

The Effects of Insertions on Mammalian Intrachromosomal Recombination

Alan R. Godwin¹ and R. Michael Liskay²

Departments of Genetics and Therapeutic Radiology, Yale University School of Medicine, New Haven, Connecticut 06510

Manuscript received July 15, 1993

Accepted for publication November 1, 1993

ABSTRACT

We examined the effects of insertion mutations on intrachromosomal recombination. A series of mouse L cell lines carrying mutant herpes simplex virus thymidine kinase (*tk*) heteroalleles was generated; these lines differed in the nature of their insertion mutations. In direct repeat lines with *different* large insertions in each gene, there was a 20-fold drop in gene conversion rate and only a five-fold drop in crossover rate relative to the analogous rates in lines with small insertions in each gene. Surprisingly, in direct repeat lines carrying the *same* large insertion in each gene, there was a larger drop in both types of recombination. When intrachromosomal recombination between inverted repeat *tk* genes with different large insertions was examined, we found that the rate of gene conversion dropped five-fold relative to small insertions, while the rate of crossing over was unaffected. The differential effects on conversion and crossing over imply that gene conversion is more sensitive to insertion mutation size. Finally, the fraction of gene conversions associated with a crossover increased from 2% for inverted repeats with small insertions to 18% for inverted repeats with large insertions. One interpretation of this finding is that during intrachromosomal recombination in mouse cells long conversion tracts are more often associated with crossing over.

GENE conversion and crossing over are often associated during meiotic recombination. Two recombination models based on the meiotic data, the Meselson-Radding model (MESELSON and RADDING 1975) and the double-strand gap repair model (SZOSTAK *et al.* 1983), explain this association by alternative cleavages (resolutions) of one, or two, Holliday intermediates, respectively. One resolution mode gives rise to simple gene conversions; the other produces gene conversions associated with crossovers.

The association between gene conversion and crossing over may be more complex for mitotic recombination. For both mammalian cells (BOLLAG and LISKAY 1988; LISKAY *et al.* 1984) and *Saccharomyces cerevisiae* (JACKSON and FINK 1981), the majority (80–90%) of mitotic intrachromosomal recombination events are simple conversions. For both species, those crossovers that do occur are often associated with gene conversions (BOLLAG and LISKAY 1988; WILLIS and KLEIN 1987). Several observations are consistent with a more complex association of mitotic gene conversion and crossing over. First, gene conversion and crossing over have been separated temporally during mitosis (ROMAN and FABRE 1983) in yeast. Those crossovers that occur during mitotic recombination in *S. cerevisiae* are predominantly associated with long conversion tracts (AGUILERA and KLEIN 1989a,b; AHN and LIVINGSTON 1986), while no such association with

conversion tract length is noted for meiotic recombination (BORTS and HABER 1987; SYMINGTON and PETES 1988). Next, UV irradiation of yeast mitotic cells induces gene conversion without inducing associated crossovers (ROMAN and JACOB 1958). Finally, *RAD52* predominantly affects simple gene conversion events in mitotic, intrachromosomal recombination (JACKSON and FINK 1981) and *hpr1* specifically increases mitotic intrachromatid crossovers (AGUILERA and KLEIN 1989a, 1990).

In mammalian cells, the mutants appropriate for examining the association between gene conversion and crossing over are not yet defined. An alternative approach is to manipulate the recombination substrate by using mutant alleles that might differentially influence gene conversion and crossing over. The results of this type of experiment can suggest common or separate intermediates for gene conversion and crossing over. In several different species large insertion mutations have been analyzed and appear to differentially impact gene conversion and crossing over (see DISCUSSION). A previous study with a mammalian intrachromosomal recombination substrate restricted to recovery of gene conversion revealed an approximately linear decrease in the rate of gene conversion with increasing insertion size (LETSOU and LISKAY 1987). Two alternative explanations for this observed decrease in gene conversion rate with increasing mutation size are that (1) gene conversion is preferentially affected by insertion mutation size or (2) large heterologies bias toward crossovers which were not recover-

¹ Current address: Department of Human Genetics, University of Utah Medical Center, Salt Lake City, Utah 84112.

² Current address: Department of Molecular and Medical Genetics, Oregon Health Sciences University, Portland, Oregon 97201.

able. A second relevant study (BOLLAG and LISKAY 1988), in fact, demonstrated an association between gene conversion and crossing over during mammalian intrachromosomal recombination.

The present report extends our findings by employing a system where the effects of insertion mutations on products of both gene conversion and crossing over can be analyzed, allowing a more careful analysis of the association between gene conversion and crossing over. We have compared recombination substrates containing small insertions in both genes, a small insertion in one gene and a large insertion in the other gene, and large insertions in both genes. In addition, we compared substrates in which the large insertions are homologous *vs.* heterologous.

MATERIALS AND METHODS

Construction of recombination substrates: A series of plasmids based on the the intrachromosomal recombination substrate, pJS3, (LISKAY *et al.* 1984, see Figure 1) were generated. Each carries a duplication of mutant herpes simplex virus thymidine kinase (*tk*) genes. The plasmids differ only in the size of the insertion mutations in the *tk* genes and in the orientation of the *tk* genes with respect to one another. The original substrate, substrate 1 (pJS3), carries a 16-bp *Xho*I linker insertion mutation at the *tk8* site in the gene flanked by *Bam*HI sites and an 8-bp *Xho*I linker insertion at the *tk26* site of the gene flanked by *Hind*III sites [(BOLLAG and LISKAY 1991; LISKAY *et al.* 1984) see Figure 1A]. Substrate 6 carries the same *tk* alleles in inverted repeat orientation (BOLLAG and LISKAY 1988).

We used two *Xho*I fragments to generate large insertion mutations. A 1.5-kb *Xho*I fragment derived from pBR322 (*Sph*I (562) to *Pvu*II (2066)) was isolated from pTK-Lin (LETSOU and LISKAY 1987). This fragment contains no homology to any other sequences in pJS3. A 1.4-kb *Xho*I fragment was generated by adding *Xho*I linkers (New England Biolabs) to a bacteriophage lambda *Cla*I (26617) to *Bss*HII (28008) restriction fragment [numbering according to DANIELS *et al.* (1983)] made blunt with the Klenow fragment. The *Hind*III (27479) and *Bam*HI (27972) sites in the lambda fragment had been destroyed previously by filling-in with Klenow.

These *Xho*I restriction fragments were cloned into the *Xho*I linker insertion sites of the *tk8* or *tk26* mutant alleles to generate substrates 2–5 and 7. Substrate 2 has the original *tk8* mutation and the 1.4-kb insertion derived from bacteriophage lambda at the *tk26* site (see Figure 1A). Substrate 3 has the 1.5-kb insertion derived from pBR322 in the *tk8* site and the original *tk26* mutation (see Figure 1A). Substrates 4 and 7 have the 1.5-kb insertion in the *tk8* site and the 1.4-kb insertion at the *tk26* site (heterologous insertions) in direct and inverted repeat orientation, respectively (see Figure 1A). Finally, substrate 5 has the same 1.5-kb pBR322 derived insertion at both the *tk8* and *tk26* sites (homologous insertions) (see Figure 1A). The 1.5-kb insertions are in direct repeat orientation in this substrate.

