communications biology

Article

https://doi.org/10.1038/s42003-024-06640-5

High-complexity of DNA double-strand breaks is key for alternative end-joining choice

Zhiyang Hou[1,2], Tianxiang Yu[1,2], Qiyi Yi[3], Yan Du[4], Libin Zhou[4], Ye Zhao[3], Yuejin Wu[1], Lijun Wu[5], Ting Wang[3] ✉ & Po Bian[3] ✉

The repair of DNA double-strand breaks (DSBs) through alternative non-homologous end-joining (alt-NHEJ) pathway significantly contributes to genetic instability. However, the mechanism governing alt-NHEJ pathway choice, particularly its association with DSB complexity, remains elusive due to the absence of a suitable reporter system. In this study, we established a unique *Escherichia coli* reporter system for detecting complex DSB-initiated alternative end-joining (A-EJ), an alt-NHEJ-like pathway. By utilizing various types of ionizing radiation to generate DSBs with varying degrees of complexity, we discovered that high complexity of DSBs might be a determinant for A-EJ choice. To facilitate efficient repair of high-complexity DSBs, A-EJ employs distinct molecular patterns such as longer microhomologous junctions and non-templated nucleotide addition. Furthermore, the A-EJ choice is modulated by the degree of homology near DSB loci, competing with homologous recombination machinery. These findings further enhance the understanding of A-EJ/alt-NHEJ pathway choice.

Exogenous and endogenous stresses cause various types of DNA damage, among which DNA double-strand breaks (DSBs) are the most deleterious[1]. To eliminate the threat, life has evolved two general pathways to repair DSBs: homologous recombination (HR) and non-homologous end-joining (NHEJ)[2,3]. HR is active only in the S/G2 phases following DNA replication[4], whereas NHEJ functions throughout the cell cycle[5–7]. NHEJ consists of two sub-pathways, canonical NHEJ (c-NHEJ) and alternative NHEJ (alt-NHEJ). C-NHEJ plays a predominant role in the rapid repair of DSBs, whereas alt-NHEJ mainly participates in the slow repair of DSBs[8–10]. Alt-NHEJ, also referred to as microhomology-mediated end-joining (MMEJ)[11,12], is a pathway that utilizes microhomology for the rejoining of DSBs[13]. However, the microhomology-mediated rejoining of DSBs across the genome often leads to chromosomal translocation, fostering genetic instability that is critical for genome plasticity, evolution, and carcinogenesis[11,14]. A mechanism similar to alt-NHEJ is present in prokaryotes, known as alternative end-joining (A-EJ)[15,16]. A-EJ operates independently of the key components of c-NHEJ, namely Ku and Ligase 4. In c-NHEJ-deficient *Escherichia coli*, A-EJ functions autonomously from HR, but hijacks certain components such as the RecBCD complex, and its DNA ligation involves Ligase-A instead of ligase-B[17]. Despite limited genetic conservation between A-EJ and alt-NHEJ, they share crucial features including the involvement of their components in DNA replication and recombination, independence from Ku and Ligase 4, a strong reliance on microhomology, and utilization of NAD^+-dependent Ligase[17].

As a backup pathway, alt-NHEJ is typically employed in cases of HR or c-NHEJ deficiencies, while other factors facilitating alt-NHEJ pathway choice remains to be elucidated. Initiation of alt-NHEJ relies on end resection that is correlated with DSB complexity[11,18]. DSB complexity is defined by the nature and number of damage in close proximity and can be classified into distinct levels[18]. Restriction endonucleases generate a simple form of DSBs without chemically altering any constituent moieties, whereas physical or chemical agents generate complex DSBs with more damage in their ends and/or close proximity[19]. Endogenous non-B DNA structures such as cruciform DNA can also lead to the formation of complex DSBs[20]. Under laboratory conditions, ionizing radiation (IR) is commonly used for generating complex DSBs, with low-linear energy transfer (LET) irradiation like X- and γ-rays for low-complexity DSBs and high-LET heavy-ion irradiation for high-complexity DSBs[19,21]. However, the role of DSB complexity in initiating alt-NHEJ repair has been rarely characterized.

Mechanisms of DSB repair are primarily investigated by measuring the repair kinetics of DSBs in various genetically deficient backgrounds[22–24]. This is typically carried out using site-specific reporter systems of plasmids

[1]Key Laboratory of High Magnetic Field and Ion Beam Physical Biology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, China. [2]University of Science and Technology of China, Hefei, China. [3]Teaching and Research Section of Nuclear Medicine, School of Basic Medical Sciences, Anhui Medical University, Hefei, China. [4]Biophysics Group, Biomedical Center, Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China. [5]Institute of Physical Science and Information Technology, Anhui University, Hefei, China. ✉e-mail: wangting@ahmu.edu.cn; bianpo@ahmu.edu.cn

