HHS Public Access

Author manuscript

DNA Repair (Amst). Author manuscript; available in PMC 2022 December 01.

Published in final edited form as:

DNA Repair (Amst). 2021 December ; 108: 103228. doi:10.1016/j.dnarep.2021.103228.

Mutagenic repair of a ZFN-induced double-strand break in yeast: Effects of cleavage site sequence and spacer size

Samantha Shaltz, **Sue Jinks-Robertson***

Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

Abstract

Double-strand breaks are repaired by error-free homologous recombination or by relatively error-prone pathways that directly join broken ends. Both types of repair have been extensively studied in *Saccharomyces cerevisiae* using enzymes HO or I-*Sce*I, which create breaks with 4-nt 3′ overhangs. In the current study, a galactose-regulated zinc-finger nuclease (ZFN) designed to cleave the Drosophila *rosy* locus was used to generate breaks with 4-nt 5′ overhangs at out-of-frame cleavage sites inserted into the yeast *LYS2* gene. Mutagenic repair was examined following selection of prototrophs on lysine-deficient medium containing galactose or surviving colonies on galactose-containing rich medium. Following cleavage of the original *rosy* spacer (ACGAAT), most Lys^+ colonies contained 1- or 4-bp insertions at the cleavage site while most survivors had either a 2-bp insertion or a large deletion. Small insertions reflected nonhomologous end joining (NHEJ) and large deletions were the product of microhomology-mediated end joining (MMEJ). Changing the original ACGAAT spacer to either AGCAAT, ACGCGT or CTATTA altered the molecular features of NHEJ events as well as their frequency relative to MMEJ. Altering the optimal 6-bp spacer size between the zinc-finger protein binding sites to 5 bp or 7 bp eliminated the effect of continuous ZFN expression on survival, but Lys^+ prototrophs were still generated. Analysis of Lys^+ revertants after cleavage of the 5-bp spacer indicated that both the position and spacing of ZFN-generated nicks were variable. Results provide insight into effects of overhang sequence on mutagenic outcomes and demonstrate ZFN cleavage of 5- or 7-bp spacers *in vivo*.

Keywords

Nonhomologous end joining; Microhomology-mediated end joining; Zinc-finger nuclease; Double-strand break; Mutagenesis; Yeast

*Correspondence to: Department of Molecular Genetics and Microbiology, Duke University Medical Center, 213 Research Dr., Box 3020 DUMC, Durham, NC 27710, USA. sue.robertson@duke.edu (S. Jinks-Robertson).

CRediT authorship contribution statement

Samantha Shaltz: Investigation, Methodology, Data curation, Writing – review & editing. **Sue Jinks-Robertson:** Conceptualization, Supervision, Methodology, Writing – original draft.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dnarep.2021.103228.

1. Introduction

DNA double-strand breaks (DSBs) are potentially toxic lesions and pathways for their repair are highly conserved in eukaryotes. Homologous recombination uses an intact duplex as a template to restore the broken region while end-joining pathways ligate broken ends back together. Direct end joining is a relatively error-prone process, with information often added or deleted from the junction. In *Saccharomyces cerevisiae* two distinct end-joining mechanisms have been defined: nonhomologous end joining (NHEJ), which does not require complementarity between the ends, and microhomology-mediated end joining (MMEJ), which requires 6–14 bp of microhomology (reviewed in [1] and [2]). Although end complementarity is not required for NHEJ, it improves both the accuracy and efficiency of repair [3]. An additional distinction between NHEJ and MMEJ is that only NHEJ requires the Ku complex, which protects broken ends from resection. Because *S. cerevisiae* relies mainly on homologous recombination for the repair of genomic DSBs, NHEJ in a chromosomal context is studied either in the absence of a repair template or in recombination-defective strains.

Most NHEJ studies in yeast have used HO or I-*Sce*I, which generates breaks with 4-nt 3′ overhangs; Zinc-Finger Nucleases (ZFNs) or Cas9 have been used to only a limited extent [4,5]. ZFNs create 4-nt 5′ overhangs that have been shown to recruit NHEJ proteins more robustly than 3′ overhangs and to be more often mutationally modified during NHEJ [5]. The default following the creation of a targeted DSB is for NHEJ to simply rejoin the ends, resulting in cycles of cleavage-ligation until a rare error-prone event renders the target sequence refractory to the enzyme. Overhang sequence dictates the spectrum of insertions/-deletions and, in contrast to breaks initiated with I-*Sce*I or HO, the sequence can be varied when using a ZFN [5]. In the current study, a galactose-inducible ZFN designed to cleave the Drosophila *rosy* locus [6] was used to examine error-prone repair as well as effects of altering the sequence and size of the spacer where the enzyme cleaves.

2. Materials and methods

2.1. Media and growth conditions

All growth was at 30 °C. Prior to ZFN induction, cells were grown non-selectively in YEP (1% yeast extract, 2% Bacto-peptone, 300 mg/ liter adenine) supplemented with 2% raffinose (YEPR). Continuous ZFN expression was induced by plating cells onto YEPGal (2% galactose) or onto selective SGal-lys synthetic medium (1.7 g/liter yeast nitrogen base, 0.5% ammonium sulfate, 2% agar, 2% galactose; all amino acids and bases except lysine). Total cell number at the time of ZFN induction was determined by plating on YEPD (2% dextrose) medium.

2.2. Strain constructions

The W303 background (*leu2–3,112 his3–11,15 trp1–1 ura3 ade2–1 CAN1 RAD5*) was used and ZFN cleavage sites (Fig. S1) were introduced into the *LYS2* locus using *delitto perfetto* [7]. A galactose-inducible ZFN (RyA and RyB heterodimer) designed to cleave the Drosophila *rosy* locus [6] was used to cut *lys2::ZFN* alleles. Genes encoding RyA and RyB

were provided in replicating plasmids pYcPlac33;nls-RyA and pYcPlac111; nls-RyB as fusions to the *GAL1* promoter (gift of Dana Carroll, University of Utah) and were subcloned into non-replicating *HIS3* vector pRS303 or *LEU2* vector pRS305 [8], respectively, for subsequent genome integration. To stabilize integrated plasmids, the duplicated portion of the *HIS3* or *LEU2* gene was replaced with a *loxP-URA3Kl-loxP* fragment amplified from plasmid pUG72 [9]. A Cre-containing plasmid was then introduced to generate a deletion marked with a single *loxP* site. A small region of sequence duplication remained flanking the *pGAL-RyA* gene at *HIS3* in order to maintain normal expression of the essential *DED1* gene, which resulted in a low background of surviving colonies that retained the ZFN cleavage site. A *ku70*Δ derivative of the strain with the original *rosy* cleavage site [*lys2::ZFN(−1,ACGAAT)* allele] was constructed by one-step allele replacement using a *loxP-TRP1-loxP* marker amplified from pSR954 [10].

