Check for updates

OPEN

Strand with mutagenic lesion is preferentially used as a template in the region of a bi-stranded clustered DNA damage site in *Escherichia coli*

Naoya Shikazono ✉ & Ken Akamatsu

The damaging potential of ionizing radiation arises largely from the generation of clustered DNA damage sites within cells. Previous studies using synthetic DNA lesions have demonstrated that models of clustered DNA damage exhibit enhanced mutagenic potential of the comprising lesions. However, little is known regarding the processes that lead to mutations in these sites, apart from the fact that base excision repair of lesions within the cluster is compromised. Unique features of the mutation frequencies within bi-stranded clusters have led researchers to speculate that the strand containing the mutagenic lesion is preferentially used as the template for DNA synthesis. To gain further insights into the processing of clustered DNA damage sites, we used a plasmid-based assay in *E. coli* cells. Our findings revealed that the strand containing a mutagenic lesion within a bi-stranded clustered DNA damage site is frequently used as the template. This suggests the presence of an, as yet unknown, strand synthesis process that is unrelated to base excision repair, and that this process plays an important role in mutagenesis. The length of the region of strand preference was found to be determined by DNA polymerase I.

DNA damage induced by ionising radiation has the potential to initiate events that can lead ultimately to various biological consequences including mutations, chromosomal aberrations and cell death. Various types of DNA damage, such as base lesions, apurinic/apyrimidinic (AP) sites, strand breaks, interstrand crosslinks and DNA-protein crosslinks, are produced by ionising radiation. Among these diverse types of DNA damage, clustered DNA damage, which is defined as two or more DNA lesions within one to two helical turns of DNA, has been proposed to be strongly relevant to the biological consequences of ionising radiation[1,2]. The biological significance of clustered DNA damage has been extensively studied, mainly with the use of chemically synthesized DNA lesions[3–6]. Various types of synthetic base lesions, such as 7,8-dihydro-8-oxoguanine (8-oxoG), thymine glycol (Tg), dihydrothymine (DHT), and 5-hydroxyuracil (hU), as well as AP sites and single-strand breaks (SSB), have been subjected to analysis.

In general, base lesions and AP sites are repaired via the base excision repair (BER) pathway[7]. In *Escherichia coli* (*E. coli*), a glycosylase, such as Fpg or Nth, initially recognizes the base lesion and excises the damaged base leaving an AP site. The AP site is further processed by an AP endonuclease (ExoIII or NfoI) or the AP-lyase activity of the glycosylase, thereby generating an SSB. A single deoxyribonucleotide (nt), in the case of short-patch repair, or multiple nt, in the case of long-patch repair, are then incorporated by DNA polymerase I (PolI) from the end of the SSB, and finally, the breaks are sealed by DNA ligase. *In vitro* studies have revealed that, although single, isolated base lesions/AP sites are repaired efficiently, the repair of lesions within a cluster is often retarded depending on the type, position and number of lesions within a cluster[3,4]. Consistent with these *in vitro* results, several groups have found an increase in mutation frequency (MF) within clustered DNA damage sites with two or more lesions in various configurations[8–16]. Studies of MF within these clustered DNA damage sites have indicated that: 1) there is a preferential order in the excision of different types of base lesions, 2) the resultant SSB

Institute for Quantum Life Science, National Institutes for Quantum and Radiological Science and Technology, 8-1-7 Umemidai, Kizugawa, Kyoto, 619-0215, Japan. ✉e-mail: shikazono.naoya@qst.go.jp

1

generated after the initial excision of a lesion retards the repair of a nearby mutagenic lesion on the complementary strand at the time of replication, and 3) eventually the SSB is either repaired or tolerated to allow replication of the damaged DNA.

Apart from this retardation of repair of the mutagenic lesions, the molecular mechanism of mutation induction within clustered DNA damage sites is still largely unknown. Some of the unique features of the MF in bi-stranded clusters, are: 1) the MF of 8-oxoG within a bi-stranded cluster is remarkably higher than that of a single 8-oxoG in the *fpg mutY* strain of *E. coli*, a strain in which the excision of 8-oxoG is considered to be minimal[10–12,15,17]; 2) the MF within a 8-oxoG+DHT bi-stranded cluster from a transformed clone is, in some cases, over 50%[11]; 3) the MF of hU within a three-lesion cluster is over 90%[14]; and 4) the presence of a persistent SSB at a distance of ~20 nt in either the 5′ or 3′ orientation increases the MF of a single 8-oxoG[18]. These observations led us and others to suggest that the mutagenic process most likely involves the preferential use of the DNA template containing the mutagenic lesion. It was speculated that a persistent SSB within a clustered DNA damage site leads to replication impairment and that the strand opposite the mutagenic lesion is likely lost.

When the template DNA is interrupted by an SSB, the replication fork may collapse and result in strand breakage[19,20]. Indeed, it has been suggested that a persistent SSB within a non-DSB cluster results in the production of a replication-dependent double-stranded (ds) DNA end[21]. In addition, rather than releasing dsDNA ends by fork collapse, the dsDNA end can be processed during the remodelling of the stalled replication fork, which potentially involves multi-step processes such as lesion skipping, fork reversal and/or template skipping[22–24]. Homologous recombination proteins appear to play a critical role in recovery from replication impairment[20,25]. In prokaryotes, the key protein in the recovery of replication stalling or collapse is RecA[20]. Another important process to consider when exploring the underlying mechanism of template strand preference is DNA synthesis. PolI, which is known to operate in BER, strongly suppresses the mutagenic potential of 8-oxoG in a bi-stranded cluster, and thus, is proposed to play an important role in the mutagenesis of clustered DNA damage sites[13]. However, it is still unclear how PolI reduces mutation induction, as previous studies strongly suggest that BER is compromised within clustered DNA damage sites.

To study the template strand preference in the region of a clustered DNA damage site, we previously developed a plasmid-based assay in *E. coli*[26]. It was found that the proportions of the two DNA strands used as the template were unaffected by the presence of a clustered DNA damage site ~1,400 nt away. This result demonstrated that the strand opposite the mutagenic lesion of the plasmid was not totally lost during replication. However, template strand preference around a clustered DNA damage site has not been directly shown, and little is known regarding the range of strand preference or the underlying mechanisms involved. Here, we examined whether: 1) the strand is displaced and synthesized around a bi-stranded cluster, and 2) the strand preference is affected by RecA and/or PolI. We found that a characteristic strand bias is observed around uracil + 8-oxoG cluster and that PolI, but not RecA, primarily determines the length of the region of template strand preference.

Results

To determine whether template strand preference is present around a DNA damage site, the 6-nucleotide (nt) XhoI recognition sequence (5′-CTCGAG-3′) was inserted within the uracil-containing strand of a plasmid as a strand-specific marker (Fig. 1A). A synthetic 8-oxoG lesion was inserted into the complementary strand opposite the uracil. After the plasmids were introduced into and propagated in *E. coli*, template strand preference was measured based on the fraction of the plasmid progeny with an uncut 8-oxoG stand, which lacked the XhoI sequence (a mismatch). In *E. coli*, a 6-nt mismatch is refractory to mismatch repair[26,27], and strand synthesis past the mismatch site would either eliminate or retain the XhoI recognition sequence, depending on which strand is used as a template. The positions of the mismatch relative to the DNA lesion in plasmids pGEM3Zf(−)NAEPBNG and pMW119f1(−)NAEPBNG are shown in Fig. 1B, C, respectively. It is notable that uracil in a clustered DNA damage site is rapidly converted into an AP site and then subsequently processed to generate an SSB in *E. coli* cells[28]. In a previous study, little loss was observed for plasmid constructs with a bi-stranded clustered DNA damage site comprising uracil and 8-oxoG (uracil + 8-oxoG) in *E. coli* transformants, and enhanced mutagenic potential was found for the 8-oxoG within the cluster, implying that the repair of 8-oxoG within the cluster is compromised but the potential strand break (derived from uracil) is repaired or tolerated in *E. coli*[10].