Cell culture and generation of experimental lines: Thymidine-kinase deficient (*tk*⁻) mouse L cells were maintained in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum or 2% fetal bovine serum and 10% newborn bovine serum in a 5% CO₂ atmosphere. Cell lines were derived by direct nuclear microinjection (CAPECCHI 1980) of *Cla*I linearized plasmid DNA (0.5 μg/ml). Trans-

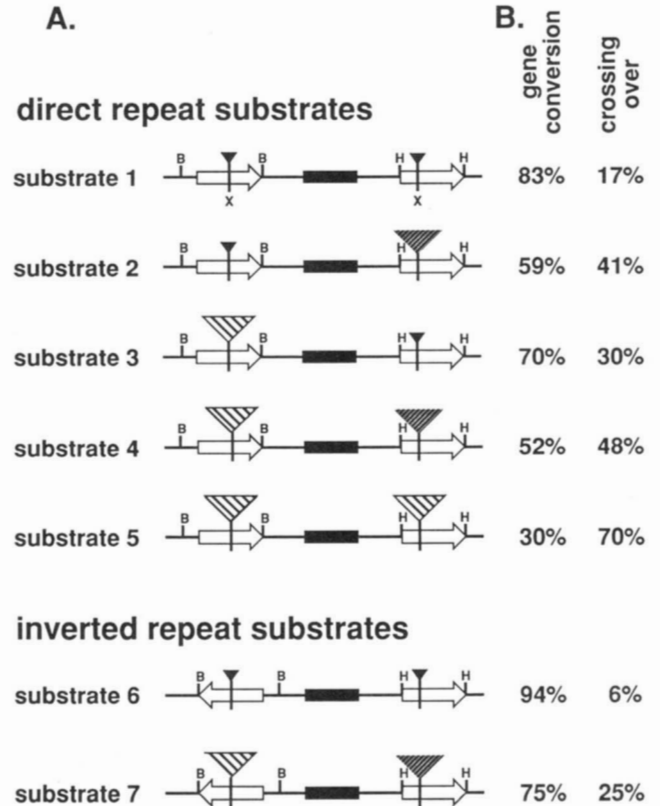


FIGURE 1.—(A) Recombination substrates. The *tk8* site is in the *Bam*HI-flanked *tk* gene and is always on the left. The *tk26* site is in the *Hind*III-flanked gene and is always on the right. The black rectangle represents the SV40 promoter driven *neo* gene, the open arrows represent the *tk* genes, the small black triangles represent the *Xho*I linker insertion mutations, the gray triangle represents the bacteriophage lambda derived insert and the striped triangle represents the pBR322 derived insertion. B represents *Bam*HI sites, H *Hind*III sites, and X *Xho*I sites. (B) Percentages of gene conversion and crossing over for each substrate. For substrates 6 and 7, the type of recombinant was determined by Southern blotting and data from recombinants showing a change in *tk* gene copy number was omitted. Crossover percentages for substrates 6 and 7 contain both simple crossovers and crossovers associated with conversions. The number of recombinants examined is 97, 64, 50, 65, 20, 191 and 40, for substrates 1–7, respectively. Data for substrates 1 and 6 are taken from BOLLAG and LISKAY (1991) and BOLLAG and LISKAY (1988), respectively.

formants were selected in 400 μg/ml G418 (Geneticin, Gibco), subcloned and tested for stability as previously described (LISKAY *et al.* 1984). Lines carrying small (*Xho*I linker insertion) mutations in both *tk* genes in direct [substrate 1 (BOLLAG and LISKAY 1991; LISKAY *et al.* 1984)] and inverted [substrate 6, (BOLLAG and LISKAY 1988)] repeat orientation had been generated and studied previously in our lab.

Southern transfer hybridization technique: Genomic DNA was isolated from *tk*⁻ parent lines and *tk*⁺ recombinants as previously described (LISKAY and EVANS 1980). Restriction enzymes were purchased from New England Biolabs or Boehringer Mannheim, and digestions were performed as recommended by the supplier. Digested DNAs (8 μg/lane) were electrophoresed on 0.8% agarose (Sigma) gels, transferred to nitrocellulose (Schleicher and Schuell) and hybridized with a 2.5-kb HSV *tk* specific, ³²P-labeled probe.

Identification of low-copy parent lines: Southern blot

analysis was performed to determine substrate copy number and integrity. Cell lines that contain the appropriately sized *HindIII* gene and a single *HindIII* junction fragment, the appropriately sized *BamHI* gene and a single *BamHI* junction fragment, and have a relative band intensity close to single copy are considered to contain a single intact copy of the recombination substrate. The single copy nature of parent lines was confirmed by obtaining G418^s recombinants by crossing over for direct repeat substrates and by Southern blot analysis of recombinants. This type of copy number analysis yielded single copy lines except for substrate 3 lines 1 and 2 (1.5 copies) and substrate 6 line 4 (4 copies). The recombination rate for substrate 6 line 4 was normalized to the rate for 1 substrate copy. An additional band that hybridized to a *tk* probe was seen in Southern blots of substrate 3 lines 1 and 2, but this band was never involved in crossovers and was never the gene corrected in gene conversion events.

Recombination analysis: Recombination rates were determined by performing Luria-Delbrück fluctuation analysis (LURIA and DELBRÜCK 1943) on colonies arising in HAT selective medium (10^{-6} M hypoxanthine, 2×10^{-6} M aminopterin, and 1.6×10^{-5} M thymidine) (SZYBALSKI *et al.* 1962). Ten or more subcultures were started from a small number (1–3) of cells, expanded in nonselective medium and then plated into HAT medium at densities of $5\text{--}8 \times 10^6$ cells in 150-cm² flasks. After 2 wk, surviving colonies were counted and a single colony was isolated from each original subculture for further analysis. A modified Luria-Delbrück equation (CAPIZZI and JAMESON 1973) was then used to calculate recombination rates.

Analysis of recombinants: For the direct repeat substrates, substrates 1–5, the type of recombinant was determined initially by G418 selection. Crossovers produce a single wild-type *tk* gene and G418 sensitivity, the result of deletion of intervening sequences including the neomycin resistance (*neo*) gene. In contrast, gene conversions retain the overall geometry of the original duplication, including the *neo* gene, and are G418 resistant. Southern blot analysis of several recombinants of both classes was undertaken. We were thus able to (1) verify the recombination class determined initially on the basis of drug resistance and (2) identify the converted *tk* allele.

For the inverted repeat substrates, substrates 6 and 7, the type of recombinant was identified on the basis of diagnostic shifts in restriction fragment mobility. Crossovers invert the orientation of the *neo* gene and generate recombinant genes flanked by *HindIII* on one end and *BamHI* on the other. Each parental line generates unique restriction fragments in crossovers due to the location of the nearest *HindIII* or *BamHI* site in genomic DNA. Gene conversions remove an insertion without altering flanking sites leading to a diagnostic shift of restriction fragment size.

