation of Cas9-sgRNA from cleaved DNA is restricted to the S phase of the cell cycle and generates three-ended DSBs, which appear to reject c-NHEJ for repair. In these three-ended DSBs, the staggered end with a long 3'-ssDNA overhang may not engage the c-NHEJ factors such as Ku70/Ku80, and the availability of sister chromatids can further promote HDR, antagonizing c-NHEJ. In addition, repair of the three-ended DSBs provides an opportunity for the DNA ends of two sister chromatids to rejoin, not only creating a palindromic chromosome with two centromeres or no centromere[37], but also leaving the third end for translocation. Both dicentric and acentric palindromic chromosomes are unstable and serves as a potential source for chromothripsis and gross chromosomal rearrangements including large deletions and insertions near the target site[32–34,38–43].

Inactivation of c-NHEJ by chemical or genetic approaches is often used to enhance HDR-mediated gene knock-in or gene correction in CRISPR/Cas9 genome editing[27–31]. However, due to weaker binding between Cas9-sgRNA and off-target sites, Cas9-sgRNA at off-target sites may be dissociated from the cleaved DNA more frequently in a spontaneous manner, exposing DSBs that are more likely to engage c-NHEJ and generating less mutagenic repair events. This strategy thus promotes mutation at off-target sites, causing stronger off-target effects. However, this stimulation of off-target effect was often ignored in CRISPR/Cas9 genome editing[27–31]. It was even thought that inactivating c-NHEJ would limit off-target activities in CRISPR/Cas9 genome editing because c-NHEJ is the primary DSB repair pathway for indel-based genome editing by CRISPR/Cas9. Therefore, we need a new strategy to inhibit c-NHEJ while causing no additional off-target effects in CRISPR genome editing.

Methods

Plasmids

The expression plasmids for truncated and mismatched sgRNAs were constructed as described[23], and the expression plasmids for SpCas9 variants eSpCas9, SpCas9-HF1 and xCas9 were constructed

previously[24,25,44]. The sgRNA target sequences and respective mutations for SpCas9 and SaCas9 are listed in **Table S2**. The HDR reporter plasmid was previously constructed[20,45]. To generate the reporter plasmid GFP-P2A-FLuc for replication fork-SpCas9 collision assays, the *P2A-Firefly luciferase* (*FLuc*) gene was fused to C-terminal of GFP in the sGEJ reporter previously established[46]. Due to an SV40 replication origin originally present in the sGEJ reporter, DNA replication can be induced by expression of SV40 large T antigen (LT) in the GFP-P2A-FLuc collision reporter.

Cell lines

HEK293 cells were cultured in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum, 1% penicillin-streptomycin and 2mM L-glutamine. The NHEJ reporter mESC, the HDR reporter mESC and the HDR reporter U2OS cells were previously established and cultured as described before[19–21,46,47]. To generate the *GFP*$^{+}$ cell lines for *GFP* knockout (KO) experiments, mESC harboring the NHEJ reporter were transfected with expression plasmids for SpCas9-g- I-SceI and *GFP*$^{+}$ cells were cloned, expanded and determined by fluorescence-activated cell sorting (FACS) using the Beckman Coulter CytoFLEX flow cytometer. Isogenic *XRCC4*$^{+/+}$ and *XRCC4*$^{-/-}$ mESC containing the HDR reporter were established previously[21]. *DNA-PKcs*$^{-/-}$ and *Ku80*$^{-/-}$ HDR reporter mESC along with isogenic wild-type clones were generated by the paired Cas9-sgRNA method as previously described[5].

Transfection and DSB repair reporter assays

Transfection of mESC was done with Lipofectamine 2000 (Invitrogen) in 24-well plates as previously described[19,45]. For U2OS or HEK293 cells transfection, 1.0×10^{5} cells were seeded on a 24-well plate and grown to 80-95% confluence. 0.8μg total DNA were transfected by Lipofectamine 2000. Cells harboring the NHEJ or HDR reporter were transfected with pcDNA3b-I-SceI or the expression plasmids for SpCas9-sgRNA or SaCas9-sgRNA as previously described[19,45].