Template strand preference around DNA lesions in a RecA- and PolI-proficient strain. Using the plasmid pGEM3Zf(−)NAEPBNG, we found that the fraction of the template strand in undamaged plasmids deviated from the theoretical 50%, as was observed in previous experiments[26,29], and that this fraction varied slightly among plasmids with mismatches at different sites (Figs. 2A, 3A, and 4A). The reason for these small deviations and variations among mismatched sites is unclear. In the present study, the strand preferences in the presence and in the absence of damage were compared at each mismatched site. In a RecA- and PolI-proficient strain, the fraction of the 8-oxoG strand used as a template did not differ regardless of the presence of a single uracil (U-1) for any mismatch sites (Fig. 2B). The fraction of the 8-oxoG strand was also unaffected by a single 8-oxoG, except for the mismatch at position +21, where the fraction was smaller than that of the undamaged construct (p < 0.01) (Fig. 2C). These results indicate that the repair of 8-oxoG has a longer patch size than that of uracil. The distinct repair patch sizes between uracil and 8-oxoG are likely attributable to the type of DNA lesion and/or the relative ratios of the participating BER enzymes[30,31]. Compared with single lesions, a bi-stranded clustered DNA damage site (uracil + 8-oxoG) had a distinct influence on the fraction of the 8-oxoG strand used as a template (Fig. 2D). At +21 and −23, the 8-oxoG strand was used much more frequently (p < 0.01) compared with fractions of the corresponding plasmids without any damage. We confirmed that the presence of a mismatch does not affect the enhanced mutagenic potential of the cluster[10] (Supplementary Fig. S1).
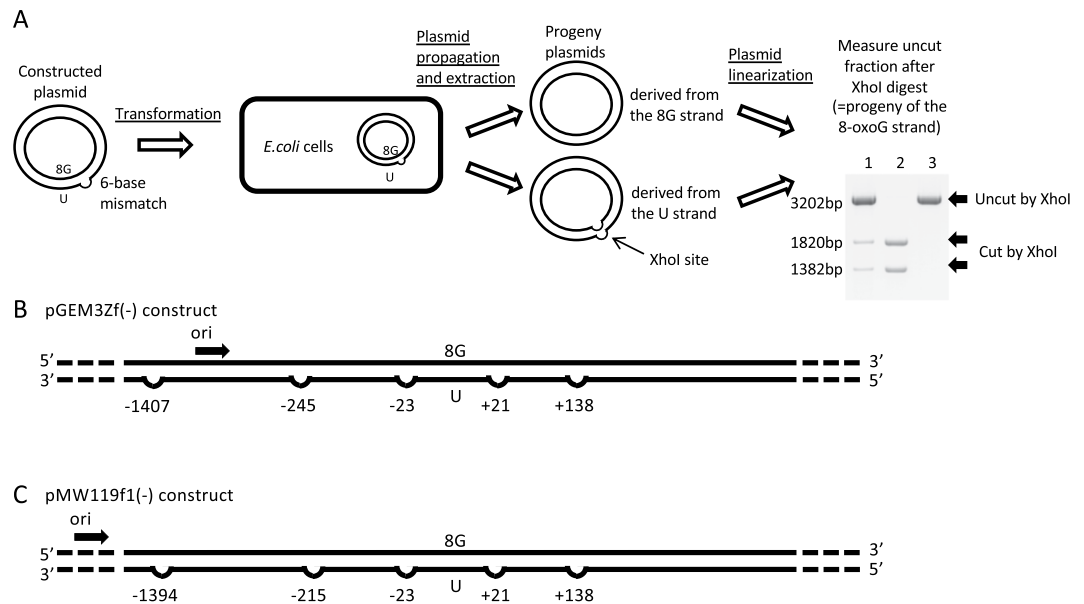
Figure 1. Schematic representations of the experiment and plasmids used to measure the level of strand bias surrounding a clustered DNA damage site. (**A**) Outline of the experiment. A plasmid was constructed that harboured a bi-stranded clustered damage site comprising uracil (U) and 8-oxoG (8G), as well as a 6-nt mismatch (XhoI site) used as a strand-specific marker. The constructed plasmids were introduced into *E. coli*, and the propagated plasmids were subsequently extracted. The level of strand preference was determined by digestion of the plasmids with XhoI, which cut the plasmids that were the progeny of the strand with the XhoI site (uracil-containing strand). XhoI-digested plasmids are shown in a representative gel image. lane 1: pGEM3Zf(−) construct with a clustered DNA damage site (uracil + 8-oxoG) and a mismatch (+21), lane 2: pGEM3Zf(−) construct without damage, but a XhoI site inserted at +21, lane 3: pGEM3Zf(−) construct without damage. The fraction of the intensity of the uncut pGEM3Zf(−) construct (3202 bp) to the intensity of all of the fragments (3202 bp + 1820 bp + 1382 bp) represents the extent of the 8-oxoG strand used as a template (See text for further details). (**B**) Mismatch sites in the pGEM3Zf(−) construct. The numbers show the separation (in nt) from 8-oxoG within the bi-stranded cluster, with − and + indicating the 5′ and 3′ direction from 8-oxoG, respectively. The black arrow designates the direction of the replication fork from the replication origin (ori). (**C**) Mismatch sites in the pMW119f1(−) construct. The numbers show the separation (in nt) from 8-oxoG within the bi-stranded cluster, with − and + indicating the 5′ and 3′ direction from 8-oxoG, respectively. The black arrow designates the direction of the replication fork from the replication origin (ori).

Template strand preference around DNA lesions in a RecA-deficient strain is similar to that in a RecA-proficient strain.

The strand preference observed around a clustered DNA damage site may have resulted from replication fork reversal and/or homologous recombination. We next examined the fraction of the 8-oxoG strand used as a template in RecA-deficient cells to determine whether the absence of RecA affected the template strand preference. We found that the fractions with a single lesion in RecA-deficient cells were quite similar to those in RecA-proficient cells (Fig. 3B, C and Fig. 2B, C) and that, overall, the template strand preference observed around a clustered DNA damage site in the RecA-deficient strain was quite similar to that in the RecA-proficient strain (Fig. 3D and Fig. 2D). These results indicate that RecA is not primarily involved in forming a bias in template strand preference around a clustered DNA damage site. We found that the mutagenic potential of clustered DNA damage is remarkably higher than that of a single lesion in the presence of a mismatch in RecA-deficient cells (Supplementary Fig. S2), which is analogous to the case in RecA-proficient cells.

The region of template strand preference around DNA lesions is larger in a PolI-deficient strain than in a PolI-proficient strain.

