DNA nicks induce mutational signatures associated with BRCA1 deficiency

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

Qian Liu  
Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine

Ruo-Dan Chen  
Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine

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

Zhi-Cheng Huang  
Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine

Xiao-Ying Yang  
Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine

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

Article

Keywords:

Posted Date: February 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1280037/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Analysis of human cancer genome sequences has revealed specific mutational signatures associated with BRCA1-deficient tumors, but the underlying mechanisms remain poorly understood. Here, we show that one-ended DNA double strand breaks (DSBs) converted from CRISPR/Cas9-induced nicks by DNA replication, not two-ended DSBs, cause more characteristic chromosomal aberrations and micronuclei in Brca1-deficient cells than in wild-type cells. BRCA1 is required for efficient homologous recombination of these nick-converted DSBs and suppresses bias towards long tract gene conversion and tandem duplication (TD) mediated by two-round strand invasion in a replication strand asymmetry. However, aberrant repair of these nick-converted one-ended DSBs, not that of two-ended DSBs in Brca1-deficient cells, generates mutational signatures such as small indels with microhomology (MH) at the junctions, unbalanced translocations and small MH-mediated TDs, resembling those in BRCA1-deficient tumors. These results suggest a major contribution of DNA nicks to mutational signatures associated with BRCA1 deficiency in cancer and the underlying mechanisms.

Highlights

1. DNA nicks, not two-ended DSBs, caused more chromosomal aberrations and micronuclei in Brca1-deficient cells than in wild-type cells;
2. Nicks induced by Cas9 nickases were converted into one-ended DSBs by DNA replication;
3. LTGC in sister chromatid recombination (SCR) was more involved in HR repair of one-ended DSBs than in that of two-ended DSBs;
4. HR defect in repair of one-ended DSBs in Brca1-deficient cells, not two-ended DSBs, accumulated mutational signatures resembling those in BRCA1-deficient tumors;
5. Repeat-directed TDs favored in Brca1-deficient cells were promoted by nicks preferentially on lagging strand template of DNA replication.
6. Repeat-directed TDs were mediated either by non-allelic SCR/LTGC or by two-round homologous strand invasion, first allelic and second non-allelic.

Introduction

Somatic mutations in human cancer genomes are generated by various mutational mechanisms involving DNA replication, damage and repair during the development of cancer. Each mechanism imprints somatic mutations with a characteristic pattern known as mutational signatures. Analysis of cancer genome sequences across cancer types has identified many distinct classes of mutational signatures. BRCA1-deficient cancers including breast, ovary and other cancers exhibit the SBS3 signature characterized by uniform distribution of all base substitution types, the ID6 signature by small 11–50 bp indels with short microhomology (MH) at junctions, small ~10 kb MH-mediated TDs, and unbalanced translocations1–7. Because BRCA1 deficiency severely impairs error-free homologous recombination (HR), one of two major repair pathways for DNA double strand breaks (DSBs), BRCA1-deficient cells become
more dependent upon alternative pathways, in particular error-prone non-homologous end joining (NHEJ), for repair of DSBs\(^8,9\). It is tempting to speculate some of these alternative repair pathways may generate mutational signatures associated with BRCA1 deficiency. In fact, several studies have implicated polymerase theta-mediated end joining (TMEJ) in generation of mutational signatures in BRCA1-deficient tumors\(^10\)\(^–\)\(^12\). However, as BRCA1 participates in different steps of HR, which repairs DSBs from various sources, and is also required for replication fork protection, it is poorly understood how individual mutational signatures are developed in BRCA1-deficient tumors\(^8,9,13\).

DSBs appear in two general forms: two-ended and one-ended. The HR function of BRCA1 is well recognized in repair of two-ended DSBs; however, two-ended DSBs rarely occur spontaneously. In contrast, collapsed or stalled replication forks often generate one-ended DSBs as the primary substrate for HR in mammalian cells\(^14\)\(^,\)\(^15\). Unlike two-ended DSBs that can be efficiently repaired by end rejoining, one-ended DSBs are more dependent on HR for repair due to lack of the second end and often trigger break-induced replication (BIR) in yeast and possibly in mammalian cells\(^16\)\(^,\)\(^17\). In fact, BRCA1 suppresses aberrant long tract gene conversion (LTGC), a potentially BIR analogue in mammalian cells\(^18\). In addition, using Escherichia coli Tus–Ter replication fork barriers to trigger site-specific replication fork stalling at a mammalian chromosome, a study found that replication fork stalling leads to collapsed forks generating both one-ended and two-ended DSBs and promotes the formation of small \(\sim\)10 kb MH-mediated TDs in Brca1-deficient mouse embryonic stem cells (mESC)\(^19\). However, given that single-strand breaks (SSBs) including nicks account for about 75% of endogenous DNA lesions detected daily in mammalian cells, it is likely that some of these SSBs could collapse replication forks and produce one-ended DSBs as a principal endogenous source\(^20\)\(^–\)\(^24\). However, while BRCA1 is also expected to participate in HR repair of such one-ended DSBs, much has yet to be learnt about the consequences of BRCA1 dysfunction.

In order to study the role of BRCA1 in repair of DSBs converted from SSBs by DNA replication and the contribution of this role to mutational signatures in BRCA1-deficient cells, we established a cellular assay system where a site-specific nick can be induced by Streptococcus pyogenes Cas9 (SpCas9) nickases (nCas9)\(^25\)\(^,\)\(^26\) and subsequently converted into a one-ended DSB by DNA replication in mammalian cells. Using this system, we demonstrated that BRCA1 is required for efficient HR of replication-dependent DSBs converted from nicks but suppresses LTGC bias during this HR repair. We also found that DNA nicks, not two-ended DSBs, induce mutational signatures such as the small indel signature, unbalanced translocations and MH-mediated TDs (\(\sim\)10 kb) in BRCA1-deficient mESC, resembling those in BRCA1-deficient tumors. In addition, nick-induced LTGC bias and repeat-directed TDs generated by two-round strand invasions are promoted in a replication strand bias in BRCA1-deficient cells. Our data suggest that DNA nicks coupled with DNA replication, not two-ended DSBs, be a major inducer to characteristic mutational signatures in BRCA1-deficient tumors.

**Results**

nCas9-induced nicks are converted into one-ended DSBs by DNA replication
We first used Cas9 or nCas9 (D10A: Cas9\textsuperscript{D}; H840A: Cas9\textsuperscript{H}), together with gB2, an sgRNA targeting short interspersed nuclear element B2 repeats\textsuperscript{25–27}, to induce a number of DSBs or nicks throughout the mouse genome, respectively (Fig. 1a). mESC transfected with nCas9-gB2 or Cas9-gB2 proliferated much slower (Extended Data Fig. 1a). Phosphorylation of Chk1 S345, Chk2 T68 and p53 S15 as well as H2AX S139 (to form “γH2AX”) was stimulated in mESC by Cas9-gB2 whereas the RPA32 S4/S8 phosphorylation remained unchanged at a basal level (Fig. 1b). Likewise, both Cas9\textsuperscript{D} and Cas9\textsuperscript{H} induced phosphorylation of H2AX, Chk1, Chk2, and p53 (Fig. 1b); however, Cas9\textsuperscript{D} induced more phosphorylation of Chk1 and RPA32 than Cas9 (Fig. 1b). The difference in this response between Cas9 and Cas9\textsuperscript{D} appears to parallel that between bleomycin and hydroxyurea (HU) or camptothecin (CPT) (Fig. 1b), indicating that Cas9\textsuperscript{D}- induced DNA damage response (DDR) may be associated with DNA replication. Likely due to the stronger nicking activity of the HNH domain\textsuperscript{28}, the DDR induced by Cas9\textsuperscript{H} is much weaker than by Cas9\textsuperscript{D}.

We also examined Cas9- or nCas9-induced focus formation of γH2AX and 53BP1, both of which serve as a DSB marker. Both Cas9 and nCas9 induced co-localized formation of γH2AX and 53BP1 foci as did bleomycin, HU and CPT (Fig. 1c and Extended Data Fig. 1b, c). Cas9\textsuperscript{D} induced even more focus formation than Cas9 in mESC, not in NIH3T3, suggesting possibly more efficient conversion of nicks into one-ended DSBs by DNA replication in mESC (Fig. 1d and Extended Data Fig. 1b, c). To further confirm that this conversion is coupled with DNA replication, we treated mESC with DNA replication inhibitors and found that this treatment had little effect on relative intensity of Cas9-induced γH2AX but abolished Cas9\textsuperscript{D}-induced γH2AX (Fig. 1e). Together, these results suggest that nCas9-induced nicks could be efficiently converted into one-ended DSBs by DNA replication.