2.3. Mutation frequencies and spectra

Independent YEPR cultures were grown to an optical density of 0.3–0.6 and plated on SGal-lys and YEPGal to select Lys^+ revertants and survivors, respectively. Frequencies were calculated by dividing the numbers of revertant or survivors by the number of colonies obtained following plating on YEPD medium. To reflect only survivors that had lost the ZFN cleavage site, the measured mean survival frequency was multiplied by the fraction of sequenced colonies in the corresponding spectrum that had lost the cleavage site. The 95% error bars were calculated by combining the 95% confidence interval for the measured survival frequency with that of the corresponding proportion of survivors that had lost the ZFN cleavage site [11]. Mean revertant and survivor frequencies are in Table S1. Mutation spectra were constructed by sequencing PCR-generated fragments spanning the ZFN recognition site (Eton Bioscience, Inc. or GeneWiz). Complete spectra for revertants and survivors are provided in Tables S2 and S3, respectively. Forward and reverse primers 1 and 2 (5′-GCCTCATGATAGTTTTTCTAACAAATACG and 5′-CCCATCACACATACCATCAAATCCAC, respectively) were used to amplify the reversion window for sequencing. Reverse primer 3 (5′-TGATCATCAGCCCTACCG) was used in conjunction with primer 1 to characterize the recurrent 1175 bp deletion. Forward and reverse primers 4 and 5 (5′-TTAATGTGAGACTTCTCTTTACCCAT and 5′-CATTATTGTTTGTTCACCGCCC, respectively) were used to amplify the junction of the 11,731 kb deletion.

2.4. ZFN induction in quiescent (Q) cells and detection of 5-bp duplications

Cells were grown in 5 ml of phosphate-limiting L medium (0.07% phosphate-free yeast nitrogen base, 10 μg/ml KH_2PO_4, 0.1% KCl, 0.5% NH_4SO_4, 2% dextrose, 1.4 g/l complete amino acid mix) for one week and then transferred to phosphate-free Q medium (0.07% phosphate-free yeast nitrogen base, 0.1% KCl, 0.5% NH_4SO_4, 1.4 g/l complete amino acid mix) supplemented with either 2% galactose or 2% glucose. After an additional three days of incubation, appropriate dilutions were plated to determine the Lys^+ frequency and genomic DNA was extracted using the MasterPure Yeast DNA Purification Kit (Lucigen). qPCR was performed (PowerTrack SYBR Green Master Mix; Applied Biosystems) using a forward primer specific for the ACGAT 5-bp duplication (5′-TGCCAAGCTACTAC*ACGAT*AC; single-

copy ACGAT italicized) and reverse primer 5′-GTCCGCAACAATGGTTACTC. *ALG9*-specific primers (forward 5′-CACGGATAGTGGCTTTGGTGAACAATTAC and reverse 5′-TATGATTATCTGGCAGCAGGAAAGAACTTGGG) were used as an internal standard. Calculations were performed as described by Schmittgen and Livak [12].

3. Results and discussion

Cleavage of DNA by a ZFN requires sequence-specific binding of two Zn-finger proteins, each of which is fused to the nuclease domain of restriction enzyme *Fok*I (reviewed in [13]). Simultaneous target-site binding promotes *Fok*I dimerization, which is required for cleavage, and cleavage is most efficient when there is a 6-bp spacer between the Zn-finger binding sites of the fusion proteins [14]. Each of the *rosy*--specific ZFN subunits (RyA and RyB) contained three zinc fingers and recognized a 9-bp target sequence flanking a 6-bp spacer of sequence 5′-ACGAAT (Fig. 1; [6]). To efficiently detect the complete fill-in of each 5′ overhang, which creates a 4-bp duplication, the 24-bp *rosy* target was inserted out-of-frame into a region of the *LYS2* gene that tolerates heterologous amino acids [15]. The resulting −1 frameshift allele is referred to as *lys2::ZFN(−1,ACGAAT)*; additional alleles with different spacers and/or reading frames are referred to using a similar nomenclature. Net +1 mutations that restored the correct reading frame were selected on lysine-deficient medium containing galactose while unbiased error-prone repair events were identified by selecting survivors on rich medium containing galactose.

3.1. ZFN-induced lys2::ZFN(−1,ACGAAT) revertants

The spontaneous reversion frequency of *lys2* frameshift alleles is ~5×10^{-9} [15,16], and this frequency increased five orders of magnitude (to 4.7×10^{-4}) when *lys2::ZFN(−1,ACGAAT)* cells were selectively plated (Table S1). Most compensatory frameshifts were small additions to the 6-bp spacer sequence and these events were of three major types (Fig. 1; see Table S2 for the complete spectrum). Among 226 revertants sequenced, 31% contained a duplication of the region flanked by the ZFN-generated nicks (+CGAA); 26% contained an additional A between the nicks (2A > 3A); and 27% contained a T insertion adjacent to the promoter-distal nick (1T > 2T). Although less frequent, there was a similar A insertion upstream of the promoter-proximal nick (1A > 2A). The frequency of galactose-induced *lys2::ZFN(−1,ACGAAT)* revertants was reduced to the background level in a *yku70*Δ background (Table S1), confirming the NHEJ dependence of these events.

The only way to generate a +CGAA event is by complete fill in of ends (Fig. 1), but there may be more than one mechanism for the other main mutation types. The frequent 2A > 3A mutation can arise, for example, by mispairing between the terminal T on the upstream overhang with the incorrect A on the downstream overhang. Removal of the 5′-CG tail/flap and filling of the flanking gaps would then generate the 2A > 3A mutation (Fig. 1). Alternatively, a 2A;> 3A change could reflect blunt-end ligation of a completely filled in left end to a right end that is processed to a single nt that is then filled in. We suggest that the frequent 1T > 2T event that is adjacent to rather than within the distal overhang likely arises by a misincorporation-realignment mechanism. If an A is inserted opposite the first template A on the distal overhang, subsequent realignment to pair with the complementary

T re-creates the 4-nt overhangs (Fig. 1). Ligation of the overhangs, followed by replication, produces a T insertion on one of the daughter duplexes. The less frequent 1A> 2A event (5′-**AA**$^{\vee}$CGAA$_{\wedge}$T) that is upstream of the ZFN-generated nick can be produced by a similar mechanism.

3.2. Survivors of continuous ZFN expression

The mean survival frequency was 1.8×10^{-3} when *lys2::ZFN(−1, ACGAAT)* cells were plated on rich, galactose-containing medium, indicating very efficient cleavage by the ZFN. Only 70% (108/155) of the surviving colonies produced a standard PCR product that spanned the cleavage site and of these, only 36 contained a mutation. Among the 36 survivors with a mutated ZFN site, 72% (26/36) had an AC insertion (+AC), which expanded an existing AC duplication, the second copy of which is nicked during ZFN cleavage. Fig. 2A presents a model for the +AC event in which a misincorporation-triggered realignment of complementary strands occurs during filling of the promoter-proximal 5′ overhang.

For those survivors that initially failed to produce a PCR product, we hypothesized that one or both of the primer-binding sites had been eliminated during repair. Using primers further away from the break site, we found that ~75% of the PCR failures reflected an 1175-bp deletion within the *LYS2* locus; the remaining survivors had an 11,731 bp deletion that eliminated *LYS2* together with the flanking *TKL2* and *RAD16* genes (Fig. 2B). The 1.2 and 11.7 kb deletions had endpoints in 13-bp and 14-bp direct repeats, respectively, a size suggestive of MMEJ rather than single-strand annealing [17]. Whereas the 11.7 kb deletion required extensive processing of both ends, the 1.2 kb deletion did not; one endpoint of the 1.2 kb deletion was within the introduced *rosy* sequence. As expected, the large deletions were Ku-independent (Table S1). Loss of Rad52 reduced the frequency of survivors 7-fold and there was a proportional reduction of large deletions from 47/83 to 19/58 ($p = 0.01$) among survivors. The complete spectrum of survivors is presented in Table S3.