In dSaCas9-sgRNA transcription blockage experiments, mouse *GFP*$^{+}$ ES cells were transfected with the expression plasmids for SpCas9-sgRNA, together with the expression plasmids for dSaCas9-sgRNA. In replication fork-dSpCas9 collision experiments, cells were transfected with the expression plasmids for I-SceI or SpCas9-sgRNA and the SV40 LT, together with the GFP-P2A-FLuc reporter plasmid as needed. If necessary, cells were treated with DNA-PKcs inhibitor NU7441 (TopScience Cat# T6276) at 6 h post-transfection. NU7441 was replaced with a fresh addition of the drug the next day. *GFP*$^{+}$ and *GFP*$^{-}$ cells were determined by FACS at 72 h and 96 h respectively post-transfection. The frequencies of NHEJ, HDR and genome editing were calculated after being corrected with background readings and normalized with transfection efficiencies as described before[19].

To evaluate the effect of Cas9 dosage on NHEJ, NHEJ reporter cells were transfected with a varying amount of Cas9-sgRNA each at 0.25μg, 0.1μg, 0.01μg, 0.001μg and 0.0001μg. Cells transfected were treated with 2.5μM NU7441 and analyzed by FACS 3 days (d) post-transfection.

GFP fluorescence measurement for CRISPRi in mESC

GFP^{+} reporter cells were transiently transfected with 0.25μg each of dCas9 and sgRNA expression plasmids in 24-well plates. Cells were analyzed at 96 h post transfection for GFP fluorescence intensity using Beckman Coulter CytExpert 2.0 normalized with mCherry transfection efficiency (TE). The GFP fluorescence intensity of cells transfected with each dCas9-sgRNA was calculated as below:

$$\mathrm{I\ (sgRNA)} = \frac{\mathrm{I\ (sgRNA}measured) - \mathrm{I\ (CTRL}measured) \times (1 - \mathrm{TE})}{\mathrm{TE}}$$

I (sgRNA): GFP intensity of cells expressing dCas9-sgRNA; I (sgRNA*measured*): GFP intensity of cells after transfecting with dCas9-sgRNA; I (CTRL*measured*): GFP intensity of cells after transfecting with dCas9-control sgRNA.

Luciferase assay

HEK293 cells were transiently transfected with GFP-P2A-Luciferase-based NHEJ reporter plasmids together with the expression plasmids for I-SceI or Cas9-sgRNA. The reporter was supplied at 0.025μg in each well of 24-well plates. At 48 h post transfection, cells were harvested and analyzed with the Dual Luciferase Reporter Assay system (Promega). All assays were done in triplicates and all values normalized for transfection efficiency against Renilla luciferase activities as internal control.

PCR targeted amplicon sequencing

For analysis of targeted genome editing at endogenous genome loci, cells were collected after NHEJ induced by Cas9-sgRNAs. Genomic DNA (gDNA) was isolated from these cells using a gDNA purification kit (Axygen). The targeted regions were PCR-amplified with respective primers listed in **Table S3**. The

Illumina deep sequencing was performed at Novogene Co. Ltd and subsequent data analysis was performed as previously described[19].

Off-target analysis

Potential off-target sites were identified using the latest version of the CRISPR Off-Target prediction website (http://crispor.tefor.net/). All potential sites were ranked by an off-target hit score, and high-ranked potential sites were selected. Off-target sites were amplified by PCR with primers listed in **Table S3** after gDNA extraction from cells transfected with Cas9-sgRNA at 3 d post-transfection. Off-target editing efficiency was determined by Illumina deep sequencing. The off-target rate was determined as the ratio of off-target to on-target mutagenesis levels.