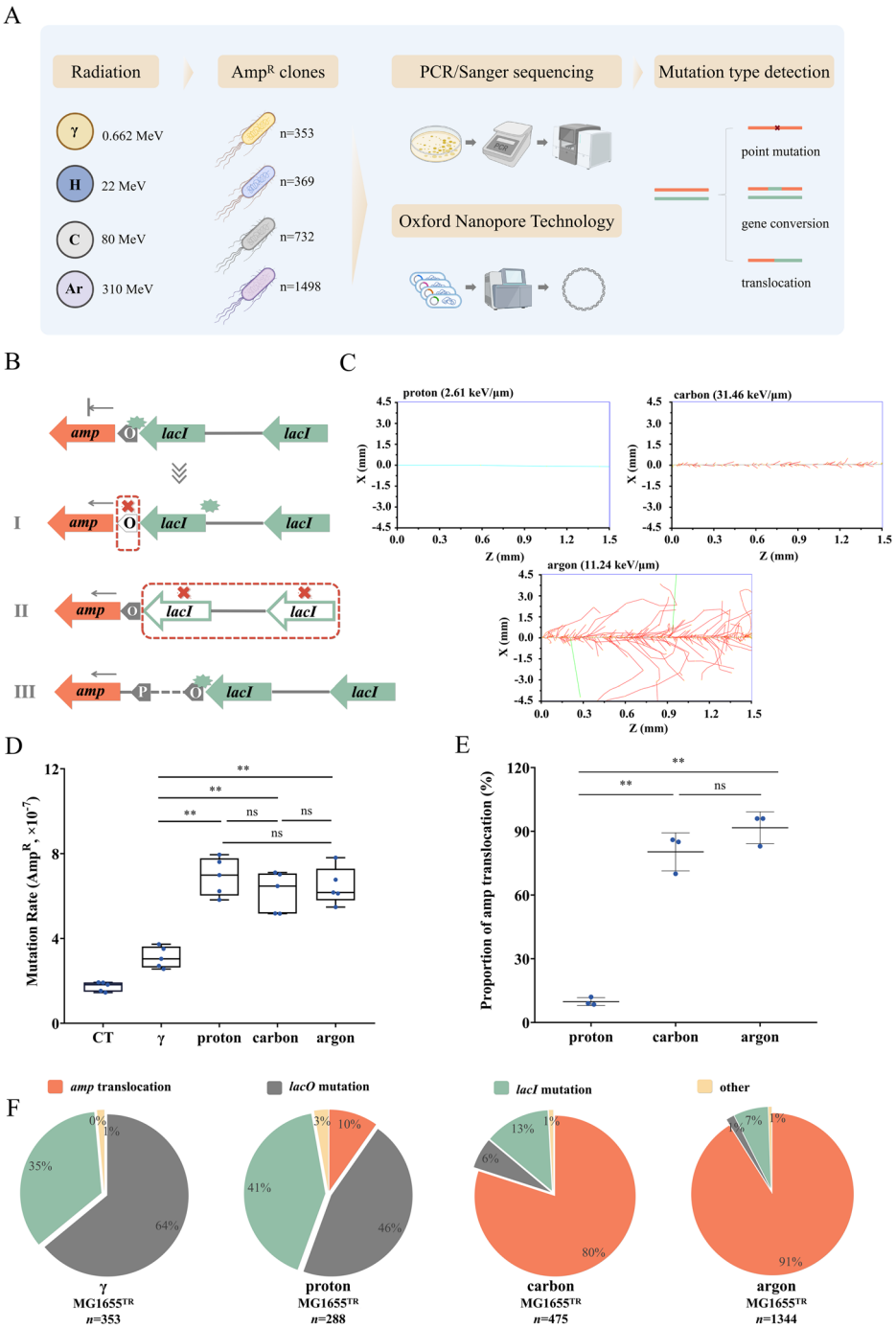

Fig. 1 | Particle irradiation initiates translocation of the *amp* gene in *Escherichia coli*. **A** An overview of the workflow, where *n* is the total number of Amp^R clones collected in this study (Created by BioRender.com). **B** Schematic diagram of the translocation reporter system. Under normal conditions, lacI (green explosive shape) specifically binds to the *lac* operator (*lacO*), hindering *lac* promoter and suppressing *amp* gene expression. Activation of *amp* gene expression can occur through three patterns: I, *lacO* mutation; II, synchronous mutations of the *lacI* genes; III, translocation of the *amp* gene. **C** Track structures of proton, carbon, and argon irradiation in water on X-Z plane. Track structures on X-Y plane are shown in Supplementary Fig. 1. **D** Mutation rates of Amp resistance (Amp^R) induced by the indicated irradiation types ($n = 5$; ns, $P \geq 0.05$; $^{**}P < 0.01$). **E** Percentages of the *amp* translocation in Amp^R clones ($n = 3$; ns, $P \geq 0.05$; $^{**}P < 0.01$). **F** Proportional distribution of the three types of *amp* activation patterns in Amp^R clones.

and chromosomes, which are suitable for endonuclease-generated DSBs with simple or chemically modified ends[22,24]. The alt-NHEJ repair of simple DSBs has been extensively characterized using thus reporter systems in prokaryotes and eukaryotes, revealing several molecular features of DSB rejoining, such as overwhelming microhomology usage, longer sequence deletion, and absence of DNA synthesis[17,22]. However, due to the stochastic distribution of complex DSBs arising from IR or other stresses throughout the genome, there is currently no available approach to molecularly characterize the alt-NHEJ repair of complex DSBs.

Here, we genetically engineered a strain of *E. coli* with a single copy of *lacI::lacI::lacO::amp* integrated into its genome, where the expression of the *amp* gene is activated by the mutation of *lacO* element, synchronous mutations of the two *lacI* genes, or translocation-mediated fusion of the *amp* gene with other coding genes. Owing to the abundance of coding sequences in the *E. coli* genome, our reporter system is capable of efficiently detecting A-EJ mediated *amp* translocation. Utilizing this unique reporter system, we characterized A-EJ repair of DSBs generated by different types of IRs (Fig. 1A).

Results

Induction of amp gene translocation by complex DSBs

In the reporter strain of *E. coli* K12 MG1655 (MG1655^{TR}), the restoration of Amp resistance (Amp^R) can occur through three different types of genetic alterations: synchronous mutations of the two *lacI* genes, mutation of the *lacO*, and ectopic expression of the *amp* gene (i.e., translocation), as schematically shown in Fig. 1B. The MG1655^{TR} strain was subjected to γ, proton, carbon, and argon irradiation with various LETs and track structures (Fig. 1C and Supplementary Fig. 1). To emphasize the radiation quality, a

dose of 100 Gy was applied for all types of irradiation. As shown in Fig. 1D, three types of particle irradiation all led to an increased proportion of Amp^R clones compared with γ-irradiation (in all cases, $P < 0.05$); however, no significant difference was observed among the particle irradiation (in all cases, $P > 0.05$) (Fig. 1E and Supplementary Table 1). Next, we detected 353 Amp^R clones generated by γ-irradiation and found that all the *amp* genes remained at their original loci. However, the *amp* genes were removed from their original loci in (9.83 ± 1.89) % (28/288), (80.33 ± 8.96) % (380/475), and (91.66 ± 7.51) % (1224/1344) of Amp^R clones generated by proton, carbon, and argon irradiation, respectively (Fig. 1E and F). Based on the Amp^R of these clones, it was concluded that the *amp* genes were not lost but were translocated to other genomic loci, correctly fusing into local encoding genes. Therefore, heavy-ion irradiation led to a higher frequency of translocation than proton irradiation (in both cases, $P < 0.01$), suggesting that high-complexity DSBs might preferentially initiate translocation. Moreover, we examined two other types of genetic alternations in the remaining Amp^R clones. As shown in Fig. 1E, γ and proton irradiation caused a higher proportion of point mutations in the *lacO*, whereas carbon and argon irradiation led to more mutations in the two *lacI* genes, including synchronous point mutations, recombinant point mutations, and structural deletions (Supplementary Table 2).

Role of A-EJ in translocation formation

Next, we determined the translocation destinations of the *amp* genes. Using thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR), we located 21, 133, and 134 translocation targets generated by proton, carbon, and argon irradiation, respectively. Molecular analysis showed that the *amp* genes were translocated across the whole genome (Fig. 2A and Supplementary Table 3), and no obvious translocation hotspot was observed for each type of irradiation, with independent translocation target in each Amp^R clone (Supplementary Table 3). The *amp* genes were fused into the local encoding genes in three patterns: upstream fusion, downstream fusion, and direct fusion with the promoter. In the three types of particle irradiation, the *amp* genes were largely fused into the upstream of targeted genes (Fig. 2B). According to the integrality of the *amp* gene in new loci, fusions are classified into two patterns: intact fusion, where no deletion occurs at the 5′ end of the *amp* gene, and missing fusion, where a small deletion occurs at the 5′ end of the *amp* gene. In the three types of particle irradiation, most of the fusions with targeted genes showed the second pattern (Fig. 2C). Among all the translocation targets, 96% (277/288) of junctions harbored 2–12 bp microhomologous sequences, 56% (154/277) of which are ≥ 3 bp (Fig. 2D and Supplementary Table 3). Relative to proton irradiation, translocations generated by carbon and argon irradiation tended to utilize shorter microhomology (2 bp). Moreover, the distributions of GC content in microhomologous junctions were not significantly different among the three types of particle irradiation (in all cases, $P > 0.05$; Fig. 2E). Surprisingly, 2–21 bp of non-templated additions were observed in translocation junctions, with 2/133 for carbon irradiation and 1/134 for argon irradiation (Supplementary Table 3). The RecBCD complex is responsible for the end resection in A-EJ[17]. We further detected the proportion of *amp* translocation in the *recB*⁻ mutant after carbon irradiation, and found a significantly reduced frequency compared with that in wild-type *E. coli* ($P < 0.05$) (Fig. 2F and Supplementary Table 1). We also conducted whole-genome sequencing on 3 Amp^R strains induced by carbon irradiation and observed a novel occurrence of translocations/deletions compared to control strains, although there was no statistically significance in the proportion of mutation types ($P > 0.05$; Fig. 2G, H). The DSBs at these loci were rejoined using microhomology (Supplementary Table 3). The extensive usage of junctional microhomology suggests an important role of A-EJ in translocation formation.