Because we previously found that PolI is involved in suppressing mutagenesis within clustered DNA damage sites[13], we next aimed to clarify the role of PolI in generating the strand preference around DNA lesions. Plasmid pGEM3Zf(−)NAEPBNG has a pMB1 replication origin and cannot propagate in a PolI-deficient background. Therefore, we used pMW119f1(−)NAEPBNG, which has a pSC101 replication origin. In our experiments, the template strand preferences in the region of the mismatches in pMW119f1(−) constructs were comparable to those in pGEM3Zf(−) constructs in the PolI-proficient cells (Fig. 2 and Supplementary Fig. S3). It should be noted, however, that the separations (in nt) of the mismatch from 8-oxoG were slightly different between the two plasmids (Fig. 1B, C). This was because identical oligonucleotides were used to generate the mismatch but the location of the oligonucleotide sequence differed between the two plasmids.

The presence of a single uracil dramatically increased the fractions of the 8-oxoG strand used as a template at −215, −23, +21 and +138, and up to over 90%, at −23, +21 and +138 in the PolI-deficient cells (Fig. 4B; $p < 0.01$), whereas the presence of a single 8-oxoG clearly decreased the fractions at −215, −23 and +21
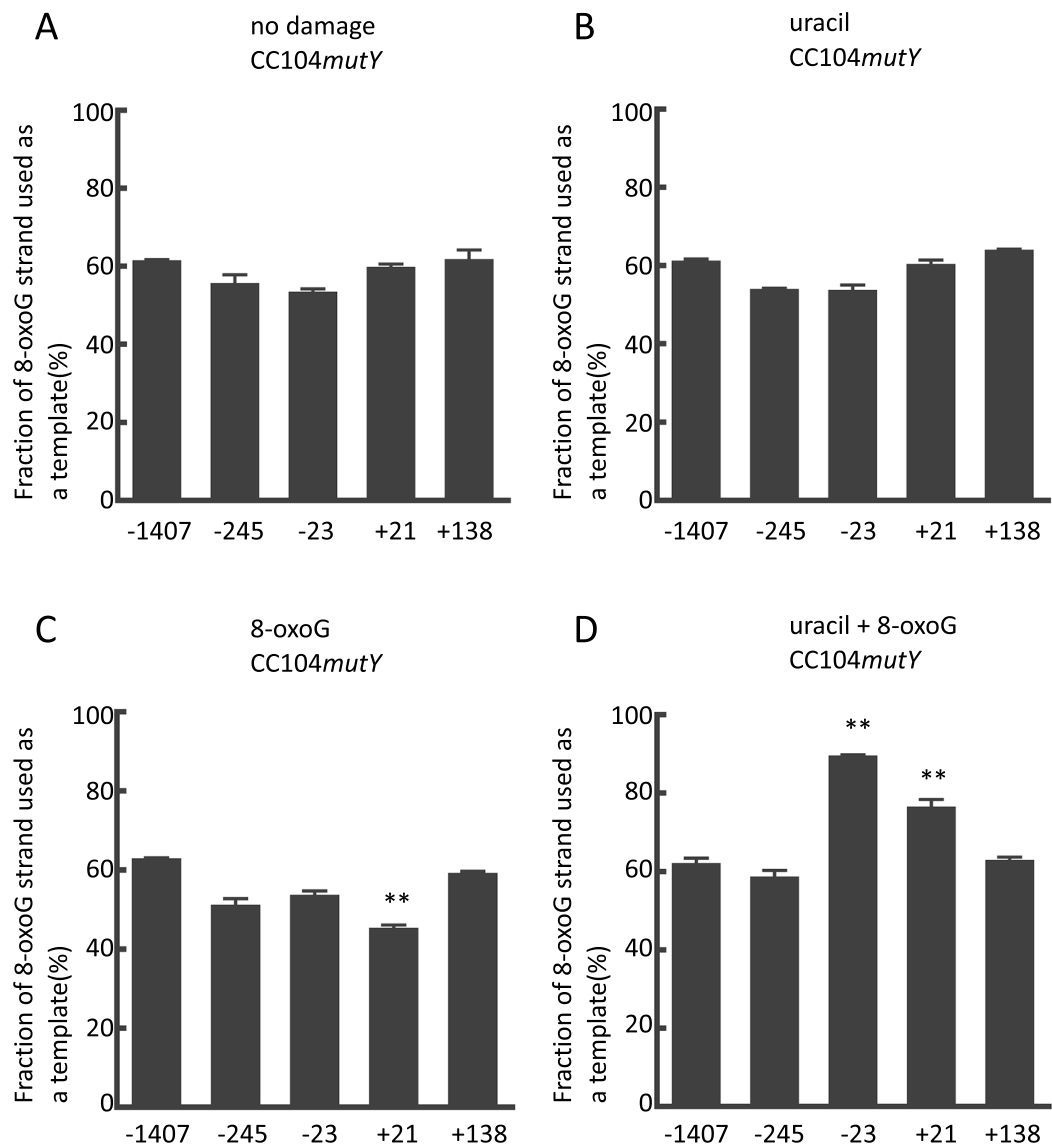
Figure 2. The fraction of the 8-oxoG strand used as a template when the pGEM3Zf(−) construct was propagated in CC104*mutY*. Panel A: no damage, Panel B: single uracil. Panel C: single 8-oxoG. Panel D: uracil + 8-oxoG bi-stranded clustered DNA damage. Along the horizontal axis, the positions of the mismatch are shown. Data represent the mean ± standard error (n = 3). Statistically significant differences between the fractions with and without DNA damage of the same mismatch site are indicated (**p < 0.01).

(p < 0.01), and was lowest at −23 (<20%; Fig. 4C). With a clustered DNA damage site, the fractions of the 8-oxoG strand used as a template at −215, −23, +21 and +138 were again significantly elevated (p < 0.01) compared with those of the undamaged plasmids (Fig. 4D). These results indicate that the range of strand preference around DNA lesions in the PolI-deficient strain is much larger than that in the PolI-proficient strain. We confirmed that, in the PolI-deficient strain, the MF within the cluster was remarkably higher than that of a single lesion with any mismatch (Supplementary Fig. S4). Consistent with a previous study[13], the MF within the clustered DNA damage site in the PolI-deficient strain was markedly higher than that in the PolI-proficient strain (Supplementary Fig. S4 and Fig. S5). We further found that the absence of Fpg in either the PolI-proficient or PolI-deficient cells did not alter the strand preference near the bi-stranded clustered DNA damage site (uracil + 8-oxoG) (Supplementary Figs. S6 and S3D, Fig. 4D), suggesting that the level of 8-oxoG removal by Fpg is, at most, quite low within the cluster.