The mutagenic end-joining events among yeast survivors were very different than those following germline cleavage in Drosophila [18]. In flies there were equivalent numbers of insertions and deletions, their sizes were highly variable, and the frequent yeast +AC event was not detected. The greater diversity/complexity of events in flies most likely reflects the absence of a yeast NHEJ-specific Pol4 equivalent as well as presence of metazoan-specific Pol theta, which supports a highly mutagenic and Ku-independent end-joining pathway [19].

It was striking that none of the net +1 alterations observed in the reversion assay were detected among surviving colonies, indicating that these events are rare in the absence of direct selection. To directly compare the frequencies of net +1 and net −1 events following ZFN cleavage, we constructed the *lys2::ZFN(+1,ACGAAT)* allele. Its reversion frequency was 7.7-fold higher than that of the original *lys2::ZFN (−1,ACGAAT)* allele (Fig. 3) and as expected, the most frequent mutation was the +AC event (86/144). Novel 1T > 3T (ACGAAT > ACTAATTT) and -A (and ACGAAT > ACGAT) mutations were also frequent (29/144 and 26/144, respectively) among Lys^{+} revertants. The absence of the 1T > 3T and rarity of -A events relative to +AC events among either *lys2::ZFN(−1,ACGAAT)* or *lys2::ZFN(+1, ACGAAT)* survivors (Table S3) suggests there are distinct differences in

repair outcomes when prototrophic selection is applied. Finally, we note that the frequency of survivors following cleavage of the *lys2::ZFN(+1, ACGAAT)* was also elevated relative to the original *lys2::ZFN(−1, ACGAAT)* allele (Fig. 4). The base immediately distal to the RyA binding site was altered during construction of the +1 allele, and it is possible that this affected protein binding.

3.3. Alteration of the 6-bp spacer sequence cleaved by the ZFN

To test the misincorporation-realignment model for the frequent +AC event observed with the original *rosy* sequence (Fig. 2A), we eliminated the ACAC repeat by changing the spacer sequence from 5′-AC*A**CG**AAT* to 5′-AC*A**GC**AAT* (spacer italicized and changes in bold). This alteration had no effect on the frequency of cleavage-site loss among survivors (Fig. 4) but as predicted, it completely eliminated +AC events (0/51; these were replaced by MMEJ). The reversion frequency of the alternative allele also was similar to that of the original allele, but there were proportionally more +CGAA events following cleavage of the original *lys2::ZFN(−1,A**CG**AAT)* allele than +GCAA events following cleavage of the alternative *lys2::ZFN(−1,A**GC**AAT)* allele (Fig. 3). Given that the only difference in the spacers was reversing the order of the C and G, the reason for this is unclear.

In addition to subtly modifying the sequence of the 5′ overhangs created by ZFN cleavage, we also changed the ends so that each contained a CG dinucleotide repeat. The reversion frequency of the resulting *lys2::ZFN(−1,ACGCGT)* allele was similar to that obtained with the ACGAAT or AGCAAT spacer. Of 45 revertants sequenced, 12 were +CGCG events and 24 were 2-bp GC deletions (-GC) that presumably reflect repeat-mediated mispairing of the 5′ overhangs (Fig. 5A). In contrast to the similar Lys^+ frequency obtained with the various spacer sequences (Fig. 3), there were approximately 10-fold more survivors following cleavage of the *lys2::ZFN(−1,ACGCGT)* allele (Fig. 4). Almost 90% (40/45) of these had insertion of a CG dinucleotide (+CG), which reflects annealing of the terminal repeat on each end, followed by gap filling and ligation (Fig. 5A).

The large excess of +GC relative to −GC events (40:2) among survivors confirms prior conclusions that end misalignment and gap filling during NHEJ is much more frequent than misalignment that additionally requires flap processing [20]. The absence of +4 events among survivors sequenced also indicates that the annealing of ends is much more frequent than the complete fill in. Finally, there was only a single putative MMEJ event detected among the *lys2::ZFN(−1,ACGCGT)* survivors analyzed, which reflects the much more efficient NHEJ observed with this allele. Overhang sequence thus not only determines error-prone changes at the break site, but also the perceived contribution of MMEJ to error-prone DSB repair (Fig. 4).

3.4. Comparison of mutagenic events at ZFN- and Top2-generated overhangs

We previously reported that mutational or chemical stabilization of the covalent yeast Top2 cleavage intermediate is associated with *de novo* duplications in a *CAN1* forward mutation assay. The most common duplication size was 4 bp, which reflects NHEJ-mediated fill-in and ligation of the 4-nt 5′ overhangs created by Top2 [21]. We noted a “TATT” hotspot for TAT as well as TATT duplications, but other types of frameshifts were rare. The 3-bp

duplication can originate through mispairing of the terminal nucleotide of each overhang following TATT cleavage (Fig. 5B). Whereas NHEJ following Top2 cleavage requires the removal of covalently linked protein from each end, the ends generated by ZFN are clean. To compare NHEJ events associated with Top2 *versus* ZFN cleavage, we changed the ZFN spacer sequence to match that of the Top2 hotspot [*lys2::ZFN(−1,CTATTA)* allele]. Among 95 Lys^+ revertants sequenced, there were 39 TATT duplications, 19 2T > 3T events, 15 2-bp deletions (TATT > TT) and 22 2C > 3C events (see Fig. 5B for possible mechanisms). The C at the 5′ border of the ZFN spacer follows a C and creates a CC dinucleotide. The CC is absent at the TATT hotspot in the *CAN1* gene, which precludes similar 2C > 3C events. Among 64 *lys2::ZFN(−1,CTATTA)* survivors that had altered the ZFN cleavage site, half were +TATT or +TAT and the remainder were a variety of different types of frameshifts. Although the data are suggestive of less flexibility in the processing/joining of the protein-linked ends following Top2 cleavage, the number of Top2-generated frameshifts detected was relatively small and the growth conditions were different from those used here. Finally, there was a much greater abundance of the 11.7 kb than the 1.2 kb MMEJ event (22 and 3, respectively) among survivors following ZFN cleavage of CTATTA spacer, while the reverse bias was observed with the other spacers examined (43 11.7 kb and 101 1.2 kb deletions total; $p < 0.0001$). The difference most likely reflects a reduction in the endpoint microhomology for the 1.2 kb deletion from 13 bp to 12 bp when the spacer sequence was changed to CTATTA. Previous work demonstrated a 10-fold increase in microhomology-mediated repair efficiency associated with each additional base pair for repeats in the 12–17 bp range [17].