Suppressive effect of homology on translocation formation

The role of A-EJ in translocation formation raises a question regarding the function of HR in repairing complex DSBs. This question can be addressed by providing additional homologous substrates for the *amp* gene and its flanking sequence. For this purpose, the strain $MG1655^{TR}$ was transfected with two different plasmids: psgRNA that harbors a 44 bp sequence of the *amp* gene, and pl4sgRNA that includes an additional 465 bp sequence of the *lacZ* gene (Fig. 3A). Under carbon irradiation, the existence of homologous sequences significantly reduced the mutation rate of Amp^R (in both cases, $P < 0.05$) (Fig. 3B and Supplementary Table 1). We molecularly analyzed 91 Amp^R clones from the strain $MG1655^{TR}$ (psgRNA) and 68 Amp^R clones from the strain $MG1655^{TR}$ (pl4sgRNA) and found that their translocations were completely suppressed, and the activation of Amp^R was largely due to mutations in the *lacO* and *lacI* genes (Fig. 3C). Similarly, in the proton-irradiated $MG1655^{TR}$ (psgRNA) and argon-irradiated $MG1655^{TR}$ (pl4sgRNA), no translocation was detected in 153 Amp^R clones, except for mutations in the *lacO* and *lacI* genes (Fig. 3C). These results suggest that when homology exists, HR may play a predominant role in repairing complex DSBs.

Competition between A-EJ and HR at low-homology DSB loci

The strains $MG1655^{TR}$ (psgRNA) and $MG1655^{TR}$ (plfsgRNA) contained multiple homologous copies of the *amp* gene and its flanking sequence. We speculated that A-EJ might compete with HR in the repair of complex DSBs at low-homology loci. In the reporter system, the two *lacI* genes were linked by a 405 bp spacer in the same orientation (Fig. 1A), providing a single homologous substrate for each other. In the above experiments, we found that the mutations of the two *lacI* genes occurred mainly in two types of patterns: 1) point mutation of one *lacI* gene after the recombinant deletion of the other; and 2) point mutation of one *lacI* gene after the translocation/direct deletion of the other (Fig. 4A).

To evaluate the competition between HR and A-EJ, we examined the ratio of recombination and translocation of the *lacI* genes in Amp^R clones of the strain $MG1655^{TR}$ (psgRNA). Under γ-irradiation, we did not detect translocation of the *lacI* gene, but observed 35% recombination (125/353). However, we observed 31% recombination (25/81) and 7% translocation (6/81) under proton irradiation, 46% recombination (42/91) and 22% translocation (20/91) under carbon irradiation, and 29% recombination (21/72) and 36% translocation (26/72) under argon irradiation (Fig. 4B and Supplementary Table 1). The distribution of the translocation targets in the genome is shown in Fig. 4C and Supplementary Table 3. Expectedly, 94% (49/52) of translocation junctions harbored microhomology (Fig. 4D and Supplementary Table 3), indicating the involvement of A-EJ in the translocation of the *lacI* gene. We further examined their ratio in the strain $MG1655^{TR}$ (pRedCas9), in which the pRedCas9 harbored the λ-Red recombinase gene and two homologous sequences of the *amp* gene (25 bp and 22 bp). The homologous sequences were mainly used to suppress the translocation of the *amp* gene. The hyperactivity of recombinase enhanced the homologous recombination of the *lacI* genes, but reduced their translocation, with 87% recombination (85/98) and 2% translocation (2/98) in carbon irradiation and 84% (69/82) recombination and 2% translocation (2/82) in argon irradiation (Fig. 4E and Supplementary Table 1). The RecBCD complex is a key component for both HR and A-EJ. In *recB*-deficient bacteria, translocation events were completely suppressed, and recombination events also exhibited a 3-fold decrease. Moreover, there was a significant different in the proportion of mutation types between wild-type and *recB* strains ($P < 0.01$) (Fig. 4F and Supplementary Table 1). These findings suggest a competition between A-EJ and HR when DSBs occur at low-homology loci.

Discussion

Based on the unique reporter system, we here investigated A-EJ-mediated translocation following various types of IR. The translocation of the *amp* gene occurs at a lower frequency (10^{-7}) compared with that in site-specific reporter system of *E. coli* (10^{-5} for both plasmid and chromosomal DSBs)[17]. While our system allows translocation targets to be distributed across the whole genome, only the *amp* gene that is properly translocated to an encoding gene can restore Amp^R (Fig. 2A). Thus, the actual level of