RESULTS

Recombination rates: Fluctuation analysis (LURIA and DELBRÜCK 1943) was performed for each independent cell line as described in MATERIALS AND METHODS to determine recombination rates (see Table 1). From previous studies the average rates of recombination are 2.8×10^{-6} and 1.2×10^{-6} for substrates 1 and 6, respectively (BOLLAG and LISKAY 1988; LISKAY *et al.* 1984). The remainder of the data were generated in the present study. The average rates for substrates 2 and 3 are 2.9×10^{-6} and 1.1×10^{-6} ,

TABLE 1
Recombination rates

Substrate	Line	Recombination rate
1 ^a	1	3.5×10^{-6}
	2	2.4×10^{-6}
	3	2.6×10^{-6}
	Mean \pm range	$2.8 \pm 0.7 \times 10^{-6}$
2	1	1.1×10^{-6}
	2	1.6×10^{-6}
	3	6.0×10^{-6}
	Mean \pm range	$2.9 \pm 3.2 \times 10^{-6}$
3	1	0.81×10^{-6}
	2	1.4×10^{-6}
	Mean \pm range	$1.1 \pm 0.3 \times 10^{-6}$
	4	
4	1	4.7×10^{-7}
	2	1.8×10^{-7}
	3	0.93×10^{-7}
	4	1.3×10^{-7}
Mean \pm range	$2.2 \pm 2.5 \times 10^{-7}$	
5	1	1.8×10^{-8}
	2	3.5×10^{-8}
	3	3.0×10^{-8}
	Mean \pm range	$2.8 \pm 1.0 \times 10^{-8}$
6 ^b	1	1.2×10^{-6}
	2	1.6×10^{-6}
	3	0.5×10^{-6}
	4	1.3×10^{-6}
Mean \pm range	$1.2 \pm 0.7 \times 10^{-6}$	
7	1	4.9×10^{-7}
	2	2.4×10^{-7}
	3	1.3×10^{-7}
	Mean \pm range	$2.9 \pm 2.0 \times 10^{-7}$

^a Data are taken from BOLLAG and LISKAY (1991).

^b Data are taken from BOLLAG and LISKAY (1988).

respectively. Thus, for these two substrates, that contain one large insertion and one small insertion, the overall rates are not different from that of substrate 1, that has two small insertion mutations. There may be a larger effect when the large insertion mutation is in the *tk8* site, but the data set is small. This larger effect agrees with the observed bias in converting the *tk8* allele over the *tk26* allele as will be discussed later.

The direct repeat substrate with large heterologous insertions (substrate 4) has an average rate of recombination of 2.1×10^{-7} . The same heteroalleles in an inverted repeat orientation (substrate 7) show an average rate of 2.9×10^{-7} . Therefore, substrates containing two large heterologous insertions show a 10-fold drop in overall recombination rate for direct repeats and a five-fold drop for inverted repeats relative to substrates with small insertions. The large homologous insertion lines (substrate 5) show an average recombination rate of only 2.8×10^{-8} , a 100-fold drop relative to the rate for the substrate with small insertions (substrate 1). More interestingly, the overall rate for the large homologous insertion substrate is 10-fold lower than for the substrate with heterologous large insertions, even though the insertions are essentially the same size.

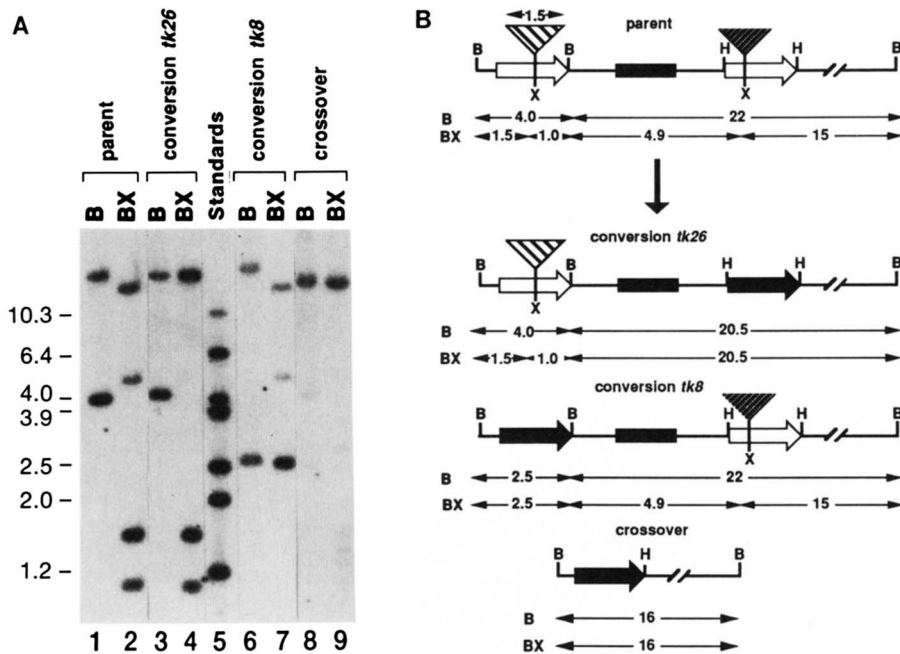


FIGURE 2.—Representative recombinants from substrate with heterologous large insertions (substrate 4). (A) Southern blot. Size of markers in kilobase pairs is shown at the left side of the blot. (B) Diagram to indicate origin of bands in Southern blot. Symbols are as described for Figure 1. Wild-type *tk* genes are denoted by black filled in arrows.

Analysis of direct repeat recombinants: Individual recombinants from each of the cell lines were examined by G418 testing and Southern blotting to discriminate gene conversions from crossovers as described in MATERIALS AND METHODS. Furthermore, Southern blot analysis revealed which mutation was converted. An example Southern blot analysis for the substrate carrying large heterologous insertions (substrate 4) as well as products of recombination is shown in Figure 2. The parent carried two mutant *tk* genes on 4.0 and 22 kb *Bam*HI restriction fragments (lane 1). Digestion with *Bam*HI and *Xho*I generates 15, 4.9, 1.5 and 1.0 kb bands due to *Xho*I cleavage at the insertion mutations (lane 2). The sizes of bands are not additive, since the *tk* probe used does not recognize the large insertions. Conversion of the *tk26* allele removed the large insertion generating a 20.5-kb *Bam*HI fragment that is resistant to *Xho*I cleavage (lanes 3 and 4). The size shift is not evident with the gel system used, but *Hind*III and *Hind*III *Xho*I digests confirmed the loss of the large insertion (data not shown). The 4.0-kb *Bam*HI fragment remained sensitive to *Xho*I cleavage (lanes 3 and 4). Conversion of the *tk8* allele removed the 1.5 kb insertion causing a shift from 4.0 to 2.5 kb (lane 6); as expected, this *tk* gene was resistant to *Xho*I cleavage (lane 7). The 22-kb *Bam*HI fragment was unaltered in size and was *Xho*I sensitive (lanes 6 and 7). Finally, a crossover generated a single *tk* gene on a 16-kb *Bam*HI fragment (lane 8), which is resistant to *Xho*I cleavage (lane 9).

Analysis of inverted repeat recombinants: Since all products of recombination from sequences oriented as inverted repeats should be G418^R, drug resistance testing was not performed. All recombi-

nants were, however, examined by Southern blot analysis. Molecular analysis of inverted repeats allowed us to distinguish classes of recombinants that we were unable to identify using direct repeat recombination substrates. These include: gene conversions associated with crossovers, intrachromatid gene conversions (a wild type and a single mutant gene with parental flanking markers) and intrachromatid double crossovers (a wild-type gene and a double mutant gene with parental flanking markers). As an example, a Southern analysis of an inverted repeat substrate and its products of recombination is shown in Figure 3. The parental line has two *tk* genes present on 4.0 and 8.5 kb *Bam*HI fragments (lane 1), which are sensitive to *Xho*I cleavage to generate 1.0, 1.5, 2.2, and 4.9 kb bands (lane 2). A conversion of *tk26* creates an *Xho*I resistant 7.1 kb *Bam*HI band (lanes 3 and 4), while the 1.0 and 1.5 kb *Bam*HI *Xho*I fragments remain (lane 4). A conversion of *tk8* generates an *Xho*I resistant 2.5 kb *Bam*HI fragment due to removal of the 1.5 kb insertion (lanes 6 and 7), while the 4.9 and 2.2 kb *Bam*HI *Xho*I bands are unaltered (lane 7). A crossover creates an *Xho*I resistant *Bam*HI band of 3.2 kb and a 9.3-kb *Bam* HI band due to inversion of the DNA between the *tk* genes (lanes 8 and 9). This 9.3-kb band is sensitive to *Xho*I and generates bands of 0.5, 1.0 and 4.9 kb upon cleavage with *Xho*I (lane 9). *Hind*III *Bam*HI digestion reveals a 2.5-kb wild-type *tk* gene (*Xho*I resistant) and a 5.0-kb band containing both large insertions (data not shown). A conversion of *tk8* associated with a crossover has the same bands as a crossover, except the 9.3-kb *Bam*HI band drops to 7.8-kb due to removal of the 1.5-kb insertion mutation (lanes 10 and 11). This 7.8-kb band cleaves to 4.9 kb