3.5. Changing the size of the ZFN cleavage-site spacer

The RyA and RyB Zn-finger proteins used here contain no linker between the zinc fingers that bind DNA and the *Fok*I cleavage domain. Similar proteins efficiently cut spacers of 5–7 bp *in vitro*, but only appeared to cleave a 6-bp spacer in Xenopus oocytes [14]. To examine whether the same is true in yeast, we constructed three additional *lys2* alleles: *lys2::ZFN(−1,ACGAT)* and *lys2::ZFN(+1,ACGAT)*, which contain a 5-bp spacer lacking the AA dinucleotide of the original ACGAAT spacer, and *lys2::ZFN(−1,ACGAAAT)*, which contains a 7-bp spacer with the 2A expanded to 3A. In contrast to the 1000-fold decrease in plating efficiency observed when strains containing the 6-bp spacers were plated non-selectively, continuous ZFN expression in the 5- or 7-bp spacer strains caused no detectable reduction in plating efficiency, which suggests inefficient ZFN cleavage (Table S1). Consistent with very poor cleavage, none of 188 survivors sequenced had an altered spacer sequence. Despite poor cleavage of the 5-bp spacer inferred from the survival assay, the corresponding Lys^+ revertant frequencies were similar to those observed with the 6-bp spacer strains (Fig. 3). We suggest that the seeming contradiction between high survival yet strong induction of revertants reflects whether or not cells continue to grow after plating on galactose medium. Even very inefficient cleavage will eventually produce a prototroph on selective medium, but rare survivors will be obscured by cells that continue to grow in the absence of selection.

Three-quarters (72/94) of the *lys2::ZFN(−1,ACGAT)* revertants had 4-bp insertions of two distinct types (22 +ACGA and 50 +CGAT events), which can be explained by alternative ZFN cleavage sites displaced from each other by 1 bp (Fig. 6A, top).

Among the remaining *lys2::ZFN(−1,ACGAT)* revertants, most (15/ 22) were various 2-bp deletions. Approximately half (87/189) of *lys2:: ZFN(+1,ACGAT)* revertants had the previously described +AC insertion that expanded the pre-existing AC dinucleotide. Unexpectedly, there was duplication of the 5-bp spacer sequence (+ACGAT) that was almost as frequent (70/189). We suggest two possible mechanisms for generating the 5-bp duplication (Fig. 6A). In the first scenario, sister chromatids are cut after DNA replication, with a 1-bp displacement between the cleavage sites. Ligation of filled ends from different chromatids can generate the observed 5-bp duplication. Alternatively, the ZFN may frequently make nicks that are 5-bp instead of 4-bp apart when the spacer size is reduced to 5 bp. End filling and ligation would then create a +ACGAT event. Although we cannot exclude that cleavage of both sister chromatids contributes to the 5-bp duplication following the plating of growing cells, this event occurred at a comparable level in nondividing, quiescent (Q) cells (Fig. 6B). The addition of galactose to Q cells stimulated the frequency of Lys^+ revertants almost 10,000-fold relative to glucose addition (4.9×10^{-2} and 3.2×10^{-6}, respectively) and a similar increase in the 5-bp duplication was detected using allele-specific qPCR analysis (Fig. 6C–D). Altogether, the data demonstrate that both the distance between and the position of ZFN-generated nicks is variable when a ZFN cleavage site contains a 5-bp spacer.

Cleavage of the 7-bp spacer of the *lys2::ZFN(−1,ACGAAAT)* allele also yielded Lys^+ revertants. That the frequency was 10-fold lower than observed with the 5- or 6-bp spacers indicates very poor cleavage of a 7-bp spacer (Fig. 3). The same general classes of net +1 events seen with the original *lys2::ZFN(−1,ACGAAT)* allele were detected with the 7-bp spacer. There were too few +4 events, however, to infer the most common distance between ZFN nicks in the 7-bp spacer or whether the positions of the nicks could vary, as they did with the 5-bp spacer. Nicks separated by 5 or possibly 6-bp followed by complete fill in of the overhangs would not have been detected as a Lys^+ revertants.

4. Conclusions

Selection of survivors or specific types of frameshifts following cleavage of various 6-bp spacers by a common ZFN suggests a hierarchy of error-prone end-joining outcomes that reflect overhang sequence. A spacer containing a dinucleotide repeat that supports out-of-register annealing between ends and requires no further processing other than gap filling was the most mutagenic. Recurrent NHEJ events explained by either out-of-register annealing that required additional 5′-end processing, or by partial end filling followed by slippage that regenerated 4-nt overhangs were less common. When alternatives were available that could be explained by partial end filling, they generally were more frequent than the complete fill-in reaction. This suggests that DNA synthesis is distributive rather than processive during end filling, as well as repetitive attempts to anneal ends and complete NHEJ. MMEJ was an alternative to NHEJ among survivors of continuous ZFN expression, but was prominent only when NHEJ was relatively inefficient. Finally, spacers of 5- or 7-bp were cleaved by the

ZFN *in vivo*, suggesting a potential source of off-target effects in addition to promiscuous Zn-finger protein binding. With the 5-bp spacer, recurrent NHEJ outcomes suggested that both the positions of and spacing between the nicks were variable; the lack of recurrent events with the 7-bp spacer suggest even greater variation in nick positions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are indebted to Dana Carroll (University of Utah) for providing the yeast-friendly ZFN components; we acknowledge Stephanie Nay for valuable contributions made during early phases of this work and we thank Tom Petes and members of SJR lab for comments on the manuscript. This work was supported by NIH grant R35 GM118077.

References

[1]. Daley JM, Palmbos PL, Wu D, Wilson TE, Nonhomologous end joining in yeast, Annu. Rev. Genet 39 (2005) 431–451. [PubMed: 16285867]

[2]. Sfeir A, Symington LS, Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? Trends Biochem. Sci 40 (2015) 701–714. [PubMed: 26439531]

[3]. Boulton SJ, Jackson SP, Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance, Nucl. Acids Res 24 (1996) 4639–4648. [PubMed: 8972848]

[4]. Lemos BR, Kaplan AC, Bae JE, Ferrazzoli AE, Kuo J, Anand RP, Waterman DP, Haber JE, CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles, Proc. Natl. Acad. Sci. USA 115 (2018), E2040–e2047. E2040-e2047. [PubMed: 29440496]

[5]. Liang Z, Sunder S, Nallasivam S, Wilson TE, Overhang polarity of chromosomal double-strand breaks impacts kinetics and fidelity of yeast non-homologous end joining, Nucleic Acids Res. 44 (2016) 2769–2781. [PubMed: 26773053]

[6]. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D, Efficient gene targeting in Drosophila with zinc-finger nucleases, Genetics 172 (2006) 2391–2403, 2391–1403. [PubMed: 16452139]

[7]. Storici F, Resnick MA, The *delitto perfetto* approach to i*n vivo* site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast, Methods Enzymol. 409 (2006) 329–345. [PubMed: 16793410]

[8]. Sikorski RS, Hieter P, A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*, Genetics 122 (1989) 19–27. [PubMed: 2659436]

[9]. Guldener U, Heinisch J, Koehler GJ, Voss D, Hegemann JH, A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast, Nucleic Acids Res. 30 (2002), 23e-23e.