**Nicks, not two-ended DSBs, induce BRCA1-linked chromosomal aberrations**

Using paired Cas9-sgRNA approach\textsuperscript{29}, we generated Brca1-deficient (Brca1\textsuperscript{m/m}) mESC in which mutated Brca1 gene encodes unstable BRCA1 lacking the C-terminal BRCT repeats (Extended Data Fig. 1d,e)\textsuperscript{30}. Cas9-gB2 and nCas9-gB2 were transfected to induce a number of two-ended DSBs and nicks in isogenic Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} mESC (Extended Data Fig. 1f). In cells expressing the dCas9-sgRNA negative control, the chromosome abnormalities spontaneously arose only in Brca1-deficient cells (Fig. 1f and Extended Data Fig. 1g,h). It is possible that DNA-bound dCas9 may act as a barrier to transcription or DNA replication at many B2 target sites, thus causing DNA damage\textsuperscript{31–34}. Either Cas9-gB2 or nCas9-gB2 induced a high level of abnormal metaphases in both Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} cells (Fig. 1f and Extended Data Fig. 1g,h). However, the level of chromosomal aberrations induced by Cas9-gB2 was similar or slightly different between Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} mESC. In contrast, chromosomal aberrations induced by Cas9\textsuperscript{D}- and Cas9\textsuperscript{H}-gB2 were much more severe in Brca1\textsuperscript{m/m} mESC (Fig. 1f and Extended Data Fig. 1g,h), suggesting that nicks (or by extension one-ended DSBs) and aberrant repair of these DNA lesions may be a major source for spontaneous chromosomal aberrations associated with BRCA1 deficiency.
Brca1-deficient MEFs have significant micronuclei formation\textsuperscript{35}. We thus analyzed micronuclei formation induced by Cas9-gB2 or nCas9-gB2. In cells expressing dCas9-gB2, 2.7% Brca1-deficient cells exhibited micronuclei formation but no Brca1-proficient cells did (Fig. 1g and Extended Data Fig. 1i). Unexpectedly, micronuclei formation in Brca1\textsuperscript{m/m} cells was less frequent than in Brca1\textsuperscript{+/+} cells after transfection with Cas9-gB2 (Fig. 1g and Extended Data Fig. 1i). In contrast, micronuclei formation induced by Cas9\textsuperscript{D} and Cas9\textsuperscript{H}-gB2 was more extensive in Brca1\textsuperscript{m/m} mESC than in Brca1\textsuperscript{+/+} cells (Fig. 1g and Extended Data Fig. 1i), suggesting that one-ended DSBs and aberrant repair of these DNA lesions may contribute to spontaneous micronuclei formation detected in Brca1-deficient cells, not in normal cells.

HR repair of nCas9-induced nicks is coupled with DNA replication

As HR between sister chromatids, i.e., sister chromatid recombination (SCR), is a predominant pathway in repair of one-ended DSBs at collapsed replication forks\textsuperscript{36,37}, defects of this HR repair may be responsible for elevated chromosomal aberrations in BRCA1-deficient cells. We used mESC harboring a single-copy “SCR-RFP” reporter integrated at the Rosa26 locus to characterize HR repair of one-ended DSBs (Fig. 2a and Extended Data Fig. 2a, see Methods for reporter details)\textsuperscript{18,38}. After transfection of reporter mESC with Cas9-sgRNAs and nCas9-sgRNAs, Cas9\textsuperscript{D} induced a significant percentage of GFP\textsuperscript{+}RFP\textsuperscript{−} cells representing short-tract gene conversion (STGC) and GFP\textsuperscript{+}RFP\textsuperscript{+} cells representing LTGC at 4 different target sites as Cas9 (Extended Data Fig. 2b). In human U2OS cells, a significant but lower STGC and LTGC were observed (Extended Data Fig. 2c). In both mESC and U2OS cells, nCas9-induced LTGC bias was respectively greater than Cas9-induced LTGC bias (Extended Data Fig. 2b,c). However, the detected LTGC bias in nCas9-induced HR was still far below 100% although HR without a second end for second-end capture or strand annealing would in theory all generate LTGC events. This implies that the second end remains available from a converging replication fork with a significant probability to facilitate termination of nCas9-induced HR by STGC, thus suppressing LTGC.

Using the SCR-dGFP reporter where no GFP homology is available in the second end (Extended Data Fig. 2d,e, see Methods for reporter details), we found the LTGC bias was increased from 2% with the SCR-RFP reporter to 10% with the SCR-dGFP reporter for Cas9-induced two-ended DSBs and from 10% to 50-70% for nCas9-induced nicks (Extended Data Fig. 2f). This again suggests that not only the non-invading second end were generated but also homologous GFP sequence of this end are critical for termination in nCas9-induced HR as in HR of two-ended DSBs. We also treated SCR-RFP reporter mESC with Aphidicolin and L-Mimosine. This replication inhibition reduced nCas9-induced STGC to a larger extent than Cas9-induced STGC (Extended Data Fig. 3a). In addition, we initiated bidirectional DNA replication in SV40 origin-containing SCR-RFP reporter by ectopic expression of SV40 large T antigen (LT) (Extended Data Fig. 3b)\textsuperscript{39}. Local DNA replication dramatically increased nCas9-induced HR by over 14 folds, but Cas9-induced HR only by 2-3 folds (Extended Data Fig. 3b). These results suggest that HR repair of nCas9-induced nicks is coupled with conversion of nicks into one-ended DSBs by DNA replication.
**BRCA1 deficiency causes replication strand-asymmetric LTGC bias in repair of one-ended DSBs**

To determine whether BRCA1 regulates HR of one-ended DSBs, we induced a site-specific DSB or nick in *Brca1*+/+ and *Brca1*mm clones (Extended Data Fig. 4a). Consistent with previous study18, I-SceI-induced STGC was impaired in *Brca1*mm clones but the LTGC bias was enhanced by *Brca1* deficiency (Extended Data Fig. 4b). Similarly, Cas9-induced STGC and LTGC at the four different breakage sites were reduced in *Brca1*mm cells (Fig. 2b and Extended Data Fig. 4c). Because the LTGC reduction was smaller than the STGC reduction, Cas9-induced LTGC bias was elevated by about 3-4 folds in cells deficient for *Brca1* (Fig. 2b and Extended Data Fig. 4c). However, while Cas9D-induced STGC was reduced in *Brca1*mm cells at all 4 different nicking sites as compared to *Brca1*+/+ cells, LTGC was reduced at the gHR1b and gHR2 site, little changed at the gHR3 site and increased at the gHR1a site in *Brca1*mm cells (Fig. 2b and Extended Data Fig. 4c). In contrast, Cas9H-induced STGC was reduced at the gHR1b and gHR2 site but little changed at the gHR3 and gHR1a site in *Brca1*mm cells but LTGC was unchanged at the gHR1b site and significantly increased at the other three sites (Fig. 2b and Extended Data Fig. 4c). As a result, in *Brca1*mm cells, the increase of Cas9H-induced LTGC bias was much greater than the increase of Cas9D-induced LTGC bias (Fig. 2b and Extended Data Fig. 4c). As Cas9D and Cas9H induced a nick respectively on target strand within the DNA-RNA hybrid and on single non-target strand at the gHR1a, gHR1b, gHR2 and gHR3 target sites, this suggests BRCA1 may suppress LTGC bias in repair of nCas9-induced nicks with an asymmetry associated with DNA replication strand (i.e., leading strand vs. lagging strand).

We used the Emx1-SCR reporter containing 6 additional nCas9 targeting sites to further test this strand asymmetry (Fig. 2c, see Methods for reporter details). Similar to the four nicking sites tested above, at the sites targeted by gEmx7, gEmx9 and gEmx12, Cas9D- and Cas9H-induced nicks could be converted into one-ended DSBs by a collision respectively with leading strand (lead collapse in Fig. 2d) and with lagging strand (lag collapse in Fig. 2d)24, leading to HR we could measure. It is opposite at the sites targeted by gEmx8, gEmx10 and gEmx11 (Fig. 2d). No matter whether LTGC bias was induced by Cas9D or Cas9H, the increase of LTGC bias coupled with lag collapse was much greater than the increase of LTGC bias with lead collapse in *Brca1*mm cells as compared to *Brca1*+/+ cells (Fig. 2e,f and Extended Data Fig. 4c,d). However, no significant difference was detected in BRCA1-mediated suppression of LTGC bias between nicks induced on target strand in the RNA-DNA hybrid and on non-target strand within the nCas9-sgRNA-DNA complex (Extended Data Fig. 5a). This suppression was also little altered between PAM on Watson strand and on Crick strand for nCas9-sgRNAs (Extended Data Fig. 5b). Together, these data suggest that the strand asymmetry in exacerbation of nCas9-induced LTGC bias by *Brca1* deficiency was associated with DNA replication, more strongly with lagging strand collision than with leading strand collision, not with conformational context of the nCas9-sgRNA-DNA complex. As LTGC causes gene amplifications (such as TDs) and SBSs, this bias may lead to asymmetric distribution of these mutations across leading and lagging strands in BRCA1-deficient cancers.
Nicks, not two-ended DSBs, stimulate NHEJ with MH-mediated small deletions in Brca1-deficient cells

In Brca1-deficient mESC, the delayed availability of the second ends would facilitate NHEJ due to inefficient HR (Extended Data Fig. 6a). Using a single-copy NHEJ reporter (Extended Data Fig. 6b, see Methods for reporter details)\(^4\), we found NHEJ repair of two-ended DSBs induced by Cas9 was highly efficient, whereas the level of Cas9\(^D\)- or Cas9\(^H\)-induced NHEJ was much smaller but detectable (Extended Data Fig. 6b). We then generated isogenic Brca1\(^{+/+}\) and Brca1\(^{1m/m}\) NHEJ reporter mESC clones (Extended Data Fig. 6c). NHEJ of two-ended DSBs was slightly increased in Brca1\(^{1m/m}\) clones (Fig. 3a). However, the frequencies of Cas9\(^D\)- and Cas9\(^H\)-induced NHEJ were greatly elevated by more than 10-folds when Brca1 is deficient (Fig. 3a), indicating that NHEJ serves as an important alternative pathway in repair of one-ended DSBs when BRCA1-mediated HR is defective.

