MOLECULAR AND CELLULAR BIOLOGY, July 1993, p. 3937–3950
0270-7306/93/073937-14$02.00/0

Vol. 13, No. 7

Substrate Length Requirements for Efficient Mitotic Recombination in *Saccharomyces cerevisiae*

SUE JINKS-ROBERTSON,* MERRILYN MICHELITCH,† AND SHARI RAMCHARAN‡

Department of Biology, Emory University, 1510 Clifton Road, Atlanta, Georgia 30322

Received 30 November 1992/Returned for modification 2 February 1993/Accepted 9 April 1993

An ectopic recombination system using *ura3* heteroalleles varying in size from 80 to 960 bp has been used to examine the effect of substrate length on spontaneous mitotic recombination. The *ura3* heteroalleles were positioned either on nonhomologous chromosomes (heterochromosomal repeats) or as direct or inverted repeats on the same chromosome (intrachromosomal repeats). While the intrachromosomal events occur at rates at least 2 orders of magnitude greater than the corresponding heterochromosomal events, the recombination rate for each type of repeat considered separately exhibits a linear dependence on substrate length. The linear relationships allow estimation of the corresponding minimal efficient processing segments, which are approximately 250 bp regardless of the relative positions of the repeats in the yeast genome. An examination of the distribution of recombination events into simple gene conversion versus crossover events indicates that reciprocal exchange is more sensitive to substrate size than is gene conversion.

Homologous recombination in eukaryotic organisms occurs in both mitosis and meiosis and is usually allelic in nature, involving like sequences at identical positions on homologous chromosomes. Recombination events can be classified as being either reciprocal or nonreciprocal in nature (for a review, see reference 23). In reciprocal recombination or crossing over, parts of chromatids are exchanged, thus altering the genetic linkages of markers on opposite sides of the exchange. Nonreciprocal recombination is usually referred to as gene conversion and involves the unidirectional transfer of information from one sequence to a homologous sequence. There is a nonrandom association of gene conversion and crossing over; if markers flanking a locus where meiotic gene conversion has occurred are examined, they will be in the exchange configuration approximately 35% of the time. The association between gene conversion and crossing over was critical to the development of the Holliday model of recombination (10) and remains a central feature of the more current Meselson-Radding (20) and double-strand-break repair (38) recombination models. The major difference in these models is in how recombination is initiated; all assume that gene conversion reflects an intermediate in the recombination process and that it is the resolution of this intermediate that results in the associated crossing over.

One of the characteristic features of eukaryotic genomes is the presence of large amounts of repetitive DNA, thus affording the opportunity for ectopic (nonallelic) as well as allelic interactions to occur. Gene conversion between dispersed repeats promotes sequence homogeneity between members of multigene families (concerted evolution [8]), but also can generate novel sequences that enhance an organism's genetic adaptability (4). While gene conversion alone does not grossly alter the genome, crossing over between dispersed repeats can result in chromosomal deletions, inversions, and translocations. Such genome rearrangements can affect an organism's reproductive fitness and may be important in speciation processes.

While there is no doubt that ectopic recombination events occur, studying these interactions between naturally occurring repeats is very difficult because, in most cases, there is no way to selectively identify the recombinants. The yeast *Saccharomyces cerevisiae* has proven to be tremendously useful in studies of this sort because it is straightforward to construct artificial duplications of defined sequence at preselected positions in the yeast genome (for a review, see reference 22). A typical ectopic duplication system utilizes auxotrophic heteroalleles that can recombine to give a selectable prototrophic phenotype. The repeats can be positioned on the same chromosome in either direct or inverted orientation with respect to each other; such repeats can undergo intrachromatid, unequal sister chromatid or unequal interhomolog interactions. Artificial repeats can also be positioned on nonhomologous chromosomes, and events involving such repeats are referred to as heterochromosomal events.

In mitotically dividing yeast cells, heterochromosomal events involving a given pair of heteroalleles occur at essentially the same rate as allelic events involving the same pair of heteroalleles (17). In meiosis, ectopic events are induced to high levels, but their rates may be somewhat less than those of the corresponding allelic events (11, 16). One paradox that has been evident since ectopic interactions were first described in *S. cerevisiae* concerns the high rate at which such events occur, particularly in meiosis. If naturally occurring repeats interacted with the same frequency as artificial repeats, it is hard to imagine how organisms with large amounts of repetitive DNA escape the associated lethal genomic rearrangements. Two hypotheses have been suggested to explain this paradox. First, one could imagine that cells might have evolved mechanisms to specifically repress ectopic recombination involving naturally occurring repeats, similar to the repression of δ-δ recombination by the yeast *TOP3* gene (42; see also reference 14). Alternatively, it might be that the structure of naturally occurring repetitive sequences significantly reduces interactions between these

* Corresponding author. Electronic mail address: BIOSJR@EMUVM1.CC.EMORY.EDU.

† Present address: Program in Biochemistry, Molecular, Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138.

‡ Present address: Grice Marine Biological Laboratory, University of Charleston, Charleston, SC 29412.

sequences (7, 24, 35). Two obvious features of natural repeats might be important: the size of the repeats (substrate length) and the degree of sequence homology between the repeats (substrate homology). Substrate size and homology might not only affect the overall rate of ectopic interactions but might also affect the frequency with which such interactions are resolved as crossovers versus noncrossovers. While transient ectopic interactions could leave behind a footprint in the form of a gene conversion event, one can imagine that only the very stable interactions would have the potential to be resolved as crossovers. The stability of a recombinational intermediate could be influenced by the length of heteroduplex formed and/or the degree of sequence identity within the heteroduplex intermediate. The purpose of the current study is to examine systematically the effect of substrate length on the rate and resolution of mitotic ectopic interactions in yeast cells.

MATERIALS AND METHODS

Media and growth conditions. YEP medium (1% yeast extract, 2% Bacto Peptone; 2.5% agar for plates) supplemented with either 2% dextrose (YPD medium) or 2% glycerol and 2% ethanol (YEP/GE medium) was used for nonselective growth of yeast strains. Synthetic dextrose (SD) complete medium (33) missing one component was used to select yeast prototrophs and score nutritional markers. SD-His, for example, was supplemented with adenine, uracil, and all standard amino acids except histidine. Synthetic medium deficient in uracil and containing 2% each galactose, glycerol, and ethanol (GGE-Ura medium) instead of dextrose was used in all Ura^+ rate measurement experiments. The agar used for the GGE-Ura plates was washed with H_2O in order to eliminate impurities that allow limited growth of Ura^- strains after selective plating. 5-Fluoroorotic acid (5-FOA) synthetic medium for selection of Ura^- segregants contained 1 mg of 5-FOA per ml (6). LB medium (1% yeast extract, 0.5% Bacto Tryptone, 1% NaCl; 1.5% agar for plates) was used for growth of *Escherichia coli* strains. Ampicillin (100 μg/ml) was added to LB for growth of plasmid-containing strains. Bacterial and yeast strains were grown at 37 and 30°C, respectively.

