Bacterial DNA polymerases participate in oligonucleotide recombination

Xin-tian Li,[1][†] Lynn C. Thomason,[1,2][†]
James A. Sawitzke,[1][†] Nina Costantino[1] and
Donald L. Court[1]*
[1]*Molecular Control and Genetics Section, Gene
Regulation and Chromosome Biology, Frederick
National Laboratory for Cancer Research, Frederick,
MD 21702, USA.*
[2]*Basic Science Program, SAIC-Frederick, Inc.,
Frederick, MD 21702, USA.*

Summary

**Synthetic single-strand oligonucleotides (oligos) with
homology to genomic DNA have proved to be highly
effective for constructing designed mutations in tar-
geted genomes, a process referred to as recom-
bineering. The cellular functions important for this
type of homologous recombination have yet to be
determined. Towards this end, we have identified
Escherichia coli functions that process the recombin-
ing oligo and affect bacteriophage λ Red-mediated
oligo recombination. To determine the nature of oligo
processing during recombination, each oligo con-
tained multiple nucleotide changes: a single base
change allowing recombinant selection, and silent
changes serving as genetic markers to determine the
extent of oligo processing during the recombination.
Such oligos were often not incorporated into the host
chromosome intact; many were partially degraded in
the process of recombination. The position and
number of these silent nucleotide changes within the
oligo strongly affect both oligo processing and
recombination frequency. Exonucleases, especially
those associated with DNA Polymerases I and III,
affect inheritance of the silent nucleotide changes in
the oligos. We demonstrate for the first time that the
major DNA polymerases (Pol I and Pol III) and DNA
ligase are directly involved with oligo recombination.**

Introduction

Homologous recombination-mediated genetic engineer-
ing, recombineering, utilizes the phage λ recombination
proteins Gam, Exo and Beta, collectively known as 'Red'
(Murphy, 1998; Zhang *et al.*, 1998; Yu *et al.*, 2000; Ellis
et al., 2001). For recombineering, a PCR product or single-
strand oligonucleotide (oligo) that contains the desired
sequence change is introduced into Red-expressing, elec-
trocompetent cells to obtain the alteration in the targeted
DNA. Oligo recombination, the focus of this paper, requires
only the λ Beta protein (Ellis *et al.*, 2001; Costantino and
Court, 2003). Beta is a single-strand DNA (ssDNA) binding
protein that can protect oligos from exonuclease degrada-
tion (Muniyappa and Radding, 1986; Karakousis *et al.*,
1998) and promote annealing of complementary ssDNAs
(Kmiec and Holloman, 1981). In a thorough screen of
E. coli host recombination functions, none is required for λ
Red-mediated oligo recombination (Ellis *et al.*, 2001;
Costantino and Court, 2003; Sawitzke *et al.*, 2011).

A number of factors affect the efficiency of oligo recom-
bination. For any target DNA sequence, either of two
complementary oligos can be used to modify it. The recom-
bination efficiency of the oligo corresponding in sequence
to the discontinuously replicated, lagging-strand is more
efficient than the leading-strand oligo (Ellis *et al.*, 2001;
Costantino and Court, 2003; Zhang *et al.*, 2003; van
Kessel and Hatfull, 2008; Swingle *et al.*, 2010; Sawitzke
et al., 2011; Van Pijkeren *et al.*, 2012). All models to explain
oligo recombination must incorporate this bias caused by
DNA replication. In our favoured model for oligo recombi-
nation, Beta protein binds, protects and anneals the oligo
to complementary regions of ssDNA existing transiently at
the DNA replication fork (Fig. 1) (Court *et al.*, 2002). The
lagging/leading-strand bias can be explained by the pres-
ence of a larger single-strand gap on the lagging-strand
(Benkovic *et al.*, 2001; Lajoie *et al.*, 2012), which allows
the lagging-strand oligo greater opportunity to anneal,
thereby resulting in a higher recombination frequency (Ellis
et al., 2001; Costantino and Court, 2003; Sawitzke *et al.*,
2011). An alternate model in which oligo recombination
initiates by Beta-mediated DNA strand invasion (Noirot and
Kolodner, 1998; Rybalchenko *et al.*, 2004), giving rise to a
D-loop, could also explain the replication bias, if the
lagging- and leading-strand oligos are treated differently
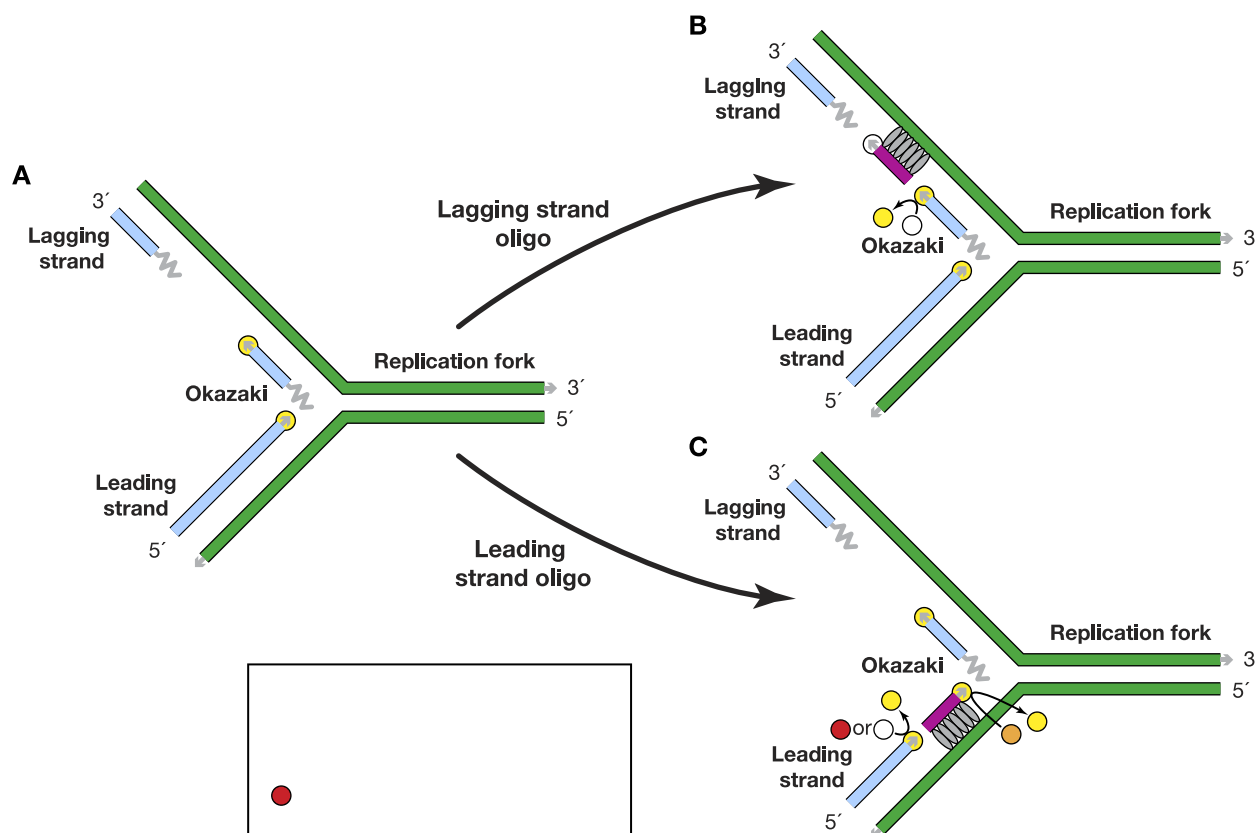
Fig. 1. Oligo processing at the replication fork.
A. A replication fork with template (green) and newly synthesized (blue) DNA is shown. Pol III (yellow circles) extends from the 3′ ends of the Okazaki fragment and the leading-strand.
B. A lagging-strand oligo (purple) is shown bound by Beta (grey ovals) to the template at a single-strand gap. Pol I, which degrades the RNA primer of the Okazaki fragment, can also degrade the 5′ end of the lagging-strand oligo after replacing Pol III. DNA polymerization by Pol I readies the recombinant oligo for DNA ligation and incorporation into the genome.
C. A leading-strand oligo (purple) is shown bound by Beta (grey ovals) to the leading-strand template at the single-strand gap ahead of the newly synthesized leading-strand. An unknown exonuclease(s) (orange circle) processes the 3′ end of the oligo. Pol I and/or an unknown exonuclease (red circle) is involved in the degradation of the 5′ end of the leading-strand oligo.

when the replication fork reaches the D-loop. Sawitzke *et al.* (2011) found that proximity of the genetic marker to the end of the oligo also affects recombination efficiency, as does the length and concentration of oligo used. The *E. coli* methyl-directed mismatch repair (MMR) system lowers the efficiency of oligo recombination at least 100-fold (Costantino and Court, 2003; Li *et al.*, 2003) by removing mismatches when an oligo is annealed to its target DNA. Thus, elimination of the MMR system results in a higher recombination frequency. Under optimal conditions, more than 50% of the cells are recombinant at the *galK* locus (Sawitzke *et al.*, 2011).

Wang *et al.* (2011) reported that not all markers in an oligo were inherited in the process of recombineering, but the molecular details of oligo recombination and this marker loss have not been elucidated. During oligo recombination, is the oligo processed while still single-stranded and/or after annealed to the target? How does marker distribution along the oligo affect recombination

frequency and marker inheritance? In principle, exonucleases that act on either ssDNA or dsDNA could affect oligo recombination. There are four major exonucleases (RecJ, ExoI, ExoVII and ExoX) that degrade ssDNA in *E. coli* (Burdett *et al.*, 2001) and decrease Red-mediated oligo recombination, but only when oligo is limiting (Sawitzke *et al.*, 2011). Several dsDNA exonucleases may also affect oligo recombination after the oligo has annealed to the target. The RecBCD enzyme has multiple nuclease activities that collectively destroy linear dsDNA (Goldmark and Linn, 1972). Other dsDNA exonucleases that could be involved in processing include Exonuclease III (Richardson *et al.*, 1964), Exonuclease IX (Allen *et al.*, 2009), and λ Exo (Little *et al.*, 1967).

Three DNA polymerases have inherent exonuclease activities. DNA Polymerase III (Pol III), the major replicative polymerase, has a 3′ → 5′ proofreading exonuclease encoded by *dnaQ* (Echols *et al.*, 1983). DNA Polymerase I (Pol I), encoded by the *polA* gene, is a multifunctional

protein involved in DNA replication and repair. Pol I has two exonuclease activities: $3' \to 5'$ proofreading and $5' \to 3'$ nick translation (Rigby *et al.*, 1977; Kornberg, 1990); the latter activity is responsible for removing RNA primers from the 5' end of Okazaki fragments on the lagging-strand. DNA Polymerase II (Pol II), encoded by *polB*, is a repair polymerase with a $3' \to 5'$ proofreading exonuclease activity (Banach-Orlowska *et al.*, 2005).

The present study used multiply marked oligos to examine the effect of exonuclease activities on recombination efficiency and the fate of the oligo. Sequence analysis of the recombinants reveals that nucleotides can be lost from either one or both ends of the oligo, and this processing occurs after the oligo is annealed to the target DNA. The number and distribution of markers on the oligo directly alter processing during incorporation and affect recombination frequency. The major DNA polymerases, Pol I and Pol III, as well as DNA ligase affect recombination frequency and marker loss.

Results and discussion

Genetic engineering *in vivo* by oligo recombination is a powerful tool that is becoming the method of choice for genome modification. This technique routinely utilizes ssDNA oligos with centrally located modifications. We recently reported (Sawitzke *et al.*, 2011) that the location of the modifying base (marker) along the length of the oligo affects the efficiency of incorporation of that base. Oligos having the genetic marker less than nine bases from an end have lower recombination efficiencies than those with the marker further from an end. To better understand how marker position affects recombination, we used oligos containing multiple markers arranged in various configurations to monitor recombination efficiency and marker recovery within recombinants.