[10]. Cho JE, Jinks-Robertson S, Deletions associated with stabilization of the Top1 cleavage complex in yeast are products of the nonhomologous end-joining pathway, Proc. Natl. Acad. Sci. USA 116 (2019) 22683–22691. [PubMed: 31636207]

[11]. Moore A, Dominska M, Greenwell P, Aksenova AY, Mirkin S, Petes T, Genetic control of genomic alterations induced in yeast by interstitial telomeric sequences, Genetics 209 (2018) 425–438. [PubMed: 29610215]

[12]. Schmittgen TD, Livak KJ, Analyzing real-time PCR data by the comparative C(T) method, Nat. Protoc 3 (2008) 1101–1108. [PubMed: 18546601]

[13]. Carroll D, Genome engineering with zinc-finger nucleases, Genetics 188 (2011) 773–782. [PubMed: 21828278]

[14]. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S, Stimulation of homologous recombination through targeted cleavage by chimeric nucleases, Mol. Cell. Biol 21 (2001) 289–297. [PubMed: 11113203]

[15]. Greene CN, Jinks-Robertson S, Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins, Mol. Cell. Biol 17 (1997) 2844–2850. [PubMed: 9111356]

[16]. Harfe BD, Jinks-Robertson S, Removal of frameshift intermediates by mismatch repair proteins in *Saccharomyces cerevisiae*, Mol. Cell. Biol 19 (1999) 4766–4773. [PubMed: 10373526]

[17]. Villarreal DD, Lee K, Deem A, Shim EY, Malkova A, Lee SE, Microhomology directs diverse DNA break repair pathways and chromosomal translocations, PLoS Genet. 8 (2012), e1003026. [PubMed: 23144625]

[18]. Bozas A, Beumer KJ, Trautman JK, Carroll D, Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila, Genetics 182 (2009) 641–651. [PubMed: 19380480]

[19]. Hanscom T, McVey M, Regulation of error-prone DNA double-strand break repair and its impact on genome evolution, Cells 9 (2020) 1657.

[20]. Wilson TE, Lieber MR, Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase beta (Pol4)-dependent pathway, J. Biol. Chem 274 (1999) 23599–23609. [PubMed: 10438542]

[21]. Stantial N, Rogojina A, Gilbertson M, Sun Y, Miles H, Shaltz S, Berger J, Nitiss KC, Jinks-Robertson S, Nitiss JL, Trapped topoisomerase II initiates formation of de novo duplications via the nonhomologous end-joining pathway in yeast, Proc. Natl. Acad. Sci. USA 117 (2020) 26876–26884. [PubMed: 33046655]

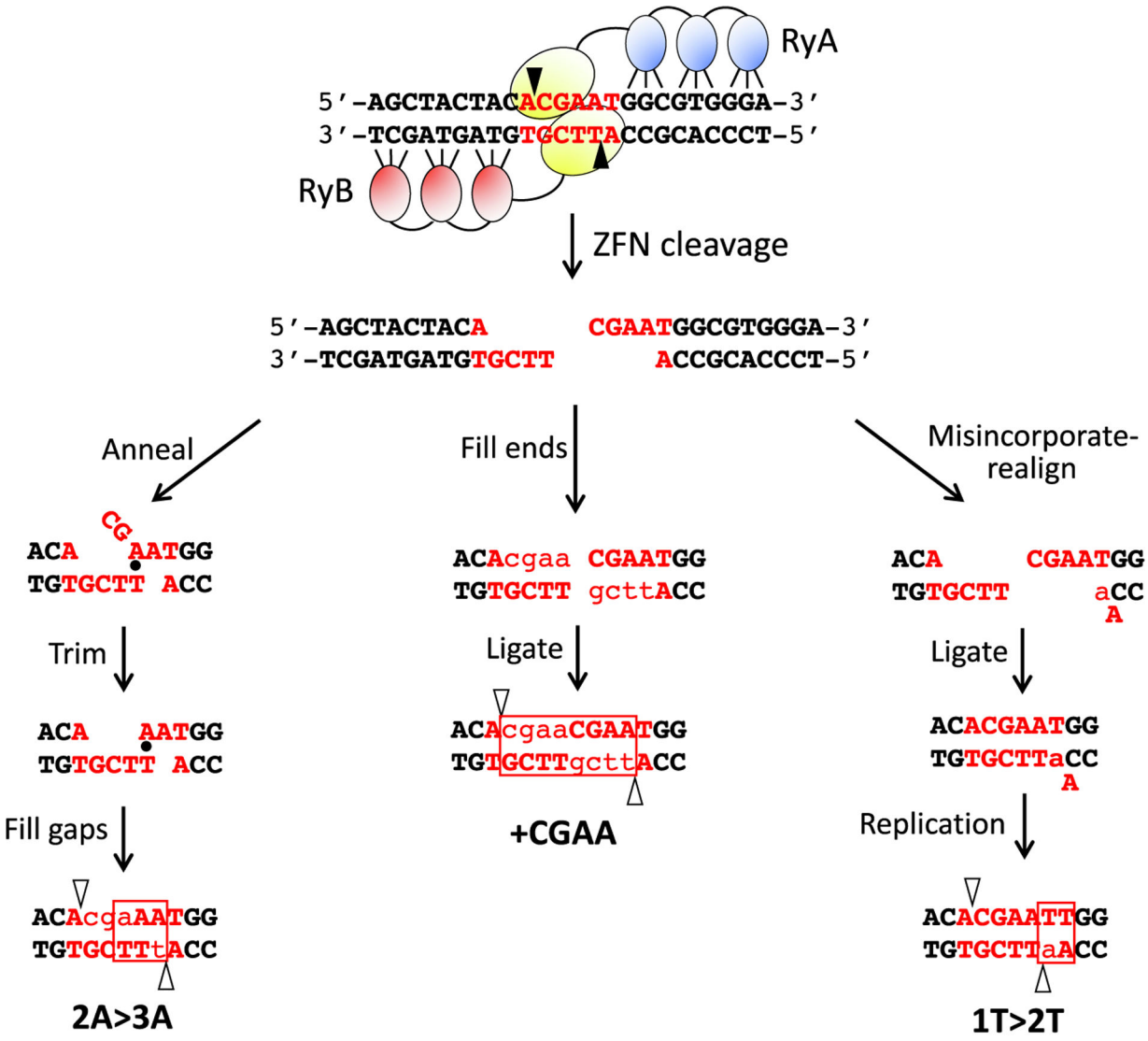

Fig. 1.

ZFN cleavage of the *lys2::ZFN(−1,ACGAAT)* allele and mutational outcomes. The sequence of the 24 bp rosy sequence inserted into *LYS2* is shown. Bases recognized by each zinc finger (blue or red ovals for RyA and RyB, respectively) are indicated; yellow ovals represent the dimerized *Fok*I cleavage domain; the 6-bp spacer is in red font; and filled triangles indicate positions of enzyme-generated nicks. Sequences added during NHEJ are in lowercase red and open triangles correspond to prior nick positions. Complete filling of ends results in duplication of the region bounded by the nicks (+CGAA). Out-of-register pairing between an A and T creates a tail/flap that must be removed and repair generates a 2A > 3A mutation. Duplication of the T downstream of the nicked region (1T > 2T) can be explained by misincorporation-realignment during filling of the downstream overhang, followed by ligation of the re-created 4-nt overhangs.