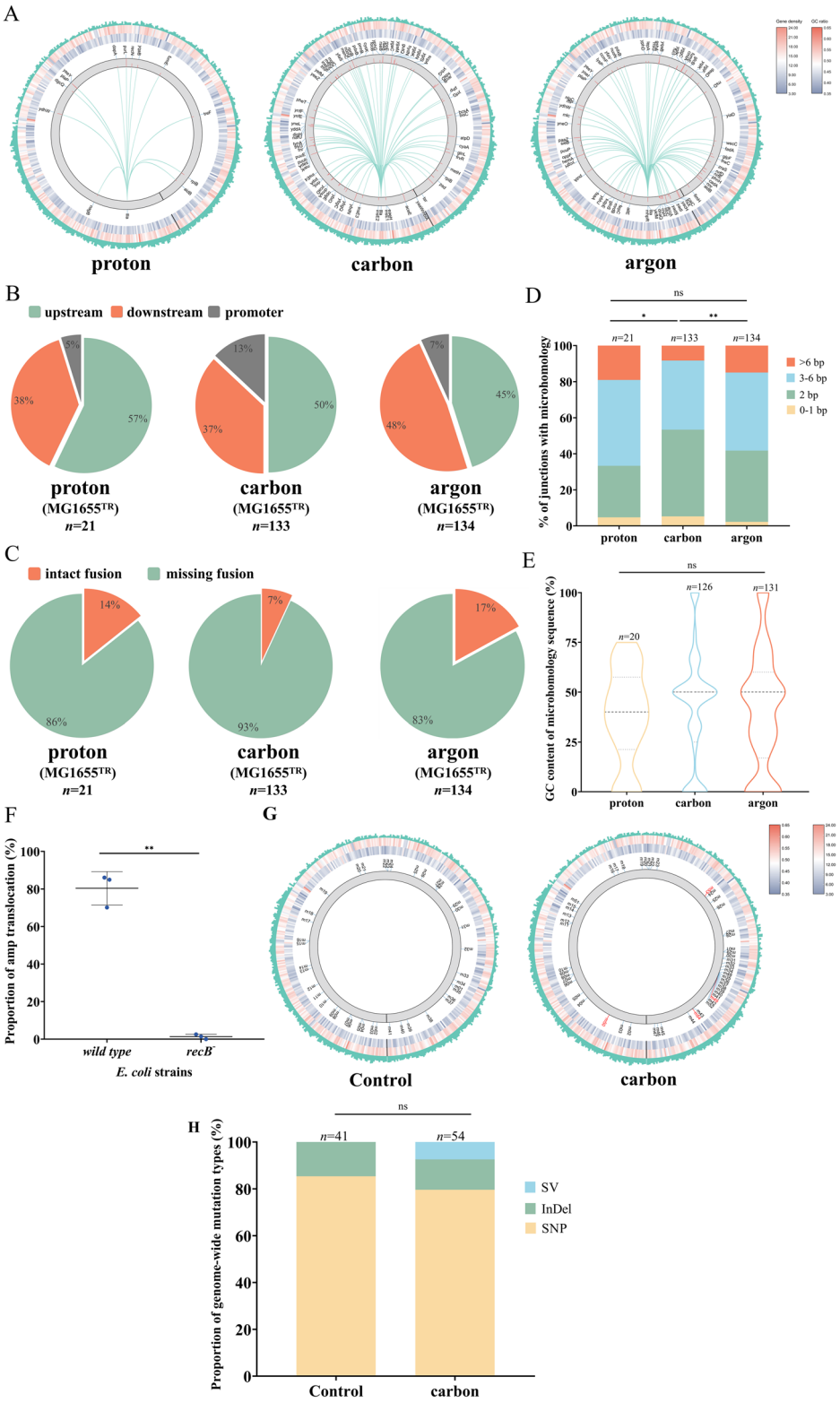

Fig. 2 | Alternative end-joining (A–EJ)-mediated translocation of the *amp* gene. A Circos plots of genome-wide translocation landscape of the *amp* gene in particle irradiation. Circles from the outside to the inside are GC skew (green), GC ratio (blue to red), gene density (blue to red), and chromosome (gray) with the micro-homology of junctions (red barplot). Individual *amp* translocations are represented as arcs originating from the llla (*lacI::lacI::lacO::amp*) site and terminating at the partner site. **B** Proportional distribution of the fusion patterns of the *amp* gene into the encoding genes in translocation targets, classified as upstream fusion, downstream fusion, and promoter fusion. **C** Proportional distribution of the intact fusion and missing fusions of the *amp* gene with encoding genes in translocation targets. **D** Proportional distribution of the lengths of micro-homologous sequences in translocation junctions (ns, $P \geq 0.05$; $^{*}P < 0.05$; $^{**}P < 0.01$). **E** Violin plot of the proportional distribution of GC content in micro-homologous junctions, which includes median (dark gray dotted line) and the lower and upper quantiles (light gray dotted line) of the data (ns, $P \geq 0.05$). **F** Proportion of *amp* translocation in *recB*$^{-}$ mutant following carbon irradiation ($^{**}P < 0.01$). **G** Circos plots of genome-wide mutation landscape in Amp^{R} strains subjected to carbon irradiation. Translocations/deletions (SV) are indicated by red fonts, while InDel and SNP are represented by black fonts. **H** Proportional distribution of three types of genome-wide mutations in bacteria irradiated by carbon ions (ns, $P \geq 0.05$).

translocation should be higher than that indicated by the current results. However, this speculation might be somewhat arbitrary because translocation preferentially targets the transcribed genomic loci[25]. While the IR-generated complex DSBs are more favorable for A-EJ than endonuclease-generated ones[11,18], their stochastic distribution also reduces the possibility of DSBs occurring at or near translocation-initiating loci. This might be another reason that our reporter system has a lower translocation frequency than site-specific reporter systems.

In this study, three types of particle irradiation were used to generate complex DSBs, but they made no significant difference in the mutation rates of Amp^{R} (Fig. 1D). Amp^{R} can be activated through three types of mutation patterns: *amp* translocation, *lacO* mutation, and synchronous mutations of the *lacI* genes. Our results show that proton irradiation preferentially mutated the *lacO* and *lacI* genes, whereas carbon and argon irradiation mainly generated the translocation of the *amp* gene (Fig. 1F). The preference of irradiation quality for mutation patterns and their common role to

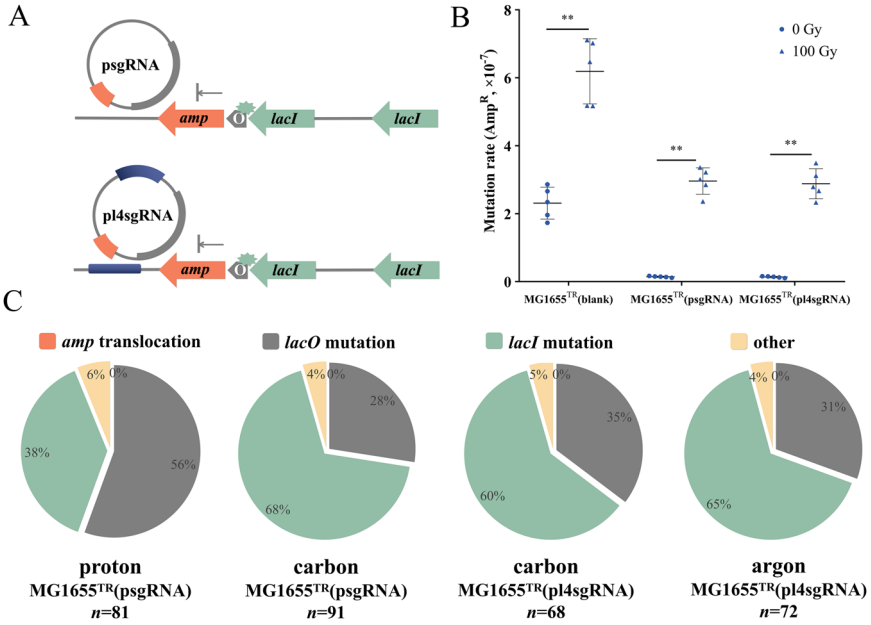

Fig. 3 | Suppressive effect of homology on *amp* translocation. A Translocation reporter strains with psgRNA plasmid that harbors a 44 bp homologous sequence of the *amp* gene (top, 44 bp sequence indicated in orange) and pl4sgRNA that harbors the 44 bp sequence and an additional 465 bp flanking sequence (bottom, 465 bp sequence indicated in blue). **B** Carbon irradiation-induced mutation rates of Amp^R in the strains $MG1655^{TR}$ (psgRNA) and $MG1655^{TR}$ (pl4sgRNA) ($n = 5$; $^{**}P < 0.01$). **C** Proportional distribution of the three patterns of the *amp* activation in Amp^R clones, in which the percentage of the *amp* translocation is zero.