Experimental design

The experimental design is illustrated in Fig. 2. We used derivatives of the MMR-defective strains, HME63 and HME68 (Costantino and Court, 2003; Thomason *et al.*, 2007a), which contain a *galK* amber mutation, *galK*$_{TYR145UAG}$. Use of MMR mutant strains ensures that marker loss from the recombining oligo is not the result of mismatch repair, but is due to other host functions. The selected genetic marker, which is always centred in the oligo, corrects the *galK* amber mutation at codon 145 and confers the ability to grow on minimal galactose medium, allowing selection of Gal⁺ recombinants. Other bases in the oligo are used as markers in the sense that they vary from wild type sequence (i.e. transitions and transversions) but, because of their placement in wobble positions, do not alter the GalK protein. These unselected markers are mispaired
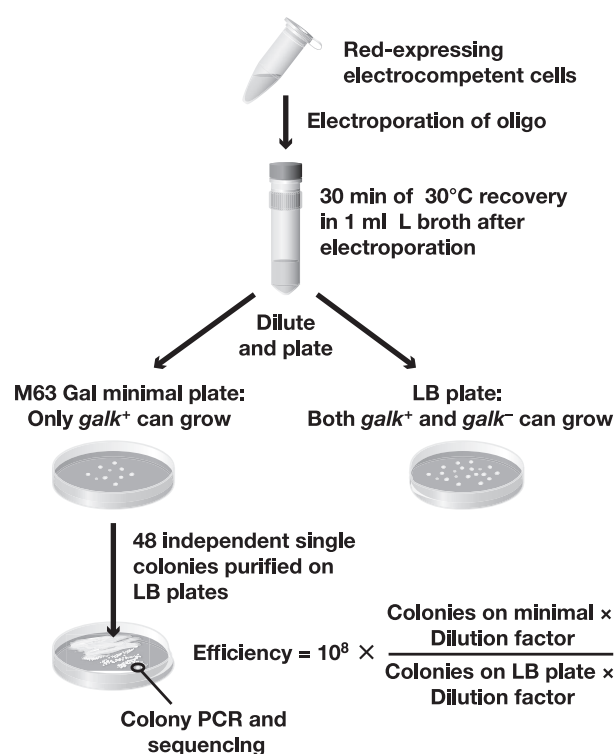
Fig. 2. Experimental outline for generating and analysing GalK⁺ recombinants. The experiments were performed as outlined. Although at least 48 independent recombinants were analysed, on occasion an individual colony yielded poor sequence and thus was not added to the data.

when the oligo is annealed to the target DNA, and their inheritance can be followed by DNA sequencing.

In these experiments, the Red functions were induced for 15 min and cells were prepared for electroporation by standard protocols (see *Experimental procedures*). The period of incubation following oligo electroporation was limited to 30 min, which allows for cell recovery (Sawitzke *et al.*, 2011).

Our previous paper demonstrates that functions in the lambda P_L operon, when induced for the 15 min time period used here, block cell division for 50 min and result in cell synchrony without loss of cell viability (Sergueev *et al.*, 2002). Hence in the experiments here, each Gal⁺ colony is derived from an independent recombination event. Sequence analysis was used to determine marker inheritance in these independent recombinant clones. Oligo concentration and the presence or absence of specific host functions have been examined for their effects on recombination frequency and marker inheritance.

Oligo design affects recombination efficiency

We found that marker placement along the oligo affected recombination frequency. Oligo 144 contains only the

A Position and number of markers affect recombination				B Marker loss (%)		C. Data
Oligo	Number of mismatches	Oligo features	Efficiency	5′	3′	Figure
144	1	5′ ——————+——————	$3.2×10^7$	NA	NA	–
XT13	7	5′ (15 21 30 39 48 54 63)	$2.2×10^7$	10	15	4A
XT36	13	5′ (3 ... 72)	$1.7×10^6$	100	100	S1 A
XT21	24	5′ (3 ... 72)	$<1.3×10^1$	NA	NA	S1 B
XT18	17	5′ (15 ... 63)	$1.3×10^6$	0	0	S1 C
XT524	24	5′ (24 ... 93)	$3.7×10^6$	0	0	S1 D
XT30	12	5′ (... 72)	$5.1×10^6$	NA	100	S1 E
XT29	13	5′ (3 ...)	$2.4×10^3$	100	NA	S1 F
XT351	18	5′ (3 ... 54)	$6.1×10^1$	100	0	S1 G
XT38	18	5′ (21 ... 72)	$6.3×10^4$	0	89	S1 H
XT37	7	5′ (3 18)	$2.3×10^7$	100	NA	S1 I
XT352	12	5′ (3 18 ... 60 72)	$1.3×10^7$	100	100	S1 J

Fig. 3. Recombination with multiply marked, lagging-strand oligos. A summary of oligos used to examine the effect of marker density and spacing is shown. Oligo 144 is 70 nt, oligo XT524 is 117 nt, and all others are 75 nt in length. Experiments were performed in the MMR mutant host, HME68. In (A) features of each oligo are shown with a diagram (to scale except XT524), indicating the number and placement of mismatches. The '+' denotes the selected base and the 'l' denotes the unselected markers. The number above markers indicates the distance in bases from the 5′ end. Recombination efficiency is normalized to 10^8 viable cells; efficiencies given here are a representative experiment from at least 3 independent experiments where variability was less than fivefold. 95% confidence limits can be found in Table S4. In (B), the percentage of recombinants that lost markers from either the 5′ or 3′ end are enumerated for each oligo. If a recombinant lost markers from both ends it is not recorded in either column, but such events happen and are shown (5′ & 3′) in Figs 4 and S1. Panel C indicates the Figure that contains the data for each of the oligos. NA: not applicable, either because no unselected markers are present on that end, or because there were no recombinants (XT21). For oligo XT21, 16 rare Gal+ colonies were recovered, 13 of which did not inherit the selected 'T' marker from the oligo, i.e. they were spontaneous revertants. The three isolates that have a TA'T' codon could either be derived from the oligo (true recombinants) or be revertants (Fig. S1B and legend). These are unlikely to be true recombinants because no other oligo markers were inherited, and they occur at the same frequency as other spontaneous revertants.

selected *galK*+ mismatch and had the highest efficiency in this study (Fig. 3A). Oligo XT13 had a similar efficiency despite having six additional silent mismatches. Oligo XT36, containing 12 silent wobble position markers six bases apart, recombined at a 20-fold reduced efficiency. In contrast, oligo XT21, with 23 silent markers spaced every third base along its entire length, showed no recombination activity. For oligo XT21, sequencing revealed that the few Gal+ colonies obtained appeared to have spontaneous mutations eliminating the amber codon (see legends to Figs 3A and S1). These four oligos differ in both density of markers and extent of homology at their respective ends, two factors which could influence recombination.

Keeping the marker density constant by using oligos with silent markers spaced every third base, we asked how the positioning of these markers along the oligo affects recombination. Oligo XT18 contains such markers flanking the selected change but leaving at least 12 bases of uninterrupted homology at each end. Recombination with this oligo is at least 10 000-fold higher than that with XT21, demonstrating the importance of homology at the oligo ends. Oligo XT524 has the same sequence as XT21

but carries an additional 21 bases of perfect homology on each end. In this respect, it is like XT18 and likewise, XT524 recombines at high efficiency (Fig. 3A).

To ask if homologies at the two ends are of equal importance, we used oligos XT30 and XT29, which are completely homologous to the target but only at the 5′ or 3′ end respectively. The 2000-fold difference in recombination frequency between these two oligos demonstrates that homology at the 5′ end is much more important than at the 3′ end (Fig. 3A). This result is consistent with our previous data showing that 5′ homology is critical for efficient recombination (Sawitzke *et al.*, 2011).

Like oligo XT29, oligo XT351 contains silent markers extending from the 5′ end, but its markers span the selected base. The efficiency of XT351 was reduced ~ 40-fold relative to oligo XT29. Like oligo XT30, oligo XT38 contains silent markers extending from the 3′ end, but its markers also span the selected base. The efficiency of XT38 was reduced ~ 80-fold relative to oligo XT30. However, recombination with XT30 and XT38 is > 1000-fold more than that with XT29 and XT351, respectively, again demonstrating the greater importance of intact 5′ homology. In general, homology just 5′ to the selected marker is important for efficient recombination. The exceptions are oligo XT18 and XT524, which compensate by having adequate homology at both the 5′ and 3′ ends.

XT37 and XT352, despite being mispaired at the 5′ ends, recombine robustly (Fig. 3A). What distinguishes them is that both oligos retain considerable homology flanking the selected marker. To further investigate why 5′ homology is important in some cases but not others, we sequenced Gal⁺ recombinants from each oligo recombination reaction to determine the fate of unselected markers.

Comparing marker loss from lagging- and leading-strand oligos

Sequencing of Gal⁺ recombinants allowed us to trace the inheritance of the silent markers and to determine whether or not they were incorporated into final recombinant products. Indeed, we found that oligos are not always incorporated in their entirety but often suffer loss of silent markers. As we will discuss below, this marker loss is caused by exonuclease degradation. Figure 4A illustrates the data collected for the lagging-strand oligo XT13 and defines the marker categories used. Among independent recombinants sequenced, all of the oligo-encoded silent markers were inherited in 67% of the recombinants, whereas 25% showed marker loss from one end or the other. The extent of marker loss is variable in different recombinants and suggests sequential but

incomplete degradation from an oligo end. In contrast, with the complementary leading-strand oligo XT14 (Fig. 4B), all the silent markers were inherited in 27% of the recombinants, whereas 57% of recombinants showed loss on one end or the other. In a few cases, markers were lost at both ends; this occurred in 2% and 7% of total recombinants for the lagging- and leading-strand oligos respectively. Recovery of these recombinants is at a rate expected if the processing events at the 5′ and 3′ ends were independent.

To examine marker loss from the ends in more detail, we used lagging- or leading-strand oligos with mismatches six bases apart (XT36 and XT418). For each oligo, independent recombinants were sequenced (Figs S1A and S2A), and per cent loss was determined for each of the 12 silent markers (Fig. 5). As shown, the extreme 5′ and 3′ markers were lost in every recombinant examined. There is a gradient of marker loss on each side of the centrally selected marker. From the 5′ ends, these gradients are similar for both oligos, whereas marker loss is greater for the leading-strand on the 3′ side of the selected marker. We conclude that lagging- and leading- strand oligos are treated differently after they enter the replication fork during the recombination event.

We also observed other rare recombinants where internal markers were replaced by chromosomal bases (see 'Other' in Figs 4, S1, S2 and Table 1). There are a total of 60 instances of these 'Other' events, and surprisingly all bases in question are derived from the chromosome. Thus, spontaneous mutagenesis is not the cause, nor is error during oligo production. Although MMR is not involved because these cells are deleted for *mutS*, another DNA repair system may be responsible. Yet, this class of recombinants was found at a similar frequency even in cells additionally mutant for very short patch repair (VSP) or the MutY repair system (data not shown). We do not understand how this class of recombinants occurs. They cannot be derived from recombination events involving more than one oligo because they appear at a similar frequency even when oligo concentration is limiting (compare Fig. 4A versus Fig. S1K).

Position of the silent markers affects marker loss

Figure 3B summarizes the markers lost by recombinants obtained from the oligos described in Fig. 3A, and allows comparison of oligos with different marker configurations. This information is derived from data like that shown in Fig. 4A for oligo XT13. As illustrated in Fig. 4A and tabulated in Fig. 3B, 25% of the oligo XT13 recombinants had lost markers from one or the other end, 10% from the 5′ end and 15% from the 3′ end, while 2% had lost markers from both ends. Recombinants with oligo XT36, which has

A HME68

Lagging-strand oligo XT13
Efficiency: 2.2 × 10⁷

B HME68

Leading-strand oligo XT14
Efficiency: 6.7 × 10⁵

C XTL70
(*polA* (5′ → 3′ exo)<>*cat*)

Lagging-strand oligo XT13
Efficiency: 2.4 × 10⁶

D XTL70
(*polA* (5′ → 3′ exo)<>*cat*)