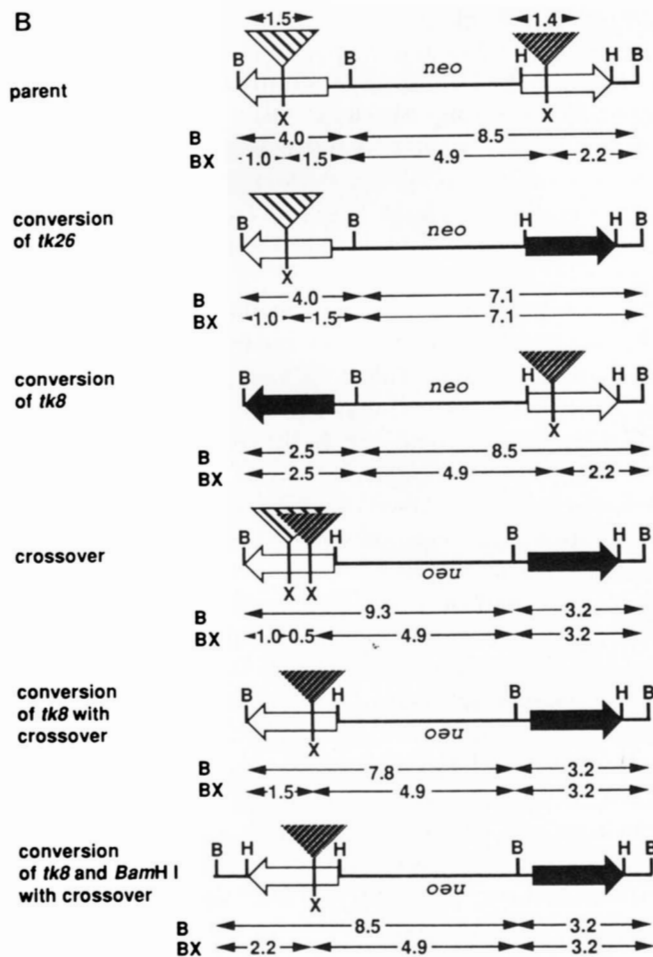
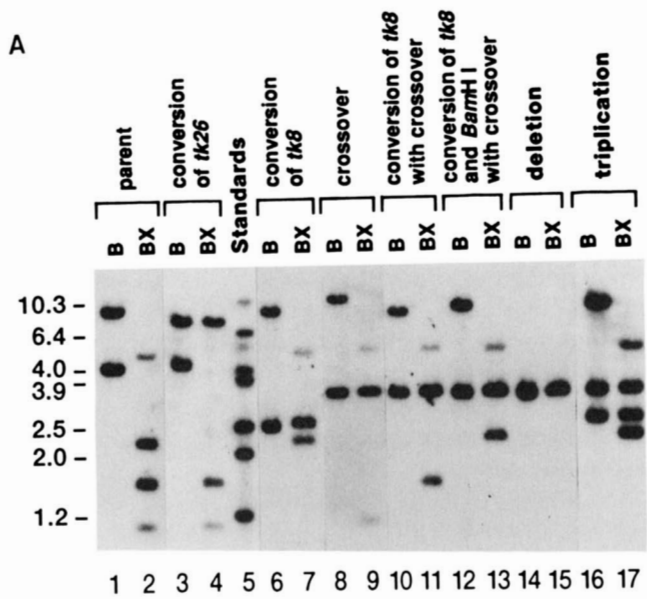
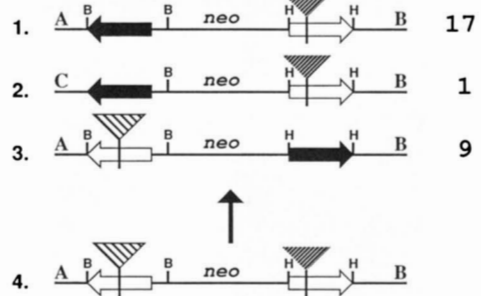


FIGURE 3.—Representative recombinants from heterologous large insertion inverted repeat substrates. (A) Southern blot. Size of markers in kilobases pairs is shown at the left of the figure. (B) Origin of bands seen in Southern blot. Symbols are as described for Figures 1 and 2.

Parental flanking markers



Recombinant flanking markers

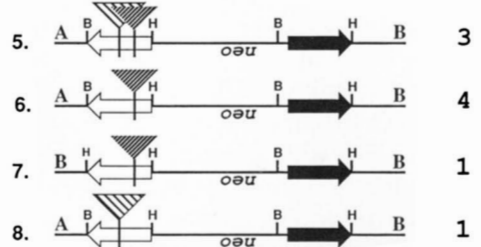


FIGURE 4.—Summary of Southern blot analysis of heterologous large insertion inverted repeat recombinants. Symbols are as described for Figures 1 and 2. A represents vector and genomic sequence normally flanking the *tk8* allele, while B represents vector and genomic sequence normally flanking the *tk26* allele. C represents new sequence which does not hybridize to flanking region A or B derived oligonucleotides. In addition, some herpes simplex virus sequence 3' to the poly A site for the *tk* gene, including the *Bam*HI site, is missing.

and 1.5 kb with *Xho*I (lane 11). The recombinant denoted “conversion of *tk8* and *Bam*HI with crossover” has the same bands as the conversion of *tk8* associated with a crossover, except the *Bam*HI fragment is 8.5 kb (lane 12) and cleaves to 4.9 and 2.2 kb fragments with *Xho*I (lane 13). This recombinant will be discussed further in the section titled *Exceptional inverted repeat recombinants*. All recombination products generated in substrate 7 lines are depicted in Figure 4. Most, 68% (27/40), of the products represented simple gene conversion events. Conversion at the *tk8* site occurred more often than at the *tk26* site (18 vs. 9 events). Three products of crossing over unassociated with gene conversions were recovered. Five products of gene conversion at the *tk8* site associated with a crossover were recovered, while only one gene conversion at the *tk26* site associated with a crossover was seen.

Exceptional inverted repeat recombinants: Two of the recombinants recovered from inverted repeats were of special interest, since in each product one of the expected flanking restriction sites was absent. Plasmid rescue of the first recombinant (see Figure 4.2) and inability to hybridize with specific oligonucleotides suggests that herpes simplex virus sequence 3'

of the end of the *tk* coding region as well as flanking material (flanking region A) has been deleted (data not shown).

