

## 

# **ABSTRACT**

 A targeted double-strand break introduced into the genome of *Saccharomyces cerevisiae* is repaired by the relatively error-prone nonhomologous-end joining (NHEJ) pathway when homologous recombination is not an option. A ZFN cleavage site was inserted out-of-frame into the *LYS2* locus of a haploid yeast strain to study the genetic control of NHEJ when the ends 28 contain 5' overhangs. Repair events that destroyed the cleavage site were identified either as Lys+ colonies on selective medium or as surviving colonies on rich medium. Junction sequences 30 in Lys<sup>+</sup> events solely reflected NHEJ and were influenced by the nuclease activity of Mre11 as well as by the presence/absence of the NHEJ-specific polymerase Pol4 and the translesion-32 synthesis DNA polymerases Pol  $\zeta$  and Pol n. Although most NHEJ events were dependent on Pol4, a 29-bp deletion with endpoints in 3-bp repeats was an exception. The Pol4-independent deletion required TLS polymerases as well as the exonuclease activity of the replicative Pol  $\delta$  DNA polymerase. Survivors were equally split between NHEJ events and 1 kb or 11 kb deletions that reflected microhomology-mediated end joining (MMEJ). MMEJ events required the processive resection activity of Exo1/Sgs1, but there unexpectedly was no dependence on 38 the Rad1-Rad10 endonuclease for the removal of presumptive 3' tails. Finally, NHEJ was more efficient in non-growing than in growing cells and was most efficient in G0 cells. These studies provide novel insight into the flexibility and complexity of error-prone DSB repair in yeast.

### 

# **INTRODUCTION**

 DNA double-strand breaks (DSBs) are potentially toxic lesions that are repaired either by homologous recombination (HR), which uses an intact duplex as a template to restore the broken region, or by nonhomologous end joining (NHEJ), which directly ligates broken ends back together. In general, NHEJ is a more error-prone process, with out-of-register annealing between ends and/or end processing creating joints with small insertions or deletions. In addition, the rejoining of ends from different breaks can generate various types of genome rearrangements and has been implicated in recurrent oncogenic translocations. Although most DSBs are pathological and reflect either replication fork collapse, abortive topoisomerase reactions or DNA damage, programmed DSBs breaks are essential in some biological processes. During meiosis, for example, the Spo11 protein creates DSBs that initiate the HR necessary for creating genetic diversity and for ensuring proper chromosome segregation (KEENEY 2008). In mitosis, mating-type switching is initiated by the HO endonuclease in yeast (HABER 2012) and RAG proteins create DSBs that initiate the NHEJ-mediated joining of immunoglobulin gene segments in vertebrates (JUNG *et al.* 2006). DSB repair pathways are highly conserved, and the yeast *Saccharomyces cerevisiae*

 has served as a model for defining relevant proteins and molecular mechanisms. Broken ends of all types are bound by the Ku (Ku70-Ku80) and MRX (Mre11-Rad50-Xrs2) complexes; both are absolutely required for NHEJ in yeast (DALEY *et al.* 2005b). Although not required for HR, 60 MRX accelerates the initiation of 5'-end resection, which creates a 3' tail that invades an homologous template and initiates HR (reviewed in SYMINGTON 2016; REGINATO AND CEJKA 62 2020). As part of the MRX complex, Mre11 nicks the 5' terminated strand and subsequent 3'>5' resection towards the break generates a free 3¢ end. MRX activity is particularly important for eliminating end-attached proteins or terminal DNA damage and additionally removes Ku from ends to prevent NHEJ. In addition to its role in short-range resection, MRX facilitates the loading

 of long-range resection activities (Exo1 and Sgs1-Dna2) to promote efficient HR. HR requires a large suite of proteins to invade/copy the repair template as well as to resolve intermediates into final products (reviewed in SYMINGTON *et al.* 2014). In addition to Ku and MRX, NHEJ in yeast requires the dedicated Dnl4 DNA ligase (TEO AND JACKSON 1997) and the Pol4 DNA polymerase (WILSON AND LIEBER 1999).

 Early NHEJ studies in yeast used transformation-based assays to assess the closure efficiency of linearized plasmids with defined end structures and to molecularly define the ligated products (reviewed in DALEY *et al.* 2005b). An advantage of this type of system is that end sequence can be manipulated *in vitro* to yield completely or partially complementary ends, or to generate completely incompatible ends (e.g., blunt ends or ends with different polarity). With ends that have complementarity, the default is simple re-ligation. Following the annealing of ends with partial complementarity, gaps flanking the annealed region must be filled before 78 ligation occurs. With 5' overhangs, the filling of associated gaps is primed from a stably base-79 paired 3' end. With 3' overhangs, however, gap filling must occur from a 3' end stabilized by at most a few base pairs and is more dependent on Pol4 (DALEY *et al.* 2005a). In contrast to 3¢ 81 overhangs, the recessed 3' ends of 5' overhangs can directly be extended in the absence of end annealing. With overhangs that lack complementarity or are incompatible, joining usually involves processing-uncovered microhomologies that flank the broken ends.

 Mitotic studies of DSB repair in a chromosomal context have relied on endonucleases that create a single, targeted DSB and repair is monitored through selection of survivors or prototrophs. HO or I-*Sce*I, which generate breaks with 4-nt 3¢ overhangs had been used extensively to study error-prone DSB repair (e.g. VILLARREAL *et al.* 2012; DENG *et al.* 2014). Zinc 88 Finger Nucleases (ZFNs), which create 4-nt 5' overhangs, and Cas9, which mostly creates blunt ends, have only rarely been used (LIANG *et al.* 2016; LEMOS *et al.* 2018; SHALTZ AND JINKS-ROBERTSON 2021). Because *S. cerevisiae* relies mainly on HR for the repair of genomic DSBs,

 NHEJ following endonuclease cleavage is studied either in the absence of a repair template or 92 by disabling recombination. Precise rejoining of the ends by the NHEJ machinery regenerates the cleavage site, resulting in repetitive cycles of cleavage-ligation until a rare error-prone event renders the target sequence refractory to cleavage. Error-prone NHEJ typically involves minor addition or deletion of sequence from the broken ends and often reflects the annealing of small microhomologies within the overhangs, although microhomology is not a requirement. The overhang sequence dictates the spectrum of insertions/deletions and this sequence cannot be varied when DSBs are initiated with I-*Sce*I or HO. By contrast, use of a ZFN to create a DSB 99 allows manipulation of the 5' overhang sequence (LIANG *et al.* 2016; SHALTZ AND JINKS-100 ROBERTSON 2021). Reflecting the different reactions that can occur at 5<sup>*'*</sup> versus 3<sup>'</sup> overhangs, the kinetics and fidelity of repair also differ (LIANG *et al.* 2016). In addition to the classical NHEJ pathway, yeast has an alternative end-joining pathway that is referred to as microhomology-mediated end joining (MMEJ). MMEJ is characterized by 104 its Ku independence and a requirement for 6-14 bp of microhomology (reviewed in SFEIR AND SYMINGTON 2015). Finally, single-strand annealing (SSA) requires more extensive homology between direct repeats and is usually considered a variant of HR. In contrast to the canonical HR pathway, however, SSA is independent of the Rad51 strand-invasion protein (IVANOV *et al.* 1996), as is MMEJ (LEE AND LEE 2007). The transition from MMEJ to SSA occurs when the microhomology reaches 15-20 bp and SSA, but not MMEJ, has strong dependency on the Rad52 strand-annealing protein (VILLARREAL *et al.* 2012). It should be noted that higher eukaryotes lack a yeast-like MMEJ pathway and instead have an alternative end-joining 112 pathway that is mediated by the Pol theta DNA polymerase (SFEIR AND SYMINGTON 2015). We previously described a system that used a galactose-induced ZFN to create a site- specific DSB in the yeast *LYS2* gene (SHALTZ AND JINKS-ROBERTSON 2021). Insertion of an out-of-frame cleavage site allowed either the selection of NHEJ-mediated repair events that

 restored *LYS2* function or of repair events that simply allowed survival. Approximately half of the latter were Ku-independent MMEJ events (SHALTZ AND JINKS-ROBERTSON 2021). In the current study, this system was used to explore the genetic control of NHEJ- and MMEJ-mediated repair events. 