Discussion
Our results indicate that, in the vicinity of a bi-stranded clustered DNA damage site comprising uracil and 8-oxoG, the 8-oxoG strand is preferentially used as a template and that the strand preference likely plays an important role in enhanced mutagenic potential of 8-oxoG. The observed bias in the template strand was characteristic to the clustering of the lesions. Sites at −23 and +21 showed a strong preference for the 8-oxoG strand
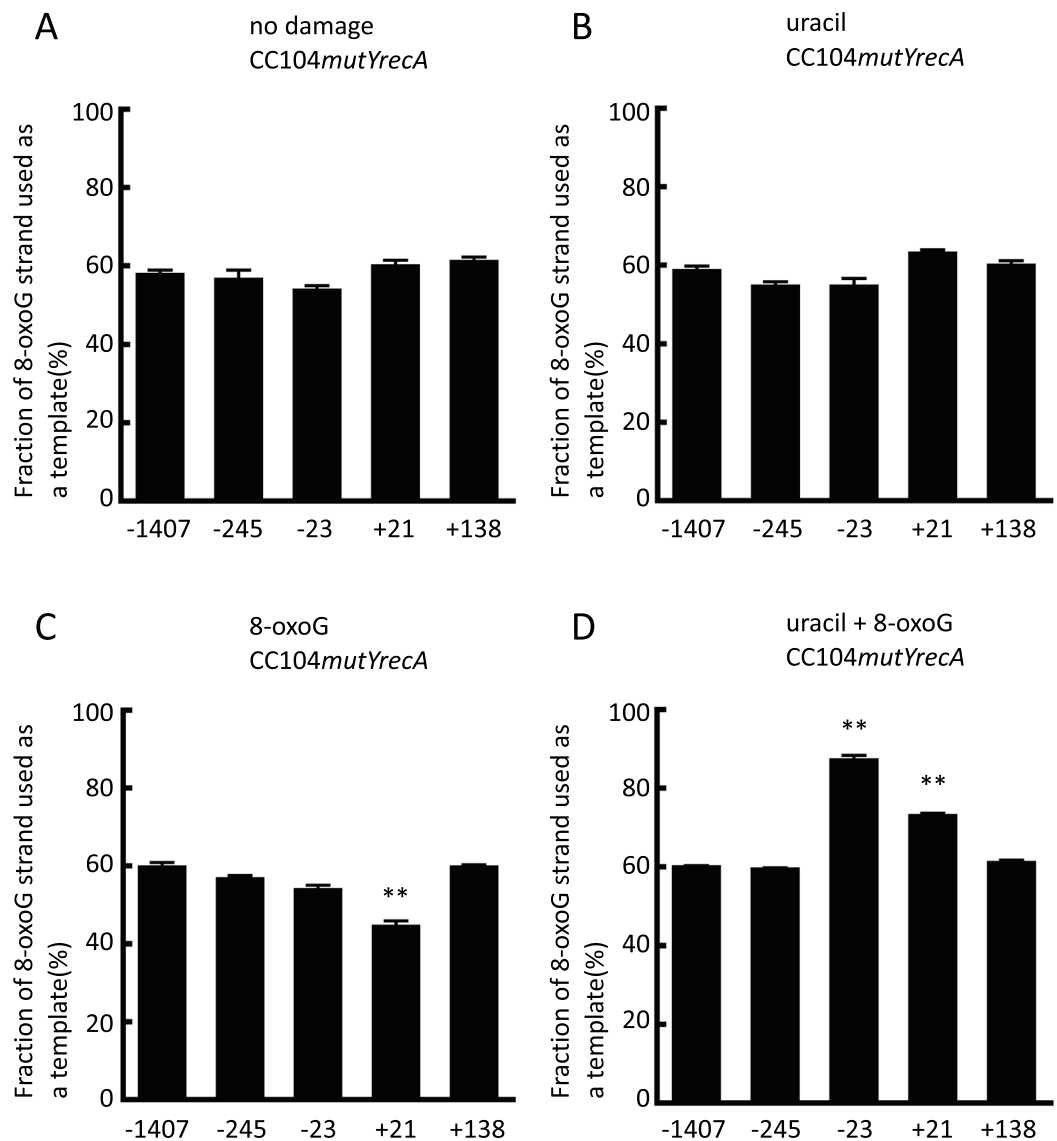
Figure 3. The fraction of the 8-oxoG strand used as a template when the pGEM3Zf(−) construct was propagated in CC104*mutYrecA*. Panel A: no damage, Panel B: single uracil. Panel C: single 8-oxoG. Panel D: uracil + 8-oxoG bi-stranded clustered DNA damage. Along the horizontal axis, the positions of the mismatch are shown. Data represent the mean ± standard error (n = 3). Statistically significant differences between the fractions with and without DNA damage of the same mismatch site are indicated (**p < 0.01).

being used as the template, which revealed that sequences in both the 5′ and 3′ directions from uracil were replaced. Previous studies have shown that the decreased MF within the bi-stranded clustered DNA damage sites is more readily observed when the separation between the lesions is increased in the 3′ rather than the 5′ direction on the complementary strand of the mutagenic lesion[11,17], which was consistent with the present finding that the level of strand replacement in the 5′ direction from uracil (+21) was lower compared with that in the 3′ direction (−23).

Previous findings indicate that, in the case of a bi-stranded cluster containing an 8-oxoG on one strand and another lesion (uracil/ 2-deoxyribonolactone/DHT/Tg) on the complementary strand, the repair of the SSB resulting from lesion processing on the complementary strand is retarded and the SSB persists until replication[10–12,14]. In canonical BER or single strand break repair in *E. coli*, the 3′ incision downstream from base damage or single strand break is most likely generated by the 5′ nuclease activity of PolI[7,31], but, to our knowledge, 5′ incision several tens of nt upstream from the damage has not been reported. These results, together with our present observations that the strand is replaced in both 5′ and 3′ directions from the lesion, support the notion that, within the clustered DNA damage site, BER is compromised and another processing pathway deals with the persistent SSB, thus leading to the characteristic template strand preference observed. As the transformation efficiency of plasmids with clustered DNA damage sites is comparable to that of plasmids with undamaged constructs[10,28], the SSB processing at a bi-stranded clustered DNA damage site (uracil + 8 G) does not lead to plasmid linearization.
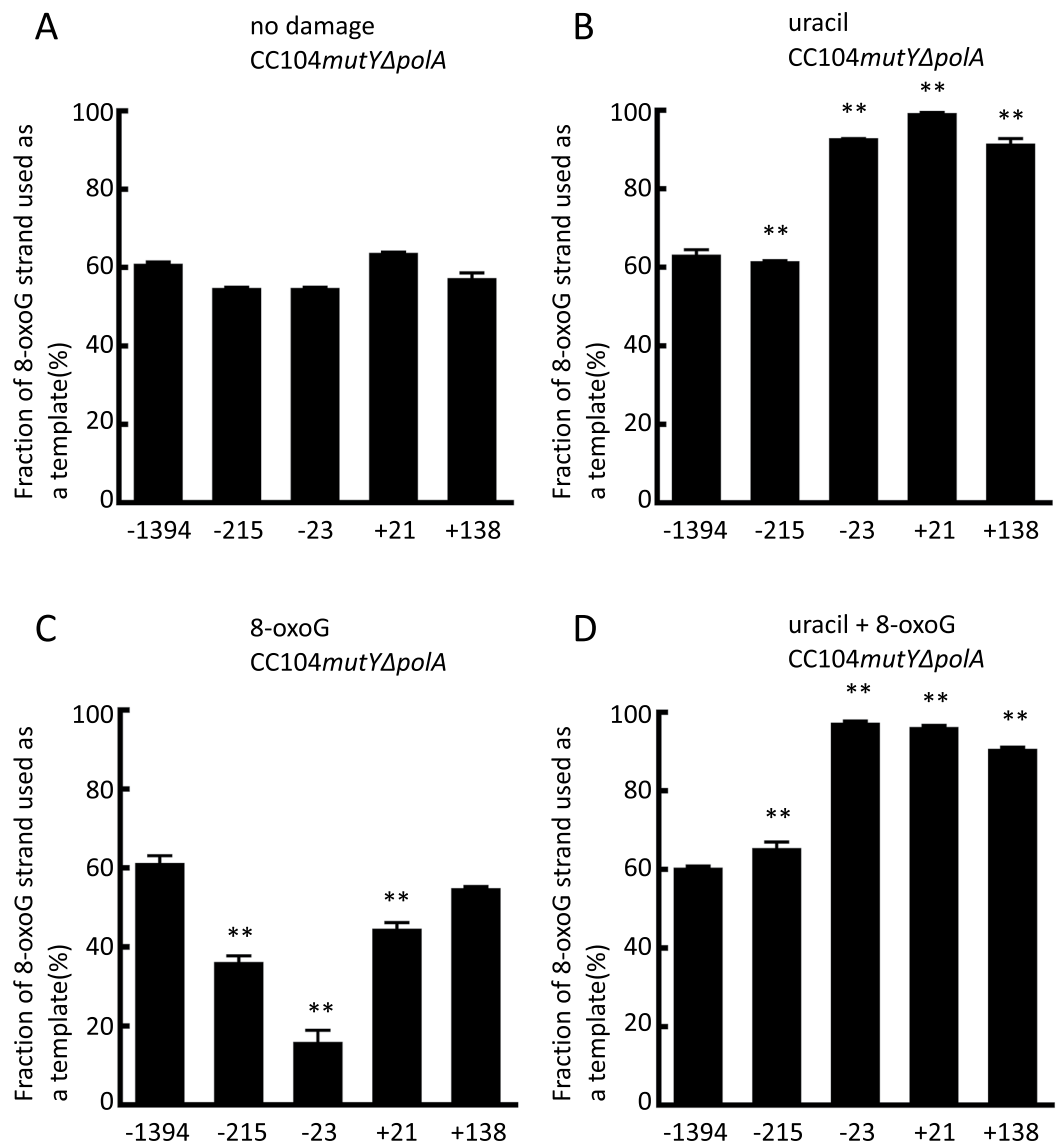
Figure 4. The fraction of the 8-oxoG strand used as a template when the pMW119f1(−) construct was propagated in CC104*mutYΔpolA*. Panel A: no damage, Panel B: single uracil. Panel C: single 8-oxoG. Panel D: uracil + 8-oxoG bi-stranded clustered DNA damage. Along the horizontal axis, the positions of the mismatch are shown. Data represent the mean ± standard error (n = 3). Statistically significant differences between the fractions with and without DNA damage of the same mismatch site are indicated (**p < 0.01).