The second recombinant (see Figures 3 and 4.7) resulted from a crossover associated with conversion of both the *tk8* site and the 3' flanking restriction site. A combination of plasmid rescue, DNA sequencing and Southern blotting data (data not shown) demonstrated that at least 700 bp of flanking region A have been replaced by sequence from flanking region B (see Figure 4.7). Two explanations may account for this recombinant. It is possible that a conversion tract extended beyond the region of homology into heterology. Such events have been documented in special cases of ectopic mitotic recombination in *S. cerevisiae* (BAILIS *et al.* 1992) and in some intrachromosomal gene conversions at the chicken IgL locus (McCORMICK and THOMPSON 1990). Alternatively, a sister chromatid exchange may have generated an isodicentric chromosome with identical chromosome arms. Although such isodicentric chromosomes are expected to be unstable, they can be maintained in human tumor cells, especially i(17q), and appear to be stabilized by inactivation of one of the centromeres (SANDBERG 1990; SCHWARTZ *et al.* 1983).

In addition, we recovered inverted repeat recombination products with altered gene copy number. One deletion and three independent triplication events were recovered. Southern blot analysis revealed that the triplications resided as a tandem array on a single chromosome (data not shown) and had the structures shown in Figure 5. The deletion carries a single 3.2-kb *XhoI* resistant *BamHI* fragment (Figure 5 and Figure 3, lanes 13 and 14). Further analysis (data not shown) revealed this recombinant is G418^R and contains a *tk* gene flanked by a *HindIII* site on one end and a *BamHI* site on the other end suggesting the product occurred by crossing over. Triplication 1 (Figure 5) has three *tk* genes present on *BamHI* fragments (Figure 3, lane 15). One of the *BamHI* fragments is 2.5 kb and resistant to *XhoI* cleavage indicative of a conversion of *tk8* (Figure 3, lanes 15 and 16). The presence of an 8.5-kb *BamHI* fragment that is cleaved into 4.9 and 2.2 kb fragments by *XhoI* demonstrates that a *tk26* allele remains. Finally, a 3.2-kb *XhoI* resistant *BamHI* fragment is present. Further analysis revealed the gene carried on the 3.2-kb fragment had recombinant flanking markers (a *HindIII* site on one end and a *BamHI* site on the other), suggesting it is the product of a crossover event. Thus, this triplication contains two wild-type *tk* genes. Analysis of the other two triplications (data not shown) revealed that triplication 2 contained a wild type *tk* gene with recombinant flanking markers and triplication 3 contained a wild-type *tk* gene lacking a flanking restriction site. A possible mechanism to explain

generation of these recombinants is discussed below.

Summary of recombinant analysis: Although there is no difference in the overall rate of recombination comparing substrate 1 with substrates 2 and 3 (see Table 1), the percentage of recombinants which are crossovers increases with substrates containing a gene carrying a large insertion (see Figure 1B). Similarly, substrates 4 and 5, carrying large insertion mutations in both *tk* genes, also show the same increase in percentage of crossovers compared with substrate 1. For substrate 5, carrying homologous large insertion mutations, crossing over now predominates. A substantial increase in the crossover percentage is again seen comparing the inverted repeat substrate carrying large insertions in both genes (substrate 7) with substrate 6.

The effects of large insertions on gene conversion and crossing over can be separately studied for each substrate by multiplying the overall recombination rate by the percentage of each type of recombinant and normalizing to the recombination rate of substrate 1 (see Table 2). This type of analysis shows that for substrates 2–5 and 7, the effect of large insertion mutations was greater on gene conversion than crossing over. In fact, for substrate 2 (with a large insertion in the *tk26* site), there may be a slight increase in crossover rate. The greater effect on overall conversion rate when the large insertion is in the *tk8* site can be explained by the bias against converting the *tk26* site.

The rates of gene conversion and crossing over are identical for substrate 4 with large heterologous insertion mutations (see Table 2). Having the same rate may be fortuitous or it may reflect a common recombination intermediate for both recombination outcomes (see DISCUSSION). The normalized rate of conversion is 0.003 and of crossing over is 0.007 for substrate 5 with large homologous insertion mutations showing conversion is more strongly affected by homology in the insertions. The large heterologous insertion substrate (substrate 4) has a 20-fold decrease in gene conversion rate and a 5-fold decrease in crossover rate relative to the substrate 1 with small insertions in both genes. The large homologous insertion substrate shows an even more striking effect on gene conversion and crossing over with reductions of 250-fold and 25-fold, respectively. The differential effects on gene conversion and crossing over are also evident comparing substrate 6 and substrate 7, as crossing over is unaffected by the large insertion mutations, but gene conversion is reduced five-fold.

Finally, analysis of which *tk* allele had been converted revealed biases (Table 3). In cases where one insertion mutation was large and the other was small (substrates 2 and 3), there is a 7 to 27-fold bias for conversion of the small insertion. This conversion