Leading-strand oligo XT14
Efficiency: 1.3 × 10⁵
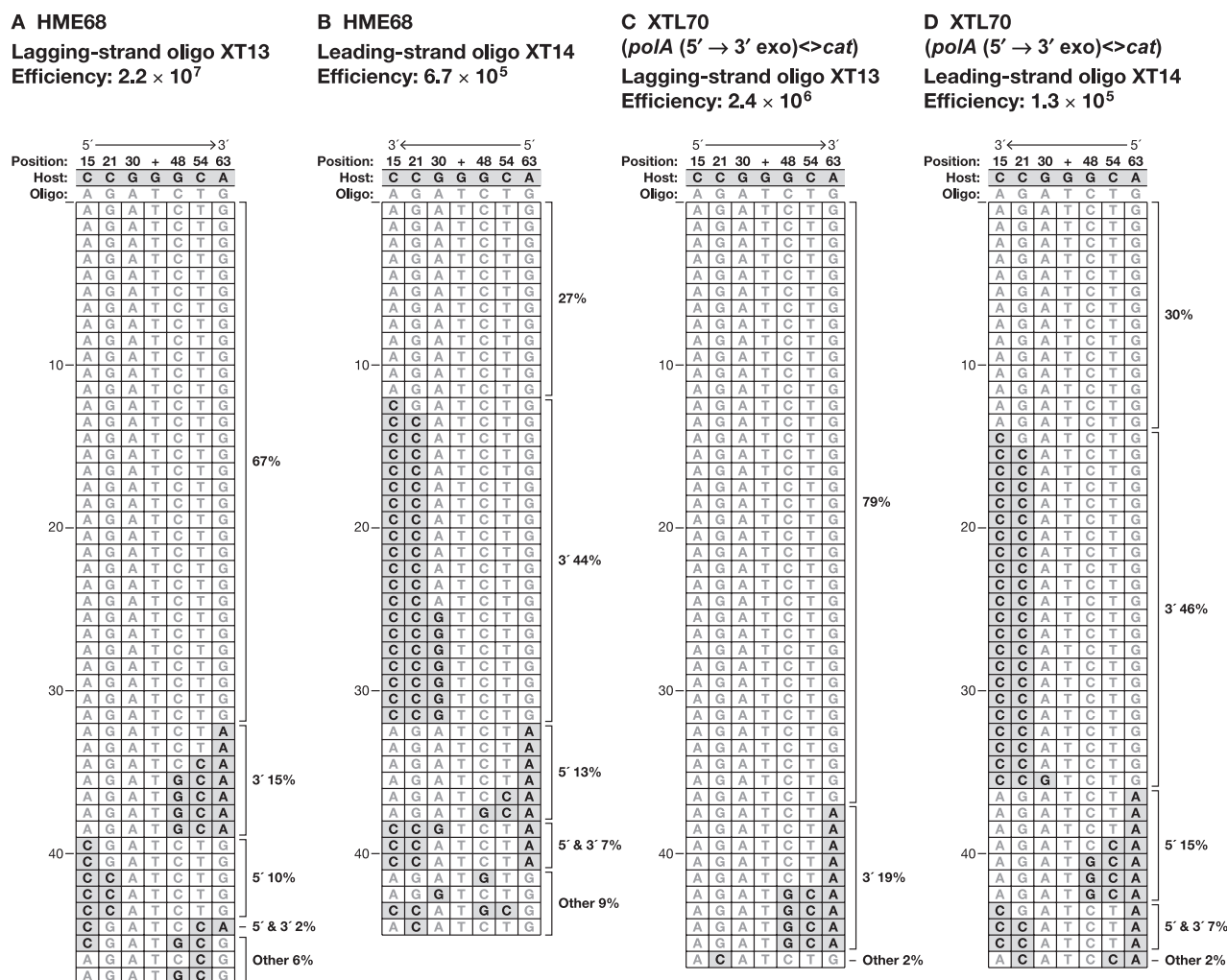
Fig. 4. Marker loss pattern in wild type and polymerase mutant cells. Marker loss is compared for the lagging- and leading-strand oligo recombinants in 'wild type', HME68 (see A and B respectively). Similar experiments are shown for the Pol I mutant derivative of HME68, XTL70, which is defective for the 5′ → 3′ exonuclease activity (C and D), the Pol III mutant, XTL76 (E and F), and the ligase mutant, XTL47 (G and H). The recombination efficiency of the illustrated experiment is shown at the top of each panel. Normal variation in recombination frequency is less than fourfold. In each panel, the 'Position' row indicates the distance of the markers from the left end of the oligo as diagrammed, with the '+' denoting the selected marker at position 39, which is the G of the TAG amber codon in the 'Host' sequence. The 'Oligo' row shows the sequence changes present on the oligo at the indicated position. The sequences of Gal⁺ recombinants are shown below the 'Oligo' row; grey shaded spaces indicate those bases that remained unchanged from the host and show where markers were lost from the oligo, white spaces indicate those markers that were inherited from the oligo. For each panel, recombinants are grouped according to their pattern of marker loss. The uppermost group of recombinants showed no marker loss. Groups with 3′, 5′, or 5′ & 3′ marker loss are indicated. The final group, 'Other', comprises recombinants with internal markers lost. In (A), (C), (E) and (G), the lagging-strand oligo XT13 (Table S2 and Fig. 3) was used for recombination. In (B), (D), (F) and (H), the leading-strand oligo XT14 (Table S2) was used. The oligos XT13 and XT14 are designed to create the same changes in the final product and have complementary sequences; their 5′ → 3′ orientation is indicated by the arrow above each panel. In all panels, the bases shown for the 'Host' are from the *galK_am* coding sequence. For simplicity of presentation, the oligo bases in (B), (D), (F) and (H) are shown as identical to those in (A), (C), (E) and (G) but in reality the leading-strand oligo (XT14) for (B), (D), (F) and (H) contains markers that are the complement of the sequence shown.

markers spaced every 6 bases extending to within 2–3 bases of the oligo ends, always lost markers from both the 5′ and 3′ ends (Figs 3B and S1A). For oligo XT36, like oligo XT13, the extent of marker loss varies in different recombinants, again suggesting sequential degradation. The Gal⁺ isolates found with oligo XT21, which contains mismatches every third base over the entire length of the oligo, are most easily explained by spontaneous mutation and not recombination (Figs 3B and S1B), implying an inability of that oligo to pair with its target. In contrast, for XT36, enough homology remains among the mismatches to allow sufficient pairing and recombination. Indeed, our unpublished *in vitro* results show that Beta is unable to anneal two complementary oligos with every third base mismatched but is able to anneal two complementary oligos with every sixth base mismatched. Similar results

E XTL76 (*dnaQ<>cat*)
Lagging-strand oligo XT13
Efficiency: 5.2 × 10⁵
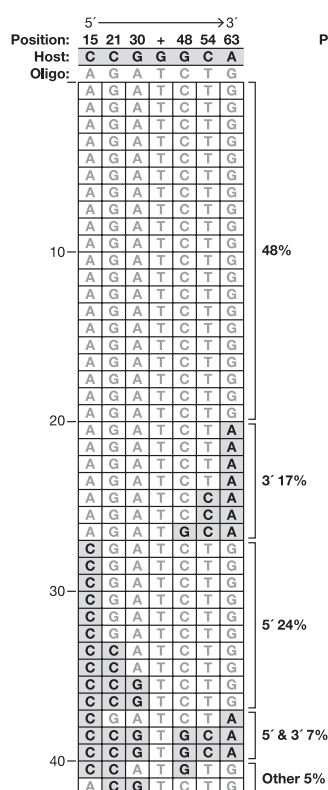
F XTL76 (*dnaQ<>cat*)
Leading-strand oligo XT14
Efficiency: 1.6 × 10⁴
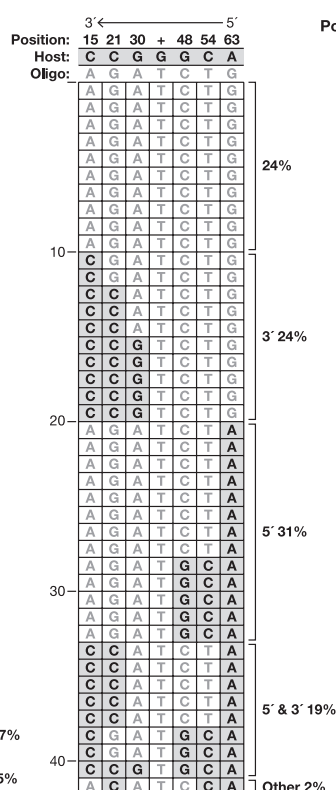
G XTL47 (*lig7ᵗˢ*)
Lagging-strand oligo XT13
Efficiency: 1.7 × 10⁵
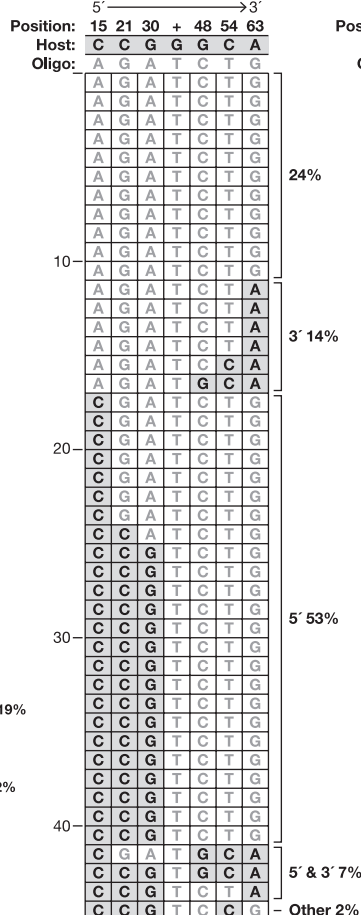
H XTL47 (*lig7ᵗˢ*)
Leading-strand oligo XT14
Efficiency: 5.3 × 10²
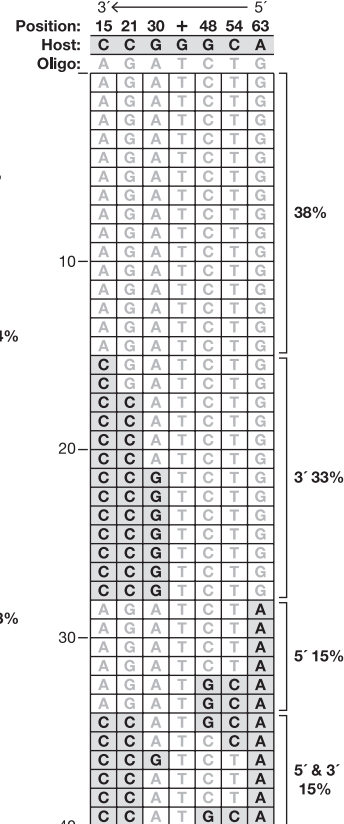
Fig. 4. *cont.*

have been found by others (Martinsohn *et al.*, 2008), showing Red-mediated recombination between 22% diverged sequences.

Recombinants obtained with oligo XT18 and XT524, which have markers every third base in the segment flanking the selected marker, inherit the entire set of silent mutations 100% of the time (Fig. 3B, Fig. S1C and D). With these oligos, sufficient homology exists at both ends to allow robust recombination despite internal mismatches at every third base. Our failure to observe any marker loss with either oligo can be explained by the idea that any degradation extending into the mismatch region prevents sufficient pairing and thus, prevents the selected recombination event. The modest reduction in recombination frequency observed with these two oligos is consistent with this idea.

It should be noted that oligo XT524 carries all 24 of the third position markers used in this study. The efficient recovery of Gal⁺ recombinants containing all the markers from this oligo shows that the 23 silent markers *per se* do not interfere with either the recombination process or GalK protein function (Fig. S1D). Oligos designed like XT18 or XT524 could be extremely useful for creation of clustered changes in a target sequence. Such changes can be screened for by direct PCR analysis or colony hybridization (Swaminathan *et al.*, 2001; Sharan *et al.*, 2009).

Oligos XT30 and XT29 have complete homology to one side of the selected marker, but contain unpaired markers every third base from the selected marker to the other end (Fig. 3A). In oligos XT351 and XT38, the mismatches extend from an end and span the selected marker (Fig. 3A). In all cases, the homology on one side must permit some pairing and recombination within the mispaired region (Fig. 3B, Fig. S1E–H). Most of the recombinants have lost markers from the mispaired end.
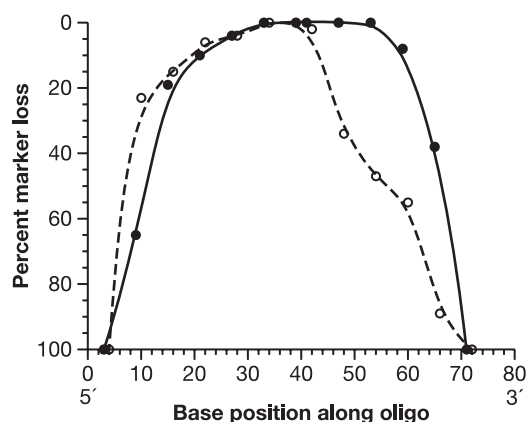
Fig. 5. Marker loss pattern of lagging-strand oligo XT36 and leading-strand oligo XT418. Loss of markers at each position along the oligos was analysed. Each oligo has one marker every 6 nucleotides, and the *x*-axis indicates those positions relative to the 5′ end of each oligo. The per cent marker loss at each position is plotted on the *y*-axis. Note that 100% marker loss is at the origin. The lagging-strand oligo XT36 (Fig. S1A) is indicated by a solid line with closed circles, and the leading-strand oligo XT418 (Fig. S2A) is indicated by a dashed line with open circles.

Results described here agree with earlier experiments (Sawitzke *et al.*, 2011), in which the selected marker was the only mispair but was positioned at different locations along the length of the oligo. There, recombination efficiency was greatly decreased if the selected marker was less than nine bases from an end. The data suggests that a minimal homology region is necessary to allow efficient pairing, and when the oligo is paired at both ends, such as with XT18, XT524 or XT13, it is readily incorporated into the newly synthesized DNA. Closely spaced mismatches at an oligo end inhibit recombination of that end, and mismatches at the 5′ end inhibit more than at the 3′ end. If the oligo has mismatches near the ends, but adequate homology closer to the selected marker (oligos XT37 and XT352), efficient recombination can occur, but the terminal mismatches are removed (Fig. 3, Fig. S1I and J).

Host functions

Sawitzke *et al.* (2011) showed that *E. coli* recombination functions had little or no effect on Red-mediated oligo recombination. Our observations of marker loss implicate exonucleases in the recombination process; therefore, we examined the role of several exonucleases to determine how they impact recombination frequency and marker loss.

Single-strand exonucleases. Several single-strand exonucleases affect oligo recombination that occurs in the absence of the λ Red proteins (Dutra *et al.*, 2007). When we eliminated four exonucleases, RecJ, ExoI, ExoVII and ExoX by deletions in a single strain XTL51 ('Quad' mutant),

we saw little effect on Red-mediated oligo recombination levels except at limiting oligo concentrations where recombination frequency was enhanced (Sawitzke *et al.*, 2011). These results suggest that the exonucleases, when present, destroy oligos and eliminate the substrate for recombination. In those experiments, the selectable marker was the only base change on the oligo, so marker loss at other positions was impossible to monitor.