Plasmid constructions. A complete list of plasmids constructed during this study is given in Table 1. The numbering system for *URA3* sequences is that of Rose et al. (26), where +1 is the *Hin*dIII site immediately upstream of the *URA3* coding sequence. Plasmid pSR77 was constructed for replacement of the endogenous *URA3* allele in an appropriate yeast strain with an engineered *ura3* allele in which the *GAL1-10* promoter (*pGAL*) was substituted for the normal *URA3* promoter (*pURA*). The relevant features of this plasmid are that it contains a *ura3* allele fused to *pGAL* and that it contains the selectable *HIS3* gene. pSR77 was constructed by the following steps. First, a 1.2-kb *Bam*HI fragment containing the *URA3* locus (the normal *Hin*dIII ends were converted to *Bam*HI sites by filling in the *Hin*dIII-generated overhangs with the Klenow fragment of DNA polymerase and ligating the resulting blunt-ended fragment into the *Hin*cII site of M13mp7) was subcloned into the *Bam*HI site of pUC9. This plasmid contained two *Pst*I sites: one in the pUC9 polylinker region upstream of *pURA* and one just downstream of *pURA* (position 205). The 200-bp *Pst*I fragment containing *pURA* was replaced with an approximately 700-bp *Pst*I fragment (from pSR88; see below) containing the *GAL1-10* transcription start sites, yielding the *pGAL-ura3* allele. The *GAL1-10* promoter fragment was inserted so that the *GAL10* promoter drives transcription of *ura3* sequences. The *Pst*I *pGAL* fragment was derived by ligating *Pst*I linkers onto a filled-in 685-bp *Eco*RI-*Bam*HI fragment from plasmid pBM150 (12); this *Pst*I fragment was then inserted into the *Pst*I site of pUC7, yielding plasmid pSR88. To provide chromosomal homology upstream of the plasmid-encoded *pGAL-ura3* allele for transplacement experiments, a 700-bp *Hin*dIII fragment immediately upstream of the genomic 1.2-kb *Hin*dIII *URA3*-containing fragment (see reference 26) was cloned into the unique pUC9-derived *Hin*dIII site upstream of *pGAL-ura3*. The final step in the construction of pSR77 was the insertion of a 1.7-kb *Eco*RI fragment encoding *HIS3* (from pSR22) into the pUC9-derived *Eco*RI site downstream of *pGAL-ura3*.

The *ura3Δ3′* alleles used in this study were derived from a series of plasmids constructed by Rose and Botstein (25). These alleles have a *Hin*dIII site at the 5′ end (the normal

TABLE 1. Plasmids used in this study

Plasmid	Description
pSR22	1.7-kb *HIS3* *Bam*HI fragment in pUC7
pSR77	*pGAL-ura3ΔTaq* plasmid containing the *HIS3* *Eco*RI fragment (from pSR22)
pSR78	*HIS3*::*uraΔ884*; *HIS3* and *ura3* transcribed in the same direction
pSR79	*HIS3*::*uraΔ884*; *HIS3* and *ura3* transcribed divergently
pSR80	*HIS3*::*uraΔ692*; *HIS3* and *ura3* transcribed divergently
pSR81	*HIS3*::*uraΔ489*; *HIS3* and *ura3* transcribed divergently
pSR82	*HIS3*::*uraΔ356*; *HIS3* and *ura3* transcribed divergently
pSR83	*HIS3*::*uraΔ284*; *HIS3* and *ura3* transcribed divergently
pSR85	*HIS3*::*uraΔ489*; *HIS3* and *ura3* transcribed in the same direction
pSR86	*HIS3*::*uraΔ284*; *HIS3* and *ura3* transcribed in the same direction
pSR87	*HIS3*::*uraΔ241*; *HIS3* and *ura3* transcribed in the same direction
pSR88	*GAL1-10* *Pst*I promoter fragment (linkers added to the 685-bp *Bam*HI-*Eco*RI fragment from pBM150) in *Pst*I site of pUC7
pSR175	*ura3ΔNco* 1.2-kb *Hin*dIII fragment in pUC9
pSR176	1.7-kb *Eco*RI *HIS3* fragment in pSR175
pSR181	*HIS3*::*pGAL-uraΔ489*
pSR185	1.2-kb *URA3* *Hin*dIII fragment in Bluescript SK+
pSR191	*ura3-Nhe* frameshift allele
pSR195	*HIS3* in pSR191
pSR215	2-kb *Hin*dIII fragment upstream of *URA3* in *Hin*cII site of pUC7
pSR216	*ura3Δ884* from pSR78 in pSR215; direct-repeat construct
pSR217	*ura3Δ884* from pSR78 in pSR215; inverted-repeat construct
pSR218	*ura3Δ489* from pSR85 in pSR215; direct-repeat construct
pSR219	*ura3Δ489* from pSR85 in pSR215; inverted-repeat construct
pSR220	*ura3Δ284* from pSR86 in pSR215; direct-repeat construct
pSR221	*ura3Δ284* from pSR86 in pSR215; inverted-repeat construct
pSR249	*ura3Δ241* from pSR87 in pSR215; direct-repeat construct
pSR250	*ura3Δ241* from pSR87 in pSR215; inverted-repeat construct

*Hin*dIII site upstream of the *URA3* coding sequence) and a *Bam*HI site at the 3′ deletion endpoint. Plasmids pRB82, pRB80, pRB79, pRB75, pRB74, and pRB72 contain *ura3Δ3′* alleles of sizes 884, 692, 489, 356, 284, and 241 bp, respectively. Individual *ura3Δ3′* alleles were inserted at a unique *Bss*HII site just downstream of the *HIS3* coding sequence on plasmid pSR22 as follows. A *Hin*dIII-*Bam*HI fragment containing the relevant *ura3Δ3′* allele was treated with the Klenow fragment of DNA polymerase to fill in the enzyme-generated 5′ overhangs. The fragment was then ligated to pSR22 that had been digested with *Bss*HII, incubated with Klenow to fill in the sticky ends, and treated with phosphatase. Insertion of the fragment into the vector regenerated the *Bam*HI site marking the endpoint of the *ura3Δ3′* deletions; the *Hin*dIII and *Bss*HII sites were destroyed. Plasmids containing the *ura3Δ3′* fragment in both orientations relative to the *HIS3* gene were isolated and will be referred to here by the size of the *ura3Δ3′* allele. *HIS3*::*ura3Δ884*, for example, contains the 884-bp fragment derived from pRB82. The *HIS3*::*ura3Δ3′* plasmids with the *URA3* and *HIS3* sequences transcribed divergently (pSR79 to pSR83 in Table 1) were used to construct the duplications on chromosomes V and XV; the *URA3* locus is transcribed toward the centromere of chromosome V, while the *HIS3* locus is transcribed away from the centromere of chromosome XV. The *HIS3*::*ura3Δ3′* plasmids with *URA3* and *HIS3* sequences transcribed in the same direction (plasmids pSR78, pSR85, pSR86, and pSR87) were used for construction of the direct and inverted repeats on chromosome V.

Plasmids for constructing the direct and inverted repeats on chromosome V were derived by inserting *Bam*HI fragments containing appropriate *HIS3*::*ura3Δ3′* constructs into a *Bgl*II site approximately 2 kb upstream of the genomic 1.2-kb *URA3* *Hin*dIII fragment. This was accomplished as follows. The *Bgl*II site is contained within a 2-kb *Hin*dIII fragment (26); this *Hin*dIII fragment was treated with Klenow enzyme and ligated to *Hin*cII-digested pUC7, creating plasmid pSR215. *Bam*HI fragments containing *HIS3*::*ura3Δ884*, *HIS3*::*ura3Δ489*, *HIS3*::*ura3Δ284*, and *HIS3*::*ura3Δ241* were ligated to *Bgl*II-digested pSR215, and inserts in both orientations were identified by appropriate restriction enzyme digestions (plasmids pSR216 to pSR221, pSR249, and pSR250).

The *HIS3*::*pGAL-ura3Δ489*-containing plasmid pSR181 was constructed by the following series of steps. First, a 285-bp *Pst*I-*Bam*HI fragment derived from pRB79 (contains *ura3Δ489*) was subcloned into *Pst*I-*Bam*HI-digested pUC9; as noted above, *Pst*I cuts just downstream of *pURA*, so that the subcloned *ura3* fragment is promoterless. Next, the 700-bp *pGAL* *Pst*I fragment (from pSR88) was inserted at the unique *Pst*I site of the resulting plasmid, creating *pGAL-ura3Δ489*. In the final step of the construction, *pGAL-ura3Δ489* was removed from the vector by digestion with *Hin*dIII-*Bam*HI and the sticky ends were filled in with the Klenow fragment of DNA polymerase. This blunt-ended fragment was inserted at the filled-in *Bss*HII site just downstream of the *HIS3* coding sequence on plasmid pSR22 (see above).