# **MATERIALS AND METHODS**

## *Media and Growth Conditions*

**All growth of yeast strains was at 30<sup>o</sup>C. For ZFN induction, cultures 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 achieved by plating appropriate dilutions onto non-selective YEPGal, which contained 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). The total number of cells at the time of plating on galactose medium was determined by plating an appropriate dilution on YEPD medium (YEP plus 2% dextrose). Following transformation during strain constructions, selection was on YEPD containing the relevant drug or on synthetic medium missing the appropriate 133 amino acid or base. Ura- derivatives during two-step allele replacement were selected on synthetic medium supplemented with 5-fluoroorotic acid. Phenotypes following tetrad dissections were determined by replica-plating onto appropriate media. All mutant genotypes were confirmed by PCR.

### *Strain constructions*

 All strains used were derived from the W303 background (*leu2-3,112 his3-11,15 trp1-1 ura3 ade2-1 CAN1 RAD5*) either by transformation or by mating and tetrad dissection. Each strain contained a galactose-inducible ZFN designed to cleave the Drosophila *rosy* locus

 (BEUMER *et al.* 2006) and a *lys2* frameshift allele containing a ZFN cleavage site. Genes encoding the ZFN constituent proteins RyA and RyB were integrated into the yeast as previously described (SHALTZ AND JINKS-ROBERTSON 2021) and ZFN cleavage sites were introduced into the *LYS2* locus using the *delitto perfetto* method (STORICI AND RESNICK 2006). Gene-deletion derivatives of SJR4848 were derived by one-step allele replacement using a plasmid-derived PCR cassette containing a selectable marker. The selectable cassettes used were *loxP-hphMX-loxP* from pSR955 (CHO AND JINKS-ROBERTSON 2019), *loxP-natMX-loxP* from pAG25 (GOLDSTEIN AND MCCUSKER 1999), *loxP-URA3Kl-loxP* from pUG72 and *loxP-TRP1-loxP* from pSR954 (CHO AND JINKS-ROBERTSON 2019). The *pol4-D367E* or *rev3-D975A* allele was introduced using *delitto perfetto*; the *mre11-D56N* or *pol3-DV* allele was introduced by two-step allele replacement using plasmid pSM444 (LLORENTE AND SYMINGTON 2004) or pY19 (JIN *et al.* 2001), respectively.

**Mutation Frequencies and Spectra**

 Independent cultures were grown in YEPR to an optical density (OD) reacofhed 0.3-0.6, at which point appropriate dilutions of "growing" cells were plated on YEPD to determine total cell number, on YEPGal to select survivors of continuous ZFN expression or on SGal-lys to 159 select Lys<sup>+</sup> revertants. To obtain a non-growing population of cells, incubation was continued for an addition day. G0 cells were isolated following the incubation of 25 ml YEPR cultures for seven days; small, unbudded G0 cells were obtained from the supernatant following low-speed centrifugation (KOZMIN AND JINKS-ROBERTSON 2013). The 95% confidence interval (CI) of the 163 Lys<sup>+</sup> frequency in each background was calculated using the standard error of the mean. Only "corrected" survival frequencies, which correspond to error-prone repair following continuous ZFN expression, are reported here and each was calculated by multiplying mean survival 166 frequency on YEPGal by the fraction of sequenced colonies in the corresponding spectrum that had lost the ZFN cleavage site. The 95% CI for each corrected survival frequency was

 calculated by combining the 95% confidence interval for the measured frequency with the 95% CI for the proportion of sequenced colonies that had lost the ZFN cleavage site (MOORE *et al.* 2018). Frequencies of revertants and survivors in all backgrounds analyzed are in Table S1. Mutation spectra in revertants and survivors were obtained by sequencing (Sanger method) a PCR-generated fragment spanning the ZFN recognition site. PCR failure was diagnostic of a large deletion that removed one or both of the primer-binding sites. The identity 174 of the deletion was confirmed using a different set of primers, followed by sequencing the product (SHALTZ AND JINKS-ROBERTSON 2021). Complete spectra in revertants and survivors in all genetic backgrounds analyzed are in Table S2 and Table S3, respectively. The distributions of mutation types in different genetic backgrounds were compared using a global Chi-square 178 test (Vassarstats.net). When a significant P value was obtained (p<0.05) comparisons of 179 individual mutation types were then done and the P value for significance was adjusted by dividing 0.05 by the total number of comparisons performed (Bonferroni correction).

# **RESULTS AND DISCUSSION**

 A galactose-regulated, heterodimeric ZFN designed to cleave the Drosophila *rosy* locus (BEUMER *et al.* 2006) was used to introduce a site-specific DSB in the *LYS2* gene (SHALTZ AND JINKS-ROBERTSON 2021). Each ZFN subunit contained three zinc fingers and recognized a 9-bp 187 target sequence flanking a spacer of sequence 5'-ACGAAT (Figure 1A). Insertion of the 24-bp *rosy* target into *LYS2* created a -1 frameshift allele and net +1 mutations that restored the correct reading frame were selected by plating exponentially growing cultures on lysine-deficient medium containing galactose. Selection of surviving colonies on galactose-containing rich medium allowed a relatively unbiased assessment of events that eliminated the ZFN cleavage site. The genetic control of revertants, which reflected canonical NHEJ, and survivors, which reflected both NHEJ and MMEJ events, are described below.

9

194

# 195 Effects of core NHEJ components on ZFN-induced Lys<sup>+</sup> revertants

**In a wild-type (WT) background, the frequency of Lys<sup>+</sup> prototrophs was 2.03**  $\times$ **10<sup>-4</sup>**  (Figure 2A) and there were three major classes of NHEJ events, each of which accounted for approximately 30% of revertants (Figure 2B; SHALTZ AND JINKS-ROBERTSON 2021). The first class (69/226) had a CGAA insertion (+CGAA) resulting from the complete fill in of each of the 200 ZFN-generated 5' overhangs (Figure 1B). The second major class (59/226) contained a 2A>3A expansion that is most easily explained by mis-annealing of the overhangs, followed by end 202 trimming, gap filling and ligation. In the third class of Lys<sup>+</sup> revertants (61/226) there was expansion to two thymines of a single thymine located immediately adjacent to the downstream ZFN-created overhang. We previously suggested that the 1T>2T event reflects misincorporation of an adenine, followed by realignment/slippage that regenerates the 4-nt overhang for direct ligation. Among the remaining events was a recurrent 29-bp deletion with endpoints in a GCC repeat (4/226; Figure 1C) as well as a variety of other minority events (33/226; seeTable S2). 208 We previously reported that the frequency of Lys<sup>+</sup> revertants decreased four orders of 209 magnitude in a *yku70*∆ background (SHALTZ AND JINKS-ROBERTSON 2021) and we observed a 210 similar reduction in *dnl4∆* and *mre11∆* strains (Table S1). Although Mre11 is required for NHEJ in yeast (MA *et al.* 2003), loss of only its nuclease activity stimulates NHEJ-mediated repair of 3¢ 212 overhangs, presumably by slowing initiation of the processive 5' end resection required for HR (LEE AND LEE 2007; DENG *et al.* 2014). Introduction of the nuclease-dead *mre11-D56N* allele 214 (MOREAU *et al.* 1999) into the ZFN system similarly stimulated Lys<sup>+</sup> revertants 8-fold (Figure 2A) but additionally was associated with an unanticipated change in the spectrum of NHEJ events (Figure 2B; p<0.0001). Thus, in addition to a general repressive effect on NHEJ, Mre11 nuclease activity altered molecular outcomes.