A persistent SSB would impede the progression of the replication fork. We found that the strand preference in a RecA-deficient strain was essentially the same as that in a RecA-proficient strain. This indicates that RecA does not play a major role in tolerating the persistent SSB and in generating the template strand preference in *E. coli*; at least, not under our experimental conditions. Although RecA was found to influence the template strand preference flanking a 1,N^6-ethenodeoxyadenosine adduct when using plasmid-based assays[29], Fuchs' group demonstrated a replication-blocking type of lesion, such as an *N*-2-acetylaminofluorene adduct, is tolerated by the RecA-dependent homology-directed gap repair within genomic DNA but not within transfected plasmid DNA[32,33]. They concluded that RecA is unable to participate in the damage tolerance of plasmid DNA as the replication fork gets uncoupled. As there was no template strand preference ~1,400 nt away from the clustered damage site in any of the strains used, we consider it unlikely that uncoupling of the fork took place in the present study. Another study demonstrated that the inactivation of RecA, RecF or RecG does not affect the survival of transformed plasmids with replication-blocking lesions (i.e. pyrimidine dimers) in a nucleotide-repair-deficient background[34]. The results led the authors to conclude that lesions in the transformed plasmid DNA are repaired within the time gained by an inefficient replication, thus obviating the need for the RecA-dependent tolerance pathway. Further investigations are required to clarify whether the reduced role of RecA in the present study was due to the type of lesion (an SSB rather than a replication block) or the transformation of the plasmid.
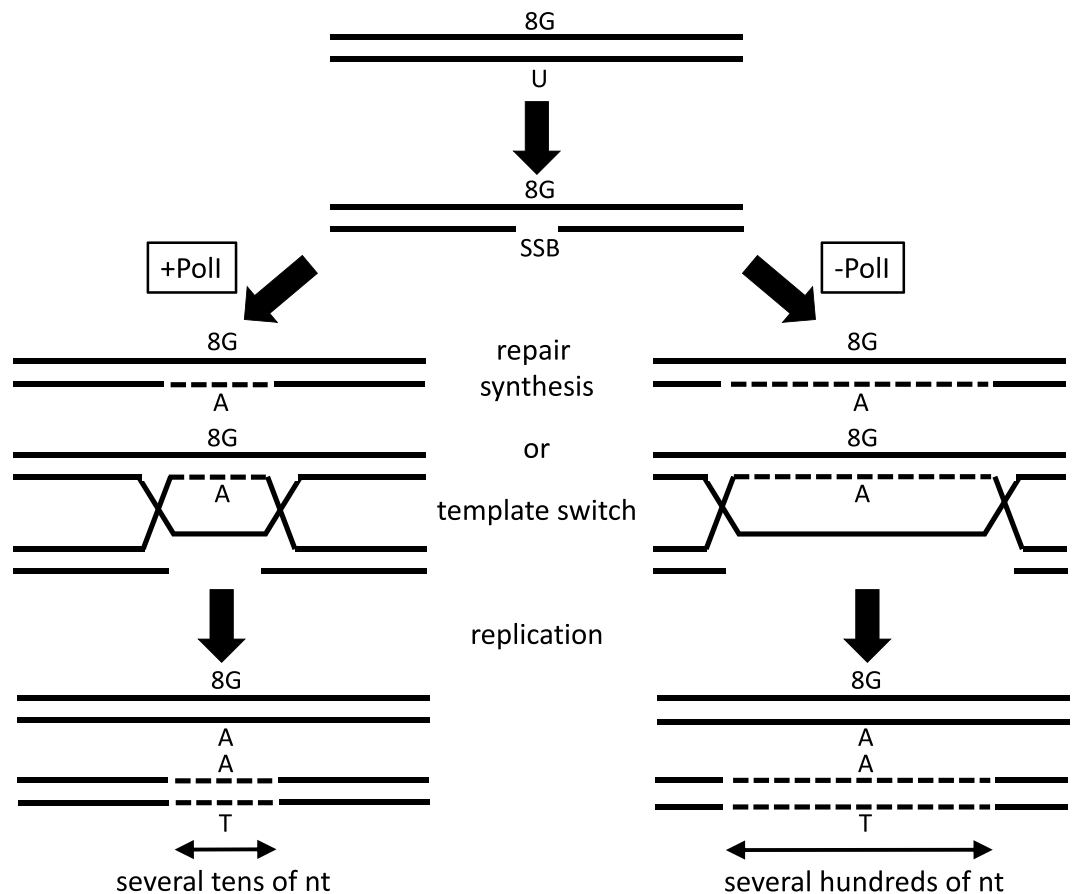
Figure 5. Possible mechanisms generating the strand bias around a bi-stranded cluster containing uracil and 8-oxoG. When PolI is present, the persistent SSB generated within the cluster is either processed by repair synthesis before replication or by template switching during replication with a patch size of several tens of nt. The strand synthesis is most likely a process unrelated to BER or NER (see Discussion). When PolI is absent, the persistent SSB generated by lesion clustering is either processed, with the use of polymerase(s) other than PolI, by repair synthesis before replication or by template switching during replication with a patch size of several hundreds of nt.