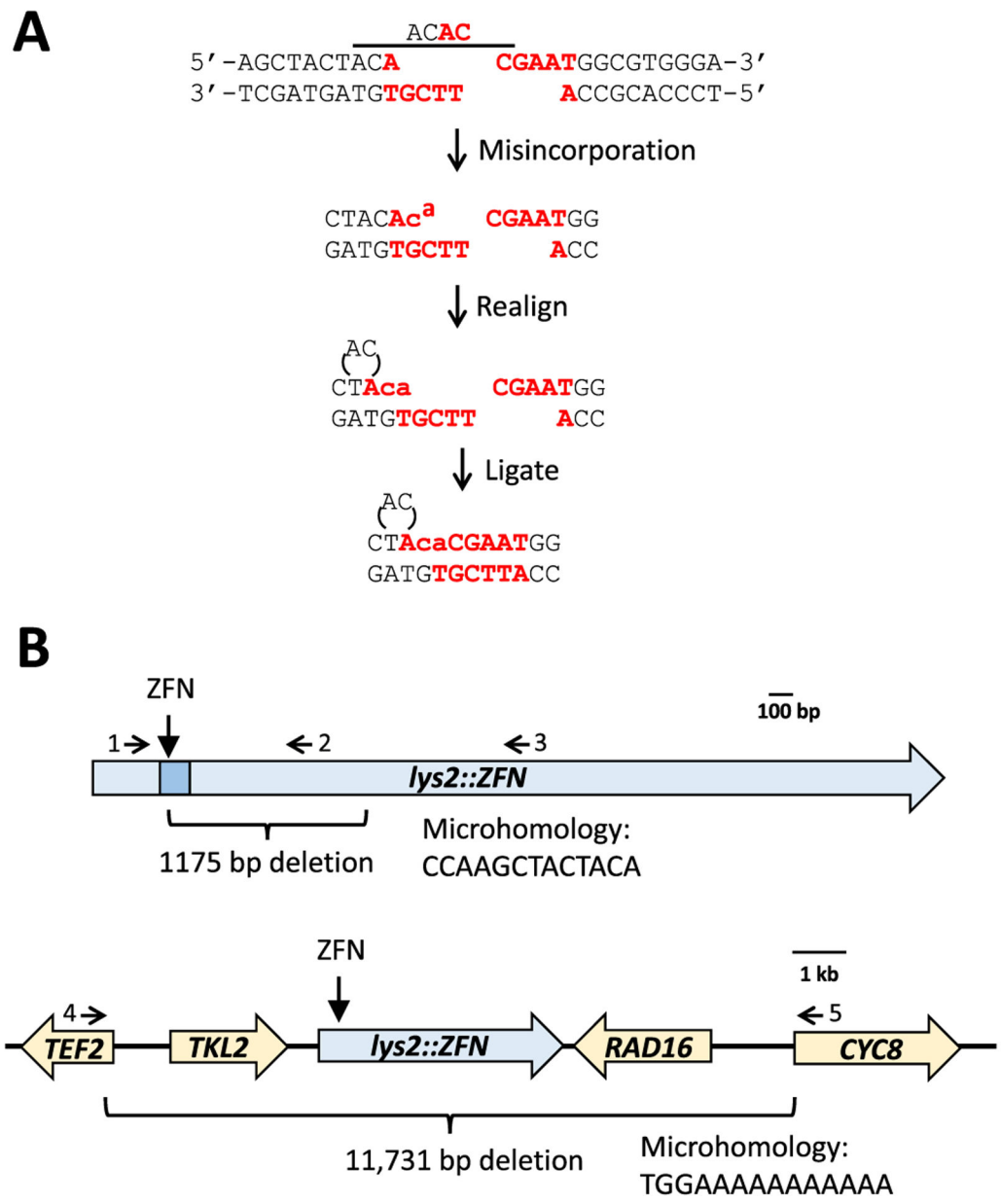

Fig. 2.

Major survivor types following cleavage of the ACGAAT spacer. (A) Model for the recurrent +AC mutation. Extension of the 3′-recessed end begins with the correct insertion of a C (inserted bases are in lowercase red) opposite the template G and this is followed by misincorporation of an A opposite the subsequent template C. The terminal mispair prevents extension and triggers realignment between complementary strands to extrude an extrahelical AC dinucleotide, which restores base pairing at the 3′ end and re-generates the 4-nt upstream overhang. Reannealing between and ligation of the overhangs complete the repair process, with replication of the top strand resulting in the +AC mutation. (B) MMEJ-mediated large deletions. A cartoon of the *LYS2* gene is at the top and the extent of the 1175 bp deletion is indicated; the reversion window near the 5′ end is indicated in dark

blue. The bottom illustrates the 11,731 bp deletion that removes the *LYS2* locus and flanking *RAD16* and *TKL2* genes. Horizontal arrows correspond to the positions of primers used for PCR amplification; endpoint microhomologies of the MMEJ events are given.