We also induced a site-specific DSB or nick at natural sites and examined the frequency and junctions of NHEJ by deep sequencing. The frequency of Cas9-induced NHEJ was not affected by Brca1 status, but nCas9-induced NHEJ was strongly stimulated by Brca1 deficiency (Fig. 3b). The deletion lengths in Cas9-induced NHEJ were shorter in Brca1\(^{1m/m}\) cells than in Brca1\(^{1+/+}\) cells (Fig. 3c). The proportions of <11 bp deletions in Brca1\(^{1m/m}\) cells were similar to Brca1\(^{1+/+}\) cells (Fig. 3d). The use of MH was reduced in Brca1\(^{1m/m}\) cells (Fig. 3e). These data are consistent with the function of BRCA1 in end resection. In contrast, in nCas9-induced NHEJ, the median deletion lengths were larger, and <11 bp deletions much less in Brca1\(^{1m/m}\) cells than in Brca1\(^{1+/+}\) cells (Fig. 3c,d). However, Brca1 deficiency increased the proportion of 11-80 bp deletions and the MH usage (Fig. 3d,e). These results together indicate that nCas9-induced small deletions in Brca1\(^{1m/m}\) mESC, not Cas9-induced small deletions, resemble those of spontaneous small deletions enriched in BRCA1-deficient tumors.

Nicks, not two-ended DSBs, promote translocations associated with BRCA1 deficiency

As translocations are accumulated as an SV signature in Brca1-deficient cancers\(^2,3\), we induced intrachromosomal and inter-chromosomal translocations in both Brca1\(^{1+/+}\) and Brca1\(^{1m/m}\) mESC. Concomitant DSBs or nicks, one at intron 14 of Em14 and the other at intron 19 of Alk, were generated in mESC respectively by Cas9 or nCas9 to induce four types of intrachromosomal translocations: Em14\(^F\)-Alk\(^F\) (Em14\(^R\), Alk\(^R\), Em14\(^F\)-Alk\(^F\) (Em14\(^R\), Alk\(^R\), Em14\(^F\)-Alk\(^F\) (Em14\(^R\), Alk\(^R\)) (Fig. 4a). Similarly, four types of inter-chromosomal translocations Rosa26\(^F\)-Ldha\(^F\) (Rosa26\(^F\)-Ldha\(^R\), Rosa26\(^F\)-Ldha\(^R\)) and Rosa26\(^R\)-Ldha\(^R\) were induced between intron 2 of Rosa26 and the other at intron 5 of Ldha (Fig. 4b). Translocations were detected by PCR and quantified by deep sequencing (Fig. 4c,d). In Brca1\(^{1+/+}\) cells, Cas9-induced translocations were efficient whereas nCas9-induced translocations were hardly detected (Fig. 4c,d). However, in Brca1\(^{1m/m}\) cells, while Cas9-induced translocations including E\(^RA\)R, E\(^RA\)F,
and \( R^L \) and \( R^F \) were induced as efficiently as in \( Brca1^{+/+} \) cells, nCas9-induced translocations of the same types were elevated by 30-60,000 times in \( Brca1^{m/m} \) cells (Fig. 4c,d). Together, these results indicate \( Brca1 \) deficiency has little effect on the generation of translocations induced by two-ended DSBs, but greatly stimulates translocations induced by DNA nicks.

Junction analysis of Cas9-induced translocations revealed that nearly all translocation events in \( Brca1^{+/+} \) mESC were generated with indels (Extended Data Fig. 7a,b). However, in \( Brca1^{m/m} \) cells, an increased fraction of translocations was generated without indels (Extended Data Fig. 7a,b). The deletions at the Cas9-induced junctions were smaller and the frequency of MH used was greatly lower in \( Brca1^{m/m} \) cells (Fig. 4e,f and Extended Data Fig. 7c,d). These results suggest that BRCA1 does not determine the frequency of Cas9-induced translocations but promotes end resection and the use of MH.

Unlike Cas9-induced translocations, nCas9-induced translocations in \( Brca1^{m/m} \) cells engaged little precise end joining (Extended Data Fig. 7a,b). The deletions at the junctions were larger than those in Cas9-induced translocations (Fig. 4e,f). Also, in \( Brca1^{m/m} \) cells, the MH usage in each type of nCas9-induced translocations was generally more frequent than in Cas9-induced translocations (Extended Data Fig. 7c,d). These data together indicate that translocations induced by DNA nicks had more extensive end resection and more frequent use of MH than translocations induced by two-ended DSBs in \( Brca1^{-} \)-deficient cells.

**BRCA1 suppresses TDs mediated by two-round homologous strand invasion**

In the SCR-RFP reporter, site-specific stalling of replication forks induced GFP^−RFP^+ cells, which represent excess \( \sim 10 \) kb MH-mediated TDs\(^{19} \). However, both spontaneous and I-SceI-induced GFP^−RFP^+ cells are products from GFP repeat-mediated TD, containing three copies of \( GFP \) with a \( \sim 10 \)-kb duplication span\(^{41,42} \). To determine which type of TDs represented by GFP^−RFP^+ cells is induced by DNA nicks and could be promoted by \( BRCA1 \) deficiency, we first analyzed the frequencies of GFP^−RFP^+ cells induced spontaneously or by nCas9 and Cas9 (Fig. 5a). Spontaneous GFP^−RFP^+ cells were scarce but increased by \( \sim 4 \)-fold in \( Brca1^{m/m} \) clones (Fig. 5b). Expression of Cas9-gHR1b generated a higher level of GFP^−RFP^+ cells, which was however unaffected by \( Brca1 \) mutation as in I-SceI-induced GFP^−RFP^+ cells (Fig. 5b)\(^{19} \).

In contrast, the frequencies of GFP^−RFP^+ cells induced by Cas9\(^H\), not those by Cas9\(^D\), were stimulated by \( Brca1 \) deficiency (Fig. 5b), indicating a strand asymmetry.

We then used the “Emx1-SCR” reporter to further determine the association of the strand asymmetry with DNA replication (Extended Data Fig. 8a). \( Brca1 \) deficiency caused varying effect on Cas9-induced GFP^−RFP^+ cells (Extended Data Fig. 8b,c). However, at the gHR1a, gHR2, gHR3, gEmx7, gEmx9 and gEmx12 sites, where nicks induced by Cas9\(^H\) (not by Cas9\(^D\)) caused lag collapse generating GFP^−RFP^+ cells, GFP^−RFP^+ cells induced by Cas9\(^H\), not those by Cas9\(^D\), were stimulated by \( Brca1 \) deficiency (Fig. 5c,d). At the
gEmx8, gEmx10 and gEmx11 sites, Cas9D-induced GFP−RFP+ cells were stimulated more than Cas9H-induced GFP−RFP+ cells by Brca1 deficiency (Fig. 5c,d). Cas9D- and Cas9H-induced GFP−RFP+ cells at these three target sites arose respectively from lag and lead collapse of DNA nicks. Thus, the combined data from these 10 sites indicated that BRCA1-mediated suppression of nCas9-induced GFP−RFP+ cells was biased towards lagging strand collision, not leading strand collision (Fig. 5e), suggesting a strand asymmetry in this BRCA1 function. Such asymmetry is not controlled by conformational context of nicked strand in the nCas9-sgRNA-DNA complex or the PAM position for Cas9 (Extended Data Fig. 8d,e).