The *ura3ΔNco* allele contained on plasmid pSR175 was constructed by filling in a unique *Nco*I site within the *URA3* coding sequence (position 432) with the Klenow fragment of DNA polymerase. This operation adds 4 bp to the coding sequence and creates an *Nsi*I site. A *HIS3*-containing 1.7-kb *Eco*RI fragment (from pSR22) was cloned into the *Eco*RI site of pSR175, yielding plasmid pSR176 for yeast transplacement experiments. The *ura3-Nhe* allele (plasmid pSR191) contains a G inserted between bases 256 and 257 in the *URA3* coding sequence. This insertion creates a unique *Nhe*I site and was constructed by oligonucleotide-mediated site-directed mutagenesis of pSR185, using a standard protocol (30). The sequence of the oligonucleotide used to prime second-strand synthesis of the single-stranded template was 5′-ACGTGCT**G**CTAGCTCATCCT-3′; the added G is in bold type, and the resulting *Nhe*I site is underlined. A *HIS3*-containing 1.7-kb *Bam*HI fragment was cloned into the *Bam*HI site of pSR191, yielding plasmid pSR195 for yeast transplacement experiments.

Yeast strain constructions. All yeast strains used in this study were derived by transformation of the haploid strain SJR114 (*MATα ade2-101 his3Δ200*) by either a spheroplast (27) or lithium acetate (31) procedure. It should be noted that the *URA3* locus in SJR114 was derived from strain FL100, as were all the plasmid-encoded *URA3* sequences used to transform this strain. All strains were initially identified on the basis of phenotype, and then the inferred genotype was confirmed by Southern blot analysis. The *pGAL-ura3* allele contained on plasmid pSR77 was introduced into strain SJR114 by a two-step plasmid integration-excision procedure (28). Plasmid pSR77 was targeted to integrate at *URA3* by digestion with *Stu*I, and His^+ transformants were selected. Excision of the plasmid (pop-outs) with retention of the *pGAL-ura3* allele on chromosome V was identified by plating transformants on 5-FOA medium, yielding strain SJR116. The *pGAL-ura3* allele originally constructed (*pGAL-ura3ΔTaq*; Klenow fill-in of the *Taq*I site at position 230) proved to be leaky and so was not used in the recombination experiments. Instead, two different *ura3* alleles were constructed in vitro and were introduced into the SJR114 background as outlined below.

The *ura3ΔTaq* mutation in SJR116 was corrected to wild type by transformation with a 900-bp *Pst*I-*Sma*I fragment containing the entire *URA3* coding sequence (but not the promoter), yielding strain SJR189 with the *URA3* gene fused to *pGAL*. It should be noted that SJR114 and SJR189 are identical with the exception of the promoter driving expression of the wild-type *URA3* gene. The *ura3ΔNco* and *ura3-Nhe* alleles were introduced by sequentially selecting for integration/pop-out of plasmids pSR176 and pSR195, respectively, following transformation of strains SJR114 and SJR189. Plasmids were linearized with *Stu*I to target integration at *URA3*; following selection of His^+ transformants, pop-outs were selected on 5-FOA. Strain SJR195 contains the *ura3ΔNco* allele, strain SJR196 contains the *pGAL-ura3ΔNco* allele, strain SJR231 contains the *ura3-Nhe* allele, and strain SJR237 contains the *pGAL-ura3-Nhe* allele.

The strains containing a second *ura3* allele on chromosome XV were constructed as follows. The *his3Δ200* allele resident at the *HIS3* locus in strains SJR195, SJR196, SJR231, and SJR237 was replaced with the various *HIS3*::*ura3Δ3′* alleles by one-step gene replacement (28). *Eco*RI digests of appropriate plasmids (pSR79 to pSR83 and pSR181) released the *HIS3*::*ura3Δ3′* alleles as linear fragments, and these digests were used to transform yeast strains, selecting His^+ transformants. The transformants contain a full-length *ura3* allele at the *URA3* locus on chromosome V and a *ura3Δ3′* allele of variable size adjacent to the *HIS3* locus on chromosome XV. No plasmid sequences are present on chromosome XV.

Derivatives of strains SJR196, SJR231, and SJR237 with the *ura3* duplications as nontandem direct or inverted repeats on chromosome V were constructed by transformation with linear *Eco*RI fragments containing *HIS3*::*ura3Δ3′* constructs flanked by DNA upstream of the *URA3* locus (plas-

TABLE 2. Yeast strains containing *ura3* duplications

Strain	Parent strain/plasmid	*ura3* alleles present	Type of duplication	Homology (bp)
SJR203	SJR195/pSR79	*ura3ΔNco/HIS3::ura3Δ884*	Heterochromosomal	884
SJR204	SJR195/pSR80	*ura3ΔNco/HIS3::ura3Δ692*	Heterochromosomal	692
SJR205	SJR195/pSR81	*ura3ΔNco/HIS3::ura3Δ489*	Heterochromosomal	489
SJR206	SJR196/pSR79	*pGAL-ura3ΔNco/HIS3::ura3Δ884*	Heterochromosomal	680
SJR207	SJR196/pSR80	*pGAL-ura3ΔNco/HIS3::ura3Δ692*	Heterochromosomal	488
SJR208	SJR196/pSR81	*pGAL-ura3ΔNco/HIS3::ura3Δ489*	Heterochromosomal	285
SJR232	SJR231/pSR79	*ura3-Nhe/HIS3::ura3Δ884*	Heterochromosomal	884
SJR233	SJR231/pSR80	*ura3-Nhe/HIS3::ura3Δ692*	Heterochromosomal	692
SJR234	SJR231/pSR81	*ura3-Nhe/HIS3::ura3Δ489*	Heterochromosomal	489
SJR235	SJR231/pSR82	*ura3-Nhe/HIS3::ura3Δ356*	Heterochromosomal	356
SJR236	SJR231/pSR83	*ura3-Nhe/HIS3::ura3Δ284*	Heterochromosomal	284
SJR238	SJR237/pSR79	*pGAL-ura3-Nhe/HIS3::ura3Δ884*	Heterochromosomal	680
SJR239	SJR237/pSR80	*pGAL-ura3-Nhe/HIS3::ura3Δ692*	Heterochromosomal	488
SJR240	SJR237/pSR81	*pGAL-ura3-Nhe/HIS3::ura3Δ489*	Heterochromosomal	285
SJR241	SJR237/pSR82	*pGAL-ura3-Nhe/HIS3::ura3Δ356*	Heterochromosomal	152
SJR242	SJR237/pSR83	*pGAL-ura3-Nhe/HIS3::ura3Δ284*	Heterochromosomal	80
SJR249	SJR195/pSR181	*ura3ΔNco/HIS3::pGAL-ura3Δ489*	Heterochromosomal	285
SJR250	SJR196/pSR181	*pGAL-ura3ΔNco/HIS3::pGAL-ura3Δ489*	Heterochromosomal	970
SJR253	SJR231/pSR181	*ura3-Nhe/HIS3::pGAL-ura3Δ489*	Heterochromosomal	285
SJR254	SJR237/pSR181	*pGAL-ura3-Nhe/HIS3::pGAL-ura3Δ489*	Heterochromosomal	970
SJR265	SJR237/pSR216	*pGAL-ura3-Nhe/HIS3::ura3Δ884*	Direct	680
SJR266	SJR237/pSR217	*pGAL-ura3-Nhe/HIS3::ura3Δ884*	Inverted	680
SJR267	SJR237/pSR218	*pGAL-ura3-Nhe/HIS3::ura3Δ489*	Direct	285
SJR268	SJR237/pSR219	*pGAL-ura3-Nhe/HIS3::ura3Δ489*	Inverted	285
SJR269	SJR237/pSR220	*pGAL-ura3-Nhe/HIS3::ura3Δ284*	Direct	80
SJR270	SJR237/pSR221	*pGAL-ura3-Nhe/HIS3::ura3Δ284*	Inverted	80
SJR286	SJR231/pSR216	*ura3-Nhe/HIS3::ura3Δ884*	Direct	884
SJR287	SJR231/pSR217	*ura3-Nhe/HIS3::ura3Δ884*	Inverted	884
SJR288	SJR231/pSR218	*ura3-Nhe/HIS3::ura3Δ489*	Direct	489
SJR289	SJR231/pSR219	*ura3-Nhe/HIS3::ura3Δ489*	Inverted	489
SJR290	SJR231/pSR220	*ura3-Nhe/HIS3::ura3Δ284*	Direct	284
SJR291	SJR231/pSR221	*ura3-Nhe/HIS3::ura3Δ284*	Inverted	284
SJR292	SJR196/pSR249	*pGAL-ura3ΔNco/HIS3::ura3Δ241*	Direct	37
SJR293	SJR196/pSR250	*pGAL-ura3ΔNco/HIS3::ura3Δ241*	Inverted	37

mids pSR216 to pSR221, pSR249, and pSR250). As in the construction of the heterochromosomal repeats, His^+ transformants were selected. Since a one-step gene replacement procedure was used, no extraneous plasmid sequences were introduced onto chromosome V. All duplication-containing yeast strains are listed in Table 2.