activate Amp^R might be a reasonable explanation for the radiation quality-independent activation of Amp^R. On the other hand, carbon irradiation has a higher LET (31.46 keV/μm) than argon irradiation (LET, 11.24 keV/μm), but generates a lower proportion of translocations (Fig. 1E). In addition to LET, another major characteristic of heavy-ion irradiation is its radial energy transfer, also known as the track structure. Argon irradiation has a wider radial dose distribution than carbon irradiation (Fig. 1C) and can generate DSBs in a larger volume. According to the 'contact first' model for the formation of rejoining-dependent translocations[26,27], the spatial distribution of argon irradiation-induced DSBs may be more favorable than that of carbon irradiation-induced DSBs for translocation formation across a wide genome range.

The usage of junctional microhomology is a prominent feature in chromosomal translocation mediated by A-EJ and alt-NHEJ[13,17]. Single-strand annealing (SSA) has also been reported to utilize microhomology with a length > 20 bp[13]. Additionally, c-NHEJ occasionally leads to chromosomal translocation by utilizing shorter microhomology[13]. However, *E. coli* lacks Ku- and Ligase-D-like proteins and is widely acknowledged as being devoid of c-NHEJ activity, making it an ideal negative control for c-NHEJ experiments in other bacteria[17,28,29]. Taking into consideration these factors along with the fact that the length of microhomology in our experiments is < 12 bp, it should be acknowledged that in *E. coli* A-EJ may be responsible for most microhomology-dependent repair events.

The deficiency of c-NHEJ in *E. coli* also enables us to specifically characterize A-EJ-mediated translocations. Using the *E. coli* reporter system, we examined the molecular feature of translocation junctions initiated by complex DSBs, such as the overwhelming usage of microhomology (≥2 bp). The distribution of microhomology length is similar between the *amp* and *lacI* translocations (Figs. 2D and 4D), indicating that the sequence context at translocation-initiating loci does not significantly affect the usage of microhomology during A-EJ repair. Notably, most of microhomology used in the rejoining of endonuclease-generated DSBs are <3 bp[17], whereas approximately half of microhomology used in the rejoining of complex DSBs are ≥ 3 bp, with 56% for the *amp* translocation and 51% for the *lacI* translocation (Figs. 2D and 4D). On the other hand, non-templated nucleotide addition is a hallmark of c-NHEJ, and its length is 2–5 bp in prokaryote *Mycobacterium*[15]. Interestingly, A-EJ has also been found to utilize some non-templated nucleotide additions (2–21 bp), albeit at a low frequency (~1%) (Supplementary Table 3). These results suggest that A-EJ might employ additional functions to facilitate the repair of high-complexity DSBs. However, it is also likely that alternative repair pathways or their associated components could potentially participate in this process.

While there is limited genetic conservation between A-EJ and alt-NHEJ, they do exhibit similarities[17]. In the mammalian cells, alt-NHEJ is usually considered to be active only when HR or NHEJ fails, such as in the cancer[12–14]. Thus, alt-NHEJ choice in normal mammalian cells is under debate owing to its tight suppression in wild-type genetic contexts. There has been a general phenomenon that for c-NHEJ-deficient mammalian cells, the relative biological effectiveness (RBE) for inactivation does not increase with the LET of irradiation[30–35]. This indicates that c-NHEJ does not participate in the repair of high-complexity DSBs, at least in no significant way. This is further supported by the fact that resection is essential for the repair of high-complexity DSBs in all phases of the cell cycle in mammalian cells[36]. Thus, it is likely that mammalian alt-NHEJ might be also highly proficient in repairing high-complexity DSBs, and can bypass the suppression by c-NHEJ and/or HR in wild-type genetic contexts in face with high-complexity DSBs. Some early work supports this speculation. For example, EM9 cells (CHO) deficient in Xrcc1, a component of alt-NHEJ, are more sensitive to carbon irradiation than γ-irradiation[35], and G1-type chromosome breaks are efficiently rejoined in normal human fibroblasts (HFL III) during 6–24 h after carbon irradiation[34]. These findings indicate a necessary role for alt-NHEJ in the repair of high-complexity DSBs in wild-type cells. While end resection-dependent c-NHEJ contributes to chromosomal translocation following DSBs generated by X-irradiation and designer nucleases[22,23], carbon irradiation-induced RBE for inactivation is approximately 1 in Ligase 4-deficent 180BR cells (human) and Ku80-deficient xrs6 cells (CHO)[32]. This indicates a dispensable role of end resection-independent and end resection-dependent c-NHEJ in repairing high-complexity DSBs in normal cells. Therefore, we propose that alt-NHEJ could be activated by high-complexity DSBs in normal mammalian cells. In addition to heavy-ion irradiation, there exists endogenous high-complexity DSBs, such as cruciform structure-derived DSBs[19]. Alt-NHEJ has been reported to participate in the repair of such DSBs[37].

Overall, the present study has successfully established a reporter system capable of detecting A-EJ repair of complex DSBs. We also demonstrated that the complexity level of DSBs plays a critical role in determining A-EJ pathway choice, as schematically shown in Fig. 5. These findings significantly contribute our understanding of A-EJ/alt-NHEJ pathway choice and provide novel mechanistic insights into the genetic instability initiated by complex DSBs.