Unlike the case of RecA, the absence of PolI led to a markedly greater range of preferential strand synthesis encompassing a DNA lesion. Even the presence of either a single uracil or a single 8-oxoG strongly affected the strand preference around the lesion in the 5′ direction as well as the 3′ direction, although this effect seemed less enhanced in the 3′ direction. Our result is consistent with a previous study that indicated a similar expanded region of strand replacement from a single, MutY-removed adenine in a PolI-deficient strain[35]. We further found that the length of strand synthesis in the region of a bi-stranded clustered DNA damage site was quite similar to that for a single uracil in PolI-deficient cells. This may not be so surprising as neither the SSB created within a cluster nor from a single uracil is likely subjected to BER. An as yet unknown SSB repair or SSB tolerance pathway seems to operate in cells even when PolI is absent.

When an obstacle to replication leads to damage avoidance, a body of evidence points towards the occurrence of a template switch, which commonly involves recombination[22,25]. Our results suggest that, in the presence of PolI, a RecA-independent template switching pathway, which operates during replication fork stalling, fork regression or fork collapse, could have potentially mediated the strand synthesis involving incorporation of adenine opposite 8-oxoG and elimination of a nearby mismatch (Fig. 5, pathway A). Growing evidence supports the idea that the presence of a nick/SSB initiates homology-dependent repair not only via the canonical RecA- or Rad51-dependent strand invasion pathways but via an alternative pathway independent of RecA or Rad51 function[36,37]. An alternative possibility regarding the emergence of strand preference is the repair synthesis of the SSB before replication. For instance, nucleotide excision repair (NER), in which the repair synthesis is mediated by PolI, could be responsible for the biased use of a template strand. SSBs derived from an AP site by Endonuclease III or Endonuclease IV are indeed known to be a substrate of NER[38]. However, we consider the involvement of NER to be unlikely, as the region of strand bias observed in the PolI-proficient strain (~50 nt) was clearly larger than the repair patch length of NER in *E. coli*, where the majority of the patch lengths are in the range of 12–20 nt[39]. Consistent with our view that NER is unrelated to the occurrence of the strand bias, the levels of strand preference at sites −23 and +21 in either UvrA- or UvrC-deficient strain, which is NER deficient, were comparable with the respective levels in an NER-proficient strain (Supplementary Figs S7 and Fig. S3D). We consider that

NAEP	5′AATTCTCTTAGTCAGGAATAT<u>GTCTC</u>TATGCTGGGAGCAAAGGCTGCA3′
BNG	5′GCCTTTGCTCCCAGCATA<u>GAGAC</u>ATATTCCTGACTAAGAG3′
U-1	5′AATTCTCTTAGTCAGGAATAT<u>GUCTC</u>TATGCTGGGAGCAAAGGCTGCA3′
8G	5′GCCTTTGCTCCCAGCATA<u>GA8GAC</u>ATATTCCTGACTAAGAG3′
X-1407/X-1394	5′GATGCTTTTCTGTGACTGGTGAGT<u>CTCGAG</u>ACTCAACCAAGTCATTCTGAGAAT3′
X-245/X-215	5′CAGTCGGGAAACCTGTCGTGCCAG<u>CTCGAG</u>CTGCATTAATGAATCGGCCAACGC3′
X-23	5′AATTCTCTTAGTCAGGAATAT<u>GTCTC</u>TATGCTGGGAGCAAAGGCTG<u>CTCGAG</u>CAGGCATGCAAGCTTGAGTATTCT3′
119X-23	5′AATTCTCTTAGTCAGGAATAT<u>GTCTC</u>TATGCTGGGAGCAAAGGCTG<u>CTCGAG</u>CAGGCATGCAAGCTTGGCGTAATC3′
X-23U-1	5′AATTCTCTTAGTCAGGAATAT<u>GUCTC</u>TATGCTGGGAGCAAAGGCTG<u>CTCGAG</u>CAGGCATGCAAGCTTGAGTATTCT3′
119X-23U-1	5′AATTCTCTTAGTCAGGAATAT<u>GUCTC</u>TATGCTGGGAGCAAAGGCTG<u>CTCGAG</u>CAGGCATGCAAGCTTGGCGTAATC3′
X+21	5′TAATACGACTCACTATAGGGCGAA<u>CTCGAG</u>TTCTCTTAGTCAGGAATAT<u>GTCTC</u>TATGCTGGGAGCAAAGGCTGCA3′
X+21U-1	5′TAATACGACTCACTATAGGGCGAA<u>CTCGAG</u>TTCTCTTAGTCAGGAATAT<u>GUCTC</u>TATGCTGGGAGCAAAGGCTGCA3′
X+138	5′CGGGCCTCTTCGCTATTACGCCAG<u>CTCGAG</u>CTGGCGAAAGGGGGATGTGCTGCA3′

Table 1. Oligonucleotides used in this study. U in the oligonucleotide sequence represents uracil, while 8G indicates 8-oxoG. The six nucleotide-long XhoI recognition sequence (5′-CTCGAG-3′), which forms a mismatch in plasmid pGEM3Zf(−)NAEPBNG and pMW119f1(−)NAEPBNG, is shown by single underline. The number after X in the oligonucleotide harbouring a mismatch denotes the separation of the mismatch in nt from the 8-oxoG, and − and + indicate the 5′ and 3′ direction from 8-oxoG, respectively. 119X−23 and 119X−23U-1 are oligonucleotides that anneal with pMW119f1(−)NAEPBNG. Alw26I recognition sequences are shown by double underline. When U-1 is annealed with 8G to form a bi-stranded uracil + 8-oxoG cluster, uracil is located 1 nt 5′ from 8-oxoG on the complementary strand.

the repair/tolerance pathway of a persistent SSB is likely independent of whether the SSB is on the leading or the lagging strand template, as the direction of the replication fork does not affect the MFs of 8-oxoG placed within or nearby bi-stranded clustered damage sites[13,18]. Elucidation of the underlying mechanism(s) of strand preference awaits further investigation. When PolI is absent, a larger repair tract is required for the processing of the persistent SSB (Fig. 5, pathway B). DNA polymerases other than PolI likely synthesized the DNA and are most probably involved in the repair or avoidance of a persistent SSB with a relatively large tract.

In summary, based on our plasmid-based assay, we found that, in the region of a clustered DNA damage site comprising uracil and 8-oxoG, the strand harbouring a mutagenic lesion (8-oxoG) is preferentially used as the template. The region of such strand preference is in the order of several tens of nt. The length of strand synthesis is determined by the function of PolI but is unaffected by RecA under our experimental conditions. Although the underlying mechanism for this strand preference has not yet been clarified, we suggest that the process is unrelated to canonical BER or NER and is of significant importance for coping with a persistent SSB *in vivo*.

Materials and Methods

Bacterial strains.   Strain CC104*mutY::Tet* was provided by Dr. Q-M. Zhang-Akiyama (Kyoto University), strain AJW377 harbouring *recA::Cm* was a gift from Dr. A. J. Wolfe (Loyola University Chicago), and strain CS5539 harbouring *uvrA::Cm* was provided by Dr. N. Goosen (Leiden University). *recA::Cm* from AJW377 and *uvrA::Cm* from CS5539 were transferred to CC104*mutY* through P1 transduction, to construct CC104*mutYrecA* and CC104*mutYuvrA*, respectively. Strains CC104*fpgmutY*, CC104*mutYΔpolA* and CC104*fpgmutYΔpolA* were also constructed by P1 transduction, as reported previously[13]. To construct strain CC104*mutYuvrC*, the *uvrC* gene of CC104*mutY* was replaced with a kanamycin-resistance cassette by the Lambda Red protocol[40]. All of the strains in the present study lacked the wild-type *mutY* gene, as these strains enable detailed monitoring of the reparability of 8-oxoG[10,13].