Measuring recombination rates. Two-day-old, freshly grown colonies were excised from YPD plates, inoculated into 5 ml of YEP/GE medium, and grown for 2 days on a roller drum at 30°C. Cells were harvested by centrifugation, washed with an equal volume of sterile H_2O, and then resuspended in 1.5 ml of sterile H_2O. Aliquots (100 μl) of a 10^{-6} dilution were plated on YPD to determine the number of CFU in the original culture; this was approximately 2×10^8 CFU/ml for all cultures examined. For the strains containing heterochromosomal repeats, at least three 100-μl aliquots of each washed, concentrated culture were plated selectively on GGE-Ura medium. For strains containing either the direct or inverted repeats, cells were diluted so that the number of Ura^+ colonies per selective GGE-Ura plate was less than 100. To achieve a cell density on the selective medium comparable to that in the heterochromosomal repeat experiments, appropriate numbers of nonreverting cells (either strain SJR195 or strain SJR231) were added to the direct- or inverted-repeat experimental cells. The median number of Ura^+ recombinants for each strain was determined by examining 10 independent cultures, and the experimentally determined median was used to calculate the recombination rate according to Lea and Coulson (15). Colony counts to estimate the number of Ura^+ recombinants per culture were done 4 days after selective plating on GGE-Ura, although prototrophs continued to appear at least until day 10 postplating. As we have shown previously, the experimental distributions of early-appearing recombinant prototrophs conform reasonably well to theoretical Lea and Coulson distributions, whereas the late-appearing recombinants more closely approximate the Poisson distribution (36). We thus assume that the early-appearing recombinants represent predominantly those events that occurred during nonselective growth prior to selective plating.

Isolation and analysis of independent Ura^+ recombinants. Independent Ura^+ recombinants from strains in which one *ura3* allele was under control of *pURA* and the other was under control of *pGAL* were isolated by patching colonies from YEP/GE onto GGE-Ura medium. To ensure the independence of recombinants, only a single Ura^+ colony was picked from each patch. Independent recombinants were purified on YPD, and then the Ura phenotype (constitutive versus galactose inducible) was assessed by plating on SD-Ura and GGE-Ura media. The phenotypically inferred type of recombination event was confirmed by Southern blot analysis of a limited number of recombinants.

For strains in which both *ura3* alleles were under control of *pURA* (or *pGAL*), single colonies grown on YPD (or YEP/GE) were patched onto SD-Ura (or GGE-Ura) medium. Ura^+ colonies were purified nonselectively on YPD,

the Ura phenotype was confirmed, and DNA was isolated by a glass bead lysis procedure (9). Southern blot analysis and/or the polymerase chain reaction (PCR) was used to distinguish gene conversion from exchange events. For Southern blot analysis, genomic DNA was digested with *Eco*RI and probed with *URA3*-specific sequences. Nonexchange recombinants have *Eco*RI fragments of approximately 13 and 16 kb, while exchange recombinants have *Eco*RI fragments of approximately 9 and 20 kb (11).

PCR buffer contained 50 mM KCl, 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 0.1% Triton X-100, and 200 μM each deoxynucleoside triphosphate. PCR was done in a 25-μl volume containing 12 ng of each primer, 0.5 U of *Taq* DNA polymerase (Promega), and approximately 100 ng of yeast genomic DNA. The 29-cycle PCR program for heterochromosomal and inverted repeats was as follows: 94°C for 3 min (initial cycle only), 94°C for 1 min, 50°C for 2 min, 72°C for 2.5 min, and 72°C for 8 min (final cycle only). For the direct-repeat DNAs, the extension phase of each cycle was lengthened to 4 min. Two sets of primers were used for characterizing the heterochromosomal and inverted-repeat recombinants: *LYS2* control primers that yield a PCR product with all DNA samples and *URA3-HIS3* primers that produce a product only if reciprocal exchange has occurred. The *LYS2* control primers were 5′-CTACCTCAGCTCGAT GTG-3′ (forward) and 5′-TTAGAAGTGCGGTTGATG-3′ (reverse), producing a product of 718 bp. The *URA3*- and *HIS3*-specific primers for the heterochromosomal repeats were 5′-TGTCAGATCCTGTAGAGACC-3′ and 5′-CTCGT TGTTGTCGTTGATGC-3′, respectively, yielding products of 1,130 and 1,600 bp for the homologous *pURA* and *pGAL* promoter constructs, respectively. For the inverted repeats, the *LYS2* control primers and the same *URA3* primer in combination with a different *HIS3* primer (5′-CCTCCACGT TGATTGTCTGC-3′) were used, yielding an exchange-specific product of 1,200 bp. Direct-repeat recombinants were characterized by using the same *URA3-HIS3* primer set as was used with the inverted repeats; the *LYS2*-specific primers were not included. Nonexchange recombinants from the direct-repeat strains yielded an approximately 3,500-bp PCR product, while exchange recombinants yielded a 1,200-bp PCR product. To confirm the accuracy of the PCR results, DNAs from approximately 10 recombinants derived from each strain were subjected to Southern blot analysis. The Southern results were in complete agreement with the PCR results.

RESULTS

A heterochromosomal system for varying substrate length. A series of isogenic haploid yeast strains with duplicated *ura3* sequences on nonhomologous chromosomes as recombination substrates was constructed. A full-length *ura3* gene containing one of two different frameshift mutations was located at the *URA3* locus on chromosome V; *uraΔNco* is a +4 frameshift allele constructed by filling in a unique *Nco*I restriction site and *ura3-Nhe* is a +1 frameshift allele constructed by site-directed mutagenesis. Substrate length was controlled by introducing adjacent to the *HIS3* locus on chromosome XV a second *ura3* allele (*ura3Δ3′*) with a defined, variable amount of coding sequence deleted from the 3′ end. The relative amounts of recombination between the duplicated *ura3* sequences were assessed by calculating the rates of Ura^+ prototroph formation.