In the GFP−RFP+ products, the second (“nested”) and third GFP copies reflect the steps of invasion and termination in repair of the damaged I-Sce-GFP, respectively (Fig. 6a). To determine the TD types that lead to GFP−RFP+ cells, we further analyzed the sequences of these two GFP copies in GFP−RFP+ products and identified three classes of invasion and termination in these GFP−RFP+ products with respect to GFP alterations: Invasion Class 1 (INV1) or Termination Class I (TER1) for the I-Sce-GFP sequence with no indels at the break point, INV2 or TER2 for the I-Sce-GFP sequence with indels at the break point, INV3 for the nested GFP with only the first half of I-Sce-GFP sequence from the break point and TER3 for the third GFP with only the second half of I-Sce-GFP from the break point (Fig. 6a).

In all 41 spontaneous GFP−RFP+ products from Brca1m/m mESC, the nested GFP and the third GFP are respectively INV1 and TER1, i.e. I-Sce-GFP with intact I-SceI site (Extended Data Fig. 9a). Two different mechanisms, i.e., typical non-allelic SCR and two-round strand invasion (TRSI) model, could mediate spontaneous generation of these GFP−RFP+ products (Extended Data Fig. 9b). Due to the longer 3’ side of the I-SceI site, spontaneous DSBs could occur with a higher probability downstream of the I-SceI site of I-Sce-GFP to induce more spontaneous GFP−RFP+ cells than spontaneous GFP+RFP+ cells via non-allelic SCR/LTGC (Extended Data Fig. 9c).

Because Cas9 induces DSBs around the I-SceI site, it is unlikely for non-allelic SCR/LTGC to generate GFP−RFP+ cells. Indeed, Cas9-induced GFP−RFP+ cells are mainly INV1 or INV2 products (Fig. 6b). This is in consistent with TRSI-mediated TD (Fig. 6c). In Brca1+/+ mESC, 19.4% Cas9-induced GFP−RFP+ products were INV1, 41.7% INV2 (only INV2/TER2) and the remaining 38.9% INV3 (Fig. 6b). In Brca1m/m mESC, 5.4% Cas9-induced GFP−RFP+ products were INV1 (only INV1/TER1), but 83.9% were INV2 (only INV2/TER2) and 10.7% INV3 (only INV3/TER2) (Fig. 6b). Surprisingly, both the nested GFP and the third GFP shared the same indels in 41 out of 47 INV2/TER2 products from Brca1m/m mESC (Extended Data Table 1). Two identical mutations could occur independently but rarely on both I-Sce-GFP of INV1/TER1 products to generate the INV2/TER2 products in TRSI-mediated TDs (Model 1 in Extended Data Fig. 9d). Thus, the I-Sce-GFP copy in the sister chromatid template could first be cleaved by Cas9 and mutated by NHEJ prior to first-round invasion. The mutation generated could then be introduced into the nested and the third GFP in INV2/TER2 products by TRSI (Model 2 in Extended Data Fig. 9d).

In addition, 43 of 44 Cas9D-induced GFP−RFP+ products and all of 53 Cas9H-induced GFP−RFP+ products were INV1/TER1 in Brca1+/+ mESC and 1 out of 44 Cas9D-induced GFP−RFP+ products INV2/TER2 at the
same gHR3 target site (Fig. 6b). Rare occurrence of INV2/TER2 in nCas9-induced GFP−RFP+ cells is consistent with lower rate of nCas9-induced targeted mutations prior to first-round strand invasion. However, only 67.5% Cas9D-induced GFP−RFP+ products were either INV1 or INV2 and 88.2% Cas9H-induced GFP−RFP+ products INV1 (only INV1/TER1) in Brca1m/m mESC (Fig. 6b). In addition, the combined frequency of Cas9-induced INV1 and INV2 was higher at 0.05% in Brca1m/m mESC than 0.033% in Brca1+/+ mESC (Fig. 6d), suggesting BRCA1 suppress TRSI-mediated TD in repair of two-ended DSBs. While the combined frequency of Cas9D-induced INV1 and INV2 was marginally smaller at 0.14% in Brca1m/m mESC than 0.16% in Brca1+/+ mESC, Cas9H induced a higher frequency of INV1 at 0.088% in Brca1m/m mESC than 0.025% in Brca1+/+ mESC, but no INV2 in both cells (Fig. 6d). This suggests a strong strand asymmetry in BRCA1-mediated suppression of nick-induced TRSI-mediated TD.

Nicks, not two-ended DSBs, induce increased MH-mediated TDs associated with BRCA1 deficiency

A small fraction of induced GFP−RFP+ cells were classified as an INV3 product, in which the left end of DSBs on I-Sce-GFP could use MH upstream of exon B in sister chromatid to mediate extensive DNA synthesis, thus creating the nested GFP that contains only the first half of I-Sce-GFP (Fig. 6e). While TER1 and TER2 are resolved by the termination step of HR (Extended Data Fig. 9d), the extended 3' end of nascent DNA could be prematurely displaced from the template of DNA synthesis upstream of I-Sce-GFP and joined with the second end generating TER3 (Fig. 6a and Extended Data Fig. 9e). In each combination of INV3 with TER1, TER2 or TER3, MH-mediated TDs span ~8-12 kb (Fig. 6e). MH analysis at the entry site for the invading ends revealed that MH is used in 13 out of 14 Cas9-induced INV3 products in Brca1+/+ cells, 4 out of 6 in Cas9-induced INV3 products and all 32 nCas9-induced INV3 products in Brca1m/m cells (Fig. 6b,f). In Brca1+/+ cells, the sizes of MH varied from 1 bp to 8 bp, but smaller in average in Cas9-induced INV3 than in nCas9-induced INV3 (Fig. 6f). These results indicate these nCas9-induced TDs are MH-mediated, resembling those associated with BRCA1 deficiency, but differ from Cas9-induced INV3.

In Brca1+/+ cells, 14 out of 36 of Cas9-induced GFP−RFP+ products were INV3, but none for nCas9-induced GFP−RFP+ products (Fig. 6b), implying MH-mediated TD in repair of one-ended DSBs is significantly suppressed in normal cells. More importantly, 38.9% Cas9-induced GFP−RFP+ products were the INV3 in Brca1+/+ cells, and 10.7% in Brca1m/m cells (Fig. 6b). The absolute frequency of Cas9-induced INV3 was reduced by 3 folds in Brca1m/m mESC (Fig. 6g), demonstrating that two-ended DSBs may not be a source for induction of TDs associated with BRCA1 mutations. Compared with no INV3 in Brca1+/+ mESC, Cas9D and Cas9H generated INV3 respectively at 32.5% and 11.8% GFP−RFP+ products in Brca1m/m cells (Fig. 6b). The absolute frequency of Cas9D- and Cas9H-induced INV3 products was significant in Brca1m/m cells, respectively at 0.066% and 0.012% (Fig. 6g), indicating BRCA1-mediated suppression of nick-induced INV3 formation. Together, these results suggest that DNA nicks, not two-ended DSBs, induce MH-mediated TDs associated with BRCA1 deficiency. In Cas9-induced MH-mediated
TDs, *Brca1* deficiency caused a greater loss of nucleotides in the invading ends and shifted the entry site for MH-mediated strand invasion further away from allelic position of the break site (Fig. 6h). This pattern of alterations was retained in nCas9-induced MH-mediated TDs in *Brca1*\textsuperscript{m/m} cells (Fig. 6h).

In the SCR-RFP reporter, BRCA1 may have a preference in suppressing nCas9-induced TDs mediated by either non-allelic SCR, TRSI or MH-mediated mechanism (Extended Data Fig. 10a). Among Cas9\textsuperscript{D}- and Cas9\textsuperscript{H}-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells, the proportions of INV3, a product of MH-mediated TD, were increased from none in *Brca1*\textsuperscript{+/+} cells to 32.5% and 11.8% in *Brca1*\textsuperscript{m/m} cells, respectively, whereas *Brca1* deficiency stimulated less or not at all nCas9-induced INV1 or INV2, a product of TRSI-mediated TD (Fig. 6b). This indicates BRCA1 control the balance between TRSI- and MH-mediated TD with a suppressive bias towards MH-mediated TD in repair of nick-induced DSBs. In *Brca1*\textsuperscript{+/+} mESC, Cas9 or nCas9 generally generated fewer GFP\textsuperscript{−}RFP\textsuperscript{+} cells than GFP\textsuperscript{+}RFP\textsuperscript{+} cells (Fig. 6i and Extended Data Fig. 10b). This suggests that non-allelic SCR/LTGC is more likely used over TRSI- and MH-mediated TD to repair Cas9- and nCas9-induced breaks. We also compared the ratios of GFP\textsuperscript{−}RFP\textsuperscript{+} cells to GFP\textsuperscript{+}RFP\textsuperscript{+} cells between *Brca1*\textsuperscript{+/+} and *Brca1*\textsuperscript{m/m} mESC. *Brca1* deficiency elevated the ratios of GFP\textsuperscript{−}RFP\textsuperscript{+} cells to GFP\textsuperscript{+}RFP\textsuperscript{+} cells induced by Cas9 at all 10 target sites (Fig. 6i), suggesting that BRCA1 promotes the choice of non-allelic SCR/LTGC over TRSI- and MH-mediated TD in repair of Cas9-induced DSBs. However, the ratios of GFP\textsuperscript{−}RFP\textsuperscript{+} cells to GFP\textsuperscript{+}RFP\textsuperscript{+} cells induced by Cas9\textsuperscript{D} or Cas9\textsuperscript{H} were not elevated by *Brca1* deficiency (Fig. 6i). This indicates a difference in BRCA1-mediated control of pathway choices between repair of nCas9-induced nicks and Cas9-induced two-ended DSBs.