Three-ended DSB repair analysis

HDR reporter mESC were transfected with Cas9-sgRNA and harvested 2 d post-transfection. For HDR reporter U2OS cells, 0.008mg of the SV40 *LT* plasmid in 0.8mg of total DNA was simultaneously transfected to initiate replication. gDNA was collected and the palindromic DNA sequences were amplified by touchdown PCR with primers listed in **Table S3.** PCR amplicons were subcloned into CE Entry vector (Vazyme C114-02) and analyzed by Sanger sequencing. Deep sequencing of PCR amplicons was also performed and their repair junctions were characterized by bioinformatics analysis.

Statistical analysis

Student's paired or unpaired t-test was used for statistical analysis of repair frequency. To allow comparison between two groups of gRNAs targeting different DNA strands, one-way ANOVA with post-hoc least significant difference (LSD) pairwise comparisons was performed. Correlation between transcription silencing and the NHEJ increase was determined by linear regression analysis.

Declarations

ACKNOWLEDGEMENTS

We thank members of the Xie lab for helpful discussions and the Core Facilities at Hua Jia Chi Campus, Zhejiang University School of Medicine, for technical support. We thank J Hu at Peking University for the gift of expression plasmids for Cas9 variants. This work was supported by the National Natural Science Foundation of China (No. 31870806 and No. 31671385 to A.Y.X., and No. 32071439 to Y.L.F), the Natural

Science Foundation of Zhejiang Province (LQ20C050004 to S.C.L) and Fundamental Research Funds for the Central Universities in China (2019QNA7031 to Y.L.F).

AUTHOR CONTRIBUTIONS

S-C.L.,Y-L.F. and X-N.S. generated DNA constructs and cell lines, conducted repair reporter assays and genome editing experiments, and performed bioinformatics analysis. R-D.C. conducted replication collision experiments and data analysis. Q.L., J-J.X., J-F.X., G-Q.C., Y.Y., S-M.X., and H.L. assisted with generation of DNA constructs and cell lines. C.L., H-D.L. and A-Y.X. assisted with bioinformatics analysis. A-Y.X. conceived the project and supervised the study. S-C.L., Y-L.F., X-N.S., R-D.C. and A-Y.X. discussed the data, and S-C.L., Y-L.F. and A-Y.X. wrote the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

References

1. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
2. Jiang, F. & Doudna, J. A. CRISPR–Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
3. Jasin, M. & Haber, J. E. The democratization of gene editing: insights from site-specific cleavage and double-strand break repair. *DNA Repair (Amst.)* **44**, 6–16 (2016).
4. Boboila, C., Alt, F. W. & Schwer, B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* **116**, 1–49 (2012).
5. Guo, T. *et al.* Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol.* **19**, 170 (2018).
6. Bétermier, M., Bertrand, P. & Lopez, B. S. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet.* **10**, e1004086 (2014).
7. Symington, L. S. & Gautier, J. Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* **45**, 247–271 (2011).
8. Feng, Y, Liu, S, Chen, R, & Xie, A-Y. Target binding and residence: a new determinant of DNA double strand break repair pathway choice in CRISPR/Cas9 genome editing. *J. Zhejiang Univ. Sci. B* **22**, 73–86 (2021).
9. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).

10. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).
11. Knight, S. C. *et al.* Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science* **350**, 823–826 (2015).
12. Ma, H. *et al.* CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.* **214**, 529–537 (2016).
13. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).
14. Brinkman, E. K. *et al.* Kinetics and fidelity of the repair of Cas9-induced double-strand DNA breaks. *Mol. Cell* **70**, 801–813.e6 (2018).
15. Verkuijl, S. A. & Rots, M. G. The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Curr. Opin. Biotechnol.* **55**, 68–73 (2019).
16. Wang, A. S. *et al.* The histone chaperone FACT induces Cas9 multi-turnover behavior and modifies genome manipulation in human cells. *Mol. Cell* **79**, 221–233.e5 (2020).
17. Clarke, R. *et al.* Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. *Mol. Cell* **71**, 42–55.e8 (2018).
18. Vrtis, K. B. *et al.* Single-strand DNA breaks cause replisome disassembly. *Mol. Cell* (2021) doi:10.1016/j.molcel.2020.12.039.
19. Feng, Y.-L. *et al.* H2AX facilitates classical non-homologous end joining at the expense of limited nucleotide loss at repair junctions. *Nucleic Acids Res.* **45**, 10614–10633 (2017).
20. Rass, E., Chandramouly, G., Zha, S., Alt, F. W. & Xie, A. Ataxia telangiectasia mutated (ATM) is dispensable for endonuclease I-SceI-induced homologous recombination in mouse embryonic stem cells. *J. Biol. Chem.* **288**, 7086–7095 (2013).
21. Xie, A. *et al.* Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair. *Mol. Cell* **28**, 1045–1057 (2007).
22. Kim, D., Luk, K., Wolfe, S. A. & Kim, J.-S. Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. *Annu. Rev. Biochem.* **88**, 191–220 (2019).
23. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
24. Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
25. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
26. Pierce, A. J., Hu, P., Han, M., Ellis, N. & Jasin, M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.* **15**, 3237–3242 (2001).

27. Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
28. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
29. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3**, e04766 (2014).
30. Yeh, C. D., Richardson, C. D. & Corn, J. E. Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* **21**, 1468–1478 (2019).
31. Canny, M. D. *et al.* Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat. Biotechnol.* **36**, 95–102 (2018).
32. Toledo, F. Mechanisms generating cancer genome complexity: back to the future. *Cancers (Basel)* **12**, 3783 (2020).
33. Umbreit, N. T. *et al.* Mechanisms generating cancer genome complexity from a single cell division error. *Science* **368**, eaba0712 (2020).
34. Kagaya, K. *et al.* Chromosome instability induced by a single defined sister chromatid fusion. *Life Sci. Alliance* **3**, e202000911 (2020).
35. Hommelsheim, C. M., Frantzeskakis, L., Huang, M. & Ülker, B. PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Sci. Rep.* **4**, 5052 (2014).
36. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
37. Tanaka, H. & Yao, M.-C. Palindromic gene amplification–an evolutionarily conserved role for DNA inverted repeats in the genome. *Nat. Rev. Cancer* **9**, 216–224 (2009).
38. Ly, P. & Cleveland, D. W. Rebuilding chromosomes after catastrophe: emerging mechanisms of chromothripsis. *Trends Cell Biol.* **27**, 917–930 (2017).
39. Zuccaro, M. V. *et al.* Allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell* **183**, 1650–1664.e15 (2020).
40. Burgio, G. & Teboul, L. Anticipating and identifying collateral damage in genome editing. *Trends Genet.* **36**, 905–914 (2020).
41. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
42. Adikusuma, F. *et al.* Large deletions induced by Cas9 cleavage. *Nature* **560**, E8–E9 (2018).
43. Leibowitz, M. L. *et al.* Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *bioRxiv* 2020.07.13.200998 (2020) doi:10.1101/2020.07.13.200998.
44. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
45. Xie, A. *et al.* Control of sister chromatid recombination by histone H2AX. *Mol. Cell* **16**, 1017–1025 (2004).

46. Xie, A., Kwok, A. & Scully, R. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat. Struct. Mol. Biol.* **16**, 814–818 (2009).
47. Chandramouly, G. *et al.* BRCA1 and CtIP suppress long-tract gene conversion between sister chromatids. *Nat Commun* **4**, 2404 (2013).