In addition to examining the effect of substrate length on the total amount of recombination, we also wanted to

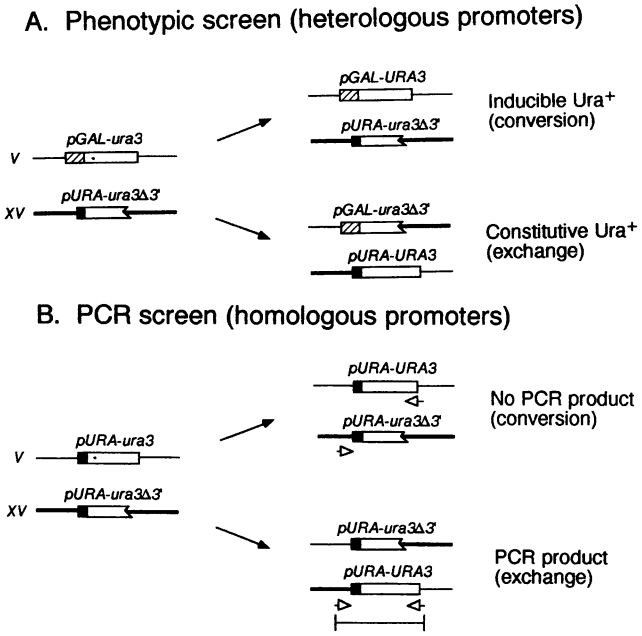
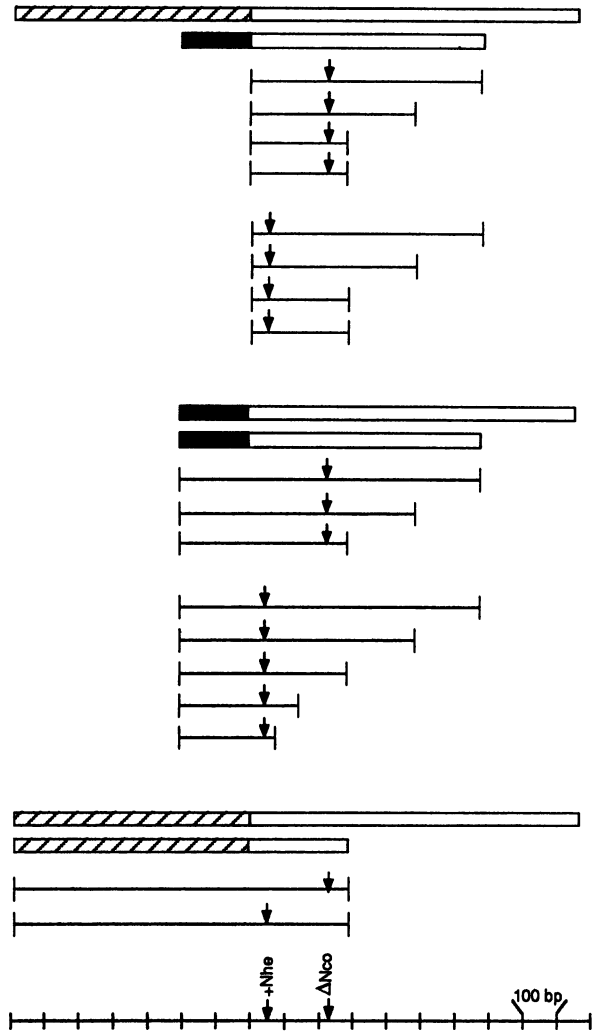


FIG. 1. Distinguishing conversion versus exchange events with *ura3* heterochromosomal repeats. Open boxes represent *ura3* sequences, solid boxes correspond to *pURA* and hatched boxes correspond to *pGAL*. The thin and thick lines represent flanking, nonhomologous sequences on chromosomes V and XV, respectively. On chromosome XV, the *HIS3* locus is downstream of and transcribed toward the *ura3Δ3′* allele. The open arrows correspond to the positions of the PCR primers used to detect crossing over.

determine whether substrate length might affect the resolution of interactions as crossover versus noncrossover events. Two different systems were used to distinguish simple gene conversion events from those events accompanied by crossing over between chromosomes V and XV. In the first system, the *ura3* repeats were under control of different promoters (Fig. 1A; the heterologous promoter system), allowing conversion and exchange events to be distinguished phenotypically. In most strains, the full-length *ura3* gene on chromosome V was fused to the galactose-inducible, glucose-repressible *GAL1-10* promoter (*pGAL*), while the *ura3Δ3′* allele on chromosome XV was under control of the normal *URA3* promoter (*pURA*). With this configuration of promoters, a gene conversion event yields a glucose-repressible *pGAL-URA3* allele; recombinants are Ura^+ on galactose-containing medium and Ura^- on glucose-containing medium. In contrast, a crossover between the *ura3* repeats on chromosomes V and XV yields a constitutive *pURA-URA3* allele.

In the second *ura3* duplication system, both *ura3* alleles were under control of the *URA3* promoter (homologous promoter system). Conversions and crossovers between the homologous promoter constructs were distinguished by using either Southern blot analysis or a PCR-based screen. In the PCR screen, one primer hybridized to *URA3* sequences on chromosome V that are downstream of the endpoints of all of the *ura3Δ3′* alleles present on chromosome XV; the second primer hybridized to *HIS3*-specific sequences on chromosome XV. As illustrated in Fig. 1B, a PCR product is produced only if both primers hybridize to a contiguous stretch of DNA, that is, only if chromosome V and XV sequences flanking the *ura3* repeats become physically



Strain	Size (bp)	Ura+ rate (10^{-8})[a]	% exchange[b]
SJR206	680	1.56	29.4 (16/93)
SJR207	488	0.75	20.0 (12/108)
SJR208	285	0.26	3.9 (2/100)
SJR249	285	0.16	6.8 (5/142)
SJR238	680	1.75	61.2 (45/102)
SJR239	488	0.64	40.0 (26/104)
SJR240	285	0.23	20.0 (8/72)
SJR253	285	0.11	23.0 (10/77)
SJR203	884	2.93	44.8 (26/90)
SJR204	692	1.46	37.9 (22/94)
SJR205	489	0.78	27.8 (15/93)
SJR232	884	2.19	70.4 (50/92)
SJR233	692	1.54	70.1 (47/87)
SJR234	489	1.13	73.1 (53/92)
SJR235	356	0.49	61.5 (40/90)
SJR236	284	0.22	37.5 (15/65)
SJR250	970	1.17	7.7 (4/100)
SJR254	970	1.49	11.3 (6/100)

FIG. 2. Recombination data obtained by using *ura3* heterochromosomal repeats. Open boxes represent *ura3* sequences, solid boxes correspond to *pURA*, and hatched boxes correspond to *pGAL*. The lines below the boxes correspond to the length of substrate homology present in each strain; vertical arrows over these lines represent the position of the *ura3ΔNco* or *ura3-Nhe* point mutation in the full-length *ura3* allele. A crossover event between the *ura3* repeats on chromosomes V and XV is assumed to yield two exchange chromatids (V:XV and XV:V reciprocal translocations) plus two nonexchange chromatids. It should be noted that viable recombinants result only when both exchange products segregate to the same daughter cell. If mitotic segregation of the recombinant chromatids is random, then one-half of the exchange events will not be detected, giving rise to inviable products containing one normal and one translocation chromosome. The observed number of exchanges thus represents only one-half of the actual number of exchanges, leading to a systematic underestimate of the total recombination rate as well as the exchange frequency unless the raw data are corrected for the exchange-associated inviability. [a]The rate of Ura^+ recombinants for each strain was corrected in order to take into account the inviability associated with one-half of the exchange events. This correction is necessary since the exchange frequency varies from strain to strain. For SJR206, for example, 16 of 93 (17%) of the viable Ura^+ recombinants represent viable crossover events. For every 100 recombinants counted, an additional 17 exchanges were presumably lost as a result of viability problems. The experimentally determined median number of SJR206 recombinants, therefore, was corrected by multiplying by 1.17. The corrected median was then used to calculate the recombination rate according to Lea and Coulson (15). [b]The percent exchange was corrected to take into account the exchange-associated inviability of recombinants. If n is the observed number of crossovers out of a total of m Ura^+ recombinants examined, then the corrected exchange frequency among the recombinants equals $2n/(m + n)$, where both the number of exchanges and the total number of recombinants are corrected for the exchange-associated lethality. The actual number of exchanges among the total number of independent Ura^+ recombinants examined is given in parentheses following the corrected percent exchange value.

linked via crossing over between the repeats. In addition to the homologous *pURA* constructs, two strains (SJR250 and SJR254; Fig. 2) with both of the *ura3* alleles fused to *pGAL* were constructed.