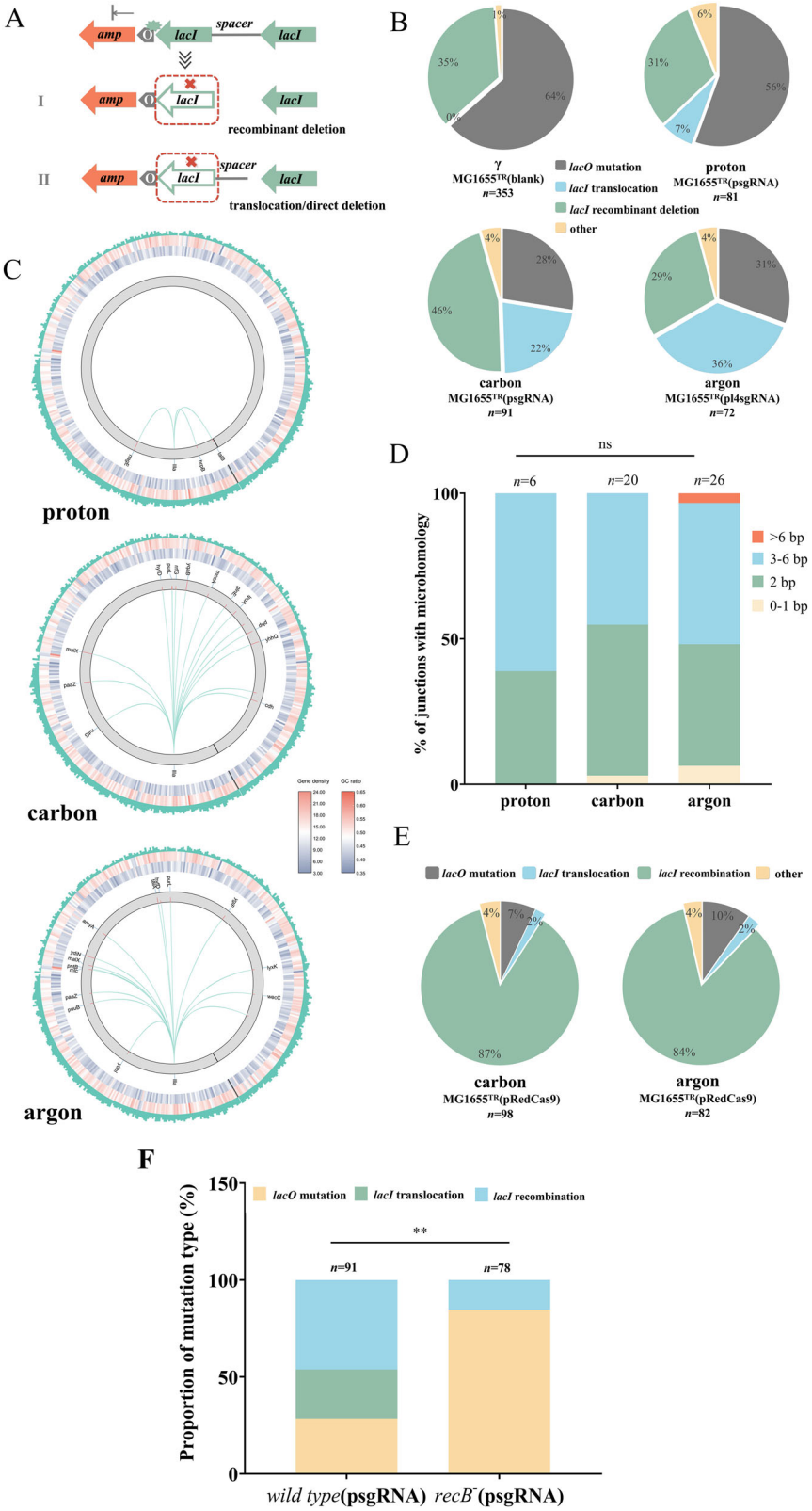

Fig. 4 | Competition between homologous recombination (HR) and alternative end-joining (A–EJ) in the repair of double-strand breaks (DSBs) at low-homology loci. A Schematic diagram of the mutation patterns of the *lacI* genes; the point mutation in one *lacI* gene is usually accompanied with recombinant deletion (I) and translocation/direct deletion (II) of the other *lacI* gene. **B** Proportional distribution of *lacO* mutation, *lacI* recombinant deletion, and *lacI* translocation in Amp^{R} clones of the strain MG1655^{TR} (psgRNA). **C** Circos plots of the genome-wide translocation landscape of the *lacI* gene in various types of irradiation. Circles from the outside to the inside are GC skew (green), GC ratio (blue to red), gene density (blue to red), and chromosome (gray) with the micro-homology of junctions (red barplot). Individual *lacI* translocations are represented as arcs originating from the llla (*lacI::lacI::lacO::amp*) site and terminating at the partner site. **D** Proportional distribution of various lengths of micro-homologous sequences in translocation junctions (ns, $P \geq 0.05$). **E** Proportional distribution of *lacO* mutation, *lacI* recombinant deletion, and *lacI* translocation in Amp^{R} clones of the strain MG1655^{TR} (pl4sgRNA), which has the hyper-active λ-Red recombinase gene. **F** Proportional distribution of translocation, recombination and *lacO* mutations in *recB*-deficient bacteria irradiated by carbon ions (**$P < 0.01$).

Materials and methods

E. coli strains and growth conditions

The *E. coli* strains utilized in the present study were all derived from *E. coli* K-12 MG1655. The strains were cultured at 37 °C or 30 °C in Luria–Bertani (LB) medium (5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl) supplemented with 50 μg/mL apramycin (cat no: A600090, Sangon Biotech, Shanghai, China), 50 μg/mL kanamycin (cat no: A600286, Sangon Biotech), or 50 μg/mL ampicillin (cat no: A610028, Sangon Biotech) when necessary[38].

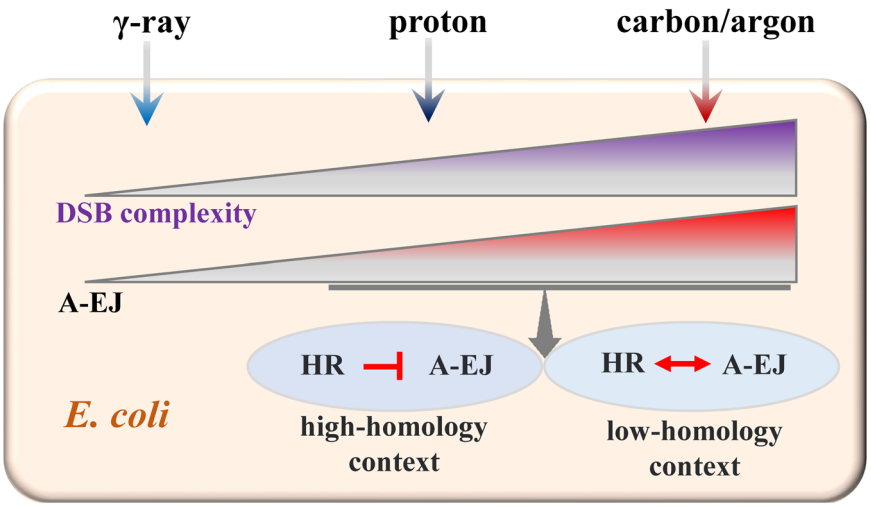

Fig. 5 | Model for the role of DSB complexity in determining the alternative end-joining (A-EJ) pathway choice. HR homologous recombination.

Establishment of translocation reporter strain of *E. coli*

DNA primers were designed using the web-based software J5 DeviceEditor (https://j5.jbei.org/) (Supplementary Table 4). The DNA fragments were PCR-amplified using KOD-Plus-Neo polymerase (cat no: KOD-401, Takara, Osaka, Japan). The restriction endonuclease BsaI (cat no: ER0291) and T4 ligase (cat no: 15224025) were purchased from ThermoFisher Scientific (Waltham, MA, USA). All plasmids used in this study were constructed using the Golden Gate method (Supplementary Table 5)[39].