10

218 In contrast to the complete absence of Lys<sup>+</sup> prototrophs in  $\frac{v}{k}a$  and *mre11* $\Delta$ 219 backgrounds, there was only a 10-fold reduction in Lys<sup>+</sup> frequency when *POL4* was deleted 220 (Figure 2A), which is consistent with most, but not all, repair of HO-generated 3¢ overhangs 221 requiring Pol4 (LEE AND LEE 2007; TSENG *et al.* 2008). A similar Pol4 dependence of NHEJ was 222 observed using plasmids with defined 3' overhangs but, in contrast to the ZFN ends generated 223 here in a chromosomal context, there was little or no Pol4 dependence with 5<sup>*'*</sup> overhangs 224 (DALEY *et al.* 2005a). Of the three major classes of Lys<sup>+</sup> events observed following ZFN 225 cleavage, only the 2A>3A event (121/183 revertants analyzed) persisted; there were no +CGAA 226 or 1T>2T events detected (Figure 2B). A feature of the 2A>3A event that distinguishes it from 227 the 1T>2T and +CGAA events is that it can occur through the annealing of 5' tails, suggesting 228 that Pol4 may be somewhat less important for gap-filling than for end-filling reactions. Of 229 particular note, the 29-bp deletion event that was rare in WT  $(\sim 2\%; 4/226)$  accounted for 26% 230 (48/183) of ZFN-initiated events in the *pol4* $\triangle$  background. When converted to frequencies, Pol4 231 loss reduced the 2A>3A frequency 5-fold but had no effect on the 29-bp deletion. Finally, the 232 relative importance of Pol4 presence versus its polymerase activity during NHEJ was examined 233 using the *pol4-D367E* allele, which encodes a catalytically inactive protein (WILSON AND LIEBER 234 1999). While the Lys<sup>+</sup> spectrum in the *pol4-D367E* background was identical to that in the *pol4*<sup> $\Delta$ </sup> 235 mutant, the frequency of Lys<sup>+</sup> revertants was 2-fold higher than in the  $\frac{p}{4\Delta}$  strain. This 236 suggests a minor structural role for Pol4 not previously detected (WILSON AND LIEBER 1999). 237 which could reflect the stimulation of Dnl4-mediated ligation by Pol4 reported *in vitro* (TSENG 238 AND TOMKINSON 2002).

239

240 **Effects of additional proteins on NHEJ-mediated repair of a ZFN-initiated break**

 $241$  The persistence of Lys<sup>+</sup> colonies in the  $\textit{pol4}\Delta$  mutant indicates involvement of other DNA 242 polymerases during repair of 5' overhangs. The Pol  $\varepsilon$  replicative polymerase has been

11



 the 2A>3A or 1T>2T events was not significantly altered (p>0.0125). Whereas individual 269 deletion of *REV3* or *RAD30* affected the spectrum but not the frequency of Lys<sup>+</sup> revertants, 270 simultaneous deletion of both was associated with a 2-fold reduction in Lys<sup>+</sup> frequency as well as a change in the NHEJ spectrum (p=0.0027). Instead of the proportional *increase* in +CGAA events observed in the single mutants, there was a specific *decrease* of this specific class in the *rev3∆ rad30∆* double mutant. While the explanation for this is not obvious, especially given the *POL4* dependence of the +CGAA event, it underscores the complexity of interactions that take place during the error-prone repair of broken ends.

276 NHEJ-mediated repair in *rad52*∆ and *sgs1*∆ *exo1*∆ backgrounds was examined while measuring potential effects on MMEJ events in surviving colonies. The Rad52 protein is essential for homologous recombination and previous studies have reported either no (FRANK- VAILLANT AND MARCAND 2002; VILLARREAL *et al.* 2012) or very small changes (DENG *et al.* 2014) in NHEJ frequency in a *rad52*D background. Sgs1 and Exo1 are redundantly required for the 281 processive resection of 5' ends (MIMITOU AND SYMINGTON 2008) and their loss has no effect on NHEJ (VILLARREAL *et al.* 2012; DENG *et al.* 2014). While we similarly detected no significant 283 change in the Lys<sup>+</sup> frequency in either a *rad52*∆ or *sgs1*∆ *exo1*∆ background, the distribution of revertant types was altered in each (Figure 3). Particularly striking was the elevation in "other" 285 NHEJ events: from 16% (37/226) in WT to 44% (41/94) in the *rad52*∆ and to 55% (52/94) in the *sgs1*D *exo1*D strain. In both backgrounds, this change reflected a significant increase in the Pol4-independent 29-bp deletion. A distinguishing feature of this specific NHEJ event is that it presumably requires at least some resection to expose complementary repeats, although it is limited by the processive resection of Sgs1/Exo1. If processive resection occurs normally, the data suggest that the 29-bp deletion may remain an option only if recombination cannot be 291 initiated (*rad52*∆ background).

13

292 Yeast tyrosyl DNA phosphodiesterase (Tdp1) resolves the 5'- and 3'-phosphotyrosyl linkages associated with stabilized topoisomerases (POULIOT *et al.* 1999; NITISS *et al.* 2006) and has a 3¢ nucleosidase activity that generates 3¢-phosphate termini (INTERTHAL *et al.* 2005). Consistent with the latter activity, deletion of *TDP1* was associated with an increase in the complete fill-in of 5¢ overhangs in a plasmid-based NHEJ assay (BAHMED *et al.* 2010). A similar effect was not observed, however, when repair of a ZFN-generated chromosomal break was examined by deep sequencing (LIANG *et al.* 2016). In our system *TDP1* deletion had no effect 299 on the frequency of Lys<sup>+</sup> revertants but did alter the spectrum of events ( $p=0.013$ ). Among the three major classes of events there was a significant change only in the proportion of +CGAA 301 events, (from 69/226 in WT to 53/112 in *tdp1* $\Delta$ ; p=0.0055). This is consistent with a subtle effect 302 on Tdp1 on the ability to fill 5' overhangs.