Effect of substrate length on heterochromosomal recombination events. The relevant features of the strains containing the heterochromosomal *ura3* repeats of variable size are summarized in Fig. 2. These strains can be divided into three groups on the basis of the promoter sequences driving expression of the *ura3* repeats. Each group can be further subdivided on the basis of whether the full-length *ura3* allele contains the*ΔNco* +4 or the *Nhe* +1 frameshift mutation. Two different *ura3* alleles were used to evaluate possible effects of the position of a mutation within the region of

homology on the rate and/or resolution of recombination events.

In the first group of strains, the *ura3* alleles are fused to different promoters so that homology at the 5′ end begins just upstream of the translation start site. The sequence homology in each strain terminates at the deletion endpoint of the *ura3Δ3′* allele adjacent to the *HIS3* locus. It should be noted that the individual strains in the strain pairs SJR208-SJR249 and SJR240-SJR253 are identical with respect to the sequences duplicated but differ with respect to the positions of the heterologous promoters. SJR208 and SJR240 have the standard configuration with the full-length *ura3* allele fused to *pGAL* and the deletion allele fused to *pURA*; SJR249 and SJR253 have the reverse configuration with the full-length *ura3* allele fused to *pURA* and the deletion allele fused to *pGAL*. These pairs of strains were constructed to control for possible gene conversion versus exchange biases that might result from the presence of the heterologous promoters. One could imagine, for example, that recombinants containing the *pGAL-URA3* allele might have a growth advantage over those containing the *pURA-URA3* allele on the galactose-containing medium used to initially identify independent recombinants. If such was the case, a strong gene conversion bias would be expected in the *pGAL-ura3/pURA-ura3Δ3′* strains (SJR208 and SJR240), while a strong exchange bias would be expected in the *pURA-ura3/pGAL-ura3Δ3′* strains (SJR249 and SJR253). The observation that the two strains within each pair yielded very similar recombination results argues that use of the heterologous promoters does not introduce a substantial bias into the data.

In the second group of strains, both *ura3* alleles are under control of *pURA*, which extends the homology between the repeats 205 bp in the 5′ direction. The repeats in strains SJR206 and SJR203, for example, have the same 3′ border of homology and the same frameshift allele, but the duplication in SJR203 has an additional 205 bp at the 5′ end relative to the duplication in SJR206. Finally, in the third group of strains, both *ura3* alleles are fused to *pGAL*.

The results of experiments with the heterochromosomal repeats are presented in Fig. 2 and can be summarized as follows. First, if one individually examines the four subgroups within the first two groups of repeats shown in Fig. 2, a common result is evident within each subgroup: as the length of homology decreases, the total rate of Ura$^+$ recombinants decreases and the proportion of events associated with crossing over decreases. The second generalization that can be made is that strains containing the *ura3-Nhe* allele consistently have a larger proportion of Ura$^+$ recombinants as exchanges than do the strains containing the *ura3ΔNco* allele. A third observation made with the first two groups is that asymmetrically located mutations close to a border of homology can be efficiently corrected by gene conversion. In strains SJR208, SJR249, and SJR205, the *ura3ΔNco* mutation is within 60 bp of the 3′ border of homology; in strains SJR238, SJR239, SJR240, and SJR253, the *ura3-Nhe* mutation is 52 bp from the 5′ border of homology; and in strain SJR236, the *ura3-Nhe* mutation is only 27 bp from the 3′ border of homology. The final general observation that applies to the first two groups of strains in Fig. 2 is that substrates of similar size produce Ura$^+$ recombinants at comparable rates. This generalization holds regardless of the particular point mutation present or of the positions of 5′ and 3′ homology borders. For example, the recombination rates between the 680-bp *ura3* repeats in strains SJR206, SJR238, SJR204, and SJR233 are all approximately 1.5×10^{-8}. As will be discussed further below, there is a very striking linear relationship between substrate length and recombination rate for the first two groups of strains.

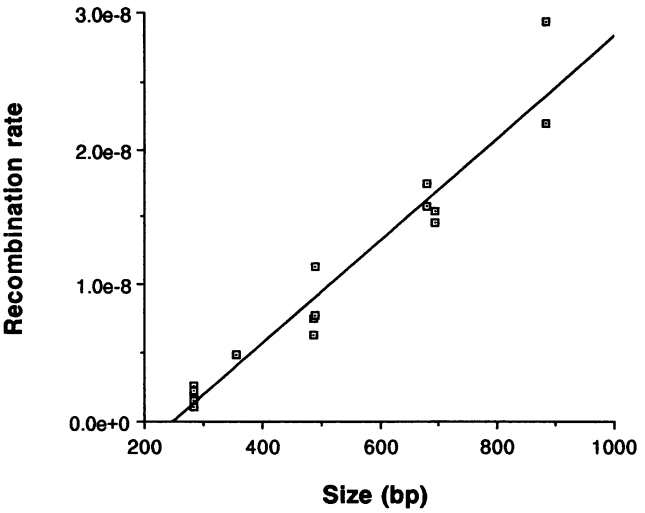


FIG. 3. Linear regression analysis of substrate size versus rate of Ura$^+$ recombinants with the heterochromosomal *ura3* repeats. The data plotted are derived from Fig. 2. Linear regression analysis was done with the computer program Cricket Graph. The equation for the line shown is R = 9.32e-9 + (3.77e-11)L, where R is the recombination rate and L is the substrate length; the correlation coefficient is 0.94. The MEPS was calculated as described by Shen and Huang (32) and is 248 bp.

In the final group of strains shown in Fig. 2, both *ura3* alleles are fused to *pGAL*. Although these 970-bp repeats are the largest repeats used, the rates of recombination events yielding Ura$^+$ prototrophs are twofold less than those seen with the 884-bp homologous *pURA* repeats (strains SJR203 and SJR232). In addition to the low recombination rates, the proportion of the events representing crossing over between chromosomes V and XV is surprisingly low; the amount of exchange is more consistent with that observed with some of the small duplications in the first two groups of strains.

The dependence of recombination rate on substrate size exhibited by the *pGAL/pURA* heterologous and *pURA* homologous promoter constructs can be used to determine the minimal efficient processing segment (MEPS) for recombination when repeats are on unlinked DNA molecules. Combining the group 1 and 2 data in Fig. 2 to plot substrate size (L) versus the rate of Ura$^+$ prototroph formation (R) yields the relationship shown in Fig. 3. Linear regression analysis was used to solve the general equation $R = a + bL$, where a is the y intercept and b is the slope. As defined by Shen and Huang (32), the MEPS is equal to $1 - a/b$. With the *ura3* repeats used here, the MEPS is calculated to be 248 bp. The SJR250 and SJR254 data were omitted from this analysis because, as noted above, the results obtained with these homologous *pGAL* constructs were not consistent with those obtained with the first two groups of strains. If one includes these data, the MEPS is calculated to be 182 bp.

In addition to the strains illustrated in Fig. 2, heterologous promoter constructs were made with substrate homologies of 152 and 80 bp (see below for similar inverted- and direct-repeat substrates). Appropriate *ura3Δ3′* alleles were put into the strain containing the *pGAL-ura3-Nhe* allele for these experiments; the *pGAL-ura3ΔNco* allele was not used since the frameshift mutation would be 3′ to the border of homology. No recombination could be detected above the