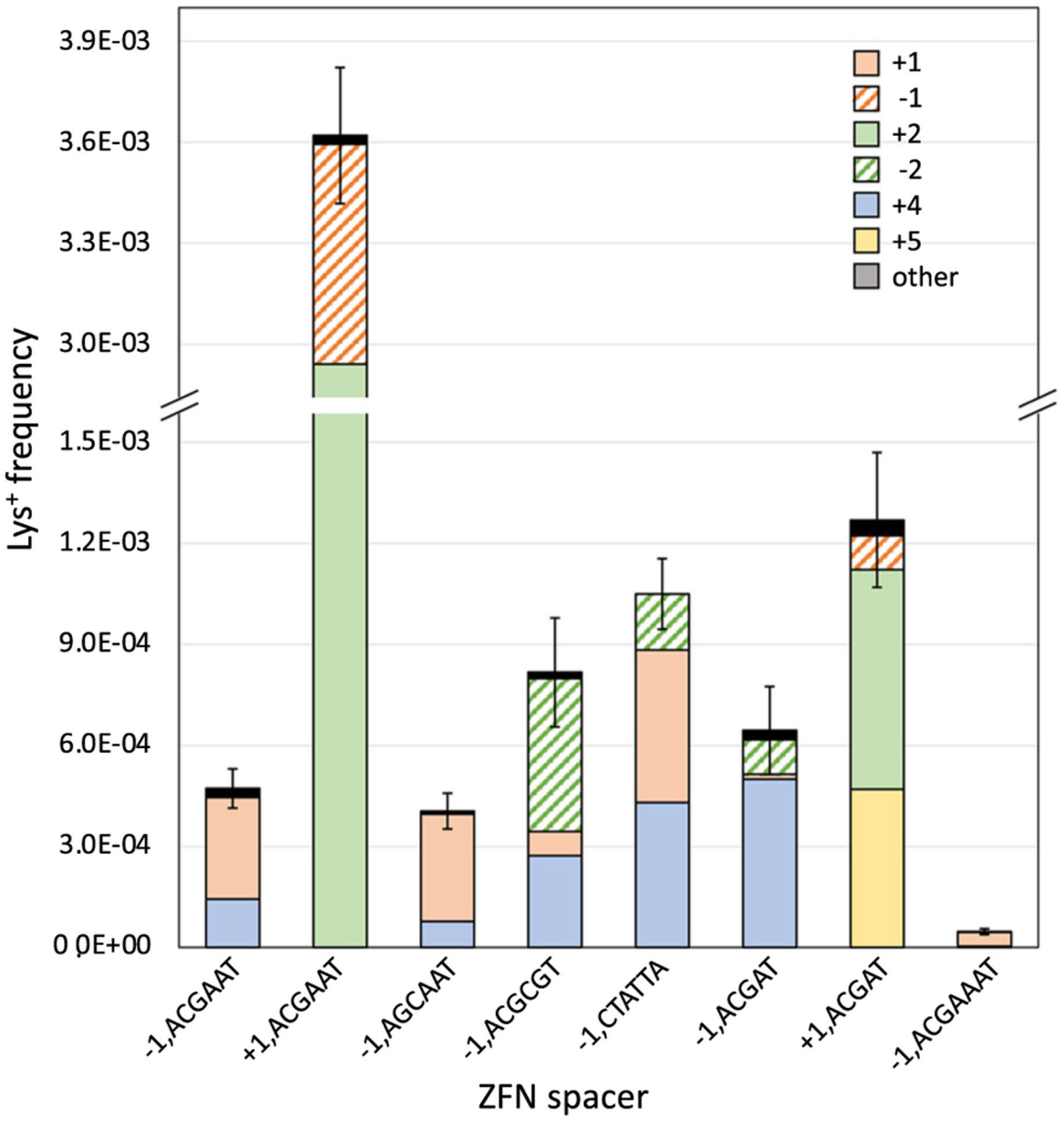

Fig. 3.

Effects of spacer sequence and size on mutational outcomes. Lys^+ frequencies and revertant types for each spacer are shown. Error bars are 95% CIs for the total frequency. Frequencies and complete spectra are in Tables S1 and S2, respectively.

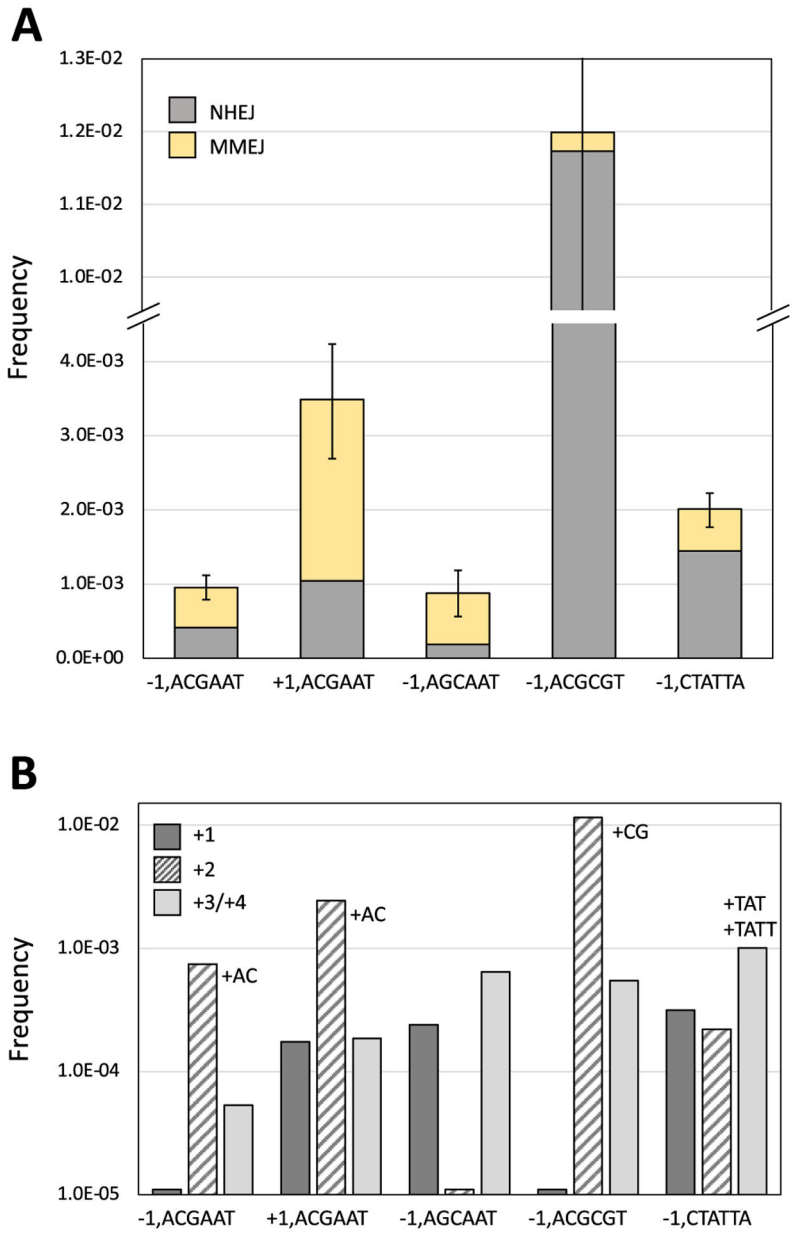

Fig. 4.

Effects of spacer sequence on survivor types. (A) Relative frequencies of NHEJ and MMEJ are shown; error bars are 95% CIs for the total frequency. (B) Frequencies of +1, +2 and +3/+4 events as a function of spacer sequence.