Figures

Figure 1

Involvement of c-NHEJ varies widely in repair of Cas9-induced DSBs. (**a**) Schematic of the NHEJ reporter. 6 sgRNAs and the sites targeted between 'Koz-ATG' and 'ATG-GFP' are indicated. Repair of Cas9-induced DSBs by c-NHEJ or a-NHEJ generates accurate NHEJ (acNHEJ) products indistinguishable from undamaged targets and mutagenic NHEJ (m-NHEJ) products represented by *GFP*$^{+}$ cells. Ku80, DNA-PKcs and XRCC4 promote c-NHEJ and suppress a-NHEJ. (**b**) Effect of DNA-PKcs inhibition and *XRCC4* deletion on SpCas9-induced NHEJ in mESC transfected with individual Cas9-sgRNA as indicated. Frequencies of SpCas9-induced *GFP*$^{+}$ cells (**left**) were measured by FACS, and relative NHEJ (**right**) was calculated by normalizing both DMSO treatment and *XRCC4*$^{+/+}$ cells to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes the standard error of the mean (S.E.M.). Columns indicate the mean ± S.E.M. of relative NHEJ. Statistical significance was analyzed by Student's t-test between each control group (DMSO or *XRCC4*$^{+/+}$) and samples, and indicated by * for P<0.05, ** for P<0.005 and *** for P<0.0005. (**c**) Model for enrichment of m-NHEJ products promoted by frequent SpCas9 recleavage with increased amount of Cas9-sgRNA transfected. With sufficient amount of SpCas9-sgRNA, acNHEJ products could be recleaved until m-NHEJ products are generated, resulting in enrichment of m-NHEJ products. (**d**) Frequencies of GFP^{+} cells (**left**) and relative NHEJ (**right**) induced by SpCas9-sgRNA at 0.001μg each, 1/250 of the regular amount (0.25μg each) transfected into mESC. Relative NHEJ was derived by normalizing DMSO treatment and *XRCC4*$^{+/+}$ cells to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative NHEJ. Student's t-test between "*XRCC4*$^{+/+}$" and "*XRCC4*$^{-/-}$": P<0.05 (*) for gEJ_W4, P<0.005 (**) for gEJ_W5 and P<0.0005 (***) for gEJ_C5. (**e**) Schematic of the HDR reporter. In cells containing the HDR reporter, repair of Cas9-induced DSBs by HDR between sister chromatids can generate *GFP*$^{+}$ cells. Ku80, DNA-PKcs and XRCC4 are expected to promote c-NHEJ and suppress HDR. (**f**) Effects of DNA-PKcs inhibition on SpCas9-induced HDR in mESC transfected with individual Cas9-sgRNA as indicated. Frequencies of SpCas9-induced *GFP*$^{+}$ cells (**left**) were measured by FACS and relative HDR (**right**) was calculated by normalizing DMSO treatment to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative HDR. Student's t-test between "DMSO"

and “NU7441”: P<0.05 (*) for gHR_C1, gHR_C2 and gHR_C3 and not significant for gHR_C4 and gHR_C5. (**g**) Effects of *DNA-PKcs*, *Ku80* or *XRCC4* deficiency on SpCas9-induced HDR in mESC transfected with individual Cas9-sgRNA as indicated. Frequencies of SpCas9-induced GFP^+ cells (**left**) were measured by FACS, and relative HDR (**right**) was calculated by normalizing both WT cells and $XRCC4^{+/+}$ cells to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative HDR. Significance was determined by Student’s t-test between each control group (WT or $XRCC4^{+/+}$) and samples, and indicated by * for P<0.05, ** for P<0.005 and *** for P<0.0005.