303

# 304 **Genetic control of the Pol4-independent 29-bp deletion**

305 The model for the 29-bp deletion, which accounted for half of revertants in the *pol4*D 306 background, requires  $5'$ >3' resection to expose GCC repeats that flank the ZFN cleavage site, 307 annealing between complementary strands of GCC repeats, and cleavage of single-strand 3' 308 tails to allow filling of flanking gaps (Figure 1C). We first examined whether TLS polymerases 309 are relevant to the Pol4-independent filling of the small gaps flanking the annealed segment. 310 *REV3*, *RAD30* or both were deleted from the *pol4*D background and there were changes in the 311 frequency and/or spectra of Lys<sup>+</sup> revertants in each relative to the  $\frac{pol4\Delta}{sub}$  single mutant (Figure 312  $\phantom{0}$  4). The Lys<sup>+</sup> frequency increased 2.0- and 3.7-fold, respectively, in the *pol4* $\Delta$  *rad30* $\Delta$  and *pol4* $\Delta$ 313 *rev3*D backgrounds but was not significantly altered in the *pol4*D *rev3*D *rad30*D triple mutant. 314 The proportion of the 29-bp deletion was reduced in the *pol4* $\triangle$  *rev3* $\triangle$  *rev3* $\triangle$  *rad30* $\triangle$ 315 backgrounds but was unaffected in the *pol4*D *rad30*D mutant. The frequencies of the 29-bp 316 deletion were estimated by multiplying Lys<sup>+</sup> frequencies and corresponding proportions of the

 29-bp deletion. The 29-bp deletion frequency was unaffected by deletion of *REV3* or *RAD30* 318 individually in the *pol4*∆ background but was reduced 10-fold in the *pol4∆ rev3∆ rad30*∆ triple 319 mutant. These data demonstrate that Pol  $\zeta$  and Pol n have redundant roles during the Pol4- independent gap filling required to generate the 29-bp deletion. Although sporadic deletions involving other short repeats were observed among revertants in all genetic backgrounds (Table S2), the GCC repeats at the endpoint of the 29-bp deletion were notable because they are close to and symmetrically flank the ZFN cleavage site. Furthermore, the frequency of the 29-bp deletion was elevated when the long-range resection pathways were eliminated (*exo1*D *sgs1*D double mutant; Figure 3). Given the repressive effect of resection on this event and the symmetry/proximity of the GCC repeats to the ZFN cleavage site, we considered the possibility that Mre11 nuclease activity might be required to expose the 328 repeats within short 3' overhangs. Mre11 nicks the 5' strand 30-35 nt from a Ku-bound end and 329 expansion of the nick into a gap by its 3<sup>'</sup>>5' exonuclease activity degrades the 5' end to displace Ku and create a short 3¢ tail (reviewed in REGINATO AND CEJKA 2020). To examine the relevance Mre11 nuclease activity to the NHEJ-dependent 29-bp deletion, the nuclease-dead *mre11-D56N* allele (LLORENTE AND SYMINGTON 2004) was introduced into the *pol4*D 333 background. The Lys<sup>+</sup> frequency increased 12-fold in the double mutant relative to the  $\text{pol}4\Delta$  single mutant and this was accompanied by a change in the distribution of revertant types (Figure 4). Almost all revertants had the 2A>3A event (91/96; p<0.0001) and only a single 29-bp deletion was observed. The *mre11-D56N* allele thus had a strong stimulatory effect on the frequency 2A>3A event, but its effect on the frequency of the 29-bp deletion was unclear. To preclude occurrence of the 2A>3A event and allow more specific focus on the 29-bp deletion in the  $pol4\Delta$  background, we changed the spacer sequence of the ZFN cleavage site

 from ACGAAT to ACGTAT. With the new ACGTAT spacer in the *POL4* background, there were similar numbers of +CGTA and 1T>2T events (43/116 and 63/116, respectively), as there were

 with the original spacer, and the 29-bp deletion remained rare (2/116). Given the absence of +4 343 and 1T>2Tevents with the original ACGAAT spacer in the  $pol4\Delta$  background, we assumed that 344 the 29-bp deletion would be present in almost all Lys<sup>+</sup> colonies derived using the new ACGTAT spacer. In the *pol4*D background with the new spacer, however, half of revertants contained the +CGTA event (54/37) and 30% had the 29-bp deletion (37/118); there still were no 1T>2T events. In the absence of Pol4, the data suggest that the end-annealing that generates the 2A>3A event may precede complete filling and obscure it occurrence.

 Elimination of Mre11 nuclease activity in a *POL4* background with the new spacer 350 increased the frequency of Lys<sup>+</sup> revertants 14-fold and altered the proportions of +CGTA and 1T>2T events (Tables S1 and S2). This mirrors the differential effects of the *mre11-D56N* allele on the NHEJ spectrum observed with the original ACGAAT spacer. Introduction of the *mre11- D56N* allele into the *pol4*∆ background resulted in a 9-fold increase in Lys<sup>+</sup> frequency relative to the *pol4*D single mutant, which was similar to the increase observed with the original spacer. In contrast to the large proportional decrease in the 29-bp deletion in the *pol4*D *mre11-D56N* background with original spacer, however, there was no reduction with the new spacer. This demonstrates that Mre11 nuclease activity is not required to expose the GCC repeats at the endpoints of the 29-bp deletion.

 In the model depicted in Figure 1C end resection allows annealing between complementary strands of the GCC repeat, thereby creating 3¢ tails that must be removed prior to gap filling and ligation. The Rad1-Rad10 nuclease is required for the removal of long 3¢ tails 362 during SSA, but tails <30 nt are efficiently removed by the exonuclease activity of Pol  $\delta$  (PAQUES AND HABER 1997). In the *pol4*D strain containing the new spacer sequence, elimination of the as a exonuclease activity of Pol δ (*pol3-DV* allele; (JIN *et al.* 2001) reduced the Lys<sup>+</sup> frequency 2.5-fold and this reflected the complete absence of the 29-bp deletion among revertants (0/95

366 revertants; p<0.0001; Figure 4). These data demonstrate that the exonuclease activity of Pol  $\delta$ is required during creation of the 29-bp deletion.

 The data presented above present a conundrum with respect to the genetic control of the NHEJ-dependent 29-bp deletion. It was antagonized by processive nuclease activity (*sgs1*D *exo1*∆ background) and yet did not appear to depend on the nuclease activity of Mre11 (*pol4*∆ *mre11-D56N* background). How then are the complementary strands of the GCC direct repeats exposed? One possibility is that the DNA melting activity of the MRX complex is responsible. This requires the ATPase activity of Rad50 (CANNON *et al.* 2013), which also is required for NHEJ in yeast (ZHANG AND PAULL 2005). The strong dependence of the 29-bp deletion on the 375 exonuclease activity of Pol δ (*pol4*Δ *pol3-DV* mutant) suggests an alternative possibility in which 376 it is the degradation of the recessed 3' ends by Pol  $\delta$  that exposes the GCC repeats. Regardless of how the complementary strands of the GCC repeats are exposed, the Ku dependence of the 29-bp deletion suggests either that Ku remains associated with the ends or that Ku re-engages the ends after displacement. Ku interacts with duplexes with 30-nt tails almost as well as with fully duplex DNA *in vitro* (FALZON *et al.* 1993); longer, 60-nt tails are not efficiently bound by Ku (FOSTER *et al.* 2011).

## **Survivors of continuous ZFN expression**

 Selection for colonies on rich medium containing galactose provides an unbiased analysis of end-joining events that confer resistance to continuous ZFN expression. As reported previously, only half of survivors (83/155) in the WT background lost the enzyme cleavage site (SHALTZ AND JINKS-ROBERTSON 2021) and these were of two major types: a Ku-dependent +AC event that expanded two copies an AC dinucleotide spanning the 6-bp ZFN spacer (36/83) and Ku-independent large deletions that removed the cleavage site (47/83). The +AC event was not detectable in the reversion assay, which selects net +1 events, and was hypothesized to arise

 by misincorporation-slippage during proximal end filling (Figure 1B). The rarity of net +1 events 392 in the survivor assay indicates that they are minor events relative to +AC. The large deletions were either 1.2 kb (36/47) or 11.7 kb (11/47) in size (Figure 5A) and because the deletion junctions were in 13- or 14-bp direct repeats, respectively, we concluded that these were MMEJ events. In the current analyses, the frequencies and profiles of survivors were examined in the 396 genetic backgrounds described above for Lys<sup>+</sup> revertants. Only those mutants that were different from the WT strain in terms of the frequency and/or spectrum of site-loss survivors are discussed (see Tables S1 and S3 for all data).