Here, those experiments were repeated using complementary lagging- and leading-strand oligos (XT13 and XT14) containing six silent markers in addition to the selected marker, and the impact of oligo concentration and single-strand exonucleases on recombination frequency and marker loss was monitored (Table 2A). Despite the additional markers, the efficiencies of recombination found with oligos XT13 and XT14 are similar to those obtained in the previous study (Sawitzke *et al.*, 2011), i.e. at high oligo concentration, there is no difference between recombination frequency for wild type and the 'Quad' mutant whereas at low oligo concentration the 'Quad' mutant has higher frequencies.

However, sequence analysis of the Gal+ recombinants from these crosses revealed a difference between wild type and the 'Quad' mutant for the pattern of marker loss for both lagging- and leading-strands, which was dependent on oligo concentration (Table 2A and relevant panels in Figs S1 and S2). At low oligo concentration, the 'Quad' mutant showed a dramatic increase in marker loss but only from the 3′ end of the lagging-strand oligo XT13 (Tables 1A and 2A; Fig. 4A, Fig. S1K, L and M). Finding an increased marker loss from the 3′ end of the lagging-strand oligo in the absence of the four exonucleases was unexpected. In explanation, we suggest that in the presence of the four single-strand exonucleases, contributions from other minor nuclease(s) might be masked. When the single-strand exonuclease activities are missing, the activity of a less aggressive 3′ exonuclease, which removes only some markers, may be revealed. An alternative explanation could involve activation of a previously silent exonuclease.

At high oligo concentration, the 'Quad' mutant showed a slight but significant decrease in marker loss but only from the 3′ end of the leading-strand oligo XT14, presumably due to the loss of nuclease activity (Tables 1B and 2A, Fig. 4B versus Fig. S2C). The fact that this effect was specific for the leading-strand indicates that the degradation occurs after the oligo anneals to its target. As Exo X is the only ssDNA exonuclease with 3′ → 5′ dsDNA activity, it was the most likely candidate (Viswanathan and Lovett, 1999). However, when tested, a mutant of Exo X (XTL532) had no effect (data not shown).

In summary, the four single-strand exonucleases do not appear to be major players in marker loss, in agreement with Mosberg *et al.* (2012), who saw a minor effect on the

Table 1A. Summary of lagging-strand marker loss data in host mutants.[a]

Strain	Genotype	Oligo[b]	5′ loss[c]	3′ loss[c]	Other[c]	Rec[d]	P-value 5′ loss	P-value 3′ loss	P-value Other
HME68	Wild type	XT13	7	8	3	48	1.0	0.48	1.0
XTL70	polA5′ → 3′ exo	XT13	0	9	1	47	**0.002**	0.72	0.51
XTL76	dnaQ<>cat	XT13	13	10	2	40	**0.02**	0.85	1.0
XTL47	lig7ᵗˢ	XT13	27	9	1	44	**<0.0001**	0.86	0.51
XTL51	Quad mutant	XT13	5	12	4	44	0.67	0.60	0.35
HME69	polA resA1	XT13	4	7	4	42	0.38	0.46	0.33
XTL92	xthA<>tet	XT13	7	10	4	43	1.00	1.00	0.34
XTL100	recBCD sbcC	XT13	7	13	2	61	0.47	0.88	0.57
XTL324	xni<>cat	XT13	7	13	5	39	0.67	0.27	0.10
SIMD90	λ Δexo	XT13	6	6	3	42	1.00	0.34	0.74
XTL372	polB<>spec	XT13	8	14	1	55	1.00	0.75	0.36
HME70	ΔrecA	XT13	7	4	4	42	0.83	0.08	0.34
HME68	Wild type	XT13/10³	5	6	4	40	0.82	0.44	0.31
XTL51	Quad mutant	XT13/10³	4	27	3	43	0.38	**<0.0001**	0.74
XTL85	polA′5′ → 3′ exo	XT13	0	9	0	46	**0.002**	0.72	0.10
			107	157	41	676			

Table 1B. Summary of leading-strand marker loss data in host mutants.[a]

Strain	Genotype	Oligo[b]	5′ loss[c]	3′ loss[c]	Other[c]	Rec[d]	P-value 5′ loss	P-value 3′ loss	P-value Other
HME68	Wild type	XT14	9	24	4	45	0.86	0.33	0.08
XTL70	polA5′ → 3′ exo	XT14	10	25	1	46	1.00	0.27	1.0
XTL76	dnaQ<>cat	XT14	23	18	1	41	**0.0006**	0.88	1.0
XTL47	lig7ᵗˢ	XT14	12	19	0	40	0.46	0.66	0.39
XTL51	Quad mutant	XT14	10	8	0	47	0.86	**0.009**	0.39
HME69	polA resA1	XT14	5	16	2	37	0.31	1.00	0.64
XTL100	recBCD sbcC	XT14	7	12	0	31	1.00	0.87	0.62
XTL324	xni<>cat	XT14	6	16	1	38	0.43	0.88	1.0
SIMD90	λ Δexo	XT14	12	23	2	45	0.60	0.49	0.68
XTL372	polB<>spec	XT14	9	20	1	52	0.49	0.79	0.40
HME70	ΔrecA	XT14	6	16	3	44	0.26	0.66	0.22
HME68	Wild type	XT14/10³	2	6	1	13	0.75	0.61	0.34
XTL51	Quad mutant	XT14/10³	11	17	3	44	0.86	0.88	0.22
			122	220	19	523			

a. Statistical calculations for a Fisher's exact test were made assuming no difference between any of the strains and conditions tested, and therefore, results were compared for each strain to a data set including all the strains (totals shown beneath relevant columns). We omitted the value for the tested strain from the total in each case. P-values shown in bold indicate significant results.
b. Oligo concentration was reduced 1000-fold for XT13/10³ and XT14/10³.
c. 5′ loss and 3′ loss indicate the number of oligos sequenced that had any loss from the 5′ or 3′ end respectively. 'Other' represents the number of rare recombinants in which an internal marker was lost.
d. The total number of recombinants examined by sequence analysis.

lagging-strand. The main effect on recombination is destruction of the oligos with a resulting reduction of recombination frequency, but only at limiting oligo concentration. Since single-strand exonucleases play only a minor role for marker loss (3′ leading-strand), and since marker loss occurs primarily after pairing, other dsDNA nucleases may be responsible.

Replicative DNA polymerases. Entry of an oligo into the replication fork, paired either to the lagging- or leading-strand DNA template, might be expected to disrupt replication elongation. During lagging-strand synthesis, the DNA polymerase holoenzyme, Pol III, is released during discontinuous replication when it encounters the RNA primer on the 5′ end of an Okazaki fragment, leaving only a nick (Li and Marians, 2000; Leu *et al.*, 2003). The Pol I polymerase will then load and remove the RNA primer by nick translation filling the gap to allow ligation. In Fig. 1B, the oligo is shown downstream of the replicating Pol III polymerase. Thus, Pol III will encounter the annealed DNA oligo, which differs in structure from the normal RNA/DNA hybrid (Horton and Finzel, 1996). Pol III is likely to dissociate from the template, due to low strand displacement activity, when it encounters the oligo and reinitiate at the next RNA primer just as it would if it had encountered an Okazaki fragment (Benkovic *et al.*, 2001; McHenry, 2011).

Table 2A. Effect of single-strand exonucleases on oligo recombination.

Genotype[a]	Oligo[b]	Strand[c]	#[d]	Efficiency[e]	% Marker loss				
					None	5′	3′	Both	Other
+	6000	Lag	48	2.2×10^7	67	10	15	2	6
+	6	Lag	44	9.4×10^3	66	11	14	0	9
+	6000	Lead	45	6.7×10^5	27	13	44	7	9
+	6	Lead	14*	9.1×10^1	36	14	43	0	7
Quad	6000	Lag	48	1.2×10^7	56	11	25	0	8
Quad	6	Lag	46	1.8×10^5	31	4	54	4	7
Quad	6000	Lead	42	2.2×10^6	53	14	29	2	2
Quad	6	Lead	47	1.1×10^4	45	13	25	11	6

Table 2B. Effect of DNA polymerase and ligase mutations on oligo recombination.

Genotype[a]	Strand[c]	#[d]	Efficiency[e]	% Marker Loss				
				None	5′	3′	Both	Other
+	Lag	48	2.2×10^7	67	10	15	2	6
	Lead	45	6.7×10^5	27	13	44	7	9
*polA resA*1	Lag	42	1.3×10^7	74	9	17	0	0
	Lead	39	3.6×10^5	49	5	33	8	5
polA (5′→3′ exo)<>*cat*	Lag	47	2.4×10^6	79	0	19	0	2
	Lead	47	1.3×10^5	30	15	46	7	2
dnaQ<>*cat*	Lag	42	5.2×10^5	48	24	17	7	5
	Lead	42	1.6×10^4	24	31	24	19	2
polB<>*spec*	Lag	56	4.1×10^7	60	13	23	2	2
	Lead	53	5.9×10^5	51	9	30	8	2
[f] *lig7*^*ts*	Lag	45	1.9×10^5	24	53	14	7	2
	Lead	40	2.5×10^3	37	15	33	15	0

a. The '+' indicates strain HME68, whereas 'Quad' is XTL51, a derivative of HME68 deleted for four single-strand exonuclease genes. For the remaining strains, all derivatives of HME63 or HME68, the relevant genotype is indicated.
b. Molecules of oligo per cell electroporated. In Table 2B, 6000 molecules per cell were used.
c. Lagging-'lag' or leading-'lead' strand oligos are XTL13 and XTL14 respectively.
d. For each data set 39–56 Gal[+] isolates were analysed. In the case denoted by the asterisk a total of 48 Gal[+] isolates were sequenced, but 34 proved to be spontaneous revertants at the target TAG codon. The sequence data for experiments shown in Table 2A are in Figs 4A, S1K, 4B, S2B, S1L, S1M, S2C and S2D. The sequence data for experiments shown in Table 2B are in Figs 4A, 4B, S1N, S2E, 4C, 4D, 4E, 4F, S1S, S2I, 4G and 4H.
e. Efficiency was calculated as Gal[+]/10[8] viable cells. Recombination efficiencies given here are a representative experiment from at least three independent experiments where variability was less than fourfold. 95% confidence limits can be found in Table S4. Controls for DNA uptake are shown in Table S5.
f. Following recovery at 30°C, recombinants were selected at 34°C instead of 32°C.

If only a nick remained (Leu *et al.*, 2003), DNA ligase could directly join the newly synthesized strand to the oligo. However, if premature release of Pol III occurred as a result of the different structure encountered, a gap would remain and another polymerase, like Pol I, would be required to fill the gap (Li and Marians, 2000; Langston *et al.*, 2009).

During leading-strand synthesis, the holoenzyme replicates continuously as the DnaB helicase opens up the DNA duplex ahead of the replication complex (Kurth and O'Donnell, 2009). The presence of an oligo between polymerase and the helicase would be expected to impede the progression of polymerase, probably causing Pol III to release its template much as it does during discontinuous replication (Fig. 1C). Other polymerases may be required to carry out repair replication until the fork is reassembled. Thus, the exonuclease activities of the replicative polymerases might be expected to impact oligo incorporation and integrity. We examine below the effects on oligonucleotide recombination frequency and marker loss in various DNA polymerase mutants.

DNA polymerase I. Pol I is used during DNA replication of the lagging-strand (Setlow *et al.*, 1972) as well as for DNA repair (Sharon *et al.*, 1975), and thus, might be expected to affect oligo recombination. Pol I has multiple activities that include a 5′ → 3′ exonuclease involved in

A HME68

B HME69

C XTL70

D XTL85
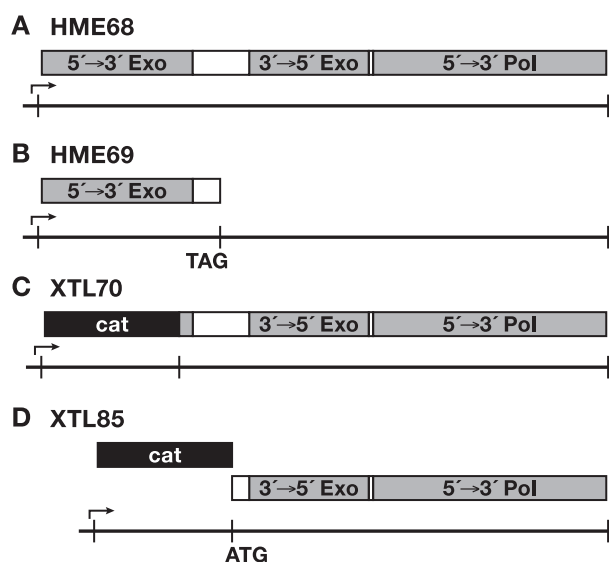
Fig. 6. Structure of the *polA* gene and mutants. This figure is approximately to scale.
A. The *polA* gene encodes the three functional domains of the Pol I protein, which are indicated by the shaded boxes: (1) a 5′ → 3′ exonuclease that mediates nick translation during lagging-strand synthesis; (2) a 3′ → 5′ exonuclease that mediates proofreading; (3) a DNA polymerase activity.
B. The *resA*1 amber mutation (TAG) is located between the two exonuclease domains.
C. The sequence that encodes the 5′ → 3′ exonuclease domain of Pol I is replaced by the *cat* open reading frame to create the *polA* (5′ → 3′ exo)◇*cat* mutation (XTL70).
D. The sequence that encodes the 5′ → 3′ exonuclease domain is replaced by the *cat* cassette with a stop codon followed by ATG, a translation restart (SD/AUG) for expression of the downstream sequences.

removal of RNA primers on Okazaki fragments, a 3′ → 5′ exonuclease proofreading activity conferring high fidelity, and a DNA polymerase activity (Fig. 6A). The *resA*1 amber mutation (Fig. 6B) results in a truncated protein lacking both 3′ → 5′ exonuclease and DNA polymerase activities but retaining 5′ → 3′ exonuclease activity (Vaccaro and Siegel, 1975; Kelley and Joyce, 1983). Recombination frequencies for the lagging- and leading-strand oligos are unaffected by the *resA*1 mutation (Table 2B). Among recombinants, the distribution of marker loss from the lagging- and leading-strand oligos in the *resA*1 strain was similar to wild type (Fig. 4A and B, Table 2B, Figs S1N and S2E).