Figure 2

Reduced target-binding affinity of Cas9-sgRNA shifts the pathway bias towards c-NHEJ in repair of Cas9-induced DSBs. (**a**) Effects of DNA-PKcs inhibition and *XRCC4* deletion on SpCas9-induced NHEJ in mESC transfected with individual SpCas9-sgRNA and its variants as indicated. Frequencies of SpCas9-induced GFP^+ cells (**left**) were measured by FACS, and relative NHEJ (**right**) was calculated by normalizing both DMSO treatment and $XRCC4^{+/+}$ cells to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative NHEJ. Statistical significance was analyzed by Student’s t-test between each control group (DMSO or $XRCC4^{+/+}$) and samples, and indicated by * for P<0.05, ** for P<0.005 and *** for P<0.0005. (**b, c**) Effects of DNA-PKcs inhibition on NHEJ-mediated genome editing at endogenous loci *Cola1* **(b)** and *Rosa26* **(c)** in mESC transfected with individual SpCas9-sgRNA and its variants as indicated. Editing efficiency (**left**) was calculated as ratios of edited reads to total reads from targeted Illumina sequencing and normalized by transfection efficiency. Relative SpCas9-induced NHEJ (**right**) was calculated by normalizing the editing efficiency with DMSO treatment to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative NHEJ. Student’s t-test between “DMSO” and “NU7441”: P <0.05 (*) for gC4-G16C and gR4-A16T, P<0.005 (**) for gC4-C1T. (**d, e**) Effects of DNA-PKcs inhibition and *DNA-PKcs*, *Ku80* or *XRCC4* deficiency on Cas9-induced HDR in mESC transfected with Cas9-gHR_C4 (**d**), Cas9-gHR_C2 (**e**) and its variants as indicated. Frequencies of Cas9-induced GFP^+ cells (**left**) were measured by FACS, and relative Cas9-induced HDR (**right**) was calculated by normalizing DMSO treatment, WT cells and $XRCC4^{+/+}$ cells to 1.0. Each circle indicates one independent experiment, each in triplicates, and the means of these independent experiments are also indicated. Error bar denotes S.E.M. Columns indicate the mean ± S.E.M. of relative HDR. Significance was determined by Student’s t-test between each control group (DMSO, WT or $XRCC4^{+/+}$) and samples and indicated by * for P<0.05, ** for P<0.005 and *** for P<0.0005.

Figure 3

DNA-PKcs inhibition and *XRCC4* deletion aggravate off-target effect in CRISPR/Cas9 genome editing. Isogenic *XRCC4*$^{+/+}$ and *XRCC4*$^{-/-}$ HDR reporter mouse ES cells were transfected with SpCas9 in complex with gPnpla3 targeting *Pnpla3* (**a**), gMertk targeting *Mertk* (**b**), gHRC1 (**c**) and gHRC2 (**d**) both targeting the HDR reporter. In an independent set of experiments, only *XRCC4*$^{+/+}$ HDR reporter mouse ES cells were used for transfection and treated with DMSO or NU7441. The indel frequencies at on-target and selected off-target sites were measured by amplicon deep sequencing and defined as the ratio of edited reads to total reads normalized by transfection efficiency. The ratio of off-target frequency to on-target frequency (i.e. ratio of off-target to on-target) indicates off-target effect. Each circle or triangle indicates one independent experiment, and the mean of these independent experiments is also indicated. Error bars indicate S.E.M. Statistical significance was detected by Student's t-test between "DMSO" and "NU7441" and between "*XRCC4*$^{+/+}$" and "*XRCC4*$^{-/-}$", and indicated by * for P<0.05 and ** for P<0.005.