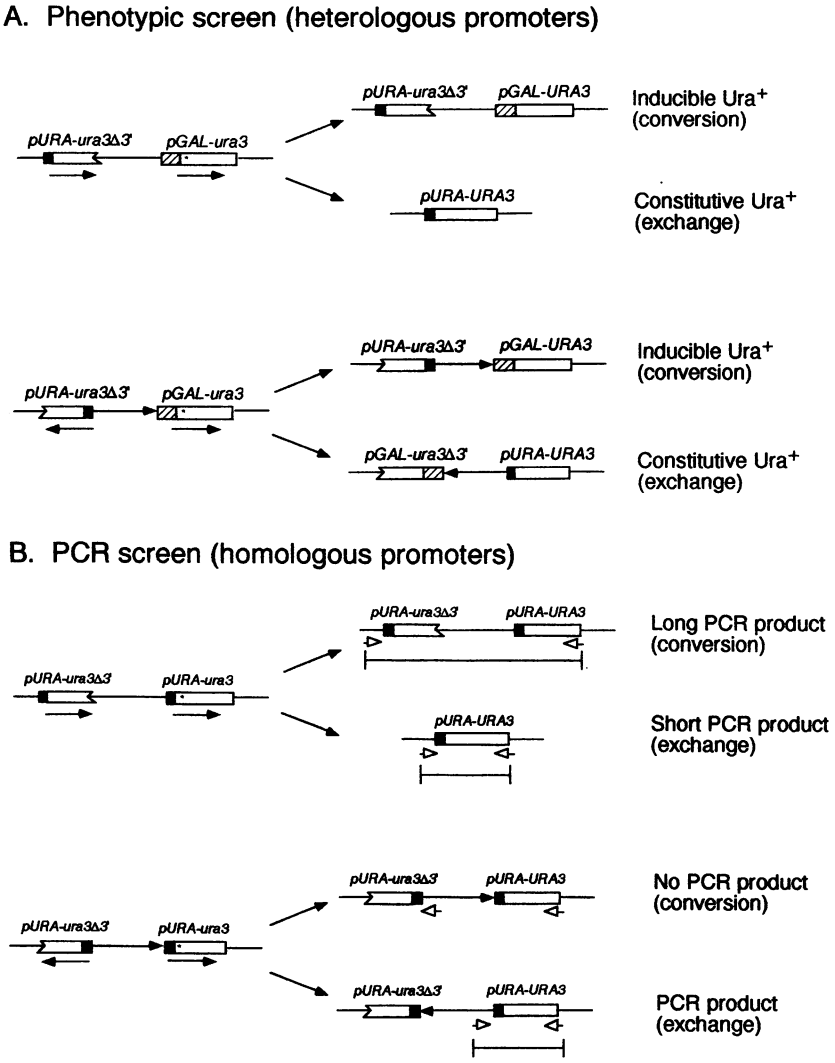
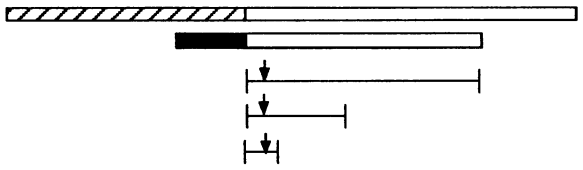


FIG. 4. Distinguishing conversion versus exchange events with *ura3* direct and inverted repeats. Open boxes represent *ura3* sequences, solid boxes correspond to *pURA*, and hatched boxes correspond to *pGAL*. The solid arrow beneath the *ura3* repeats represents the direction of transcription; the solid arrow between the inverted repeats is included to emphasize the inversion of this region when crossing over occurs between the inverted repeats; the open arrows correspond to the positions of the PCR primers used to detect crossing over between the repeats. In the case of the direct repeats, the *HIS3* locus is upstream of *pURA-ura3Δ3'*; there are approximately 2 and 2.5 kb between the homologous and heterologous promoter repeats, respectively. In the case of the inverted repeats, the *HIS3* locus is also upstream of *pURA-ura3Δ3'*, so that it is between the duplicated *ura3* sequences. As a result, there are approximately 3.3 and 3.8 kb between the inverted repeats with homologous and heterologous promoters, respectively.

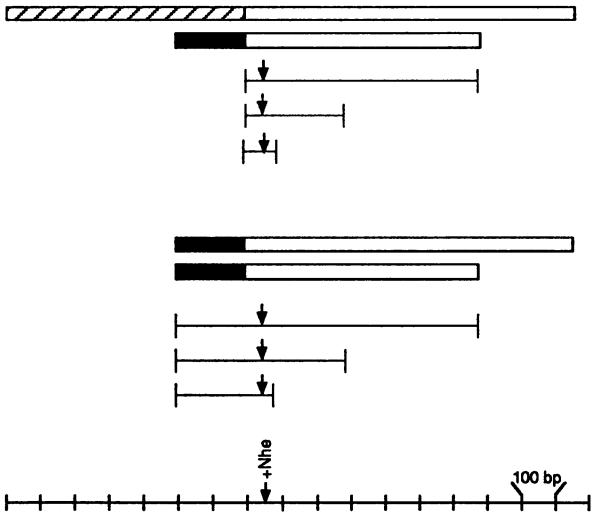
background reversion frequency of the *ura3-Nhe* allele ($<5 \times 10^{-10}$).

Effect of substrate size on recombination between nontandem direct and inverted repeats. The experiments described above examine the effect of substrate length on mitotic recombination when the interacting repeats are on nonhomologous chromosomes and hence are physically unlinked. We also constructed two additional series of isogenic strains in which the same *ura3* repeats were situated as either nontandem direct or inverted repeats on the same chromosome. As described above for the heterochromosomal repeats, strains with the repeats under control of the heterologous *pGAL/pURA* promoters were constructed as well as homologous promoter strains with both repeats transcribed from *pURA*. Following selection of Ura^+ recombinants, conversion versus crossover events in the heterologous and homologous promoter strains were assessed by using the previously described phenotypic and PCR screens, respectively. Figure 4 illustrates these screens and also the consequences of intrachromatid conversion versus exchange. Intrachromatid (or sister chromatid) conversion does not alter the sequences between either the direct or inverted *ura3* repeats. In contrast, intrachromatid exchange between direct repeats results in the deletion of all sequences between the repeats, while intrachromatid exchange between inverted repeats results in the inversion of sequences between the interacting repeats. It should be noted that sister chromatid exchange can occur between direct repeats and would, with the orientation of the alleles used here, yield a product indistinguishable from the intrachromatid exchange product. In contrast, a sister chromatid exchange involving the inverted repeats would yield acentric and dicentric chroma-

A. Direct repeats



B. Inverted repeats



Strain	Size (bp)	Ura+ Rate (10^{-8})	% exchange
SJR265	680	495	95.7 (90/94)
SJR267	285	127	94.6 (105/111)
SJR269	80	5	98.5 (130/132)
SJR286	884	1210	86.0 (86/100)
SJR288	489	321	87.0 (87/100)
SJR290	284	29	87.0 (87/100)
SJR266	680	258	69.6 (78/112)
SJR268	285	27	44.1 (49/111)
SJR270	80	0.5	20.0 (16/81)
SJR287	884	563	47.0 (47/100)
SJR289	489	175	30.0 (30/100)
SJR291	284	11	42.3 (41/97)

FIG. 5. Recombination data obtained with use of *ura3* direct and inverted repeats on chromosome V. Open boxes represent *ura3* sequences, solid boxes correspond to *pURA*, and hatched boxes correspond to *pGAL*. The lines below the boxes correspond to the length of substrate homology present in each strain; vertical arrows over these lines represent the position of the *ura3-Nhe* point mutation in the full-length *ura3* allele. The actual number of exchanges among the total number of independent Ura$^+$ recombinants examined is given in parentheses following the percent exchange value.

tids; the resulting lethality precludes detection of these events. Intrachromatid exchanges were readily detected with both the direct- and inverted-repeat constructs, indicating that the 2-kb genomic region immediately upstream of the *URA3* locus is dispensable for growth.

The relevant features of the direct- and inverted-repeat strains are summarized in Fig. 5, which also presents the recombination data obtained with these strains. The strains are divided into four groups: direct repeats with heterologous promoters, direct repeats with homologous promoters, inverted repeats with heterologous promoters, and inverted repeats with homologous promoters. The full-length *ura3* gene in all of the strains shown in Fig. 5 contains the *ura3-Nhe* allele. As seen with the heterochromosomal repeats, the rate of Ura$^+$ recombinant formation is directly proportional to the length of sequence homology for both the direct and inverted repeats. While the rates of recombination between direct and inverted repeats of a given size are similar, the rates of recombination for the direct repeats are consistently higher than those for the inverted repeats.