 Because NHEJ and MMEJ proportions are similar among survivors, the complete loss of either pathway would be expected to reduce the survival frequency only 2-fold. Indeed, as 401 reported previously in the *yku7*0 $\Delta$  background (SHALTZ AND JINKS-ROBERTSON 2021), the survival frequency was reduced 2-fold in the NHEJ-deficient *pol4*D (Figure 5B-C) and *dnl4*D (Tables S1 and S3) backgrounds and only large deletions were detected. Although NHEJ events were also absent in an *mre11*D strain, there was 5.9-fold reduction in overall survival frequency rather than the expected 2-fold (Figure 5B). Survivor frequencies were also low in *rad52*D single and *sgs1*D *exo1*D double-mutant backgrounds, with reductions of 6.7- and 16.7- fold, respectively, relative to WT. Impaired growth was a common feature of these three mutant backgrounds, and we suggest that this may be responsible for the large decrease in survivor frequency. We favor this interpretation because, in contrast to the survivor assay, the NHEJ  $f_1$  frequency as measured by Lys<sup>+</sup> reversion was not affected in either the *rad52* $\Delta$  or *sgs1* $\Delta$  *exo1* $\Delta$  background (Figure 3). It is possible that repetitive cleavage by a ZFN impairs viability more when cells can continue to divide (nonselective rich medium) than when cells are plated under conditions where error-prone repair must restore prototrophy before cells can begin dividing. 414 In addition to reduced survival, the spectrum of events among survivors was altered in the *rad52*D and *sgs1*D *exo1*D backgrounds relative to WT (p=0.0013 and p<0.0001,

416 respectively). There was a reduction in the proportion of large deletions in the *rad52* $\triangle$  mutant, (from 47/83 to 19/58; p=0.009), suggestive of a partial requirement of Rad52 for the MMEJ- mediated large deletions. In an earlier study that systematically examined the effect of repeat size on microhomology-mediated end joining (VILLARREAL *et al.* 2012), Rad52 promoted end joining for repeats 15 bp or larger, suggesting a variation of SSA as the mechanism, and strongly inhibited end-joining between repeats 12 bp or smaller. The sizes of the repeats at the MMEJ endpoints in the current study (13-14 bp) are in the transition zone for Rad52 423 dependence. In the  $sgs1\Delta$   $exo1\Delta$  double mutant the proportion of large deletions was reduced 10-fold: from 47/83 in WT to only 2/39. This reduction likely reflects the extensive resection required to expose the junction repeats of large deletions and is consistent with results obtained with similar large deletions (VILLARREAL *et al.* 2012). This is in contrast to the suppressive effect of Sgs1/Exo1 resection (as well as Mre11 presence) with respect to Ku-independent deletions between 12-bp repeats very close to I-*Sce*I-generated ends (DENG *et al.* 2014). The resection- related suppression is similar to that described above for the Pol4-independent 29-bp deletion, although the 29-bp deletion is Ku- and Mre11-dependent (see above). Because the nuclease activity of Mre11 inhibits NHEJ, an elevation in the NHEJ- dependent +AC event was expected in an *mre11-D56N* background. There, however, was neither a change in survivor frequency nor spectrum relative to WT. This result provides further support for an influence of Mre11 nuclease activity on specific NHEJ outcomes, as inferred from 435 the variable effect of the *mre11-D56N* allele on Lys<sup>+</sup> revertant types in the WT and  $\frac{pol4\Delta}{D}$ 436 backgrounds. Finally, there was 2.7-fold increase in survivor frequency in a rev7∆ background that specifically reflected an increase in MMEJ; no effect on frequency or spectra was observed upon loss of *REV3* (Table S1). These data more strongly support a potentially suppressive role of Rev7 on processive resection in yeast, as inferred previously from the slightly increased 440 frequency of Lys<sup>+</sup> revertants.

441 The Rad1-Rad10 endonuclease is required to remove 3' tails during SSA (IVANOV AND HABER 1995) and a similar importance for Rad1 in MMEJ was inferred through analysis of HO- initiated 2-kb deletions between 18-bp direct repeats (VILLARREAL *et al.* 2012). The strong Rad52 dependence for deletions between 18-bp, however, suggests that the events were a variation of SSA. We thus re-examined the requirement for Rad1 for the large, NHEJ- independent deletions in our system. With the smaller (13- and 14-bp) repeats, there was no 447 significant change in either the frequency or the distribution of survivor types in a  $rad1\Delta$ 448 background (Figure 5; see Tables S1-S3 for similar *rad10*∆ data). To potentially identify the 449 relevant structure-specific nuclease that removes 3' tails during MMEJ, we examined survival in *mus81*D single and *rad1*D *mus81*D double mutants, as well as in an *slx4*D background. Mus81 can process 3¢ flaps (SCHWARTZ AND HEYER 2011) while Slx4 is a scaffold for multiple structure- specific endonucleases (CUSSIOL *et al.* 2017). There was no proportional decrease in large deletions in any of these additional mutant backgrounds (Table S3). Either a different nuclease is relevant, or there is functional redundancy between the nucleases examined. Just as there is a size-related transition in Rad52 dependence for microhomology-mediated deletions (VILLARREAL *et al.* 2012), there may be a similar transition in terms of a requirement for Rad1- Rad10 in 3¢-tail removal. One interesting possibility is that Rad52-driven annealing dictates whether subsequent tail removal is dependent on Rad1-Rad10. 

# **The physiological state of cells affects repair of ZFN-induced DSBs**

 NHEJ-mediated ligation of a linearized plasmid is more efficient when non-growing (NG) 462 cells are transformed than when growing cells are transformed (KARATHANASIS AND WILSON 2002). To examine the effects of growth state on error-prone repair of a chromosomal DSB, 464 either non-growing (NG) or isolated G0 cells were plated in the presence of galactose. The Lys<sup>+</sup> frequency in NG cells was 2.4-fold higher than in growing cells and there were proportionally

 fewer +CGAA events among revertants (Figure 6A). When the event types were converted to frequencies, however, the +CGAA frequency was unaltered while the frequencies of 2A>3A, 1T>2T and other NHEJ events were each elevated about 3-fold. There was a further stimulation 469 of the Lys<sup>+</sup> frequency when G0 cells were plated: 6.8- and 2.8-fold relative to growing and NG cells, respectively. Interestingly, there was a very large proportional increase in the +CGAA class of events in G0 cells that corresponded to a 10-fold increase in its frequency relative to growing/NG cells; the frequencies of the other classes changed little relative to NG cells. In terms of surviving colonies that had lost the ZFN cleavage site, the overall frequency also was elevated relative to growing cells: 2.4-fold in NG cells and 5.4-fold in G0 cells (Figure 6B). The spectrum of survivors was also altered. As expected of events that require extensive resection, the proportion of large deletions was greatly reduced: from 0.56 (47/83) in growing cells to 0.16 (15/92) and 0.05 (4/79) when NG and G0 cells, respectively, were plated. Given the large increase in survival, however, the large-deletion frequency was reduced only about 2-fold in NG and G0 cells. In terms of frequency, the +AC event was elevated 4.7-fold in NG cells and 8.7-fold in G0 cells; the frequency of "other" events was similarly elevated. The data demonstrate that NHEJ-mediated repair of a ZFN-generated DSB is more efficient in NG/G0 cells than in growing cells, and that physiological state additionally affects how the resulting ends are modified during error-prone repair.