Figure 4

Transcription has no effect on c-NHEJ in repair of Cas9-induced DSBs. (**a**) dSpCas9-mediated transcriptional silencing. Mouse ES cells containing pPGK-GFP expression cassette were transfected with expression plasmids for dSpCas9 in complex with either of 16 sgRNAs as well as non-targeting U6 sgRNA as control (gCtrl) as indicated. The mean fluorescence intensity of GFP in cells was determined at 4 d post-transfection. (**b**) Involvement of c-NHEJ in SpCas9-induced DSBs. Mouse ES cells containing pPGK-GFP expression cassette were transfected with a low amount of individual SpCas9-sgRNA expression plasmids (0.001μg SpCas9, 0.001μg sgRNA, 1/500 of total DNA each) as shown and treated with NU7441 at 6 h post-transfection. Relative NHEJ was calculated by normalizing DMSO treatment to 1.0. Columns indicate the mean ± S.E.M. of relative NHEJ. (**c**) Analysis of strand bias in transcriptional silencing (**top**) and c-NHEJ involvement (**bottom**) between transcription template strand and transcription non-template strand targeted by dSpCas9-sgRNA or SpCas9-sgRNA. Transcriptional silencing and c-NHEJ involvement were defined as the percentage of GFP fluorescence intensity reduced by dSpCas9-sgRNA and the percentage of SpCas9-induced NHEJ stimulated by NU7441, respectively. Each circle indicates the mean ± S.E.M. of three independent experiments, each in triplicate, for individual dSpCas9-sgRNA or SpCas9-sgRNA. Error bars indicate S.E.M. (**d**) Correlation between dSpCas9-mediated transcriptional silencing and c-NHEJ involvement in repair of SpCas9-induced DSBs. Each circle indicates the level of dSpCas9-mediated transcriptional silencing and stimulation of SpCas9-induced NHEJ by NU7441 at the same target. Two sgRNAs gG_W5 and gG_C4 are indicated by arrows. (**e**) Mouse ES cells containing the pPGK-GFP expression cassette were transfected with expression plasmids for dSaCas9 in complex with either of 6 sgRNAs as well as non-targeting U6 sgRNA as control (gCtrl). The GFP fluorescence was determined at 4 d post-transfection. Sites in the 5' non-coding region of the pPGK-GFP reporter are indicated for targeting by dSaCas9 in complex with individual sgRNAs as shown on the top. (**f**) Impact of transcription blockage on DNA-PKcs involvement in SpCas9-mediated *GFP* gene editing. *GFP*$^{+}$ cells were co-transfected with SpCas9-sgRNA (SpCas9 and sgRNA at 0.125mg and 0.0005mg respectively) and

dSaCas9-sgRNA (0.125mg each) and frequencies of SpCas9-induced *GFP*$^{-}$ cells measured by FACS at 4 d post-transfection. Transcription blockage by dSaCas9-gSaG$_W$1 and dSaCas9-gSaG$_W$2 was determined by significant reduction in the mean fluorescence intensity of GFP in *GFP*$^{+}$ cells transfected with dSaCas9-sgRNA as indicated in the inset. Columns indicate the mean ± S.E.M. of at least three independent experiments, each in triplicates.

Figure 5

Replication adjacent to targets of Cas9-sgRNA suppresses c-NHEJ in repair of Cas9-induced DSBs. (**a**) Impact of local replication on DNA-PKcs involvement in NHEJ induced by I-SceI and SpCas9-gEJ$_W$10. 293 cells were co-transfected with plasmids for a Firefly luciferase-based NHEJ reporter (FLuc: Firefly luciferase; pCAG: CAG promoter; SV40 ori: SV40 DNA replication origin), Renilla luciferase as an internal control and I-SceI or SpCas9-gEJ$_W$10, together with SV40 *LT* expression plasmids or empty vector, and treated with NU7441. The NHEJ reporter contains an SV40 origin that can be bound by SV40 LT to initiate DNA replication during DNA cleavage induced by I-SceI and SpCas9-gEJ$_W$10. Firefly luciferase assays were performed at 2 d post-transfection and NHEJ induced by I-SceI and SpCas9-gEJ$_W$10 was calculated as the ratio of Firefly luciferase activity to Renilla luciferase activity as indicated. Columns indicate the mean ± S.E.M. of at least three independent experiments, each in triplicates. The percentage of NHEJ reduction by NU7441 is indicated above each column. Statistical significance was analyzed by Student's t-test between "DMSO" and "NU7441" and indicated by * for P<0.05 and *** for P<0.0005. (**b**) Impact of local replication on DNA-PKcs involvement in HDR induced by I-SceI or SpCas9-gHR$_C$2. HDR reporter U2OS cells were transfected with I-SceI or SpCas9-gHR$_C$2 expression plasmids, together with SV40 *LT* expression plasmids or empty vector, and treated with NU7441. The SV40 promoter driving *I-SceI-GFP* in the HDR reporter contains an SV40 origin that can be bound by SV40 LT to initiate DNA replication during DNA cleavage induced by I-SceI and SpCas9-gHR$_C$2. The SV40 *LT* expression plasmid was titrated for transfection as indicated. Columns indicate the mean ± S.E.M. of at least three independent experiments, each in triplicates. The fold of the increase in the percentage of *GFP*$^{+}$ cells by NU7441 is shown above each column. Significance was analyzed by Student's t-test between "DMSO" and "NU7441" and indicated by * for P<0.05 and n.s for not significant. (**c**) Analysis schematic for HDR and NHEJ in repair of the same SpCas9-induced DSBs in the HDR reporter. After SpCas9-mediated DSB induction around the I-SceI site, repair by HDR and NHEJ generates "wtGFP" and "mutant GFP" products respectively. These two repair outcomes can be determined by nested PCR followed by deep sequencing, and the HDR preference can be calculated as the ratio of HDR reads to total edited reads, i.e. HDR/(HDR+NHEJ). (**d**) Effect of Cas9-sgRNA target residence on the pathway choice between HDR and NHEJ in repair of Cas9-induced DSBs. HDR reporter mESC (**left**) were transfected with expression plasmids for SpCas9 and gHR$_C$4 or its two sgRNA mutants as indicated. HDR reporter U2OS cells (**right**) were transfected with expression plasmids for SpCas9-gHR$_C$2 and SV40 *LT* or empty vector. At 3 days post-transfection, gDNA was isolated, the frequencies of HDR and total edited were determined by nested PCR and deep sequencing,