**Discussion**

Defective HR repair of two-ended DSBs due to *BRCA1* mutation is often regarded as a cause of characteristic mutational signatures in *BRCA1*-deficient tumors. However, by comparing the repair outcomes for directly two-ended DSBs induced by Cas9 between *Brca1*\textsuperscript{+/+} and *Brca1*\textsuperscript{m/m} cells, this study demonstrated that this type of DSBs do not induce increased accumulation of mutational signatures resembling those in *BRCA1*-deficient tumors. Instead, as one of the most frequent endogenous DNA lesions in mammalian cells, DNA nicks could be converted into one-ended DSBs by DNA replication and become a major inducer of mutational signatures associated with *BRCA1* deficiency. By exploiting long duration of Cas9-sgRNA post-cleavage target residence\textsuperscript{43–45}, together with fast doubling time and short G1 phase of mESC, we increased the probability of a collision between nCas9-induced nicks and DNA replication forks, readily generating one-ended DSBs. Due to *BRCA1* mutation, one-ended DSBs that are supposed to be efficiently repaired by BRCA1-mediated SCR could be left unrepaired, increasing the opportunity for alternative repair pathways\textsuperscript{9,13}. In addition to providing an approach to study repair of nick-converted one-ended DSBs in a site-specific manner, we found alternative repair of these DSBs accumulates mutational signatures resembling those in *BRCA1*-deficient tumors.

First, nCas9-induced NHEJ was stimulated by up to 20 folds in *BRCA1*-deficient cells, indicating that one-ended DSBs could be repaired by NHEJ when HR is defective. The junctions of such NHEJ had more
end processing in BRCA1-deficient cells, suggesting that this alternative repair pathway does not require BRCA1-mediated end resection. The deletion length and MH usage of these NHEJ products exhibit a pattern similar to the ID6 signature in BRCA1-deficient tumors\textsuperscript{5,7}. In contrast, BRCA1-deficient cells do not preferentially accumulate Cas9-induced NHEJ products, which also show an indel pattern distinct from the ID6 signature. Second, our data demonstrated that translocations between two distant nicks, not those between two distant two-ended DSBs, were stimulated by BRCA1 deficiency. This indicates that some unrepaired one-ended DSBs accumulated in BRCA1-deficient cells could be repaired by MH-mediated translocations. While two-ended DSBs could lead to balanced translocation, this is unlikely for nick-induced one-ended DSBs owing to lack of two simultaneous ends, thus helping explain the characteristic SV signature of unbalanced translocations in BRCA1-deficient tumors. In consistent with the role of BRCA1 in fork protection, not in end resection, the junctions of translocations in BRCA1-deficient cells exhibited more end processing in repair of one-ended DSBs. Thirdly, analysis of spontaneous, Cas9- and nCas9-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells revealed that only a small fraction of Cas9- and nCas9-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells resemble MH-mediated \~10 kb TDs associated with BRCA1 deficiency. nCas9-induced MH-mediated TDs were greatly stimulated in BRCA1-deficient cells whereas Cas9-induced MH-mediated TDs were instead inhibited, suggesting that DNA nicks, not two-ended DSBs, induce BRCA1-linked TDs.

Human genome contains many directly oriented repeats. These repeats can potentially lead to DNA rearrangements such as deletions between repeats, TDs of intervening segments and repeat triplications, destabilize the genome and cause human diseases\textsuperscript{14}. In this study, we used the SCR-RFP reporter to analyze repeat-directed TDs of the intervening RFP exon B and A cassette induced by DNA nicks and found that this type of TDs can be further distinguished by the underlying “LTGC” mechanisms including typical non-allelic SCR/LTGC and TRSI-mediated TDs (Extended Data Fig. 10a). Cas9- and nCas9-induced GFP\textsuperscript{+}RFP\textsuperscript{+} cells are known to be products of repeat-directed TD by typical non-allelic SCR. In contrast, while allelic homology is expected to be the most preferred template for HR, the invading ends could be prematurely dissociated from homologous templates after allelic strand invasion followed by limited DNA synthesis in yeast\textsuperscript{46,47}. The displaced strands that restore the damaged site could reinvade into non-allelic GFP in the sister chromatid template, allowing us to capture and detect the preceding allelic strand invasion in LTGC-mediated TD products (i.e., GFP\textsuperscript{−}RFP\textsuperscript{+} cells). Thus, after detecting the INV1 and INV2 structures in majority of Cas9- and nCas9-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells in Brca1\textsuperscript{+/+} mESC, we proposed the TRSI mechanism for generation of these GFP\textsuperscript{−}RFP\textsuperscript{+} TDs. In addition, only minority of Cas9- and nCas9-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells are products of MH-mediated TD, and even none of nCas9-induced GFP\textsuperscript{−}RFP\textsuperscript{+} cells are generated in Brca1\textsuperscript{+/+} cells. This suggests that MH-mediated TD in Brca1\textsuperscript{+/+} mESC is strongly suppressed in repair of nCas9-induced DSBs but not Cas9-induced DSBs. Nevertheless, these repair choices dictated by strand invasions are applicable to DSB repair in human genome, particularly in the presence of directly oriented repeats, leading to different types of TDs and triplications\textsuperscript{48}. 
Does BRCA1 function in these choices? In Cas9-induced HR, Brca1 deficiency not only elevated the ratios of GFP−RFP+ cells to GFP+RFP+ cells but also lowered the proportion of MH-mediated TDs among GFP−RFP+ cells. In contrast, Brca1 deficiency hardly affected the ratios of GFP−RFP+ cells to GFP+RFP+ cells but increased the proportion of MH-mediated TDs among GFP−RFP+ cells from none to a detectable level in nCas9-induced HR. This suggests that BRCA1 suppresses MH-mediated TD proportionally the most among non-allelic SCR/LTGC, TRSI- and MH-mediated TD in HR repair of nCas9-induced DSBs.

Our data also revealed a replication strand asymmetry in BRCA1-mediated suppression of the LTGC bias and TRSI/MH-mediated TD in repair of nCas9-induced DSBs. Although DNA nicks can encounter either the leading strand or the lagging strand of DNA replication, being converted into one-ended DSBs, only the collision with the lagging strand promotes preferential accumulation of nCas9-induced GFP+RFP+ cells (i.e., non-allelic SCR/LTGC) or GFP−RFP+ cells (i.e., TRSI/MH-mediated TD) in BRCA1-deficient cells. It is unclear how this replication strand bias is regulated by BRCA1, but this regulation may involve distinct BRCA1 functions in end resection, strand invasion and homologous pairing with second ends. As mutational signatures in human cancer genomes recurrently exhibit a bias towards replication strand, our data indicate a replication strand asymmetry in inducing SV mutational signature in human cancer genome. Thus, it is of value to determine the contribution of DNA nicks on either replication strand to the mutational processes leading to these asymmetric mutational signatures.

In clinical applications of CRISPR/Cas9 genome editing, Cas9-induced two-ended DSBs are considered as a serious safety issue because two-ended DSBs can cause undesired on-target and off-target rearrangements. Two-ended DSBs also induce p53-mediated DDR that suppresses genome editing. Consequently, pre-existing inactivation of this pathway is preferentially selected in CRISPR/Cas9 genome editing, potentially increasing a cancer risk. In contrast, nCas9-based applications such as base editing and prime editing generate no two-ended DSBs and are thought to avoid DSB-related safety issues. However, as shown in this study, nCas9-induced nicks, in particular in context of persistent nCas9-sgRNA target residency, can be readily converted into one-ended DSBs by DNA replication, activate the DDR and cause chromosomal rearrangements, chromosomal aberrations and micronuclei formation in normal cells. These by-products may become even more prevalent in HR-defective cells such as Brca1-deficient cells. Therefore, besides safety concerns related to Cas9-induced two-ended DSBs, caution should also be taken in nCas9-based applications in cycling cells. In particular, while stronger target binding or longer target residence duration may help improve the efficiency of nCas9-based genome editing, this strategy may increase the replication-driven conversion of nCas9-induced nicks to one-ended DSBs. The conflicting needs should be balanced in order to improve nCas9-based genome modifications.