The effect of the 5′ → 3′ exonuclease activity of Pol I was determined by deleting that portion of the *polA* coding sequence and replacing it with a chloramphenicol resistance (*cat*) open reading frame expressed from the native *polA* promoter (Fig. 6C). This substitution fused the chloramphenicol acetyltransferase protein to the Klenow fragment containing the 3′ → 5′ exonuclease and the polymerase domain. With this mutant Pol I, the recombi-

nation frequency is reduced ninefold for the lagging-strand and fivefold for the leading-strand oligos (Table 2B). Marker loss at the 5′ end of the lagging-strand oligo is completely eliminated (Tables 1A and 2B, Fig. 4A versus C). The marker loss for the leading-strand is unchanged relative to wild type (Table 2B, Fig. 4B versus D). To determine whether these effects are due to poor functioning of the Klenow fragment in the fusion construct, a second construct (XTL85) was made that contained the chloramphenicol cassette but with a stop codon as well as a strong Shine–Dalgarno sequence and a start codon for the downstream Klenow fragment of *polA* (Fig. 6D). Similar recombination frequencies and lagging-strand marker loss results were obtained with this mutant construct, which does not make a fusion protein (Fig. S1U). None of these mutants support replication of pBR322 (Table S5) as reported previously (Kogoma and Maldonado, 1997). We were able to generate a deletion mutant of the entire coding sequence of *polA*, however, growth was so poor that it could not be propagated in culture, in agreement with a previous result (Baba *et al.*, 2006) and in contrast to another (Joyce and Grindley, 1984).

Our results demonstrate that the 5′ → 3′ exonuclease activity of Pol I affects both oligo recombination frequency and marker loss. This is most apparent with the lagging-strand oligo, and implicates the 5′ → 3′ exonuclease activity in marker loss from the 5′ end of that oligo. We propose that during a recombination event with a lagging-strand oligo annealed at the replication fork, Pol III is replaced by Pol I and the 5′ → 3′ exonuclease activity of Pol I may remove part of the 5′ end of the oligo via nick translation during Pol I replication from the 3′ end of the upstream Okazaki fragment (Fig. 1B). Although the Pol I nick translation activity does not affect 5′ marker loss on the leading-strand, it has a modest effect on the recombination frequency (Tables 2B and S4). DNA Polymerase III molecules replicate the leading- and lagging-strands co-ordinately as a complex (Maki *et al.*, 1988) while Pol I is thought to act primarily on the lagging-strand (Setlow *et al.*, 1972). Without the Pol I 5′ → 3′ exonuclease activity to remove the RNA primers on Okazaki fragments, synthesis of the lagging-strand may be slowed, thereby perturbing the replication complex and consequently leading-strand synthesis. Thus, recombination frequencies of both the leading- and lagging-strands may be affected in this mutant. Our recombination results are also consistent with the Pol I 5′ → 3′ exonuclease activity directly aiding a discontinuous type of replication on the leading-strand; results described below and by others (Wang, 2005; Langston *et al.*, 2009; Lia *et al.*, 2012) suggest that such an interaction might occur.

We note that 5′ marker loss from the lagging-strand still occurs in a *polA resA*1 mutant lacking the Klenow fragment (Fig. S1N). This implies that the remaining portion of

the protein, the 5′ → 3′ exonuclease, still localizes correctly either by itself or associated with another protein such as Pol III.

DNA polymerase III. Pol III is the major replicative DNA polymerase in *E. coli*. This polymerase forms a dimer or trimer with the leading- and lagging-strands each being replicated by one of the two active polymerase units with a possible third held in reserve (Maki *et al.*, 1988; McInerney *et al.*, 2007; Reyes-Lamothe *et al.*, 2010; Lia *et al.*, 2012). Among the different protein subunits composing the holoenzyme are the 3′ → 5′ proof-reading exonucleases encoded by the *dnaQ* gene (Echols *et al.*, 1983). When we replaced the *dnaQ* gene with a *cat* cassette, there was a greater than 40-fold reduction in recombination with both the lagging- and leading-strand oligos (Tables 2B and S4) in contrast to a small effect when recombination was assayed on a plasmid (Huen *et al.*, 2006). Thus, Pol III integrity is important for efficient oligo recombination on both the lagging- and leading-strands, indicating that the DNA replication machinery is critical for oligo recombination.

Among recombinants isolated from the *dnaQ* mutant, marker loss is altered for both strands. For the lagging-strand oligo, marker loss from the 5′ end increased compared with wild type, but the 3′ end of the oligo was unaffected (Tables 1A and 2B; Fig. 4A and E). The latter result is probably a consequence of the oligo being positioned between Pol III and the downstream Okazaki fragment (Fig. 1B) with the oligo 3′ end shielded from Pol III. Pol I would be expected to fill in the gap between the oligo 3′ end and the downstream Okazaki fragment irrespective of the *dnaQ* allele. The low level of processing at the 5′ end in wild type (10%) and the increased processing of the 5′ end in the *dnaQ* mutant (24%) can be explained as follows. If the wild type Pol III usually replicates up to the oligo before falling off, a nick directly repairable by DNA ligase would be left in the DNA and no marker loss would occur. The data suggest that Pol III dissociates prematurely at least 10% of the time, leaving a gap for Pol I entry and removal of 5′ markers from the oligo (Table 2B, Fig. 4A). In the *dnaQ* mutant, the less processive Pol III enzyme dissociates more often (Studwell and O'Donnell, 1990), leaving a gap between the newly synthesized strand and the 5′ end of the oligo at least 24% of the time (Table 2B, Fig. 4E).

Among leading-strand recombinants isolated from the *dnaQ* mutant, marker loss from the 5′ end increased substantially relative to wild type (Tables 1B and 2B, Fig. 4B and F). In a mechanism similar to that proposed for the lagging-strand oligo, Pol I may act to fill in the gap between the prematurely terminated leading-strand and the oligo, explaining the increase in 5′ marker loss in the *dnaQ* mutant (Fig. 1C). We tried to test the role of Pol I by

generating a double mutant defective for *dnaQ* and lacking the 5′ → 3′ exonuclease of Pol I; however, we could not propagate the strain. It is also possible that an unknown 5′ → 3′ exonuclease is involved in the degradation of the 5′ end of the leading-strand oligo.

DNA polymerase II. The SOS-inducible repair polymerase Pol II also has a 3′ → 5′ exonuclease proof-reading activity (Langston *et al.*, 2009). At the replication fork, Pol II has been shown to replace Pol III and correct mistakes on either the leading- or lagging-strand (Banach-Orlowska *et al.*, 2005). We replaced the *polB* gene, encoding Pol II, with a spectinomycin cassette. This mutation had no effect on recombination frequency, and although marker loss among recombinants was slightly reduced on the 3′ end of the leading-strand oligo (from 44% to 30%), this decrease was not significant (Tables 1 and 2B, Figs 4A, S1S, 4B, S2I).

In summation, we have determined that the 5′ → 3′ exonuclease activity of Pol I is responsible for all marker loss observed on the 5′ end of lagging-strand oligos. Although we see some reduction in 3′ marker loss when the other polymerases are mutated, we never eliminated 3′ marker loss on the leading- or lagging-strands and redundant functions may be responsible. In order to look at potential redundant functions, we tried to construct double mutants such as *dnaQ polB* and *dnaQ polA*; however, these double mutants were inviable. It is also possible some other 3′ → 5′ exonuclease is responsible as diagrammed in Fig. 1C.

DNA ligase. We expect that DNA ligase is necessary to connect the annealed oligo with the growing DNA, either leading- or lagging-strand. As DNA ligase is an essential function, a temperature-sensitive mutation, *lig7ts* (Gottesman *et al.*, 1973) was used to examine the effect of ligase deficiency on recombination. Indeed, when the *lig7ts* cells are incubated at the semi-restrictive temperature of 34°C after electroporation, the recombination frequency is reduced 100-fold for the lagging-strand and 260-fold for the leading-strand oligos, demonstrating that ligase activity is necessary for oligo recombination (Tables 2B and S4, Fig. 4G and H). Five times as many lagging-strand recombinants experienced marker loss on the 5′ end of the oligo and degradation from this end was more extensive under these semi-restrictive conditions (Tables 1 and 2B, Fig. 4A versus G). The straightforward interpretation of these data is that when ligase activity is limiting, the annealed lagging-strand oligo is not promptly joined to the upstream Okazaki fragment, leaving it more vulnerable to exonuclease digestion at the 5′ end. We were unable to test whether the observed marker loss was due to the Pol I 5′ → 3′ exonuclease, as the *lig7ts polA* (5′ → 3′ exo)<>*cat* double mutant could not be made.

Other recombinases and dsDNA exonucleases. Several other genes encoding recombinases and/or dsDNA exonucleases showed little effect on recombination frequencies or marker loss distribution (Table 1) where tested. These include *recA*, *recBCD sbcC*, *xthA* (Exo III) and *xni* (Exo IX) (Figs S1T, S2J, S1P, S2F, S1O, S1Q and S2G). We also asked whether phage λ Exo influenced recombination frequencies or marker loss but saw no effect in a strain expressing only Beta (Datta *et al.*, 2008) (Figs S1R and S2H, Table S3).

Conclusions

In the current study, we analysed phage λ Red-mediated recombination of synthetic oligos carrying a subset of mispaired bases (markers). The spacing, position, and marker density were all important factors in determining both recombination efficiency and marker loss. Homology on the 5′ end was especially important for efficient recombination. This is likely because the 3′ end can be extended by polymerase(s) and ligated, whereas mismatches at the 5′ end would be difficult to ligate to upstream DNA (Sawitzke *et al.*, 2011). In addition, mismatches on the 5′ end might preferentially be removed by nick translation and/or a flap endonuclease. Several well-studied dsDNA exonucleases including RecBCD, Exo III, SbcCD and λ Exo had little effect on oligo recombination or marker loss distributions, nor did the homologous recombination enzyme RecA. In contrast, the replicative DNA polymerases, Pol I and Pol III, had major effects on recombination and marker loss. A ligase conditional mutation decreased the recombination efficiency substantially for both strands and increased the oligo marker loss on the lagging-strand. These data present a more complete picture of what occurs during λ Red-mediated oligo recombination by demonstrating that replicative polymerases process the recombination intermediates. The experiments also illustrate a novel method of studying DNA replication and repair at the replication fork.

Experimental procedures

Bacterial strains and oligos

Bacterial strains used in this study are listed in Table S1. Gene replacements were made by dsDNA recombineering (Yu *et al.*, 2000); construction details including oligo sequences will be supplied upon request. The sequences for oligos used in recombination experiments are listed in Table S2. Multiple mutants were made by either recombineering or P1 transduction (Thomason *et al.*, 2007b). All strains used in the experiments were deleted for *mutS*. Oligos were purchased from Integrated DNA Technologies (IDT) and were desalted but not further purified.

Recombineering

Recombineering was performed as described (see: http://redrecombineering.ncifcrf.gov/; Sawitzke *et al.*, 2007; Thomason *et al.*, 2007c). Overnight cultures were grown from single colonies, diluted ~ 80-fold in L broth and grown until OD_{600} was ~ 0.4. However, as *dnaQ mutS* cells are extremely mutagenic, cells were taken directly from a frozen glycerol stock and grown to OD_{600} ~ 0.4 without growing an overnight. The *dnaQ* glycerol was prepared by picking hundreds of the smallest colonies, avoiding the many large colonies (suppressors) from a L plate incubated overnight at 30°. The unstable nature of this genotype is reflected in the large confidence interval seen in Table S4.

In strain construction of gene replacements by drug markers, there was a 2 h recovery, after which cells were plated on appropriate antibiotic plates (30 μg ml⁻¹ ampicillin, 10 μg ml⁻¹ chloramphenicol, 12.5 μg ml⁻¹ tetracycline, 30 μg ml⁻¹ kanamycin, 30 μg ml⁻¹ spectinomycin). For *galK_{am}* gene correction, Gal⁺ recombinants were selected on minimal M63 agar containing galactose (0.2%) and biotin (0.01%) (Ellis *et al.*, 2001; Costantino and Court, 2003) after a 30 min recovery.