Oligonucleotides and plasmids.   The 5′-phosphorylated oligonucleotides were purchased from Tsukuba Oligo Service (Tsukuba, Japan) (Table 1). NAEP and U-1 are identical except for the presence of uracil in U-1, and BNG and 8G are identical except for the 8-oxoG in 8G. NAEP and BNG are complementary sequences, and NAEP + BNG double-stranded oligonucleotides were inserted between the EcoRI and PstI sites of pGEM3Zf(−) to construct pGEM3Zf(−)NAEPBNG. Plasmid pMW119f1(−)NAEPBNG was constructed by replacing the HindIII–NdeI fragment of pMW119 with the HindIII–NdeI fragment of pGEM3Zf(−)NAEPBNG. Oligonucleotides X-1407/X-1394, X-245/X-215, NAEPX-23, U-1EPX-23, NAEPX + 21, U-1EPX + 21 and X + 138 contained the 6-nt XhoI recognition sequence (5′-CTCGAG-3′), which formed a mismatch in plasmid pGEM3Zf(−)NAEPBNG. Analogously, X-1407/X-1394, X-245/X-215, 119NAEPX-23, 119U-1EPX-23, NAEPX + 21, U-1EPX + 21 and X + 138 were used to create a mismatched site in pMW119f1(−)NAEPBNG.

Construction of a plasmid with DNA damage and a mismatch.   Plasmids with DNA damage and a mismatch were constructed according to the method described previously[26], with slight modifications. In brief, to the cultures of CJ236 harbouring pGEM3Zf(−)NAEPBNG or pMW119f1(−)NAEPBNG grown in the presence of uridine (75 μg/mL), VCM13 helper phages were added and phage particles were precipitated overnight. Uracil-containing circular single-stranded (ss) DNA was extracted from the pelleted phage particles with lysis buffer (10 mM MOPS, pH 6.5, 1% Triton X-100, 500 mM guanidine-HCl) and a Qiagen HiSpeed Plasmid Midi kit (Qiagen, Tokyo, Japan). The oligonucleotides (50 pmol) were annealed with 20 pmol of uracil-containing circular

ssDNA. The 1st strand was synthesized from the annealed primer in a reaction volume of 540 µL in the presence of 50 mM Tris–HCl, pH 7.5, 5 mM MgCl$_2$, 1 mM DTT, 0.5 mM ATP, 0.185 mM dNTPs, 5% PEG (average MW 8,000), 0.2 U pyrophosphatase (New England Biolabs, Tokyo, Japan), 50 U native T7 DNA polymerase (New England Biolabs) and 15 U T4 DNA ligase (Thermo Fisher Scientific, Tokyo, Japan) at 33 °C for 30 min, followed by heat inactivation of the enzymes at 75 °C for 20 min. After digestion of the uracil-containing template strand at 37 °C for 1 h with 3U UDG, 50 U Exonuclease I and 50U Exonuclease III (all from New England Biolabs), the resultant ss circular 1st strand was purified with a QIAquick Gel Extraction Kit (Qiagen). Twenty pmol of the 5′-phosphorylated primer was annealed to the 1st strand, and the 2nd strand was synthesised in the same reaction mixture used for 1st strand synthesis but in a reaction volume of 270 µl. The final product (circular dsDNA) was ethanol-precipitated and gel-purified with a QIAquick Gel Extraction Kit.

Transformation and analysis of amplified plasmids. The constructed plasmids (6.6 ng) were electroporated into E. coli cells with a Bio-Rad E.coli pulser set at 1.8 kV, 200 Ω and 25 µF. Plasmids from the overnight culture were extracted with a QIAprep Spin Miniprep Kit (Qiagen). To quantitate the fraction of the strand used as a template, plasmids were first linearized with ScaI, except for plasmids synthesized with X-1407/X-1394 (which were linearized with HindIII), and the linearized fragments were further digested with XhoI. The complete digestion with XhoI was confirmed by independent digestion with another restriction endonuclease, of which the recognition sequence was located opposite the XhoI recognition site in the constructed plasmid. After digestion with the restriction enzymes, plasmids were electrophoresed in an agarose gel, and the relative intensities of the ethidium bromide-stained DNA fragments were measured with a gel imaging system (FluoroChem Imager, ProteinSimple, Tokyo, Japan). The fraction of the strand used as a template was estimated based on the intensity of the uncut fragment relative to the intensity of all of the fragments (both cut and uncut fragments). Transformation was repeated three times for each plasmid construct, and the statistical significance of the differences in the fractions of uncut fragments between the plasmids was determined with the two-tailed t-test.

Data availability
All the data generated or analysed during this study are included in this article and its supplementary information files.

References
1. Ward, J. F. The complexity of DNA damage: relevance to biological consequences. *Int. J. Radiat. Biol.* **66**, 427–432 (1994).
2. Goodhead, D. T. Initial events in the cellular effects of ionizing radiations: clustered damage in DNA. *Int J Radiat Biol.* **65**, 7–17 (1994).
3. Shikazono, N., Noguchi, M., Fujii, K., Urushibara, A. & Yokoya, A. The yield, processing, and biological consequences of clustered DNA damage induced by ionizing radiation. *J Radiat Res* **50**, 27–36 (2009).
4. Eccles, L. J., O'Neill, P. & Lomax, M. E. Delayed repair of radiation induced clustered DNA damage: friend or foe? *Mutat Res* **711**, 134–141, https://doi.org/10.1016/j.mrfmmm.2010.11.003 (2011).
5. Sage, E. & Harrison, L. Clustered DNA lesion repair in eukaryotes: relevance to mutagenesis and cell survival. *Mutat Res* **711**, 123–133, https://doi.org/10.1016/j.mrfmmm.2010.12.010 (2011).
6. Sage, E. & Shikazono, N. Radiation-induced clustered DNA lesions: Repair and mutagenesis. *Free Radic Biol Med* **107**, 125–135, https://doi.org/10.1016/j.freeradbiomed.2016.12.008 (2017).
7. Friedberg, E. C., *et al* G. C.W.R. D.R. A.T. DNA repair and mutagenesis. Second edn, (ASM press, 2005).
8. Malyarchuk, S., Youngblood, R., Landry, A. M., Quillin, E. & Harrison, L. The mutation frequency of 8-oxo-7,8-dihydroguanine (8-oxodG) situated in a multiply damaged site: comparison of a single and two closely opposed 8-oxodG in Escherichia coli. *DNA Repair (Amst)* **2**, 695–705 (2003).
9. Malyarchuk, S., Brame, K. L., Youngblood, R., Shi, R. & Harrison, L. Two clustered 8-oxo-7,8-dihydroguanine (8-oxodG) lesions increase the point mutation frequency of 8-oxodG, but do not result in double strand breaks or deletions in Escherichia coli. *Nucleic Acids Res* **32**, 5721–5731 (2004).
10. Pearson, C. G., Shikazono, N., Thacker, J. & O'Neill, P. Enhanced mutagenic potential of 8-oxo-7,8-dihydroguanine when present within a clustered DNA damage site. *Nucleic Acids Res* **32**, 263–270 (2004).
11. Shikazono, N., Pearson, C., O'Neill, P. & Thacker, J. The roles of specific glycosylases in determining the mutagenic consequences of clustered DNA base damage. *Nucleic Acids Res* **34**, 3722–3730, https://doi.org/10.1093/nar/gkl503 (2006).
12. Bellon, S., Shikazono, N., Cunniffe, S., Lomax, M. E. & O'Neill, P. Processing of thymine glycol in a clustered DNA damage site: mutagenic or cytotoxic. *Nucleic Acids Res* **37**, 4430–4440, https://doi.org/10.1093/nar/gkp422 (2009).
13. Shikazono, N. *et al.* Significance of DNA polymerase I in *in vivo* processing of clustered DNA damage. *Mutat Res* **749**, 9–15, https://doi.org/10.1016/j.mrfmmm.2013.07.010 (2013).
14. Sedletska, Y., Radicella, J. P. & Sage, E. Replication fork collapse is a major cause of the high mutation frequency at three-base lesion clusters. *Nucleic Acids Res* **41**, 9339–9348, https://doi.org/10.1093/nar/gkt731 (2013).
15. Cunniffe, S., O'Neill, P., Greenberg, M. M. & Lomax, M. E. Reduced repair capacity of a DNA clustered damage site comprised of 8-oxo-7,8-dihydro-2'-deoxyguanosine and 2-deoxyribonolactone results in an increased mutagenic potential of these lesions. *Mutat Res* **762**, 32–39, https://doi.org/10.1016/j.mrfmmm.2014.02.005 (2014).
16. Cunniffe, S., Walker, A., Stabler, R., O'Neill, P. & Lomax, M. E. Increased mutability and decreased repairability of a three-lesion clustered DNA-damaged site comprised of an AP site and bi-stranded 8-oxoG lesions. *Int J Radiat Biol* **90**, 468–479, https://doi.org/10.3109/09553002.2014.899449 (2014).
17. Noguchi, M., Urushibara, A., Yokoya, A., O'Neill, P. & Shikazono, N. The mutagenic potential of 8-oxoG/single strand break-containing clusters depends on their relative positions. *Mutat Res* **732**, 34–42, https://doi.org/10.1016/j.mrfmmm.2011.12.009 (2012).
18. Shikazono, N. & Akamatsu, K. Mutagenic potential of 8-oxo-7,8-dihydroguanine (8-oxoG) is influenced by nearby clustered lesions. *Mutat Res* **810**, 6–12, https://doi.org/10.1016/j.mrfmmm.2018.05.001 (2018).
19. Kuzminov, A. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci USA* **98**, 8241–8246, https://doi.org/10.1073/pnas.131009198 (2001).
20. Michel, B., Sinha, A. K. & Leach, D. R. F. Replication Fork Breakage and Restart in Escherichia coli. Microbiol Mol Biol Rev 82, https://doi.org/10.1128/MMBR.00013-18 (2018).