The construction of translocation reporter system was described in our previous study[40]. In brief, specific fragments of *lacZ* and *mhpR* genes were utilized as N20 sequences for insertion into a single guide RNA-encoded plasmid (psgRNA), resulting in the generation of plasgRNA and pmlsgRNA plasmids. The *lacI* and *amp* genes were precisely integrated into their respective genome loci using CRISPR-Cas9 editing with these two plasmids. *E. coli* strain MG1655 was initially transformed with plasmids of pRedCas9 and plasgRNA, followed by induction of *Cas9* expression at 30 °C with 2 g/L arabinose. After screening for knock-in strains through PCR amplification and DNA sequencing, the plasmids were eliminated through multiple-generation culture at 37 °C. Subsequently, the resulting strain was transformed with both pRedCas9 and pmlsgRNA. Following similar induction and screening procedures, a transgenic strain carrying the *lacI:lacI:lacO::amp* cassette, was successfully established and designated as MG1655^{TR}. Similarly, we introduced a *recB*$^{-}$ mutation into MG1655^{TR} by utilizing a prsgRNA with the *recB* gene as the N20 sequence. Additionlly, We constructed another psgRNA-derived plasmid pl4sgRNA, which consisted a 44 bp *amp* gene sequence and an additional 465 bp *lacZ* gene sequence.

Irradiation of *E. coli* strains

For irradiation treatments, a single colony of the specified *E. coli* strains was randomly selected and inoculated into 5 mL LB medium followed by incubatiion at 37 °C until reaching an approximate concentration of 10^8 colony forming units (CFU)/mL. Subsequently, 1 mL of the dispersed *E. coli* solution was transferred to a 35 mm *petri* dish with a depth of approximately 0.3 mm for various types of irradiation treatment. The γ-irradiation was performed using a Biobeam Cs137 irradiator (cat no: GM2000, Gamma-Service Medical, Leipzig, Germany). A 22-MeV proton beam generated at the Beijing Tandem Accelerator Nuclear Physics National Laboratory, China, was utilized for proton irradiation, with an LET of 2.61 keV/μm. Carbon irradiation was conducted at 80 MeV/nucleon with an LET of 31.46 keV/μm while argon irradiation at 310 MeV/nucleon with an LET of 11.24 keV/μm at the Heavy Ion Research Facility in Lanzhou, China.

Simulation of track structures of particle irradiation

TOPAS-nBio, an extension of TOPAS (version 3.8) was used to simulate the track structures of particle irradiation[41]. The process was as follows: (1) Monte Carlo geometry: To ensure the accuracy of particle trail structure simulation, the size of the beam incident surface of water phantom was set to 0.9 mm × 0.9 mm and the depth to 1.5 mm. In the simulation, the water phantom was placed in a cube with a side length of 4.0 mm, which was introduced to maintain electron equilibrium conditions. (2) Material settings: the filling material for the above geometries was water. (3) Ion source setting: the beam source used was set as a point source and placed at the isocenter of the water phantom. The number of incident particle was set to one. (4) Physical process settings: the process of track structure simulation of the single-energy ion beam adopted the g4em-livermore physical list, which is based on the common databases of EPDL97, EADL, and EEDL. Electron impact ionization was calculated theoretically according to Seltzer's method, which combines the modified Weizsacker-Williams formula for soft collisions and the modified Möller formula for hard collisions[42].

Detection of ampicillin (Amp)-resistant clones

After irradiation, *E. coli* were gradiently diluted, and 100 μL of the resulting solutions were respectively plated on control LB agar plates and the selection plates containing 100 μg/mL of ampicillin (Amp). Following overnight incubation at 37 °C, colonies were manually counted using a dissecting microscope (ES-18BZL, Motic, Xiamen, China) to determine the number of CFUs. The Amp^R colones collected from ampicillin selection plates were further verified by individual culturing in LB liquid medium supplemented with 100 μg/mL of ampicillin. Mutation frequencies were determined by dividing the number of verified Amp^R clones by those on the control LB plates. The presented data represent the average values obtained from five independent experiments.

Location and characterization of amp translocation targets

To determine whether the *amp* gene remained at the original loci, we designed three pairs of PCR primers for the *amp* gene and its flanking sequences (Supplementary Table 4). In the Amp^R clones, no amplicon for any primer pairs was regarded as *amp* translocation. The translocation targets in the genome were located using a TAIL-PCR instrument (T100 Thermal Cycler, BIO-RAD, Hercules, USA), in which, a set of three nested primers were specially designed at the 5′ end of the *amp* gene. The special primers and arbitrary degenerate primers are listed in Supplementary Table 4. The third round amplicons were further sequenced by Sangon Bioteck Company (Shanghai, China), and the translocation targets were located through an alignment between the amplicon and *E. coli* genome using a National Center for Biotechnology Information (NCBI) alignment tool. The R package "circlize" (v0.4.15) was used to generate the Circos plots

of translocation, and the sequence analysis software DNAMAN (https://www.lynnon.com/dnaman.html) was used to determine the sequence characteristics of the translocation junctions.

Detection of recombination and translocation of the *lacI* genes

PCR was used to determine the recombination of the *lacI* genes and to preliminarily identify the translocation/direct deletion of the *lacI* gene. The primers used are listed in Supplementary Table 4. In Amp^R clones, the recombination was determined when one *lacI* gene and the whole spacer sequence were deleted but the genomic sequence flanking the deleted *lacI* gene remained intact. In the clones containing either the entire spacer or its fragment, the tail-PCR was used to further determine the *lacI* translocation (Supplementary Table 3).

Identification of mutations across genome of strains

The strains were subjected to whole-genome sequencing using Oxford Nanopore technology conducted by Biomarker Technologies Company (Beijing, China). The filtered reads were assembled using Canu software (1.5 version), followed by cyclizing assembly of genome using Circlator (1.5.5 version). The GenBlastA program (1.0.4 version) was employed to scan the entire genomes for identification of genetic alternations.

Statistical analysis

For the mutation rate of Amp^R clones and the proportion of *amp* translocations, the results are presented as the mean ± standard deviation. Differences between two groups were evaluated by Student's *t*-test for normally distributed data or Mann–Whitney *u*-test for non-normally distributed data (SPSS, 23.0 version, IBM, New York USA). The normality of data was determined using the Kolmogorov–Smirnov test. Differences in length distribution of micro-homology and proportion of mutation types were evaluated by Chi-Square test using the SPSS. Significance was considered at $P < 0.05$, indicated as $^*P < 0.05$ and $^{**}P < 0.01$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data generated in this study are available within the article and its supplementary data files, and majority of the datasets utilized are publicly available at the zenodo website (https://doi.org/10.5281/zenodo.12622604)[43]. All other data are available from the corresponding authors on reasonable request.