Methods And Materials

Design of SCR-RFP, SCR-dGFP, Exm1-SCR and NHEJ reporters
The SCR-RFP reporter contains two mutant GFP copies in tandem, 5' truncated GFP (TrGFP) and I-Sce-GFP interrupted by insertion of an 18-bp recognition site for the rare cutting endonuclease I-SceI\textsuperscript{18,38}. Artificial RFP exons B and A along with respective introns are placed between TrGFP and I-Sce-GFP. Site-specific two-ended DSBs induced by I-SceI or Cas9 could be repaired by non-allelic SCR using TrGFP of sister chromatid as a homologous template. STGC generate GFP$^+$RFP$^-$ cells, and LTGC generates GFP$^+$RFP$^+$ cells by GFP repeat-mediated TD of the RFP exon B and A cassette. Therefore, the frequency of GFP$^+$RFP$^-$ cells and GFP$^+$RFP$^+$ cells induced represent the level of STGC and LTGC in SCR, respectively. If the nick induced in the reporter is converted into a one-ended DSB by collision with DNA replication, the solitary end of this DSB could use TrGFP of sister chromatid as a homologous template to start non-allelic SCR. This HR could be terminated by the second end generated by the converging replication fork via second-end capture or homologous annealing, leading to STGC or LTGC. Similar to Cas9-induced HR, nCas9-induced STGC and LTGC would generate GFP$^+$RFP$^-$ cells and GFP$^+$RFP$^+$ cells, respectively. In repair of Cas9- or nCas9-induced DSBs, TD of the RFP exon B and A cassette could also be mediated by TRSI or MH in the SCR-RFP reporter, possibly generating GFP$^-$RFP$^+$ cells.

To determine the participation of homologous sequence in the second ends for nCas9-induced HR, the GFP sequence downstream of the I-SceI site of I-Sce-GFP in the SCR-RFP reporter in reporter mESC was deleted to generate the SCR-dGFP reporter by paired Cas9-sgRNA approach\textsuperscript{29}. The second end without GFP sequence is likely incapable of timely engaging second-end capture or homologous annealing for STGC in nCas9-induced HR as in Cas9-induced HR, thus favoring LTGC. To determine the replication strand asymmetry in SCR of one-ended DSBs converted from nicks encountering the leading strand or the lagging strand of DNA replication, a 44-bp human Emx1 sequence was inserted at 5' of the I-SceI site of the SCR-RFP reporter by CRISPR/Cas9-induced ssODN-based homology-directed repair\textsuperscript{44} to generate 6 additional nCas9 targeting sites, where Cas9$^D$ or Cas9$^H$ can induce a nick colliding with either replication strand.

The NHEJ reporter contains a GFP expression cassette driven by the phosphoglycerate kinase (PGK) promoter\textsuperscript{60}. No in-frame GFP is translated due to an upstream, out-of-frame translation start site Kozak-ATG ('Koz-ATG') flanked by two sequentially positioned I-SceI sites. A site-specific DSB can be induced between "Koz-ATG" and the ATG-GFP coding region by Cas9, allowing repair by mutagenic NHEJ to generate indels at the repair junction\textsuperscript{40}. In theory, only a third of indels can lead to GFP$^+$ cells, and the frequency of induced GFP$^+$ cells thus represents the relative efficiency of NHEJ.

**Plasmids and chemical reagents**

Plasmid px330 was originally obtained from Addgene (Cat #42230) and the SpCas9 cassette was cloned into pcDNA3b-Hyg-based expression vector. Plasmids expressing nCas9 (D10A and H840A) were generated by site-directed mutation from pcDNA3b-Cas9 using KOD-Plus-Neo Kit (TOYOBO). Plasmids of
sgRNAs were constructed from the U6-sgRNA vector as described previously\(^4\). The sgRNA target sequences are listed in Extended Data Table 2. The newly constructed plasmids were confirmed by Sanger sequencing. Chemical treatments were performed with Aphidicolin (CAS 38966-21-1, Sigma) at 5 \(\mu g/ml\), L-Mimosine (S7446, Selleck) and Hydroxyurea (S1896, Selleck) at 2 mM, Bleomycin (S1214, Selleck) at 20 \(\mu g/ml\), Olaparib (S1060, Selleck) at 2 \(\mu M\) and Camptothecin (S1288, Selleck) at 1 \(\mu M\).

**Cell lines and cell culture**

mESC were grown in ESC medium supplied with 20% fetal bovine serum (Gibco), 1% penicillin-streptomycin (Gibco), 2 mM L-glutamine (Gibco), 0.1 mM b-mercaptoethanol (Sigma), 0.1 mM non-essential amino acid (Gibco), 1 mM sodium pyruvate (Gibco) and 1000U/mL leukemia inhibitory factor (Millipore) on either MEF feeders or gelatinized plates. Mouse NIH3T3 cells and human U2OS cells were cultured in high glucose DMEM containing 10% fetal bovine serum, 1% penicillin-streptomycin and 2 mM L-glutamine. Brca1-deficient cell lines were generated by deletion of BRCA1 BRCT domain using paired Cas9-sgRNAs\(^{29}\). Briefly, 2×10\(^5\) mESC were transfected with the expression plasmids for paired sgRNAs and Cas9 in a 24-well plate, and were seeded onto MEF feeder cells at a 10 cm plate for single clones without any antibiotic selection. Single clones were picked, expanded and verified by PCR along with Sanger sequencing and Western blot for Brca1-deficient clones. PCR primers were listed in Extended Data Table 3.

**Transfection and reporter assays**

Transfection of mESC was performed with Lipofectamine 2000 (Invitrogen) in a 24-well plate as previously described\(^{61}\). Total 2×10\(^5\) mESC harboring the HR/NHEJ reporter were transfected with 0.5 \(\mu g\) total DNA. For U2OS and NIH3T3 cells, 1.0×10\(^5\) cells were seeded on a 24-well plate and total 0.8 \(\mu g\) DNA were transfected by Lipofectamine 2000. To initiate DNA replication in HR reporter U2OS cells, 0.16 \(\mu g\) of the SV40 LT expression plasmid in 0.8 \(\mu g\) of total DNA was co-transfected with expression plasmids for I-SceI or CRISPR nuclease. For chemical treatment, small molecule inhibitors were added at 6 h post-transfection, and replaced with fresh ones next day for a continued treatment for the rest of the experiment. Transfected or treated cells were analyzed for \(GFP^+RFP^-\), \(GFP^+RFP^+\) and \(GFP^-RFP^+\) frequencies using the Beckman Coulter CytoFLEX flow cytometer at least 3 days post-transfection. The HR/NHEJ frequencies were calculated after being corrected with background readings and normalized with transfection efficiencies as described before\(^{60,61}\).

**Generation of Emx1-SCR reporter mESC**
A 138-nt ssODN repair template (Extended Data Table 3) was designed to contain a short sequence of human Emx1 as additional cleavage sites for Cas9 flanked by homologous sequences from the HR reporter and synthesized by TsingKe Biological Technology. The ssODNs along with the Cas9-sgRNA expression plasmids were transfected into mESC and Cas9-induced DSB at specific site in the HR reporter allowed site-specific knock-in of short Emx1 sequence. After 72 h post-transfection, cells were seeded on MEF feeder cells for formation of single colonies without selection. gDNA was extracted with a purification kit (Axygen), and site-specific knock-in of short Emx1 sequence verified by PCR and Sanger sequencing to confirm the establishment of the Emx1-SCR reporter in mESC. Brca1-deficient cell lines carrying the Emx1-SCR reporter were generated using paired Cas9-sgRNA approach as described above.

**Analysis of metaphase spread and micronucleus formation**

24 h post transfection, cells were treated with colchicine (NSC 757, Selleck) at 0.2 μg/ml for 2 h to arrest in metaphase. The cells in metaphase were harvested and wash with cold PBS twice, and then incubated in 0.075 M KCl at 37°C for 20 min. Hypotonic cells were then centrifuged for 10 min at 1000 rpm and fixed in Cornoy's fixative (methanol:acetic acid = 3:1) for 20 min at room temperature twice. Cells were resuspended with 0.5 mL Cornoy's fixative and the suspension was dropped on a lean slide to release chromosomes. These slides were stained with 1% Giemsa solution for 5 min and mounted with neutral balsam (Sangon Biotech). All metaphase spread plates were imaged at a magnification of 1000 using a fluorescence microscope (Leica DM4000). Different types of chromosomal aberrations such as chromatid breaks and radial chromosomes were scored and calculated as percentage of chromosomal aberrations.

For micronucleus preparation, cells were grown on glass coverslips in a 6-well plate and washed gently with cold PBS twice after 24 h post transfection. Cells were fixed with 4% paraformaldehyde and stained with DAPI. All plates were imaged at a magnification of 1000 using a fluorescence microscope (Leica DM4000). Micronuclei were defined as discrete DNA aggregates juxtaposed to primary nuclei in cells and the frequency of micronucleated cells was calculated.