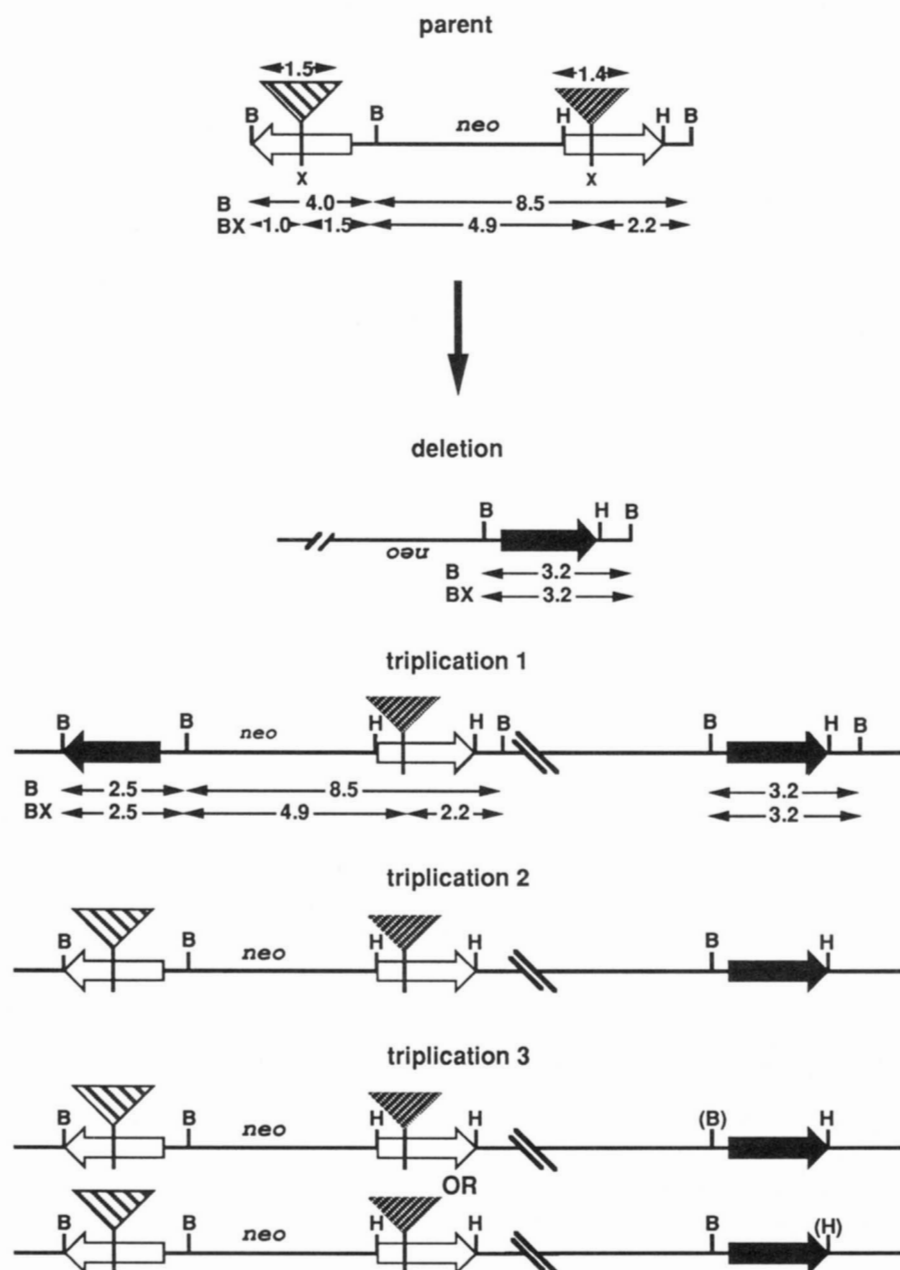


FIGURE 5.—Deletion and triplication products from the heterologous large insertion-inverted repeat substrate. Symbols are as described for Figures 1 and 2. We could not distinguish between the two possibilities depicted for triplication 3. Restriction fragment sizes are for the deletion and triplication products seen in Figure 3A (lanes 14–17). The deletion consists of a single wild-type copy of *tk*.

difference is consistent with the observed 10-fold decrease in overall recombination rate for substrate 4 (heterologous large insertions) relative to substrate 1 (see Table 1). Finally, for substrates harboring identically sized mutations at the *tk8* and *tk26* sites (substrates 4, 5, 6 and 7), there was an approximately two-fold bias for conversion at the *tk8* site. Only substrate 1 did not conform to this bias.

DISCUSSION

To study the mechanism of intrachromosomal recombination, we have generated a series of heteroalleles with different sized insertion mutations. These *tk*-based substrates can produce both gene conversion and crossover products. For each substrate, gene con-

version is preferentially inhibited by large insertions. Homologous large insertion mutations reduced recombination more than heterologous large insertions. Conversions of large insertions are preferentially associated with crossovers. Finally, several "aberrant" recombinants from an inverted repeat substrate are consistent with dicentric chromosomes, formed by sister chromatid exchange, followed by chromosome breakage.

Comparison with other systems: Our results are consistent with those of several other studies. Large insertion heterologies affect gene conversion more than crossing over during meiotic recombination at the *D. melanogaster rosy* locus (HILLIKER *et al.* 1988), and large insertions bias toward crossovers at the expense of gene conversions during meiotic recom-

TABLE 2

Relative rates of gene conversion and crossing over^a

Substrate	Overall rate	Gene conversion rate	Crossover rate
1	1.0	0.8 (1.0)	0.2 (1.0)
2	1.0	0.6 (0.8)	0.4 (2.0)
3	0.4	0.3 (0.3)	0.1 (0.8)
4	0.08	0.04 (0.05)	0.04 (0.2)
5	0.01	0.003 (0.004)	0.007 (0.04)
6	0.4	0.4 (0.5)	0.03 (0.2)
7	0.1	0.08 (0.09)	0.03 (0.2)

^a All numbers without parentheses are normalized to the overall rate of recombination for substrate 1 (2.8×10^{-6}). The rate of gene conversion for each substrate was obtained by multiplying the fraction of recombinants that were conversions by the overall rate of recombination of the substrate. The crossover rate was obtained analogously. Numbers in parentheses in the gene conversion column are normalized to the gene conversion rate for substrate 1 (2.3×10^{-6}). Numbers in parentheses in the crossing over column are normalized to the crossover rate for substrate 1 (4.8×10^{-7}).

bination at the *b2* locus of *Ascobolus immersus* (LANGIN *et al.* 1988a,b). Similarly, a decrease in recombinant recovery with increasing insertion size was noted for extrachromosomal recombination in mouse L cells (BRENNER *et al.* 1985). Interestingly, heterologies as large as 12.3 kb have no effect on gene targeting efficiency (MANSOUR *et al.* 1990). Three possible explanations for this discrepancy are: (1) that gene targeting of large heterologies occurs by double crossovers, (2) that sufficient flanking homology overcomes the deleterious effects of large insertions or (3) that the influence of large heterologies on gene targeting is different than on chromosome-chromosome recombination due to different topological constraints.

Preferential association of crossovers with long conversion tracts: BOLLAG and LISKAY (1988) found the percentage of gene conversions associated with crossovers is 2% for inverted repeats carrying small insertion mutations. We found that 18% of gene conversions of large insertion mutations were associated with crossovers. The greater association of gene conversions with crossovers for large *vs.* small insertion mutations, suggests that conversions of large insertions have a greater likelihood for crossing over. Similar studies in *S. cerevisiae* comparing mitotic gene conversion of point mutations *vs.* large insertions (VINCENT and PETES 1989) or deletions (CHERNOFF *et al.* 1984) showed conversions of the large heterology were more often associated with crossing over. It was postulated that longer (more stable) heteroduplexes were needed to convert the large insertion or deletion

TABLE 3

Distribution of gene conversion events^a

Substrate	Conversion of <i>tk8</i>	Conversion of <i>tk26</i>
1 ^b	36	28
2	27	1
3	3	21
4	19	8
5 ^c	6	0
6 ^c	124	54
7	18	9

^a A standard chi-squared test showed these results are different from a 1:1 ratio at a 1% level of significance (or better) for substrates 2, 3, 5 and 6 and at a 10% level of significance for substrates 4 and 7. The "bias" for substrate 1 is not statistically significant.

^b Data for substrate 1 are taken from BOLLAG and LISKAY (1991).

^c Data for substrate 6 are taken from BOLLAG and LISKAY (1988).

heterologies, and that these longer heteroduplexes had a greater association with crossing over (VINCENT and PETES 1989). Studies in which mitotic conversion tract length could be measured (AGUILERA and KLEIN 1989a,b; AHN and LIVINGSTON 1986) show that long conversion tracts are more often associated with crossing over. These studies are consistent with the suggestion that short tract conversion events are not associated with crossovers, while longer conversion tracts often resolve as a crossover (AGUILERA and KLEIN 1989b; CARPENTER 1984).

Using substrates containing small insertion mutations, the average conversion tract length has been estimated to be less than 358 bp in mammalian cells (LISKAY and STACHELEK 1986). Thus, the majority of gene conversions appear to be short and may not be able to proceed through a large insertion. If we assume that only longer conversion tracts are stable enough to proceed through insertion heterologies, then we selectively recover products of longer conversion tracts using substrates containing large insertion mutations. Selective recovery of long conversion tract products results in the observed greater association with crossing over. In yeast, the product of the excision repair gene *RAD1* and the DNA helicase encoded by *RAD3* are involved in determining conversion tract length (AGUILERA and KLEIN 1989b). In addition, mutations in the *RAD1* gene reduce all size classes of gene conversions associated with crossovers (AGUILERA and KLEIN 1989b). It is possible that proteins, such as that encoded by *RAD1*, can only act to promote associated crossovers after a certain length of heteroduplex is formed. Perhaps having the same rates of gene conversion and crossing over for the heterologous large insertion directly repeated *tk* genes (substrate 4) reflects an equal probability of resolution as a gene conversion or a crossover event once sufficient heteroduplex has been formed. It will be interesting to see whether mammalian cells also express proteins that determine conversion tract length.