21. Harper, J. V., Anderson, J. A. & O'Neill, P. Radiation induced DNA DSBs: Contribution from stalled replication forks? *DNA Repair (Amst)* **9**, 907–913, https://doi.org/10.1016/j.dnarep.2010.06.002 (2010).

22. Marians, K. J. Lesion Bypass and the Reactivation of Stalled Replication Forks. *Annu Rev Biochem* **87**, 217–238, https://doi.org/10.1146/annurev-biochem-062917-011921 (2018).

23. Syeda, A. H., Hawkins, M. & McGlynn, P. Recombination and replication. *Cold Spring Harb Perspect Biol* **6**, a016550, https://doi.org/10.1101/cshperspect.a016550 (2014).

24. McGlynn, P. & Lloyd, R. G. Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell* **101**, 35–45, https://doi.org/10.1016/S0092-8674(00)80621-2 (2000).

25. Cortez, D. Replication-Coupled DNA Repair. *Mol Cell* **74**, 866–876, https://doi.org/10.1016/j.molcel.2019.04.027 (2019).

26. Takahashi, M., Akamatsu, K. & Shikazono, N. A polymerization-based method to construct a plasmid containing clustered DNA damage and a mismatch. *Anal Biochem* **510**, 129–135, https://doi.org/10.1016/j.ab.2016.07.007 (2016).

27. Parker, B. O. & Marinus, M. G. Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. *Proc Natl Acad Sci USA* **89**, 1730–1734, https://doi.org/10.1073/pnas.89.5.1730 (1992).

28. Shikazono, N. & O'Neill, P. Biological consequences of potential repair intermediates of clustered base damage site in Escherichia coli. *Mutat Res* **669**, 162–168, https://doi.org/10.1016/j.mrfmmm.2009.06.004 (2009).

29. Pandya, G. A., Yang, I. Y., Grollman, A. P. & Moriya, M. Escherichia coli responses to a single DNA adduct. *J Bacteriol* **182**, 6598–6604 (2000).

30. Fortini, P. & Dogliotti, E. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair (Amst)* **6**, 398–409, https://doi.org/10.1016/j.dnarep.2006.10.008 (2007).

31. Sung, J. S. & Mosbaugh, D. W. Escherichia coli uracil- and ethenocytosine-initiated base excision DNA repair: rate-limiting step and patch size distribution. *Biochemistry* **42**, 4613–4625, https://doi.org/10.1021/bi027115v (2003).

32. Laureti, L., Demol, J., Fuchs, R. P. & Pages, V. Bacterial Proliferation: Keep Dividing and Don't Mind the Gap. *PLoS Genet* **11**, e1005757, https://doi.org/10.1371/journal.pgen.1005757 (2015).

33. Pages, V., Mazon, G., Naiman, K., Philippin, G. & Fuchs, R. P. Monitoring bypass of single replication-blocking lesions by damage avoidance in the Escherichia coli chromosome. *Nucleic Acids Res* **40**, 9036–9043, https://doi.org/10.1093/nar/gks675 (2012).

34. Jeiranian, H. A., Courcelle, C. T. & Courcelle, J. Inefficient replication reduces RecA-mediated repair of UV-damaged plasmids introduced into competent Escherichia coli. *Plasmid* **68**, 113–124, https://doi.org/10.1016/j.plasmid.2012.04.002 (2012).

35. Radicella, J. P., Clark, E. A., Chen, S. & Fox, M. S. Patch length of localized repair events: role of DNA polymerase I in mutY-dependent mismatch repair. *J Bacteriol* **175**, 7732–7736, https://doi.org/10.1128/jb.175.23.7732-7736.1993 (1993).

36. Maizels, N. & Davis, L. Initiation of homologous recombination at DNA nicks. *Nucleic Acids Res* **46**, 6962–6973, https://doi.org/10.1093/nar/gky588 (2018).

37. Verma, P. & Greenberg, R. A. Noncanonical views of homology-directed DNA repair. *Genes Dev* **30**, 1138–1154, https://doi.org/10.1101/gad.280545.116 (2016).

38. Snowden, A., Kow, Y. W. & Van Houten, B. Damage repertoire of the Escherichia coli UvrABC nuclease complex includes abasic sites, base-damage analogues, and lesions containing adjacent 5' or 3' nicks. *Biochemistry* **29**, 7251–7259, https://doi.org/10.1021/bi00483a013 (1990).

39. Van Houten, B., Gamper, H., Hearst, J. E. & Sancar, A. Analysis of sequential steps of nucleotide excision repair in Escherichia coli using synthetic substrates containing single psoralen adducts. *J Biol Chem* **263**, 16553–16560 (1988).

40. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640–6645, https://doi.org/10.1073/pnas.120163297 (2000).

Acknowledgements

Author contributions

N. S. and K. A. designed the study. N. S. performed the experiments. N. S. and K. A. analysed and interpreted the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-66651-0.

Correspondence and requests for materials should be addressed to N.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.