**Western blot and Immunofluorescence**

For western blot, cells were harvested after 24 h post transfection or 2 h post treatment with chemicals and lysed with RIPA buffer for 30 min. Cell extraction were separated by SDS-PAGE electrophoresis and analyzed by western blot with corresponding antibodies. The following primary antibodies used in this study were anti-GAPDH (EM1101, 1:1000) and anti-RPA32 (ET7109-41, 1:1000) from HuaBio; anti-γH2AX (#05-636, 1:1000) and anti-H2AX (#07-627, 1:1000) from Millipore; anti-p-CHK1 (#2348, 1:1000), anti-p-CHK2 (#2661, 1:1000), anti-p-P53 (#9284, 1:1000), anti-CHK2 (#2662, 1:1000) and anti-P53 (#2524, 1:1000) from Cell Signaling Technology; anti-CHK1 (sc-8408, 1:1000) from Santa Cruz; anti-p-RPA32...
(ab87277, 1:1000) and anti-53BP1 (ab21083, 1:500) from Abcam; anti-BRCA1 antibody was a gift from the Lu Lab.

For immunofluorescence staining, cells were grown on glass coverslips in a 24-well plate and fixed with 4% paraformaldehyde after culturing for 24 h with or without chemical treatment. Cells were permeabilized with 0.1% Triton X-100, followed by blocking in 5% bovine serum albumin in PBS. Cells were probed with primary antibodies and Alexa Fluor 488/594-conjugated secondary antibodies (Jackson ImmunoResearch), stained with DAPI and imaged by a fluorescence microscope (Leica DM4000).

**Cell proliferation assays**

Cells were transfected and seeded into a 96-well plate. Cell proliferation activity was measured at indicated days according to manufacturer’s instructions (CellTiter-AQueous MTS assay, Promega). Briefly, 10 μL of MTS reagents were added directly to the wells, and cell plates were incubated at 37°C for 1 h. Absorbance was measured at 490 nm on a SpectraMax M5 reader (Molecular Devices).

**PCR amplification and NHEJ quantification by Illumina deep sequencing**

To analyze the indel patterns in NHEJ products, cells were collected after NHEJ induced by Cas9 or nCas9. Briefly, at 72 h after cell transfection with expression plasmids for Cas9- or nCas9-sgRNA, cells were harvested and gDNA was isolated as described above. The targeted regions of less than 300 bp in gDNA were PCR-amplified with primers listed in Extended Data Table 3. Next-generation sequencing was performed at Novogene Co. Ltd (Beijing). PCR products purified with PCR Clean-up kit (Axygen) were end-repaired, adenylated at 3’ ends, ligated with adapters, purified, and amplified by the second round of PCR to incorporate the P7 and P5 Illumina adapters according to the manufacturer’s protocols (Yeasen, Hieff NGS Ultima DNA Library Prep Kit for Illumina). Sequences were analyzed to determine the editing efficiency and identify the indel pattern at repair junctions using DBS-Aligner as described previously.

**Detection of translocation events**

To detect the intra-chromosomal and inter-chromosomal translocation events, cells were transfected with expression plasmids for Cas9- or nCas9-sgRNA and transfection with pcDNA3β-GFP was used as transfection efficiencies. After 72 h, cells were harvested and gDNA isolated for PCR. Translocation events with junctions, as well as the Gapdh region as internal PCR control were amplified from each sample and mixed in the 10:1 ratio in volume for next-generation sequencing. The relative frequency of each translocation event was calculated as the ratio of translocation reads to Gapdh reads, and normalized by transfection efficiency. Translocation breakpoints were determined by DBS-Aligner.
Analysis of TD breakpoints in RFP+ cells

Cells transfected were sorted for spontaneous RFP+ cells and induced RFP+ cells after 72 h transfection by FACS using Beckman Moflo Astrios EQ. Sorted cells were seeded on a gelatinized 6-well plate for expansion and also on MEF feeder cells for single clone isolation. gDNA from each RFP+ single clone was isolated. Breakpoint junctions at the invasion site and the termination site were PCR-amplified with primers listed in Extended Data Table 3 and determined by Sanger sequencing. Breakpoint sequences from individual clones were analyzed and aligned to determine particular TD types.

Declarations

Data availability

All large number of microscope datasets from this work are available from the corresponding author on reasonable request, citing the experiments of interest. Deep sequencing raw datasets are available in the Sequence Read Archive (SRA).

Acknowledgements

This work is funded 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 LZ22C050001 to Y.L.F). 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.

Authors’ contributions

A-Y.X. conceived the project and supervised the study. Y-L.F. and Q.L. generated DNA constructs and cell lines, and performed experiments and bioinformatics analysis. R-D.C. conducted replication collision experiments. S-C.L. assisted with bioinformatics analysis. R-D.C., S-C.L., Z-C.H. and X-Y.Y. assisted with generation of DNA constructs and cell lines. A-Y.X., Y-L.F., Q.L. and R-D.C. analyzed and discussed the data. Y-L.F. and A-Y.X. made tables and figures, and A-Y.X. and Y-L.F. wrote the manuscript.

Competing interests

The authors declare that they have no competing interests.


**Figures**

Figure 1

**Nicks induce increased chromosomal aberrations in Brca1-deficient cells.**

a, Schematic of DSB and nick induction at short interspersed nuclear element (SINE) B2 repeats by SpCas9 nucleases. Cas9<sup>D</sup>: Cas9 D10A; Cas9<sup>H</sup>: Cas9 H840A. The PAM is highlighted by red line.

b, Immunoblot of WCEs in mESC treated with DNA damage-inducing agents or transfected with expression plasmids for dCas9-gB2, Cas9-gB2, Cas9<sup>D</sup>-gB2 and Cas9<sup>H</sup>-gB2. Transfection efficiencies were similar at about 70%.

c, Representative images of 53BP1 and gH2AX focus formation in mESC transfected with expression plasmids for dCas9-gB2, Cas9-gB2, Cas9<sup>D</sup>-gB2 and Cas9<sup>H</sup>-gB2 as indicated.

d, Quantification of 53BP1 foci in c. Each dot represents the number of 53BP1 foci in each cell.
**Figure 2**

**BRCA1 suppresses nick-induced LTGC bias in a strand-asymmetric manner.**

* a, Schematic of the SCR-RFP reporter and HR repair products induced by Cas9 and nCas9. Cas9 induces two-ended DSBs whereas nCas9-induced nicks could be converted into one-ended DSBs by DNA replication. STGC and LTGC in non-allelic SCR induced by Cas9 or nCas9 generate GFP+RFP− cells and GFP+RFP+ cells with a TD span size at ~10 kb as indicated, respectively. Only one direction of one-ended DSBs is expected to result in GFP+RFP+ cells and the direction of DNA replication generating such one-ended DSBs is indicated. Grey boxes: TrGFP and I-Sce-GFP; green box: wtGFP; circles A and B: two artificial RFP exons. Four sgRNAs targeting the I-SceI site are denoted.

* b, Relative STGC (top), LTGC (middle) and LTGC bias (bottom) induced by Cas9 or nCas9 together with each of 4 sgRNAs as indicated in isogenic Brca1+/+ and Brca1m/m mESC. Relative STGC, LTGC and LTGC bias are derived from the original measurements in Extended Data Fig. 4c by normalizing STGC, LTGC and LTGC bias in Brca1+/+ cells to 100%.
c, Schematic for generation of the Emx1-SCR reporter in mESC. The Emx1-SCR reporter was generated by knock-in of a 44-nt hEmx1 sequence as indicated at 5' of the I-SceI site of the SCR-RFP reporter in mESC. Among 6 sgRNAs targeting the Emx1-I-SceI site, gEmx7, gEmx9 and gEmx12 target the PAM on the Watson strand, and gEmx8, gEmx10 and gEmx11 on the Crick strand as indicated.

d, Schematic for generation of one-ended DSBs by collision of nCas9-induced nicks with leading strand or lagging strand of DNA replication. Collapse of DNA replication forks by collision with leading strand or lagging strand generates one-ended DSBs and is respectively termed lead collapse (I) or lag collapse (II). nCas9-sgRNAs that cause lead collapse and lag collapse are denoted.

e, Alteration of LTGC bias between Brca1+/+ and Brca1m/m cells. The replication strand that encounters nCas9-induced nicks, potentially leading to production of GFP^+RFP^+ cells, is indicated. The fold of increase in LTGC bias between Brca1+/+ and Brca1m/m cells is stated above each pair of columns for each indicated nCas9-sgRNA.

f, Combined analysis of strand asymmetry in alteration of LTGC bias by Brca1 deficiency. The replication strand that encounters nCas9-induced nicks, potentially leading to production of GFP^+RFP^+ cells as LTGC events, is indicated.