Effect of homologous insertions: Having homologous large insertions (substrate 5) causes the greatest reduction in recombination rate. Although the large insertion mutations used for substrates 4 and 5 are of comparable size, there is a 10-fold lower recombination rate with homologous large insertions. Homology in the inserts had the greatest effect on gene conversion, which is reduced 10-fold relative to large heterologous inserts in both genes, while crossing over is only reduced 5-fold. Formally, recombination between inserts might be competing with recombination between the *tk* genes. One explanation is that when heteroduplex forms between *tk* sequences in the homologous large insertion lines, the large insertions loop out. This single strand loop may activate a second round of recombination with the homologous insert to yield a final recombination product that is nonproductive in our system, *i.e.*, is not *tk*⁺. The competition proposal could be further tested by verifying that lines carrying the bacteriophage lambda-derived large insertion in both *tk* genes show the same reduction in recombination rate and by reducing the amount of homology by replacing increasing amounts of one copy of the pBR322-derived large insertion with pieces of the bacteriophage lambda-derived insertion. If the homology in the insertions is the important factor, progressively decreasing the amount of homology should progressively increase the recombination rate.

Conversion bias: In addition to the effects of large insertions, a bias toward conversion at the *tk8* site is seen. Earlier experiments from our lab revealed a bias toward conversion of the *Xho*I linker at the *tk8* site *vs.* the *tk26* site in orientation I direct repeats (LISKAY *et al.* 1984), orientation II direct repeats (BOLLAG and LISKAY 1991) and inverted repeats (BOLLAG and LISKAY 1988). In these substrates the relative orientation and distance of the mutation sites vary with respect to elements contained in each recombination substrate, such as the SV40 enhancer/promoter or origin. Therefore, the bias appears to result from the site of the mutations in the *tk* gene. The present study shows that with large insertions there is also a bias for conversion at the *tk8* site. The bias is evident for both homologous and heterologous large insertions, suggesting the nature of the mutation is not responsible for the bias of conversion at the two sites. Rather, the DNA sequence immediately flanking the mutations and/or the relative position of the mutations within the *tk* gene appears to be critical.

One possible explanation for the bias is that the length of homology flanking the mutation is important for conversion efficiency. The closer proximity of the *tk26* site to the 5' border of homology (519 bp) may be a factor. Alternatively, a cold spot for gene conversion/hot spot for crossing over due to the local

DNA sequence at the *tk26* site could explain the conversion bias.

Exceptional inverted repeat recombinants: Based upon simple predictions, inverted repeat recombination to yield *tk*⁺ genes, reciprocal or nonreciprocal, should not change *tk* gene number, yet we recovered four recombinants which increased or decreased copy number. In addition, three recombinants have lost a flanking restriction site. One explanation for all of these events is a chromatid bridge-fusion-break cycle as recently proposed for amplification (TOLEDO *et al.* 1992; MA *et al.* 1993). Amplification of the Chinese hamster ovary cell dihydrofolate reductase locus frequently initiates by chromosome breakage followed by chromatid fusion or by unequal sister chromatid exchange (Ma *et al.* 1993). In our system the fusion-break cycle would be initiated by a sister-chromatid exchange between the inverted repeat *tk* genes generating a dicentric chromosome. The preponderance of sister chromatid over intrachromatid crossovers for direct repeats in our system (BOLLAG and LISKAY 1991) supports this proposal. In addition, three of the triplication/deletion events have wild-type *tk* genes with recombinant flanking markers indicative of crossing over. The exceptional recombinant (Figure 4.7), that has the same flanking region after both *tk* genes, might now lie on an isodicentric chromosome, that was produced by a sister chromatid exchange. For both prokaryotes (GOLUB and LOW 1983) and eukaryotes (RAY *et al.* 1989) chromosome breaks or damage as far away as 8.7 kb can stimulate recombination, so chromosome break-activated recombination might explain the observation of a triplication with two wild-type *tk* genes. One particular conversion of *tk8*, which deleted part of the herpes simplex virus sequence and flanking material (Figure 4.2), could also be explained by chromosome breakage. For this recombinant, one would invoke that the break occurred within the viral sequence or that subsequent degradation deleted material to this site. Inverted repeats (of as many as several hundred kilobases pairs) flanking an internal region of 200–1000 bp are an initial intermediate for amplification (HYRIEN *et al.* 1987, 1988; RUIZ and WAHL 1988) and deletions often are coupled to amplification (NALBANTOGLU and MEUTH 1986; WINDLE *et al.* 1991), so we may have inadvertently generated a recombination substrate prone to amplification and deletion events. Finally, since all but one of the exceptional recombination events are derived from one parental line, the integration site of this recombination substrate may predispose to the recovery of these exceptional events.

Conclusions: In summary, we have found for both direct and inverted repeats, that large insertion mutations preferentially reduce gene conversion relative to crossing over in mammalian intrachromosomal re-

combination. Homology within the insertion mutations has the largest effect on recovery of recombinants. Analysis of inverted repeats suggests that the conversion tracts required to convert large insertions are preferentially associated with crossing over. As long conversion tracts are probably necessary to convert large heterologies, we conclude that long conversion tracts are more likely to be associated with crossing over during mammalian mitotic recombination.

The authors would like to thank H. KLEIN, T. PETES and members of the LISKAY laboratory for useful discussions during the course of this work. We are also indebted to J. STACHELEK and E. TOBIN for excellent technical assistance and A. LETSOU for critical reading of this manuscript. This work was supported by National Institutes of Health grants GM32741-08 and GM45413-01 to R.M.L. A.R.G. was supported by a National Science Foundation Graduate Fellowship.