### **CONCLUSIONS**

 • Following ZFN cleavage, half of error-prone repair events reflected small insertions/deletions at the cleavage site (NHEJ) and half were large deletions (MMEJ).

• In contrast to the complete dependence of NHEJ on Ku, Mre11 and Dnl4, loss of Pol4

reduced NHEJ only 10-fold. Although most NHEJ events were dependent on Pol4, a

recurrent 29-bp deletion was Pol4-independent. This deletion was suppressed by

- 491 processive 5' end resection and required the 3'>5' exonuclease activity of the replicative
- 492 DNA polymerase  $\delta$ .
- Mre11 nuclease activity suppressed NHEJ and additionally affected the spectrum of events
- but had no effect on MMEJ.
- 495 Pol  $\zeta$  and Pol  $\eta$  altered NHEJ outcomes in both the presence and absence of Pol4.
- 496 Absence of the Rev3 or Rev7 component of Pol $\zeta$  had different effects on DSB repair. Of
- particular note, the suppressive effect of Rev7 on MMEJ is consistent with a modulation of
- resection.
- Long single-strand tails created by resection must be removed to complete MMEJ, but
- neither Rad1, Mus81 nor Sxl4 was required for this step.
- NHEJ was increased and MMEJ was decreased when ZFN cleavage occurred in non-
- dividing cells.
- 

# **ACKNOWLEDGEMENTS**

- This work was supported by National Institutes of Health grant R35GM118077 to SJR.
- 

# **REFERENCES**

- Bahmed, K., K. C. Nitiss and J. L. Nitiss, 2010 Yeast Tdp1 regulates the fidelity of nonhomologous end joining. Proc. Natl. Acad. Sci. USA 107**:** 4057-4062.
- Beumer, K., G. Bhattacharyya, M. Bibikova, J. K. Trautman and D. Carroll, 2006 Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics 172**:** 2391-1403.
- Cannon, B., J. Kuhnlein, S. H. Yang, A. Cheng, D. Schindler *et al.*, 2013 Visualization of local DNA unwinding by Mre11/Rad50/Nbs1 using single-molecule FRET. Proc. Natl. Acad. Sci. USA 110**:** 18868-18873.
- Cho, J. E., and S. Jinks-Robertson, 2019 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**:** 22683-22691.
- Cussiol, J. R., D. Dibitetto, A. Pellicioli and M. B. Smolka, 2017 Slx4 scaffolding in homologous recombination and checkpoint control: lessons from yeast. Chromosoma 126**:** 45-58.
- Daley, J. M., R. L. Laan, A. Suresh and T. E. Wilson, 2005a DNA joint dependence of pol X family polymerase action in nonhomologous end joining. J. Biol. Chem. 280**:** 29030- 29037.

- Daley, J. M., P. L. Palmbos, D. Wu and T. E. Wilson, 2005b Nonhomologous end joining in yeast. Annu. Rev. Genet. 39**:** 431-451.
- Deng, S. K., B. Gibb, M. J. de Almeida, E. C. Greene and L. S. Symington, 2014 RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. Nat. Struct. Mol. Biol. 21**:** 405-412.
- Falzon, M., J. W. Fewell and E. L. Kuff, 1993 EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. J. Biol. Chem. 268**:** 10546-10552.
- Foster, S. S., A. Balestrini and J. H. Petrini, 2011 Functional interplay of the Mre11 nuclease and Ku in the response to replication-associated DNA damage. Mol Cell Biol 31**:** 4379- 4389.
- Frank-Vaillant, M., and S. Marcand, 2002 Transient stability of DNA ends allows nonhomologous end joining to precede homologous recombination. Mol. Cell 10**:** 1189- 1199.
- Goldstein, A. L., and J. H. McCusker, 1999 Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast 15**:** 1541-1553.
- Haber, J. E., 2012 Mating-type genes and *MAT* switching in *Saccharomyces cerevisiae*. Genetics 191**:** 33-64.
- Hatkevich, T., D. E. Miller, C. A. Turcotte, M. C. Miller and J. Sekelsky, 2021 A pathway for error-free non-homologous end joining of resected meiotic double-strand breaks. Nucleic Acids Res. 49**:** 879-890.
- Hirano, Y., and K. Sugimoto, 2006 ATR homolog Mec1 controls association of DNA polymerase ζ-Rev1 complex with regions near a double-strand break. Curr. Biol. 16**:** 586-590.
- Interthal, H., H. J. Chen and J. J. Champoux, 2005 Human Tdp1 cleaves a broad spectrum of substrates, including phosphoamide linkages. J. Biol. Chem. 280**:** 36518-36528.
- Ivanov, E. L., and J. E. Haber, 1995 *RAD1* and *RAD10*, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 15**:** 2245-2251.
- Ivanov, E. L., N. Sugawara, J. Fishman-Lobell and J. E. Haber, 1996 Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. Genetics 142**:** 693-704.
- Jin, Y. H., R. Obert, P. M. Burgers, T. A. Kunkel, M. A. Resnick *et al.*, 2001 The 3'-->5' exonuclease of DNA polymerase δ can substitute for the 5' flap endonuclease Rad27/Fen1 in processing Okazaki fragments and preventing genome instability. Proc. Natl. Acad. Sci. USA 98**:** 5122-5127.
- Johnson, R. E., L. Prakash and S. Prakash, 2012 Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. Proc. Natl. Acad. Sci. USA 109**:** 12455-12460.
- Jung, D., C. Giallourakis, R. Mostoslavsky and F. W. Alt, 2006 Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu. Rev. Immunol. 24**:** 541- 570.
- Karathanasis, E., and T. E. Wilson, 2002 Enhancement of *Saccharomyces cerevisiae* end-565 joining efficiency by cell growth stage but not by impairment of recombination. Genetics 161**:** 1015-1027.
- Keeney, S., 2008 Spo11 and the formation of DNA double-strand breaks in meiosis. Genome Dyn. Stab. 2**:** 81-123.
- Kozmin, S. G., and S. Jinks-Robertson, 2013 The mechanism of nucleotide excision repair-mediated UV-induced mutagenesis in nonproliferating cells. Genetics 193**:** 803-817.
- Lee, K., and S. E. Lee, 2007 *Saccharomyces cerevisiae* Sae2- and Tel1-dependent single- strand DNA formation at DNA break promotes microhomology-mediated end joining. Genetics 176**:** 2003-2014.