Each symbol represents the mean of at least three independent experiments for one sgRNA and statistics is performed by two-tailed Student’s t-test in b, f. Columns indicate the mean ± S.E.M and statistical significance is detected by two-tailed Student’s t-test in e. *, P<0.05 and ***, P<0.001.

Figure 3

Brca1 deficiency stimulates nick-induced NHEJ with longer deletions and more MH usage at repair junctions.

a, Cas9- or nCas9-induced NHEJ in multiple isogenic Brca1+/+ and Brca1m/m mESC clones containing an NHEJ reporter indicated on the top. The NHEJ reporter contains a stronger artificial Kozak-ATG translation start site ‘Koz-ATG’ that causes protein synthesis of out-of-frame GFP. NHEJ of Cas9- or nCas9-induced breaks in the reporter could generate GFP^+ cells by GFP reframing. Solid green circles and squares indicate independent clones.

b, Cas9- or nCas9-induced NHEJ at intron 5 of mouse Ldha and intron 2 of Rosa26 locus in Brca1+/+ and Brca1m/m mESC. The NHEJ-mediated editing efficiency is calculated as ratios of edited reads to total reads from targeted Illumina sequencing and normalized by respective transfection efficiencies. Solid circles and squares indicate independent experiments (n=5).
Analysis of NHEJ junctions for deletion length with the median length indicated (c), distribution of deletions with three size ranges (d) and the usage of MH (e) between Brca1+/+ and Brca1m/m mESC.

Statistics is performed by One-way ANOVA with Tukey's multiple comparison test for a, b. Columns indicate the mean ± S.E.M from five independent experiments, and statistical significance is detected by two-tailed Mann–Whitney test for c–e. *, P<0.05; **, P<0.01 and ***, P<0.001.

**Figure 4**

**BRCA1 suppresses nick-induced unbalanced translocation.**

a, b, Schematic for induction of intrachromosomal translocations between the Alk and Eml4 loci (a) or interchromosomal translocations between the Rosa26 and Ldha loci (b) in mESC. Red lines indicate the Cas9 or nCas9 target sites flanked by PCR primers denoted in arrowed line for translocation detections.

c,d, PCR products and quantification of Eml4–Alk translocations (c) and Rosa26–Ldha translocations (d) in Brca1+/+ and Brca1m/m mESC transfected with dCas9, Cas9 or nCas9. Schematics of Eml4–Alk translocations E\textsuperscript{RA}R and E\textsuperscript{RA}F (c) and Rosa26–Ldha translocations R\textsuperscript{RL}F and R\textsuperscript{FL}F (d) are shown on top. PCR products of translocations and Gapdh as an internal PCR control (middle) were mixed in the ratio of 10:1 in volume for Illumina sequencing. Translocations were quantified as relative translocation levels, which are ratios of translocation reads to Gapdh reads and normalized by transfection efficiencies (bottom). The fold of increase between Brca1+/+ and Brca1m/m cells is indicated above each column.

e,f, Deletion length at the junctions of Eml4–Alk translocations E\textsuperscript{RA}R and E\textsuperscript{RA}F (e) and Rosa26–Ldha translocations R\textsuperscript{RL}F and R\textsuperscript{FL}F (f) between Brca1+/+ and Brca1m/m mESC. The median length is indicated. Each dot represents 500 reads.

Columns indicate the mean ± S.E.M from three independent experiments. Statistical significance is detected by two-tailed Student’s t-test. To compare the deletion distribution between each sample, two-tailed Mann–Whitney test is performed. *, P<0.05; **, P<0.01 and ***, P<0.001.

**Figure 5**

**BRCA1 suppresses nick-induced TDs in a strand-biased manner**

a, Schematic of GFP\textsuperscript{−}RFP\textsuperscript{+} products generated from the SCR-RFP reporter.
b, Spontaneous, Cas9-gHR1b- and nCas9-gHR1b-induced GFP–RFP+ products in three of Brca1+/+ and Brca1m/m clones containing the SCR-RFP reporter.

c, d, Frequency of GFP–RFP+ cells induced by Cas9D (c) or Cas9H (d) together with one of 9 sgRNAs in Brca1+/+ and Brca1m/m cells. The replication strand that encounters nCas9-induced nicks is indicated. The fold of increase between Brca1+/+ and Brca1m/m cells is stated above each pair of columns for indicated nCas9-sgRNA.

e, Combined analysis of strand asymmetry in altering the frequency of nick-induced GFP–RFP+ cells by Brca1 deficiency. The replication strand that encounters nCas9-induced nicks is indicated.

Each dot indicates one independent experiment and statistics is performed by One-way ANOVA with Tukey’s multiple comparison test in b. Columns indicate the mean ± S.E.M from three independent experiments and statistics is performed by two-tailed Student’s t-test in c, d. Each symbol represents the mean of at least three independent experiments for one sgRNA and statistics is performed by two-tailed Student’s t-test in e. *, P<0.05; **, P<0.01 and ***, P<0.001.

Figure 6

Nick-induced TDs suppressed by BRCA1 are mediated either by two-round strand invasion or by MH.

a, Structural classification of TD products induced by Cas9-gHR3 and nCas9-gHR3. After induction of a two-ended or one-ended DSB by Cas9-gHR3 and nCas9-gHR3, results in TDs (i.e., GFP–RFP+ cells), three types of invasion junctions, i.e., INV1, INV2 and INV3 (blue dotted frame), can be determined and classified by PCR with primers F1 and R1 as indicated by blue arrows and three types of termination junctions, i.e., TER1, TER2 and TER3 (green dotted frame), with primers F2 and R2 as indicated by green arrows. The straight black line and curve red lines within I-Sce-GFP indicate an intact and mutated I-SceI site, respectively. The dotted square and circle in INV3 and TER3 denote deletion.

b, Number (in bold) of Cas9- or nCas9-induced TD events with each combined type of invasion and termination junctions in Brca1+/+ and Brca1m/m cells. The portion of each combined type in GFP–RFP+ cells analyzed is shown in parenthesis.

c, TRSI-mediated TD model. The first homologous strand invasion into allelic I-Sce-GFP restores the I-SceI site and the second homologous strand invasion into TrGFP allows TD of the RFP exon B and A cassette.

d, Frequency of Cas9- and nCas9-induced TRSI-mediated TDs in Brca1+/+ and Brca1m/m cells. TRSI-mediated TDs are characterized by GFP–RFP+ cells with INV1 or INV2. The frequency is calculated as the
overall frequency of GFP\textsuperscript{−}RFP\textsuperscript{+} cells \times the proportion of INV1 or INV2 in GFP\textsuperscript{−}RFP\textsuperscript{+} cells analyzed.

e, MH-mediated TD model. MH-mediated invasion that occurs at upstream of exon B allows TD of the RFP exon B and A cassette but generates a nested GFP that contains only the first half of I-Sce-GFP.

f, MH distribution at invasion junctions of Cas9- and nCas9-induced MH-mediated TDs in Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} cells. MH-mediated TDs are characterized by GFP\textsuperscript{−}RFP\textsuperscript{+} cells with INV3. The number for MH-mediated TD events analyzed is indicated under each column.

g, Frequency of Cas9- and nCas9-induced MH-mediated TDs in Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} cells. The frequency is calculated as the overall frequency of GFP\textsuperscript{−}RFP\textsuperscript{+} cells \times the proportion of INV3 in GFP\textsuperscript{−}RFP\textsuperscript{+} cells analyzed.

h, The processed length of Cas9- and nCas9-induced breaks (x-axis) and distance of MH-mediated invasion point to allelic position of Cas9- and nCas9-induced breaks in sister chromatid (y-axis) in MH-mediated TDs in Brca1\textsuperscript{+/+} and Brca1\textsuperscript{m/m} cells. One circle denotes one TD event.

i, Effect of Brca1 deficiency on ratios of GFP\textsuperscript{−}RFP\textsuperscript{+} cells to GFP\textsuperscript{+}RFP\textsuperscript{+} cells induced by Cas9 and nCas9 at each of 10 independent sites as indicated.

Columns indicate the mean ± S.E.M from three independent experiments and statistics is performed by two-tailed Student’s t-test in d, h. Each symbol represents the ratio at one site and statistics is performed by two-tailed Student’s t-test in i. ns, P>0.05; *, P<0.05; **, P<0.01.

Supplementary Files

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

- ExtendedDataTable1JunctionanalysisofRFPpositiveTDs.xlsx
- ExtendedDataTable2sgRNAusedinthisstudy.docx
- ExtendedDataTable3Primersandoligousedinthisstudy.docx
- FigS1.pdf
- FigS2.pdf
- FigS3.pdf
- FigS4.pdf
- FigS5.pdf
- FigS6.pdf
- FigS7.pdf
• FigS8.pdf
• FigS9.pdf
• SupplementaryInformation.docx