LITERATURE CITED

- AGUILERA, A., and H. L. KLEIN, 1989a Genetic and molecular analysis of recombination events in *Saccharomyces cerevisiae* occurring in the presence of the hyper-recombining mutation *hpr1*. *Genetics* **122**: 503-517.
- AGUILERA, A., and H. L. KLEIN, 1989b Yeast intrachromosomal recombination: long gene conversion tracts are preferentially associated with reciprocal exchange and require the *RAD1* and *RAD3* gene products. *Genetics* **123**: 683-694.
- AGUILERA, A., and H. L. KLEIN, 1990 *HPR1*, a novel yeast gene that prevents intrachromosomal excision recombination, shows carboxy-terminal homology to the *Saccharomyces cerevisiae* *TOP1* gene. *Mol. Cell. Biol.* **10**: 1439-1451.
- AHN, B.-Y., and D. M. LIVINGSTON, 1986 Mitotic gene conversion lengths, coconversion patterns, and the incidence of reciprocal recombination in a *Saccharomyces cerevisiae* plasmid system. *Mol. Cell. Biol.* **6**: 3685-3693.
- BAILIS, A. M., L. ARTHUR and R. ROTHSTEIN, 1992 Genome rearrangement in *top3* mutants of *Saccharomyces cerevisiae* requires a functional *RAD1* excision repair gene. *Mol. Cell. Biol.* **12**: 4988-4993.
- BOLLAG, R. J., and R. M. LISKAY, 1988 Conservative intrachromosomal recombination between inverted repeats in mouse cells: association between reciprocal exchange and gene conversion. *Genetics* **119**: 161-169.
- BOLLAG, R. J., and R. M. LISKAY, 1991 Direct-repeat analysis of chromatid interactions during intrachromosomal recombination in mouse cells. *Mol. Cell. Biol.* **11**: 4839-4845.
- BORTS, R. H., and J. E. HABER, 1989 Length and distribution of meiotic gene conversion tracts and crossovers in *Saccharomyces cerevisiae*. *Genetics* **123**: 69-80.
- BRENNER, D. A., A. C. SMIGOCKI and R. D. CAMERINI-OTERO, 1985 Effect of insertions, deletions, and double-strand breaks on homologous recombination in mouse L cells. *Mol. Cell. Biol.* **5**: 684-691.
- CAPECCHI, M. R., 1980 High efficiency transformation by direct nuclear microinjection of DNA into cultured mammalian cells. *Cell* **22**: 479-488.
- CAPIZZI, R. L., and W. JAMESON, 1973 A table for the estimation of the spontaneous mutation rate of cells in culture. *Mutat. Res.* **17**: 147-148.
- CARPENTER, A. T. C., 1984 Meiotic roles of crossing-over and of gene conversion. *Cold Spring Harbor Symp. Quant. Biol.* **49**: 23-29.
- CHERNOFF, Y. O., O. V. KIDGOTKO, O. DEMBERELIJN, I. L. LUCHNIKOVA, S. P. SOLDATOV, *et al.*, 1984 Mitotic intragenic recombination in the yeast *Saccharomyces*: marker-effects on conversion and reciprocity of recombination. *Curr. Genet.* **9**: 31-37.
- DANIELS, D., J. SCHROEDER, W. SZYBALSKI, F. SANGER, A. COULSON, *et al.*, 1983 Appendix II. Complete annotated lambda sequence. pp. 519-676 in *Lambda II*, edited by R. W. HENDRIX, J. W. ROBERTS, F. W. STAHL and R. A. WEISBERG. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
- GOLUB, E. I., and K. B. LOW, 1983 Indirect stimulation of genetic recombination. *Proc. Natl. Acad. Sci. USA* **80**: 1401-1405.
- HILLIKER, A. J., S. H. CLARK and A. CHOVIK, 1988 Genetic analysis of intragenic recombination in *Drosophila*. pp. 73-90 in *The Recombination of Genetic Material*, edited by K. B. Low. Academic Press, Inc., New York.
- HYRIEN, O., M. DEBATISSE, G. BUTTIN and B. ROBERT DE SAINT VINCENT, 1987 A hotspot for novel amplification joints in a mosaic of *Alu*-like repeats and palindromic A+T rich DNA. *EMBO J.* **6**: 2401-2408.
- HYRIEN, O., M. DEBATISSE, G. BUTTIN and B. ROBERT DE SAINT VINCENT, 1988 The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification. *EMBO J.* **7**: 407-417.
- JACKSON, J. A., and G. R. FINK, 1981 Gene conversion between duplicated genetic elements in yeast. *Nature* **292**: 306-311.
- LANGIN, T., V. HAEDENS and J. L. ROSSIGNOL, 1988a Hybrid DNA extension and reciprocal exchanges: alternative issues of an early intermediate during meiotic recombination? *Genetics* **119**: 337-344.
- LANGIN, T., H. HAMZA, V. HAEDENS and J. L. ROSSIGNOL, 1988b Reciprocal exchanges instigated by large heterologies in the *b2* gene of *Ascolobolus* are not associated with long adjacent hybrid DNA stretches. *Genetics* **119**: 329-336.
- LETSOU, A., and R. M. LISKAY, 1987 Effect of the molecular nature of mutation on the efficiency of intrachromosomal gene conversion in mouse cells. *Genetics* **117**: 759-769.
- LISKAY, R. M., and R. J. EVANS, 1980 Inactive X chromosome DNA does not function in DNA-mediated cell transformation for the hypoxanthine phosphoribosyltransferase gene. *Proc. Natl. Acad. Sci. USA* **77**: 4895-4898.
- LISKAY, R. M., and J. L. STACHELEK, 1986 Information transfer between duplicated chromosomal sequences in mammalian cells involves contiguous regions of DNA. *Proc. Natl. Acad. Sci. USA* **83**: 1802-1806.
- LISKAY, R. M., J. L. STACHELEK and A. LETSOU, 1984 Homologous recombination between repeated chromosomal sequences in mouse cells. *Cold Spring Harbor Symp. Quant. Biol.* **49**: 183-189.
- LURIA, S. E., and M. DELBRUCK, 1943 Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491-511.
- MA, C., S. MARTIN, B. TRASK and J. L. HAMLIN, 1993 Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. *Genes Dev.* **7**: 605-620.
- MANSOUR, S. L., K. R. THOMAS, C. DENG and M. R. CAPECCHI, 1990 Introduction of a *lacZ* reporter gene into the mouse *int-2* locus by homologous recombination. *Proc. Natl. Acad. Sci. USA* **87**: 7688-7692.
- MCCORMICK, W. T., and C. B. THOMPSON, 1990 Chicken IgL variable region gene conversions display pseudo gene donor preference and 5' to 3' polarity. *Genes Dev.* **4**: 548-558.
- MESELSON, M. S., and C. M. RADDING, 1975 A general model for genetic recombination. *Proc. Natl. Acad. Sci. USA* **72**: 358-361.
- NALBANTOGLU, J., and M. MEUTH, 1986 DNA amplification-deletion in a spontaneous mutation of the hamster *aprt* locus: structure and sequence of the novel joint. *Nucleic Acids Res.* **14**: 8361-8371.
- RAY, A., N. MACHIN and F. W. STAHL, 1989 A DNA double

- chain break stimulates triparental recombination in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **86**: 6225-6229.
- ROMAN, H., and F. FABRE, 1983 Gene conversion and associated reciprocal recombination are separable events in vegetative cells of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **80**: 6912-6916.
- ROMAN, H., and F. JACOB, 1958 A comparison of spontaneous and ultraviolet-induced allelic recombination with reference to the recombination of outside markers. Cold Spring Harbor Symp. Quant. Biol. **23**: 155-160.
- RUIZ, J. C., and G. M. WAHL, 1988 Formation of an inverted duplication can be an initial step in gene amplification. Mol. Cell. Biol. **8**: 4302-4313.
- SANDBERG, A. A., 1990 *The Chromosomes in Human Cancer and Leukemia*. Elsevier Science Publishing Co., Inc., New York.
- SCHWARTZ, S., C. G. PALMER, D. D. WEAVER and J. PRIEST, 1983 Dicentric chromosome 13 and centromere inactivation. Hum. Genet. **63**: 332-337.
- SYMINGTON, L. S., and T. D. PETES, 1988 Expansions and contractions of the genetic map relative to the physical map of yeast chromosome III. Mol. Cell. Biol. **8**: 595-604.
- SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN and F. W. STAHL, 1983 The double-strand-break repair model for recombination. Cell **33**: 25-35.
- SZYBALSKI, W., E. H. SZYBALSKA and G. RAGNI, 1962 Genetic studies with human cell lines. Natl. Cancer Inst. Monogr. **7**: 75-88.
- TOLEDO, F., D. LE ROSCOUET, G. BUTTIN and M. DEBATISSE, 1992 Coamplified markers alternate in megabase long chromosomal inverted repeats and cluster independently in interphase nuclei at early steps of mammalian gene amplification. EMBO J. **11**: 2665-2673.
- VINCENT, A., and T. D. PETES, 1989 Mitotic and meiotic gene conversion of Ty elements and other insertions in *Saccharomyces cerevisiae*. Genetics **122**: 759-772.
- WILLIS, K. K., and H. L. KLEIN, 1987 Intrachromosomal recombination in *Saccharomyces cerevisiae*: Reciprocal exchange in an inverted repeat and associated gene conversion. Genetics **117**: 633-643.
- WINDLE, B., B. W. DRAPER, Y. YIN, S. O'GORMAN and G. M. WAHL, 1991 A central role for chromosome breakage in gene amplification, deletion formation, and amplicon integration. Genes Dev. **5**: 160-174.

Communicating editor: S. JINKS-ROBERTSON