- Lemos, B. R., A. C. Kaplan, J. E. Bae, A. E. Ferrazzoli, J. Kuo *et al.*, 2018 CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. Proc. Natl. Acad. Sci. USA 115**:** E2040-e2047.
- Liang, Z., S. Sunder, S. Nallasivam and T. E. Wilson, 2016 Overhang polarity of chromosomal double-strand breaks impacts kinetics and fidelity of yeast non-homologous end joining. Nucleic Acids Res. 44**:** 2769-2781.
- Llorente, B., and L. S. Symington, 2004 The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. Mol. Cell. Biol. 24**:** 9682-9694.
- Ma, J. L., E. M. Kim, J. E. Haber and S. E. Lee, 2003 Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. Mol. Cell. Biol. 23**:** 8820-8828.
- Mimitou, E. P., and L. S. Symington, 2008 Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature 455**:** 770-774.
- Moore, A., M. Dominska, P. Greenwell, A. Y. Aksenova, S. Mirkin *et al.*, 2018 Genetic control of genomic alterations induced in yeast by interstitial telomeric sequences. Genetics 209**:** 425-438.
- Moreau, S., J. R. Ferguson and L. S. Symington, 1999 The nuclease activity of Mre11 is required for meiosis but not for mating type switching, end joining, or telomere maintenance. Mol. Cell. Biol. 19**:** 556-566.
- Nitiss, K. C., M. Malik, X. He, S. W. White and J. L. Nitiss, 2006 Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage. Proc. Natl. Acad. Sci. U.S.A. 103**:** 8953-8958.
- Paques, F., and J. E. Haber, 1997 Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 17**:** 6765-6771.
- Pouliot, J. J., K. C. Yao, C. A. Robertson and H. A. Nash, 1999 Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. Science 286**:** 552-555.
- Reginato, G., and P. Cejka, 2020 The MRE11 complex: A versatile toolkit for the repair of broken DNA. DNA Repair 91-92**:** 102869.
- Schwartz, E. K., and W. D. Heyer, 2011 Processing of joint molecule intermediates by structure- selective endonucleases during homologous recombination in eukaryotes. Chromosoma 120**:** 109-127.
- Sfeir, A., and L. S. Symington, 2015 Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? Trends Biochem Sci 40**:** 701-714.
- Shaltz, S., and S. Jinks-Robertson, 2021 Mutagenic repair of a ZFN-induced double-strand break in yeast: Effects of cleavage site sequence and spacer size. DNA Repair 108**:** 103228.
- Storici, F., and M. A. Resnick, 2006 The *delitto perfetto* approach to i*n vivo* site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol. 409**:** 329-345.
- Symington, L. S., 2016 Mechanism and regulation of DNA end resection in eukaryotes. Crit. Rev. Biochem. Mol. Biol. 51**:** 195-212.
- Symington, L. S., R. Rothstein and M. Lisby, 2014 Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. Genetics 198**:** 795-835.
- Teo, S. H., and S. P. Jackson, 1997 Identification of *Saccharomyces cerevisiae* DNA ligase IV: involvement in DNA double-strand break repair. EMBO J. 16**:** 4788-4795.
- Tseng, H. M., and A. E. Tomkinson, 2002 A physical and functional interaction between yeast Pol4 and Dnl4-Lif1 links DNA synthesis and ligation in nonhomologous end joining. J. Biol. Chem. 277**:** 45630-45637.

- Tseng, S. F., A. Gabriel and S. C. Teng, 2008 Proofreading activity of DNA polymerase Pol2 mediates 3'-end processing during nonhomologous end joining in yeast. PLoS Genet. 4**:** e1000060.
- Villarreal, D. D., K. Lee, A. Deem, E. Y. Shim, A. Malkova *et al.*, 2012 Microhomology directs diverse DNA break repair pathways and chromosomal translocations. PLoS Genet. 8**:** e1003026.
- Wilson, T. E., and M. R. Lieber, 1999 Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase beta (Pol4)-dependent pathway. J. Biol. Chem. 274**:** 23599-23609.
- Xu, G., J. R. Chapman, I. Brandsma, J. Yuan, M. Mistrik *et al.*, 2015 REV7 counteracts DNA double-strand break resection and affects PARP inhibition. Nature 521**:** 541-544.
- Zhang, X., and T. T. Paull, 2005 The Mre11/Rad50/Xrs2 complex and non-homologous end-joining of incompatible ends in *S. cerevisiae*. DNA Repair 4**:** 1281-1294.
- 
- 
- 

# **FIGURE LEGENDS**

- **Figure 1.** ZFN cleavage of spacer ACGAAT and error-prone repair events. (**A**) The sequence of
- the 24 bp *rosy* sequence inserted into *LYS2* is shown, with the three bases recognized by each
- zinc finger (blue or red ovals for RyA and RyB, respectively) indicated. The yellow ovals
- represent the dimerized *Fok*I domains; the 6-bp spacer is in red font and triangles indicate
- positions of enzyme-generated nicks that create 4-bp 5¢ overhangs. (**B**) Sequences added
- during repair events are in lowercase red. Only the first three classes (+CGAA, 2A>3A and
- 1T>2T) generate Lys<sup>+</sup> revertants; the fourth class (+AC) was only observed only among
- survivors. Complete filling of ends duplicates the region bounded by the ZFN nicks (+CGAA)
- while out-of-register pairing between an A and T in the overhangs generates the 2A>3A
- mutation. Duplication of the T (1T>2T) adjacent to the distal overhang can be generated by
- misincorporation-realignment, followed by ligation of the re-created 4-nt overhangs. The +AC
- event can be generated by a similar misincorporation-realignment mechanism during initial
- filling of the proximal overhang. (**C**) Mechanism for the NHEJ-dependent 29-bp deletion that is
- 653 frequent among Lys<sup>+</sup> revertants in the *pol4∆* background. Resection of the 5<sup>*'*</sup> ends allows

 pairing between GCC repeats (red) that flank the DSB. Subsequent removal of 3¢ tails and filling of flanking gaps (red lowercase) completes repair.

**Figure 2.** Frequencies and distributions of Lys<sup>+</sup> revertants in strains with altered core NHEJ 658 proteins. (A) Mean frequencies of Lys<sup>+</sup> colonies; error bars are 95% confidence intervals for the mean. (**B**) Distributions of the five major NHEJ types among revertants. Overall distributions were compared to WT using a global 2 x 5 contingency chi square test (p values are above each spectrum) and if p<0.05, then individual mutation types were compared using the Bonferroni correction to determine significance (p<0.05/5); asterisks indicate a significant proportional class increase/decrease.

 **Figure 3.** Altered frequencies and distributions of Lys<sup>+</sup> revertants in mutant backgrounds. Mean 666 frequencies of Lys<sup>+</sup> colonies are shown; error bars are 95% confidence intervals for the means. The overall distribution of the five major NHEJ types in each mutant background was compared to WT using a 2 x 5 contingency chi square test (ns, not significant). If the global p value (above each spectrum) was less than 0.05, individual mutation classes were compared using the Bonferroni correction to determine significance (p<0.05/5); asterisks indicate a significant proportional increase/decrease.

**Figure 4.** Genetic control of the Pol4-independent, 29-bp deletion. (A) Mean frequencies of Lys<sup>+</sup> colonies; error bars are 95% confidence intervals for the means. (**B**) The overall distribution of the three major NHEJ types in each mutant background. With the ACGAAT spacer only the 2A>3A, the 29-bp deletion, and random "other" events were observed; there were no 1T>2T or +CGAA events. Changing the ACGAAT spacer to ACGTAT eliminated the 2A>3A event (orange cross-hatched), which was replaced by +CGTA (blue). Each distribution was compared to the corresponding WT using a 2 x 3 contingency chi square test (ns, not significant). If the global p

value (above each spectrum) was less than 0.05, individual mutation classes were compared