Control for DNA uptake

For strains in which oligo recombination defects were observed, efficiency of DNA uptake was tested in two ways. Cells were transformed with either pBR322 (New England Biolabs) or mini-λ Spec (Court *et al.*, 2003). The mini-λ transformation was done under the same conditions as the recombineering experiments to allow expression of Int. The mini-λ has no origin of replication and can only be maintained after integration into the *attB* site on the bacterial chromosome using Int. Results are shown in Table S5.

Sequencing and analysis

Sequencing was done by SAIC-Frederick, Inc. Sequencing results were analysed with Sequencher version 4.8.

References

Allen, L.M., Hodskinson, M.R., and Sayers, J.R. (2009) Active site substitutions delineate distinct classes of eubacterial flap endonuclease. *Biochem J* **418:** 285–292.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., *et al.* (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2:** 2006–0008.

Banach-Orlowska, M., Fijalkowska, I.J., Schaaper, R.M., and Jonczyk, P. (2005) DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. *Mol Microbiol* **58:** 61–70.

Benkovic, S.J., Valentine, A.M., and Salinas, F. (2001) Replisome-mediated DNA replication. *Annu Rev Biochem* **70:** 181–208.

Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S.T., and Modrich, P. (2001) *In vivo* requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc Natl Acad Sci USA* **98:** 6765–6770.

Costantino, N., and Court, D.L. (2003) Enhanced levels of λ Red-mediated recombinants in mismatch repair mutants. *Proc Natl Acad Sci USA* **100:** 15748–15753.

Court, D.L., Sawitzke, J.A., and Thomason, L.C. (2002) Genetic engineering using homologous recombination. *Annu Rev Genet* **36:** 361–388.

Court, D.L., Swaminathan, S., Yu, D., Wilson, H., Baker, T., Bubunenko, M., *et al.* (2003) Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene* **315:** 63–69.

Datta, S., Costantino, N., Zhou, X., and Court, D.L. (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc Natl Acad Sci USA* **105:** 1626–1631.

Dutra, B.E., Sutera, V.A., and Lovett, S.T. (2007) RecA-independent recombination is efficient but limited by exonucleases. *Proc Natl Acad Sci USA* **104:** 216–221.

Echols, H., Lu, C., and Burgers, P.M. (1983) Mutator strains of *Escherichia coli*, *mutD* and *dnaQ*, with defective exonucleolytic editing by DNA polymerase III holoenzyme. *Proc Natl Acad Sci USA* **80:** 2189–2192.

Ellis, H.M., Yu, D., DiTizio, T., and Court, D.L. (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci USA* **98:** 6742–6746.

Goldmark, P.J., and Linn, S. (1972) Purification and properties of the *recBC* DNase of *Escherichia coli* K-12. *J Biol Chem* **247:** 1849–1860.

Gottesman, M.M., Hicks, M.L., and Gellert, M. (1973) Genetics and function of DNA ligase in *Escherichia coli*. *J Mol Biol* **77:** 531–547.

Horton, N.C., and Finzel, B.C. (1996) The structure of an RNA/DNA hybrid: a substrate of the ribonuclease activity of HIV-1 reverse transcriptase. *J Mol Biol* **264:** 521–533.

Huen, M.S., Li, X.T., Lu, L.Y., Watt, R.M., Liu, D., and Huang, J.D. (2006) The involvement of replication in single stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res* **34:** 6183–6194.

Joyce, C.M., and Grindley, N.D. (1984) Method for determining whether a gene of *Escherichia coli* is essential: application to the *polA* gene. *J Bacteriol* **158:** 636–643.

Karakousis, G., Ye, N., Li, Z., Chiu, S.K., Reddy, G., and Radding, C.M. (1998) The Beta protein of phage λ binds preferentially to an intermediate in DNA renaturation. *J Mol Biol* **276:** 721–731.

Kelley, W.S., and Joyce, C.M. (1983) Genetic characterization of early amber mutations in the *Escherichia coli polA* gene and purification of the amber peptides. *J Mol Biol* **164:** 529–560.

van Kessel, J.C., and Hatfull, G.F. (2008) Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Mol Microbiol* **67:** 1094–1107.

Kmiec, E., and Holloman, W.K. (1981) β protein of bacteriophage λ promotes renaturation of DNA. *J Biol Chem* **256:** 12636–12639.

Kogoma, T., and Maldonado, R.R. (1997) DNA polymerase I in constitutive stable DNA replication in *Escherichia coli*. *J Bacteriol* **179:** 2109–2115.

Kornberg, A. (1990) The private life of DNA polymerase I. *Methods Enzymol* **182:** 783–788.

Kurth, I., and O'Donnell, M. (2009) Replisome dynamics during chromosome duplication. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Böck, A., Curtiss, R., III, Kaper, J.B., Karp, P.D., Neidhardt, F.C., Nyström, T., *et al.* (eds). Washington, DC: ASM Press, pp. 1–47.

Lajoie, M.J., Gregg, C.J., Mosberg, J.A., Washington, G.C., and Church, G.M. (2012) Manipulating replisome dynamics to enhance lambda Red-mediated multiplex genome engineering. *Nucleic Acids Res* **40:** e170.

Langston, L.D., Indiani, C., and O'Donnell, M. (2009) Whither the replisome: emerging perspectives on the dynamic nature of the DNA replication machinery. *Cell Cycle* **8:** 2686–2691.

Leu, F.P., Georgescu, R., and O'Donnell, M. (2003) Mechanism of the *E. coli* τ processivity switch during lagging-strand synthesis. *Mol Cell* **11:** 315–327.

Li, X., and Marians, K.J. (2000) Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. *J Biol Chem* **275:** 34757–34765.

Li, X.T., Costantino, N., Lu, L.Y., Liu, D., Watt, R.M., Cheah, K.S., *et al.* (2003) Identification of factors influencing strand bias in oligonucleotide-mediated recombination in *Escherichia coli*. *Nucleic Acids Res* **31:** 6674–6687.

Lia, G., Michel, B., and Allemand, J.F. (2012) Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science* **335:** 328–331.

Little, J.W., Lehman, I.R., and Kaiser, A.D. (1967) An exonuclease induced by bacteriophage λ. I. Preparation of the crystalline enzyme. *J Biol Chem* **242:** 672–678.

McHenry, C.S. (2011) DNA replicases from a bacterial perspective. *Annu Rev Biochem* **80:** 403–436.

McInerney, P., Johnson, A., Katz, F., and O'Donnell, M. (2007) Characterization of a triple DNA polymerase replisome. *Mol Cell* **27:** 527–538.

Maki, H., Maki, S., and Kornberg, A. (1988) DNA polymerase III holoenzyme of *Escherichia coli*. IV. The holoenzyme is an asymmetric dimer with twin active sites. *J Biol Chem* **263:** 6570–6578.

Martinsohn, J.T., Radman, M., and Petit, M.A. (2008) The λ Red proteins promote efficient recombination between

diverged sequences: implications for bacteriophage genome mosaicism. *PLoS Genet* **4:** e1000065.

Mosberg, J.A., Gregg, C.J., Lajoie, M.J., Wang, H.H., and Church, G.M. (2012) Improving lambda red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. *PLoS ONE* **7:** e44638.

Muniyappa, K., and Radding, C.M. (1986) The homologous recombination system of phage λ. Pairing activities of β protein. *J Biol Chem* **261:** 7472–7478.

Murphy, K.C. (1998) Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol* **180:** 2063–2071.

Noirot, P., and Kolodner, R.D. (1998) DNA strand invasion promoted by *Escherichia coli* RecT protein. *J Biol Chem* **273:** 12274–12280.

Reyes-Lamothe, R., Sherratt, D.J., and Leake, M.C. (2010) Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328:** 498–501.

Richardson, C.C., Lehman, I.R., and Kornberg, A. (1964) A deoxyribonucleic acid phosphatase-exonuclease from *Escherichia coli*. II. Characterization of the exonuclease activity. *J Biol Chem* **239:** 251–258.

Rigby, P.W., Dieckmann, M., Rhodes, C., and Berg, P. (1977) Labeling deoxyribonucleic acid to high specific Activity *in vitro* by nick translation with DNA polymerase I. *J Mol Biol* **113:** 237–251.

Rybalchenko, N., Golub, E.I., Bi, B., and Radding, C.M. (2004) Strand invasion promoted by recombination protein β of coliphage λ. *Proc Natl Acad Sci USA* **101:** 17056–17060.

Sawitzke, J.A., Thomason, L.C., Costantino, N., Bubunenko, M., Datta, S., and Court, D.L. (2007) Recombineering: *in vivo* genetic engineering in *E. coli*, *S. enterica*, and beyond. *Methods Enzymol* **421:** 171–199.

Sawitzke, J.A., Costantino, N., Li, X.T., Thomason, L.C., Bubunenko, M., Court, C., and Court, D.L. (2011) Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J Mol Biol* **407:** 45–59.

Sergueev, K., Court, D., Reaves, L., and Austin, S. (2002) *E. coli* cell-cycle regulation by bacteriophage lambda. *J Mol Biol* **324:** 297–307.

Setlow, P., Brutlag, D., and Kornberg, A. (1972) Deoxyribonucleic acid polymerase: two distinct enzymes in one polypeptide. I. A proteolytic fragment containing the polymerase and 3′→5′ exonuclease functions. *J Biol Chem* **247:** 224–231.

Sharan, S.K., Thomason, L.C., Kuznetsov, S.G., and Court, D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* **4:** 206–223.

Sharon, R., Miller, C., and Ben-Ishai, R. (1975) Two modes of excision repair in toluene-treated *Escherichia coli*. *J Bacteriol* **123:** 1107–1114.

Studwell, P.S., and O'Donnell, M. (1990) Processive replication is contingent on the exonuclease subunit of DNA polymerase III holoenzyme. *J Biol Chem* **265:** 1171–1178.

Swaminathan, S., Ellis, H.M., Waters, L.S., Yu, D., Lee, E.C., Court, D.L., and Sharan, S.K. (2001) Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* **29:** 14–21.

Swingle, B., Markel, E., Costantino, N., Bubunenko, M.G., Cartinhour, S., and Court, D.L. (2010) Oligonucleotide recombination in Gram-negative bacteria. *Mol Microbiol* **75:** 138–148.

Thomason, L.C., Costantino, N., Shaw, D.V., and Court, D.L. (2007a) Multicopy plasmid modification with phage λ Red recombineering. *Plasmid* **58:** 148–158.

Thomason, L.C., Costantino, N., and Court, D.L. (2007b) *E. coli* genome manipulation by P1 transduction. *Curr Protoc Mol Biol* **Chapter 1:** Unit 1 17.

Thomason, L., Court, D.L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., and Oppenheim, A. (2007c) Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **Chapter 1:** Unit 1 16.

Vaccaro, K.K., and Siegel, E.C. (1975) Increased spontaneous reversion of certain frameshift mutations in DNA polymerase I deficient strains of *Escherichia coli*. *Mol Gen Genet* **141:** 251–262.

Van Pijkeren, J.P., Neoh, K.M., Sirias, D., Findley, A.S., and Britton, R.A. (2012) Exploring optimization parameters to increase ssDNA recombineering in *Lactococcus lactis* and *Lactobacillus reuteri*. *Bioengineered* **3:** 209–217.

Viswanathan, M., and Lovett, S.T. (1999) Exonuclease X of *Escherichia coli*. A novel 3′-5′ DNase and DnaQ superfamily member involved in DNA repair. *J Biol Chem* **274:** 30094–30100.

Wang, H.H., Xu, G., Vonner, A.J., and Church, G. (2011) Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. *Nucleic Acids Res* **39:** 7336–7347.

Wang, T.C. (2005) Discontinuous or semi-discontinuous DNA replication in *Escherichia coli*? *Bioessays* **27:** 633–636.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* **97:** 5978–5983.

Zhang, Y., Buchholz, F., Muyrers, J., and Stewart, A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* **20:** 123–128.