Received: 30 January 2024; Accepted: 26 July 2024;
Published online: 03 August 2024

References

1. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian. *cells*. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
2. San Filippo, J., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* **77**, 229–257 (2008).
3. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
4. Symington, L. S. & Gautier, J. Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* **45**, 247–271 (2011).
5. Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.* **4**, 712–720 (2003).
6. Weterings, E. & Chen, D. J. The endless tale of non-homologous end-joining. *Cell Res.* **18**, 114–124 (2008).
7. Lieber, M. R., Gu, J., Lu, H., Shimazaki, N. & Tsai, A. G. Nonhomologous DNA end joining (NHEJ) and chromosomal translocations in humans. *Subcell. Biochem.* **50**, 279–296 (2010).
8. Wang, H. et al. DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.* **65**, 4020–4030 (2005).
9. Lee-Theilen, M., Matthews, A. J., Kelly, D., Zheng, S. & Chaudhuri, J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat. Struct. Mol. Biol.* **18**, 75–79 (2011).
10. Boboila, C., Alt, F. W. & Schwer, B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* **116**, 1–49 (2012).
11. Frit, P., Barboule, N., Yuan, Y., Gomez, D. & Calsou, P. Alternative end-joining pathway (s): bricolage at DNA breaks. *DNA repair* **17**, 81–97 (2014).
12. Caracciolo, D., Riillo, C., Di Martino, M. T., Tagliaferri, P. & Tassone, P. Alternative non-homologous end-joining: error-prone DNA repair as cancer's achilles' heel. *Cancers* **13**, 1392 (2021).
13. Chang, H. H., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
14. Bunting, S. F. & Nussenzweig, A. End-joining, translocations and cancer. *Nat. Rev. Cancer* **13**, 443–454 (2013).
15. Aniukwu, J., Glickman, M. S. & Shuman, S. The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Gene Dev* **22**, 512–527 (2008).
16. Dupuy, P., Sauviac, L. & Bruand, C. Stress-inducible NHEJ in bacteria: function in DNA repair and acquisition of heterologous DNA. *Nucleic Acids Res.* **47**, 1335–1349 (2018).
17. Chayot, R., Montagne, B., Mazel, D. & Ricchetti, M. An end-joining repair mechanism in Escherichia coli. *Proc. Natl Acad. Sci. USA* **107**, 2141–2146 (2010).
18. Yajima, H. et al. The complexity of DNA double strand breaks is a critical factor enhancing end-resection. *DNA Repair* **12**, 936–946 (2013).
19. Schipler, A. & Iliakis, G. DNA double-strand–break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. *Nucleic Acids Res.* **41**, 7589–7605 (2013).
20. Wang, G. & Vasquez, K. M. Dynamic alternative DNA structures in biology and disease. *Nat. Rev. Genet*. **24**, 211–234 (2023).
21. Hagiwara, Y. et al. Clustered DNA double-strand break formation and the repair pathway following heavy-ion irradiation. *J. Radiat. Res.* **60**, 69–79 (2019).
22. Ghezraoui, H. et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell* **55**, 829–842 (2014).
23. Biehs, R. et al. DNA double-strand break resection occurs during non-homologous end joining in G1 but is distinct from resection during homologous recombination. *Mol. Cell* **65**, 671–684.e675 (2017).
24. Dutta, A. et al. Microhomology-mediated end joining is activated in irradiated human cells due to phosphorylation-dependent formation of the XRCC1 repair complex. *Nucleic Acids Res.* **45**, 2585–2599 (2017).
25. Chiarle, R. et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* **147**, 107–119 (2011).
26. Meaburn, K. J., Misteli, T. & Soutoglou, E. *Seminars in Cancer Biology*: p. 80-90 (Elsevier, 2007).
27. Roukos, V. et al. Spatial dynamics of chromosome translocations in living cells. *Science* **341**, 660–664 (2013).
28. Wilson, T. E., Topper, L. M. & Palmbos, P. L. Non-homologous end-joining: bacteria join the chromosome breakdance. *Trends Biochem. Sci.* **28**, 62–66 (2003).

29. Malyarchuk, S. et al. Expression of Mycobacterium tuberculosis Ku and Ligase D in Escherichia coli results in RecA and RecB-independent DNA end-joining at regions of microhomology. *DNA Repair* **6**, 1413–1424 (2007).
30. Blakely, E. A., Ngo, F. Q., Curtis, S. B. & Tobias, C. Heavy-ion radiobiology: cellular studies. *Adv. Radiat. Biol.* **11**, 295–389 (1984).
31. Lett, J., Cox, A. & Bergtold, D. Cellular and tissue responses to heavy ions: basic considerations. *Radiat. Environ. Biophys.* **25**, 1–12 (1986).
32. Eguchi-Kasai, K. et al. The role of DNA repair on cell killing by charged particles. *Adv. Space Res.* **18**, 109–118 (1996).
33. Weyrather, W. K., Ritter, S., Scholz, M. & Kraft, G. RBE for carbon track-segment irradiation in cell lines of differing repair capacity. *Int. J. Radiat. Biol.* **75**, 1357–1364 (1999).
34. Okayasu, R. et al. Repair of DNA damage induced by accelerated heavy ions in mammalian cells proficient and deficient in the non-homologous end-joining pathway. *Radiat. Res.* **165**, 59–67 (2006).
35. Tokuyama, Y., Furusawa, Y., Ide, H., Yasui, A. & Terato, H. Role of isolated and clustered DNA damage and the post-irradiating repair process in the effects of heavy ion beam irradiation. *J. Radiat. Res.* **56**, 446–455 (2015).
36. Averbeck, N. B. et al. DNA end resection is needed for the repair of complex lesions in G1-phase human cells. *Cell Cycle* **13**, 2509–2516 (2014).
37. Lu, S. et al. Short inverted repeats are hotspots for genetic instability: relevance to cancer genomes. *Cell Rep.* **10**, 1674–1680 (2015).
38. Tuttle, A. R., Trahan, N. D. & Son, M. S. Growth and maintenance of Escherichia coli laboratory strains. *Curr. Protoc.* **1**, e20 (2021).
39. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA assembly design automation software. *ACS Synth. Biol.* **1**, 14–21 (2012).
40. Hou, Z. Y. et al. Enhancement of repeat-mediated deletion rearrangement induced by particle irradiation in a recA-dependent manner in Escherichia coli. *Biology* **12**, 1406 (2023).
41. McNamara, A. et al. Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries. *Phys. Med.* **33**, 207–215 (2017).
42. Kyriakou, I. et al. Influence of track structure and condensed history physics models of Geant4 to nanoscale electron transport in liquid water. *Phys. Med.* **58**, 149–154 (2019).
43. Hou, Z. Y. et al. High-complexity of DNA double-strand breaks is key for alternative end-joining choice [data set]. *Zenodo* https://doi.org/10.5281/zenodo.12622604 (2024).

Acknowledgements

We thank Prof. Changhao Bi for providing the strains of *E. coli* and related plasmids, Dr. Li Liu for her assistance in the plasmid constru rradiation. This work was supported by the National Natural Science Foundation of China (12135016 to P.B., and 12075275 to P.B.) and Innovative Center of Radiation Application (KFZC2021010401 to T.W.).

Author contributions

P.B. and T.W. conceived and supervised the project. P.B., L.Z., L.W., Y.W., and Y.Z. designed the experiments; Z.H., T.W., T.Y., and Y.D. performed the experiments; T.W., Z.H., Q.Y, and P.B. analyzed the data; The paper was written by P.B. and Z.H. with contributions and final approval of all co-authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-06640-5.

Correspondence and requests for materials should be addressed to Ting Wang or Po Bian.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Valeria Naim and Christina Karlsson Rosenthal. A peer review file is available.