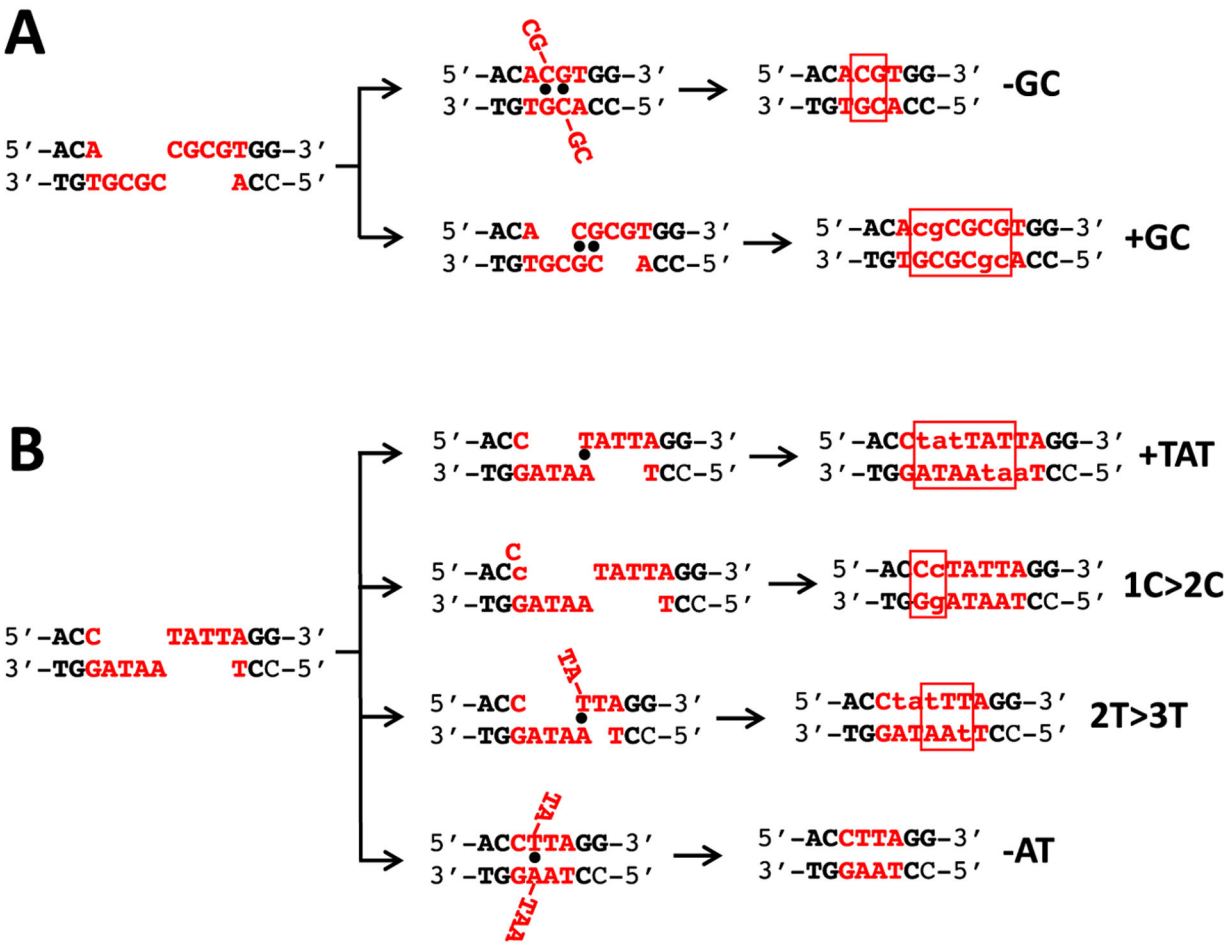

Fig. 5.

Models for common mutation types in modified 6-bp spacers. (A) Cleavage of the −1,ACGCGT spacer (red) creates repetitive overhangs. Misalignment between the overhangs generates -GC (top) or +GC (bottom) events. (B) Cleavage of the −1,CTATTA spacer that matches a putative Top2 hotspot for 4-bp duplications. Possible mechanisms for generating other major mutation types are illustrated.

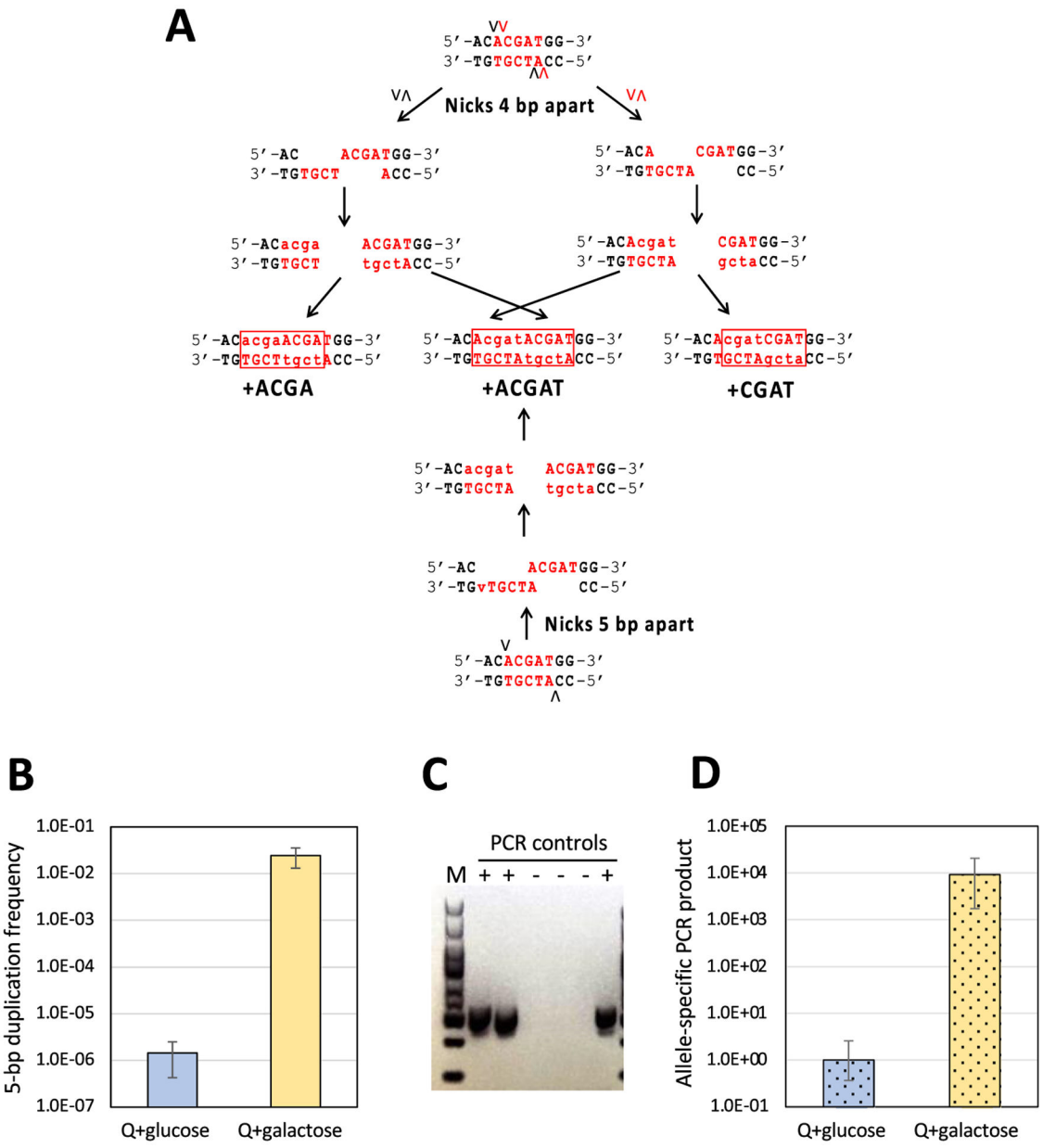

Fig. 6.

Duplications detected following cleavage of the 5-bp ACGAT spacer (red). (A) Alternative cleavage sites with 4-bp spacing are indicated by black and red carets. Complete filling in of the resulting ends (lowercase red) duplicates either ACGA (left) or CGAT (right) following cleavage of the −1,ACGAT allele. Joining of ends derived from the complete fill in of different overhangs can result in the observed +ACGAT event (middle) detected following cleavage of the +1,ACGAT allele. Alternatively, +ACGAT can reflect nicks that are 5-bp apart (bottom; black carets), followed by complete fill-in of the resulting 5-nt overhangs. (B) Mean Lys^+ frequencies following parallel transfer of Q cells to Q medium containing glucose or galactose. Data are from eight cultures and error bars are 95% CIs. (C) Allele-specific PCR detection of the 5-bp duplication. Lanes labeled “+” and “−” contain

template DNA from Lys^+ revertants that have or do not have the 5-bp duplication. (D) The 5-bp duplication-specific PCR product in Q cells incubated in glucose or galactose medium was quantitated by pPCR. Data are the average of three biological replicates and were normalized to the $\Delta\Delta C_T$ value in Q+glucose medium. Error bars are the standard error of the mean.