Zhang, Y., Muyrers, J., Rientjes, J., and Stewart, A.F. (2003) Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol Biol* **4:** 1.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

A. HME68

Lagging-strand oligo XT36

Efficiency: 1.7×10^6

5′ ———————————→ 3′

Position:	3	9	15	21	27	33	+	42	48	54	60	66	72
Host:	G	C	C	C	G	G	G	T	G	C	A	C	T
Oligo:	T	G	A	A	A	A	T	C	A	T	G	A	C

B. HME68

Lagging-strand oligo XT21

Efficiency: $<1.3 \times 10^1$

5′ ———————————→ 3′

Position:	3	6	9	12	15	18	21	24	27	30	33	36	37	38	+	42	45	48	51	54	57	60	63	66	69	72
Host:	G	A	C	G	C	A	C	A	G	G	G	T	T	A	G	T	G	G	G	C	C	A	A	C	G	T
Oligo:	T	G	G	C	A	C	A	C	A	A	A	C	T	A	T	C	A	A	A	T	A	G	G	A	C	C

Supplemental Figure 1

C. HME68
Lagging-strand oligo XT18
Efficiency: 1.3×10^6

5′ ——————————————→ 3′

Position:	15	18	21	24	27	30	33	36	+	42	45	48	51	54	57	60	63
Host:	C	A	C	A	G	G	G	T	G	T	G	G	G	C	C	A	A
Oligo:	A	C	A	C	A	A	A	C	T	C	A	A	A	T	A	G	G

(grid of 30 rows, each: A C A C A A A C T C A A A T A G G; row markers at 10, 20, 30)

D. HME68
Lagging-strand oligo XT524
Efficiency: 3.7×10^6

5′ ——————————————→ 3′

Position:	24	27	30	33	36	39	42	45	48	51	54	57	+	63	66	69	72	75	78	81	84	87	90	93
Host:	G	A	C	G	C	A	C	A	G	G	G	T	G	T	G	G	G	C	C	A	A	C	G	T
Oligo:	T	G	G	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C

(grid of ~52 rows, each: T G G C A C A C A A A C T C A A A T A G G A C C; row markers at 10, 20, 30, 40, 50)

Supplemental Figure 1

E. HME68
Lagging-strand oligo XT30
Efficiency: 5.1 × 10^6

5′ —————————→ 3′

Position:	+	42	45	48	51	54	57	60	63	66	69	72
Host:	G	T	G	G	G	C	C	A	A	C	G	T
Oligo:	T	C	A	A	A	T	A	G	G	A	C	C

	T	C	A	A	A	T	A	G	A	C	G	T
	T	C	A	A	A	T	C	A	A	C	G	T
	T	C	A	A	A	T	A	G	A	C	G	T
	T	C	A	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
10	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
20	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
30	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
40	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T
	T	T	G	G	G	C	C	A	A	C	G	T

F. HME68
Lagging-strand oligo XT29
Efficiency: 2.4 × 10^3

5′ —————————→ 3′

Position:	3	6	9	12	15	18	21	24	27	30	33	36	+
Host:	G	A	C	G	C	A	C	A	G	G	G	T	G
Oligo:	T	G	G	C	A	C	A	C	A	A	A	C	T

	G	A	C	C	A	C	C	C	A	A	A	C	T
	G	A	C	C	A	C	C	C	A	A	A	C	T
	G	A	C	G	A	C	C	C	A	A	A	C	T
	G	A	C	G	C	A	C	C	A	A	A	C	T
	G	A	C	G	C	A	C	A	A	A	A	C	T
	G	A	C	G	C	A	C	A	G	G	A	C	T
	G	A	C	G	C	A	C	A	G	G	A	C	T
	G	A	C	G	C	A	C	A	G	G	A	C	T
	G	A	C	G	C	A	C	A	G	G	G	C	T
10	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
20	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
30	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
	G	A	C	G	C	A	C	A	G	G	G	T	T
40	G	A	C	G	C	A	C	A	G	G	G	T	T

Supplemental Figure 1

G. HME68
Lagging-strand oligo XT351
Efficiency: 6.1×10^1

5′ ——————————————→ 3′

Position:	3	6	9	12	15	18	21	24	27	30	33	36	+	42	45	48	51	54
Host:	G	A	C	C	C	A	C	A	G	G	G	T	G	T	G	G	G	C
Oligo:	T	G	G	G	A	C	A	C	A	A	A	C	T	C	A	A	A	T

G	G	G	G	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	G	G	G	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	G	G	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	A	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	C	C	A	C	A	A	A	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	A	A	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	A	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	G	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	G	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	G	C	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	G	T	T	C	A	A	A	T
G	A	C	C	C	A	C	A	G	G	G	T	T	C	A	A	A	T

10

H. HME68
Lagging-strand oligo XT38
Efficiency: 6.3×10^4

5′ ——————————————→ 3′

Position:	21	24	27	30	33	36	+	42	45	48	51	54	57	60	63	66	69	72
Host:	C	A	G	G	G	T	G	T	G	G	G	C	C	A	A	C	G	T
Oligo:	A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C

A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C	11%
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	C	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	C	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	A	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	G	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	89%
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	G	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	A	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	T	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	A	A	C	C	A	A	C	G	T	
A	C	A	A	A	C	T	C	A	G	G	C	C	A	A	C	G	T	

10
20
30
40

Supplemental Figure 1

Supplemental Figure 1

L. XTL51 (Quad mutant)
Lagging-strand oligo XT13
Efficiency: 1.2×10^7

M. XTL51 (Quad mutant)
1/1000 Lagging-strand oligo XT13
Efficiency: 1.8×10^5

N. HME69 (*polA resA1*)
Lagging-strand oligo XT13
Efficiency: 1.3×10^7

Supplemental Figure 1

O. XTL92 (*xthA<>tet*)
Lagging-strand oligo XT13
Efficiency: 4.4 × 10^7

5′ ⟶ 3′
Position: 15 21 30 + 48 54 63
Host: C C G G G C A
Oligo: A G A T C T G

57%

3′ 19%

5′ 13%

5′ & 3′ 2%

Other 9%

P. XTL100 (*recBCD<>kan sbcC<>amp*)
Lagging-strand oligo XT13
Efficiency: 1.3 × 10^7

5′ ⟶ 3′
Position: 15 21 30 + 48 54 63
Host: C C G G G C A
Oligo: A G A T C T G

68%

3′ 18%

5′ 8%

5′ & 3′ 3%

Other 3%

Supplemental Figure 1

Q. XTL324 (*xni<>cat*)
Lagging strand oligo XT13
Efficiency: 2.6×10^7

R. SIMD90 (λ Δexo)
Lagging strand oligo XT13
Efficiency: 1.8×10^7

S. XTL372 (*polB<>spec*)
Lagging-strand oligo XT13
Efficiency: 4.1×10^7

Supplemental Figure 1

T. HME70 (ΔrecA)
Lagging-strand oligo XT13
Efficiency: 2.8 × 10^7

U. XTL85 (polA 5´-3´ exo<>cat non-fusion)
Lagging-strand oligo XT13
Efficiency: 4.3 × 10^6

Supplemental Figure 1

A. HME68
Leading-strand oligo XT418
Efficiency: 1.0 × 10⁶
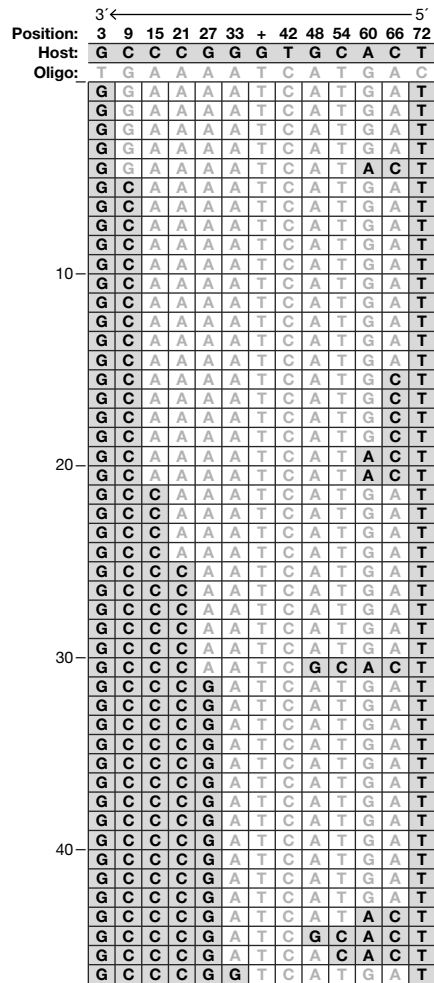
B. HME68
Leading-strand oligo 1/1000 XT14
Efficiency: 9.1 × 10¹
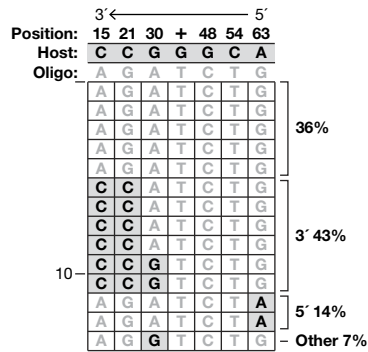
C. XTL51 (*Quad mutant*)
Leading-strand oligo XT14
Efficiency: 1.5 × 10⁶

Supplemental Figure 2
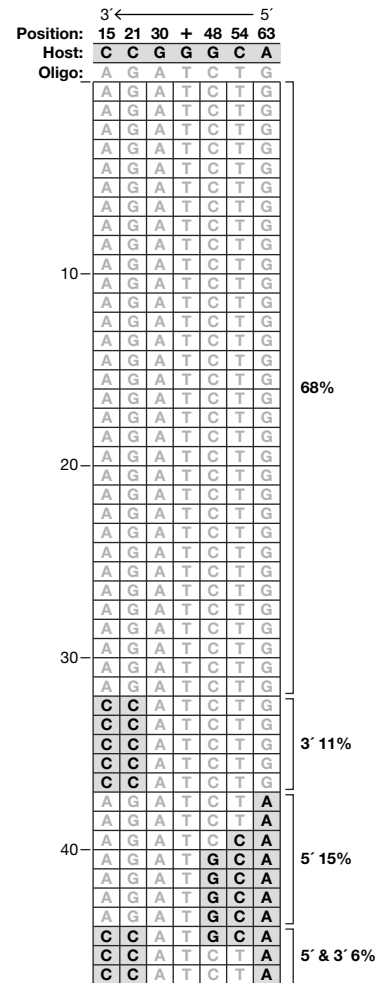
D. XTL51 (*Quad mutant*)
Leading-strand oligo 1/1000 XT14
Efficiency: 1.1×10^4

E. HME69 (*polA resA1*)
Leading-strand oligo XT14
Efficiency: 3.6×10^5

F. XTL100 (*recBCD<>kan sbcC<>amp*)
Leading-strand oligo XT14
Efficiency: 1.6×10^5

Supplemental Figure 2

G. XTL324 (*xni<>cat*)
Leading-strand oligo XT14
Efficiency: 2.5×10^4

3´ ←——————→ 5´
Position: 15 21 30 + 48 54 63
Host: C C G G G C A
Oligo: A G A T C T G

44%

3´ 39%

5´ 13%

5´ & 3´ 2%
Other 2%

H. SIMD90 (λ Δ*exo*)
Leading-strand oligo XT14
Efficiency: 2.2×10^6

3´ ←——————→ 5´
Position: 15 21 30 + 48 54 63
Host: C C G G G C A
Oligo: A G A T C T G

32%

3´ 38%

5´ 15%

5´ & 3´ 11%

Other 4%

I. XTL372 (*polB<>spec*)
Leading-strand oligo XT14
Efficiency: 5.9×10^5

3´ ←——————→ 5´
Position: 15 21 30 + 48 54 63
Host: C C G G G C A
Oligo: A G A T C T G

51%

3´ 30%

5´ 9%

5´ & 3´ 8%

Other 2%

Supplemental Figure 2

J. HME70 (Δ*recA*)

Leading-strand oligo XT14

Efficiency: 8.4 × 10⁵

Efficiency: 8.4×10^5

3′ ⟵⟶ 5′

Position:	15	21	30	+	48	54	63
Host:	C	C	G	G	G	C	A
Oligo:	A	G	A	T	C	T	G

A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	**47%**
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	
A	G	A	T	C	T	G	

(positions 10, 20 marked on left axis)

C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	**3′ 34%**
C	C	A	T	C	T	G	
C	C	A	T	C	T	G	
C	C	G	T	C	T	G	
C	C	G	T	C	T	G	
C	C	G	T	C	T	G	
C	C	G	T	C	T	G	
C	C	G	T	C	T	G	
C	C	G	T	C	T	G	

(position 30 marked on left axis)

A	G	A	T	C	T	A	
A	G	A	T	C	T	A	
A	G	A	T	C	T	A	**5′ 13%**
A	G	A	T	C	T	A	
A	G	A	T	G	C	A	
A	G	A	T	G	C	A	

(position 40 marked on left axis)

A	G	A	T	G	C	G	
A	C	G	T	C	T	G	**Other 6%**
C	C	A	T	G	T	G	

Supplemental Figure 2

Bacterial DNA Polymerases Participate in Oligonucleotide Recombination

Supplementary Information

Xin-tian Li[1]*, Lynn C. Thomason[1,2]*, James A. Sawitzke[1]*, Nina Costantino[1], and Donald L. Court[1.]

[1] Molecular Control and Genetics Section, Gene Regulation and Chromosome Biology, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

[2] Basic Science Program, SAIC-Frederick, Inc., Frederick MD 21702, USA

* These authors contributed equally to this manuscript.

Figure S1: Marker loss pattern as affected by host functions and the position of markers in the lagging-strand oligos

The efficiency of recombination for each lagging-strand oligo is shown at the top of each panel. In each panel, the "Position" row indicates the distance of the markers from the left end of the oligo diagrammed, with the "+" denoting the selected marker at position 39, which is the G of the TAG amber codon in the "Host" sequence (with the exception of panel D, oligo XT524, in which the "+" is at position 60). The "Oligo" row shows the sequence changes present on the oligo at the indicated position. The sequences of Gal⁺ recombinants are shown below the "Oligo" row; grey shaded spaces indicate those bases that remained unchanged from the host and show where markers were lost from the oligo, white spaces indicate those markers that were inherited

from the oligo. For each panel, recombinants are grouped according to their pattern of marker loss. The uppermost group of recombinants showed no marker loss. Groups with 3', 5', or 5'&3' marker loss are indicated. The final group, "Other", comprises recombinants with internal markers lost. Note that of all the panels, only the oligos used for E, F, and I have all their markers on one side or the other of the selected base. In panels K and M, the oligo concentration was 1000-fold less than in all other panels.