and the HDR bias was calculated as the ratio of HDR to total edited. Columns indicate the mean ± S.E.M. of three independent experiments. Significance was analyzed by Student's t-test and indicated by * for P<0.05. (**e**) Detection schematic for three ends generated by a collision between a converging DNA replication fork and residing Cas9-sgRNA at cleaved target. As the ends of two sister chromatids generated by replication can be ligated *via* NHEJ, the presence of this palindromic sister chromatid NHEJ products would indicate generation of such three ends. A single PCR primer targeting one sister chromatid could in theory amplify these products. Due to potential interference of extensive palindromic sequences in PCR amplification, three primers with different distance to the end, TF1 (426bp), TF2 (336bp) and TF3 (231bp), were screened in pairs as indicated. (**f**) PCR detection of palindromic sister chromatid ligation. HDR reporter mESC were transfected with SpCas9-gHR$_C$4 expression plasmids or empty vector as indicated. HDR reporter U2OS cells were additionally co-transfected with SV40 *LT* expression plasmids or control vectors. At 2 d post-transfection, gDNA was isolated and PCR was performed with the primer pair of TF1 and TF2. M: 2kb DNA ladder. (**g**) Repair junction of palindromic sister chromatid NHEJ products. The PCR products from (**f**) were subcloned and detected by Sanger sequencing, and the sequencing data were pooled from two independent experiments. Only two types of products (#1 and #2) were detected with the size and position of deletion (del) and insertion (ins) as indicated. *T: insertion of a thymidine nucleotide. (**h**) Impact of Cas9-sgRNA target residence on local repair pathway choice. Due to varied binding affinity and residence duration, the Cas9-sgRNA complex bound to different targets or even to the same target at different times may be released from the cleaved target differently. Spontaneous or transcription-mediated release of Cas9-sgRNA from the cleaved DNA generates a DSB with two conventional ends, suitable for engaging c-NHEJ factors such as DNA-PKcs, Ku70/Ku80 and XRCC4. However, persistent target binding of Cas9-sgRNA may increase the probability of locally encountering a DNA replication fork, resulting in dissociation of Cas9-sgRNA from the cleaved DNA and a DSB with three ends, a blunt end at the leading strand, a 3'-overhanging end at the lagging strand and a blunt end away from the replication fork. The configurations of these ends and the S-phase exposure of the DSB may bias DSB repair from c-NHEJ towards either a-NHEJ or HDR. RNAP: RNA polymerase.

Supplementary Files

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

- ManuscriptSupplentaryFiguresandTablesLiuetal022221.pdf