- 
- 











Strain genotype



**B**









### ND – None detected

Gray shading indicates that the strain was not analyzed

### **Table S2. Spectra in Lys+ revertants**

### **SJR4848 (-1,ACGAAT) Lys+ revertants** Mutation # T A C **A C G A A T** G G C +CGAA 69 T A C **A C G A A C G A A T** G G C 2A>3A 59 T A C **A C G A A A T** G G C 1T>2T 61 T A C **A C G A A T T** G G C 1A>2A 14 T A C **A A C G A A T** G G C 1G>2G 5 T A C **A C G G A A T** G G C 1C>2C 4 T A C **A C C G A A T** G G C 2G>3G 1 T A C **A C G A A T G** G G C +CGAACCC 1 T A C **A C G A A C C C C G A A T** G G C +CGAATGG 1 T A C **A C G A A T G G C G A A T** G G C +GAAT 1 T A C **A C G A A T G A A T** G G C -AC 1 T A C **- - G A A T** G G C -GAATG 2 T A C **A C - - - - -** G C -TGGCG 1 T A C **A C G A A - - - -** -23 bp 1 **- - - - - - - - - - - -** None -29 bp 4 **- - - - - - - - - - - -** GCC -56 bp 1 **- - - - - - A A T** G G C AAT *TOTAL 226*

### **SJR4883 (-1,ACGAAT) ku70Δ Lys+ revertants**

#### **SJR4880 (-1,ACGAAT) dnl4Δ Lys+ revertants**

#### **SJR5247 (-1,ACGAAT) mre11Δ Lys+ revertants**





### **SJR4876 (-1,ACGAAT) pol4Δ Lys+ Revertants**



#### **SJR5097 (-1,ACGAAT) pol4-D367E Lys+ Revertants**



#### **SJR5169 (-1,ACGAAT) pol4Δ mre11-D56N Lys+ Revertants**



### **SJR5196 (-1,ACGAAT) pol4Δrad1Δ Lys+ Revertants**



### **SJR52422 (-1,ACGAAT) pol4Δrev3Δ Lys+ Revertants**



#### **SJR4882 (-1,ACGAAT) pol4Δrad30Δ Lys+ Revertants**



### **SJR4006 (-1,ACGAAT) pol4Δrev3Δrad30Δ Lys+ Revertants**



### **SJR4884 (-1,ACGAAT) rev3Δ Lys+ Revertants**



### **SJR5098 (-1,ACGAAT) rev3-D975A Lys+ Revertants**





### **SJR5193 (-1,ACGAAT) rev7Δ Lys+ Revertants**



### **SJR4963 (-1,ACGAAT) rad30Δ Lys+ Revertants**



### **SJR4964 (-1,ACGAAT) rev3Δrad30Δ Lys+ Revertants**



### **SJR5013 (-1,ACGAAT) sgs1Δexo1Δ Lys+ Revertants**





### **SJR4885 (-1,ACGAAT) tdp1Δ Lys+ Revertants**



### **SJR5199 (-1,ACGAAT) rad1Δ Lys+ Revertants**



### **SJR5200 (-1,ACGAAT) rad10Δ Lys+ Revertants Combined**





### **SJR5202 (-1,ACGAAT) mus81Δrad1Δ Lys+ Revertants**



### **SJR5240 (-1,ACGAAT) slx4Δ Lys+ Revertants**



### **SJR5224 (-1,ACGTAT) Lys+ Revertants**



# **SJR5227 (-1,ACGTAT) mre11-D56NLys+ Revertants**





### **SJR5225 (-1,ACGTAT) pol4Δ Lys+ Revertants**



### **SJR5226 (-1,ACGTAT) pol4Δ mre11-D56NLys+ Revertants**



### **SJR5248 (-1,ACGTAT) pol4Δ mre11Δ Lys+ revertants**



# **Table S3. Spectra in survivors SJR4848 (-1,ACGAAT) Survivors**



-1175 bp 36 T A C **A - - - - - - - -** C C A A G C T A C T A C A -11,731 bp 11 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

#### **SJR4883 (-1,ACGAAT) ku70Δ Survivors**



#### **SJR4880 (-1,ACGAAT) dnl4Δ Survivors**



#### **SJR5247 (-1,ACGAAT) mre11Δ Survivors**



#### **SJR5223 (-1,ACGAAT) mre11-D56N Survivors**



**SJR4876 (-1,ACGAAT) pol4Δ Survivors**



#### **SJR5097 (-1,ACGAAT) pol4-D367E Survivors**



 $CCAAGCTACTACA$ -11,731 bp 2 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

#### **SJR5169 (-1,ACGAAT) pol4Δ mre11-D56N Survivors**



 $CCAAGCTACTACA$ -11,731 bp 8 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

#### **SJR5196 (-1,ACGAAT) pol4Δrad1Δ Survivors**



-1175 bp 38 T A C **A - - - - - - - -** C C A A G C T A C T A C A was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: [https://doi.org/10.1101/2023.05.04.539391;](https://doi.org/10.1101/2023.05.04.539391) this version posted May 4, 2023. The copyright holder for this preprint (which



#### **SJR5242 (-1,ACGAAT) pol4Δrev3Δ Survivors**



-1175 bp 20 T A C **A - - - - - - - -** C C A A G C T A C T A C A -11,731 bp 10 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR4882 (-1,ACGAAT) pol4Δrad30Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 10 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR4884 (-1,ACGAAT) rev3Δ Survivors**



-1175 bp 22 T A C **A - - - - - - - -** C C A A G C T A C T A C A -11,731 bp 12 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

#### **SJR5098 (-1,ACGAAT) rev3-D975A Survivors**



 $CCAAGCTACTACA$  $T G G A A A A A A A A A A A$ 

#### **SJR5193 (-1,ACGAAT) rev7Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 18 **- - - - - - - - - - - - - - -** T G G A A A A A A A A A A A

#### **SJR4963 (-1,ACGAAT) rad30Δ Survivors**



 $CCAAGCTACTACA$  $T G G A A A A A A A A A A A A$ 

 $\frac{1}{28}$  T  $\frac{1}{28}$  T  $\frac{1}{28}$  T  $\frac{1}{28}$  T  $\frac{1}{28}$  T  $\frac{1}{28}$  T  $\frac{1}{28}$  C  $\frac{1}{28}$   $\frac{1}{28}$ Maintain Cut Site **1985** | and altable under aCC-BY-NC-NIC-NIC-4.0 International license.<br>TOTAL sequenced 91 was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: [https://doi.org/10.1101/2023.05.04.539391;](https://doi.org/10.1101/2023.05.04.539391) this version posted May 4, 2023. The copyright holder for this preprint (which





 $CCAAGCTACTACA$ -11,731 bp 11 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5013 (-1,ACGAAT) sgs1Δexo1Δ Survivors**



 $CCAAGCTACTACA$ 

#### **SJR5194 (-1,ACGAAT) rad52Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 1 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR4885 (-1,ACGAAT) tdp1Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 14 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5199 (-1,ACGAAT) rad1Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 9 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5200 (-1,ACGAAT) rad10Δ Survivors combined**



 $CCAAGCTACTACA$ -11,731 bp 17 **- - - - - - - - - - - -** T G G A A A A A A A A A A A



 $CCAAGCTACTACA$ -11,731 bp 6 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5202 (-1,ACGAAT) mus81Δrad1Δ Survivors**



#### **SJR5240 (-1,ACGAAT) slx4Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 0 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5224 (-1,ACGTAT) Survivors**

### SJR5227 (-1,ACGTAT) mre11-D56N Survivors

#### **SJR5225 (-1,ACGTAT) pol4Δ Survivors**

#### SJR5226 (-1, ACGTAT) pol4Δ mre11-D56N Survivors

### **SJR5248 (-1,ACGTAT) pol4Δmre11Δ Survivors**



 $CCAAGCTACTACA$ -11,731 bp 0 **- - - - - - - - - - - -** T G G A A A A A A A A A A A

### **SJR5241 (-1,ACGTAT) pol4Δ pol3-DV Survivors**



 $CCAAGCTACTACA$  $T G G A A A A A A A A A A A$