In panel B with oligo XT21, all three bases comprising the amber codon are included and left white so that changes found in that codon can be seen. Note that three isolates contain the TAT sequence at the selected codon. All of these Gal$^+$ colonies are unlikely to be true recombinants because no other markers from the oligo were inherited, and they occurred at the same frequency as spontaneous revertants.

In some experiments, either no recombination occurred (panel B), or all markers were inherited (panels C and D), or terminal markers were always lost (panels A, E, F, G, I and J). In these cases no percentages are shown. For each panel, the relevant genotype examined and oligos used are as follows. A: "wild type" (HME68), oligo XT36 (data used for Figure 5); B: wild type, oligo XT21; C: wild type, oligo XT18; D: wild type, oligo XT524; E: wild type, oligo XT30; F: wild type, oligo XT29; G: wild type, oligo XT351; H: wild type, oligo XT38; I: wild type, oligo XT37; J: wild type, oligo XT352; K: wild type, oligo XT13 1/1000 dilution; L: (Quad) *recJ<>amp xonA<>kan xseA<>tet exoX<>spec*, oligo XT13; M: (Quad) *recJ<>amp xonA<>kan xseA<>tet exoX<>spec*, oligo XT13 1/1000 dilution; N: *polA resA*1, oligo XT13; O: *xthA<>tet*, oligo XT13; P: *recBCD<>kan sbcC<>amp*, oligo XT13; Q: *xni<>cat*, oligo XT13; R: λ Δ*exo*, oligo XT13; S: *polB<>spec,* oligo XT13; T: Δ*recA*, oligo XT13. U: *polA(5'→3'exo)<>cat* non-fusion, oligo XT13

Figure S2: Marker loss pattern as affected by host functions and the position of markers in the lagging-strand oligos

The efficiency of recombination for each leading-strand oligo is shown at the top of each panel. See Figure S1 legend for details defining panel features. For each panel, the genotype examined and oligos used are as follows. A: wild type, oligo XT418 (data used for Figure 5); B: wild type, oligo XT14 1/1000 dilution; C: (Quad) *recJ<>amp xonA<>kan xseA<>tet exoX<>spec*, oligo XT14; D: (Quad) *recJ<>amp xonA<>kan xseA<>tet exoX<>spec*, oligo XT14 1/1000 dilution; E: *polA resA*1, oligo XT14; F: *recBCD<>kan sbcC<>amp*, oligo XT14; G: *xni<>cat*, oligo XT14; H: λ Δ*exo*, oligo XT14; I: *polB<>spec,* oligo XT14; J: Δ*recA*, oligo XT14.

Supporting information

Table S1. Bacterial strains

Strains	Genotype	Source or derivation
HME6	W3110 Δ(*argF-lac*)*U*169 [λ *c*I857Δ(*cro-bioA*)] *galK*$_{TYR145UAG}$	(Ellis *et al.*, 2001)
HME63	HME6 *mutS*<>*amp*	(Costantino & Court, 2003)
HME68	HME6 *mutS*<>*cat*	(Thomason *et al.*, 2007c)
HME69	HME68 *polA resA*1	this study
HME70	HME6 *mutS*<>*cat* Δ(*srlA-recA*)::*Tn*10	(Thomason *et al.*, 2007c)
SIMD90	HME68 [λ (*int-c*III)<>*bet c*I857Δ(*cro-bioA*)]	(Datta *et al.*, 2008)
XTL47	HME68 *lig7^{ts} ptsH*<>*tet*	this study
XTL51	HME68 *recJ*<>*amp xonA*<>*kan xseA*<>*tet exoX*<>*spec*	this study
XTL70	HME63 *polA*(5'→3'*exo*)<>*cat* fusion	this study
XTL76	HME63 *dnaQ*<>*cat*	this study
XTL85	HME63 *polA*(5'→3'*exo*)<>*cat* non-fusion	this study
XTL92	HME63 *xthA*<>*tet*	this study
XTL100	HME68 *recBCD*<>*kan sbcC*<>*amp*	this study
XTL324	HME63 *xni*<>*cat*	this study
XTL372	HME63 *polB*<>*spec*	this study
XTL532	HME68 *exoX*<>*spec*	this study

Table S2. Oligos used for recombineering experiments

Oligo	Sequence[1]	Strand
NC144	AAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTA**CC**ATCTGCCGCTGGACGGCGCACAAATCGCGCTTAA	Lag
XT13	CTGGAAGTCGCGGTaGGAACgGTATTGCAaCAGCTTTA**T**CATCTGCCcCTGGAtGGCGCACAgATCGCGCTTAAC	Lag
XT14	GTTAAGCGCGATcTGTGCGCCaTCCAGgGGCAGATG**A**TAAAGCTGtTGCAATACcGTTCCtACCGCGACTTCCAG	Lead
XT18	CTGGAAGTCGCGGTaGGcACaGTcTTaCAaCAaCTcTA**T**CAcCTaCCaCTaGAtGGaGCgCAgATCGCGCTTAAC	Lag
XT21	CTtGAgGTgGCcGTaGGcACaGTcTTaCAaCAaCTcTA**T**CAcCTaCCaCTaGAtGGaGCgCAgATaGCcCTcAAC	Lag
XT29	CTtGAgGTgGCcGTaGGcACaGTcTTaCAaCAaCTcTA**T**CATCTGCCGCTGGACGGCGCACAAATCGCGCTTAAC	Lag
XT30	CTGGAAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTA**T**CAcCTaCCaCTaGAtGGaGCgCAgATaGCcCTcAAC	Lag
XT36	CTtGAAGTgGCGGTaGGAACaGTATTaCAGCAaCTTTA**T**CAcCTGCCaCTGGAtGGCGCgCAAATaGCGCTcAAC	Lag
XT37	CTtGAgGTgGCcGTaGGcACCGTATTGCAGCAGCTTTA**T**CATCTGCCGCTGGACGGCGCACAAATCGCGCTTAAC	Lag
XT38	CTGGAAGTCGCGGTCGGAACaGTcTTaCAaCAaCTcTA**T**CAcCTaCCaCTaGAtGGaGCgCAgATaGCcCTcAAC	Lag
XT351	CTtGAgGTgGCcGTaGGcACaGTcTTaCAaCAaCTcTA**T**CAcCTaCCaCTaGAtGGCGCACAAATCGCGCTTAAC	Lag
XT352	CTtGAgGTgGCgGTaGGcACCGTATTGCAGCAGCTTTA**T**CATCTGCCGCTGGACGGCGCgCAgATaGCcCTcAAC	Lag
XT418	GTTgAGCGCtATTTGcGCGCCaTCCAGtGGCAGgTG**A**TAAAGtTGCTGtAATACtGTTCCtACCGCcACTTCaAG	Lead
XT524	GGGTTAAGTTCTTCCGCTTCACTtGAgGTgGCcGTaGGcACaGTcTTaCAaCAaCTcTA**T**CAcCTaCCaCTaGAtGGaGCgCAgATaGCcCTcAACGGTCAGGAAGCAGAAAACCAG	Lag

[1] Lower case letters indicate mismatches with respect to the wild type sequence. The underlined bold uppercase letter in each sequence denotes the selected base.

Table S3. Effect of λ exonuclease on oligo recombination

				% Marker Loss				
Strain	Activity missing	Strand[1]	Efficiency[2]	None	5'	3'	Both	Other
HME68[3]	None	lag	2.2×10^7	67	10	15	2	6
HME68[3]	None	lead	6.7×10^5	27	13	44	7	9
SIMD90	λ Exo	lag	1.8×10^7	71	9	9	4	7
SIMD90	λ Exo	lead	2.2×10^6	32	15	38	11	4

[1] Lagging-strand is oligo XT13 and leading-strand is oligo XT14.

[2] Efficiency calculated as $Gal^+/10^8$ viable cells. Recombination efficiencies given here are a representative experiment from at least 3 independent experiments where variability was less than 4-fold.

[3] Data are from Table 3 and are here to simplify direct comparisons. 42-48 independent colonies were sequenced.

Table S4. 95% confidence intervals for data in Figure 3 and Table 1. [1]

Figure 3 Data:

Strain	Oligo	Mean	Confidence Intervals
HME68	144	3.7×10^7	$2.8 \times 10^7 – 4.5 \times 10^7$
	XT13	3.0×10^7	$1.6 \times 10^7 – 4.3 \times 10^7$
	XT36	1.0×10^6	$3.3 \times 10^5 – 1.7 \times 10^6$
	XT21	4.5×10^1	$3.7 \times 10^1 – 5.2 \times 10^1$
	XT18	1.2×10^6	$1.1 \times 10^6 – 1.3 \times 10^6$
	XT524	4.0×10^6	$2.9 \times 10^6 – 5.1 \times 10^6$
	XT30	3.6×10^6	$2.0 \times 10^6 – 5.1 \times 10^6$
	XT29	2.4×10^3	$2.3 \times 10^3 – 2.6 \times 10^3$
	XT351	5.8×10^1	$3.9 \times 10^1 – 7.7 \times 10^1$
	XT38	3.4×10^4	$4.3 \times 10^3 – 6.4 \times 10^4$
	XT37	1.8×10^7	$1.2 \times 10^7 – 2.3 \times 10^7$
	XT352	1.3×10^7	$1.2 \times 10^7 – 1.3 \times 10^7$

Table 1A Data:

Strain	Oligo[2]	Mean	Confidence Intervals
HME68 (+)	lag	3.0×10^7	$1.6 \times 10^7 – 4.3 \times 10^7$
+	lag/10^3	1.0×10^4	$6.4 \times 10^3 – 1.4 \times 10^4$
+	lead	6.6×10^5	$5.6 \times 10^5 – 7.5 \times 10^5$
+	lead/10^3	8.8×10^1	$7.0 \times 10^1 – 1.1 \times 10^2$
Quad	lag	2.7×10^7	$5.0 \times 10^6 – 4.9 \times 10^7$
Quad	lag/10^3	3.4×10^5	$1.8 \times 10^5 – 5.0 \times 10^5$
Quad	lead	2.6×10^6	$1.1 \times 10^6 – 4.1 \times 10^6$
Quad	lead/10^3	1.7×10^4	$5.3 \times 10^3 – 2.9 \times 10^4$

Table 1B Data:

Strain	Oligo[2]	Mean	Confidence Intervals
+	lag	3.0×10^7	$1.6 \times 10^7 – 4.3 \times 10^7$
+	lead	6.6×10^5	$5.6 \times 10^5 – 7.5 \times 10^5$
*polA resA*1	lag	1.3×10^7	$6.2 \times 10^6 – 2.1 \times 10^7$
*polA resA*1	lead	3.0×10^5	$2.2 \times 10^5 – 3.7 \times 10^5$
polA (5'→3')<>*cat*	lag	3.1×10^6	$2.6 \times 10^6 – 3.5 \times 10^6$
polA (5'→3')<>*cat*	lead	1.4×10^5	$6.0 \times 10^4 – 2.2 \times 10^5$
dnaQ<>*cat*	lag	1.9×10^5	$0 – 4.1 \times 10^5$
dnaQ<>*cat*	lead	1.0×10^4	$5.6 \times 10^3 – 1.4 \times 10^4$
polB<>*spec*	lag	3.5×10^7	$2.8 \times 10^7 – 4.1 \times 10^7$
polB<>*spec*	lead	4.4×10^5	$1.2 \times 10^5 – 7.6 \times 10^5$
lig7^{ts}	lag	1.9×10^5	$1.7 \times 10^5 – 2.1 \times 10^5$
lig7^{ts}	lead	3.1×10^2	$2.2 \times 10^2 – 4.0 \times 10^2$

[1] 95% confidence limits were calculated assuming a normal distribution. At least 3 repeats are included in all calculations.

[2] "Lag" = oligo XT13. "Lead" = oligo XT14. Oligo concentration was reduced 1000-fold for XT13/10^3 and XT14/10^3.

Table S5. Transformation efficiency

Strain	Genotype	pBR322[1]	Mini-λ[2]
HME68		1.3×10^5	5.1×10^2
HME69	*polA resA1*	<2	5.9×10^2
XTL47	*lig7*ᵗˢ	1.8×10^5	4.7×10^2
XTL70	*polA* (5'-3') <>*cat* fusion	<3	5.2×10^2
XTL85	*polA* (5'-3') <>*cat* non-fusion	<8	5.5×10^2
XTL76	*dnaQ<>cat*	2.0×10^5	3.2×10^2
XTL372	*polB<>spec*	2.2×10^5	n.d.

[1] Transformation with 1ng of pBR322 selecting either Tet^R or Amp^R as appropriate. This amounts to ~1 plasmid per cell and is in the linear range for transformation. Efficiencies are normalized with 10^8 viable cells. These values are an average of 3 experiments and the maximum standard deviation was <56%.

[2] Transformation of mini-λ selecting Spec^R. Efficiencies are normalized with 10^8 viable cells. This site-specific integration reaction is independent of host replication and thus assays DNA uptake only (Court *et al.*, 2003). These are an average of 3 experiments and the maximum standard deviations was <76%.