Plotting the substrate size versus the rate of Ura$^+$ prototroph production for the direct and inverted repeats separately yields a linear relationship for each (Fig. 6). The MEPS for the direct repeats is calculated to be 271 bp, while

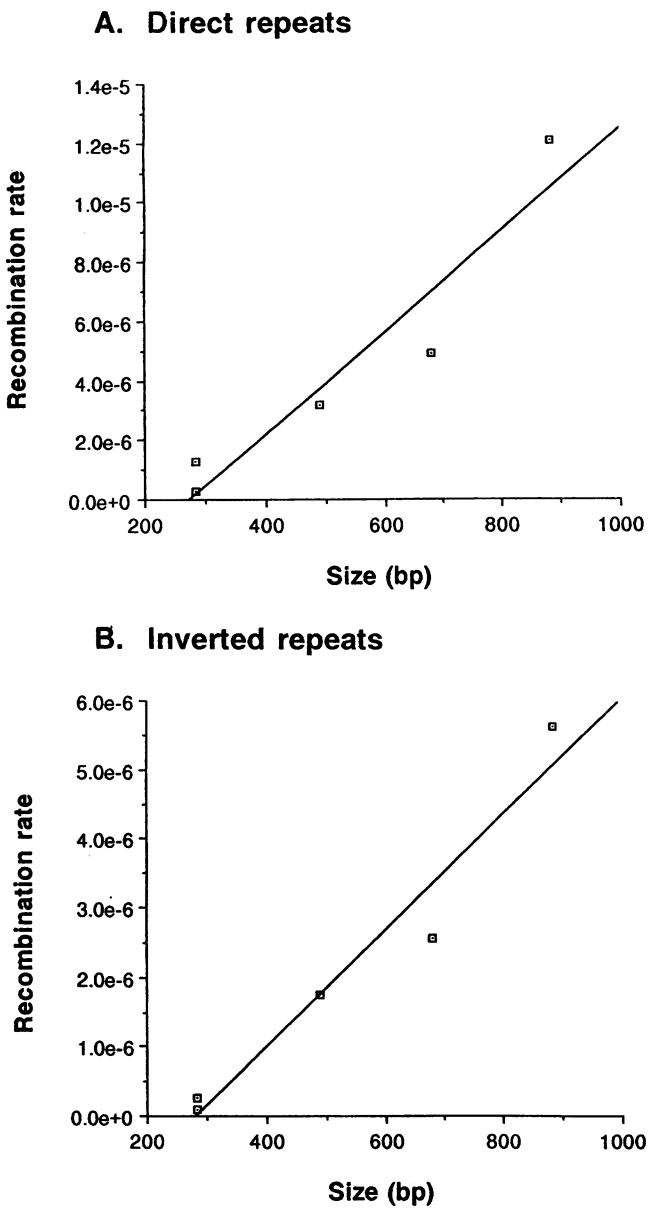


FIG. 6. Linear regression analysis of substrate length versus rate of Ura$^+$ recombinants with the direct and inverted *ura3* repeats. The data plotted were derived from Fig. 5 and were analyzed as described for Fig. 3. The relationship between the direct-repeat size and recombination rate is described by the equation R = 4.64e-6 + (1.71e-8)L; the correlation coefficient is 0.91. The relationship between the inverted-repeat size and recombination rate is described by the equation R = 2.35e-6 + (8.43e-9)L; the correlation coefficient is 0.95. The calculated MEPS for the direct and inverted repeats are 271 and 280 bp, respectively.

that for the inverted repeats is 280 bp. The calculated MEPS predicts that recombination should be relatively inefficient with repeats smaller than approximately 275 bp. The observed rapid decline in recombination rate when the substrates are reduced from 285 to 80 bp is consistent with this prediction. It should be noted that the data obtained with the 80-bp duplications were omitted for the linear regression analyses shown in Fig. 6. If these data are included, the MEPS for the direct and inverted repeats are 197 and 205 bp, respectively. Since the 80-bp data fall well below the MEPS, regardless of whether they are included in the linear regression analysis, their omission is justified.

In addition to the repeats shown in Fig. 5, 37-bp direct and inverted repeats with *pGAL-ura3ΔNco* as the full-length repeat were constructed. Only crossover events can be detected with these duplications since the 3′ border of homology is 5′ of the *ura3*ΔNco lesion. Our failure to detect any recombination with either of the 37-bp repeats above the background reversion frequency of *ura3ΔNco* ($<5 \times 10^{-10}$) indicates that the rate of recombination is at least 2 orders of magnitude less than that seen with the 80-bp duplications. This observation is again consistent with the notion that the recombination rate plummets when substrate lengths are less than approximately 275 bp.

A final point to be made about the direct and inverted repeats concerns the distribution of independent Ura$^+$ recombinants into conversion versus crossover events. The inverted repeats behave similarly to the heterochromosomal repeats in this respect, with the proportion of events that are exchanges decreasing as the substrate size decreases. In contrast, the proportion of direct-repeat recombination events that are exchanges does not change as a function of substrate size. It is also noteworthy that greater than 85% of recombination events involving the direct repeats are exchanges, suggesting that the mechanism of recombination between the direct repeats may differ from that involving the inverted and heterochromosomal repeats.

DISCUSSION

We have examined the effect of substrate length on mitotic recombination, using isogenic haploid yeast strains containing artificial *ura3* repeats ranging in size from 80 to 960 bp. In each strain, one of the *ura3* alleles was full length and contained a defined frameshift mutation; the other allele had variable amounts of coding sequence deleted from the 3′ end. The repeats were constructed so that no extraneous plasmid sequences were introduced into the yeast genome and were positioned either on nonhomologous chromosomes (heterochromosomal repeats) or as direct or inverted repeats on the same chromosome (intrachromosomal repeats). Recombination rates were estimated by measuring the rate of Ura$^+$ prototroph formation in each strain. In addition, independent Ura$^+$ recombinants were isolated, and each was classified as either a noncrossover (conversion) or crossover (exchange) event. In the discussion that follows, no explicit assumptions are made concerning the initiation of mitotic recombination events (i.e., a single-strand nick versus a double-strand break). We do, however, assume the formation of a heteroduplex DNA intermediate, an assumption that is compatible with all current recombination models.

The results obtained with the heterochromosomal repeats (Fig. 2) clearly demonstrate that the total rate of recombination decreases as the substrate size decreases. One of the most striking aspects of these data is that similarly sized substrates produce Ura$^+$ recombinants at comparable rates, regardless of the absolute positions of the homology borders, the position or identity of the frameshift allele, or the distribution of events into conversions versus exchanges (see below for further discussion). One can thus combine the data in Fig. 2 and plot substrate size versus recombination rate to obtain the linear relationship shown in Fig. 3. As originally pointed out by Shen and Huang (32), such a linear relationship can be used to determine what they termed the

minimal efficient processing segment, or MEPS. They suggested that any segment of DNA can be considered as an overlapping series of MEPS; the longer the segment, the more MEPS it contains and hence a linear relationship between substrate size and recombination rate is observed. We have interpreted our data in a similar manner and have estimated the MEPS for yeast repeats on nonhomologous chromosomes to be approximately 250 bp.

One aspect of the heterochromosomal data that deserves mention is the observation that recombination rates seem to be related only to substrate size. For example, although the recombination substrates in strains SJR238 and SJR204 are approximately 680 bp in length, the endpoints of the homologous segments differ, the natures of the point mutations differ, and the relative positions of the point mutations within the homologous segments differ (Fig. 2). In addition, mutations within 30 bp of a homology border appear to convert as efficiently as centrally located mutations. These results can be rationalized if one assumes that the MEPS reflects the homology necessary to efficiently initiate an interaction and that once an event is initiated, the accompanying heteroduplex will extend some average length. If the size of the duplication being examined is less than this average heteroduplex extension length, then regardless of where within the duplication an event initiates, heteroduplex will extend essentially all the way to one (or possibly both) of the homology borders. There will thus be little bias against inclusion of an asymmetrically positioned marker in the heteroduplex that forms. If, however, the substrate is substantially larger than the average length of heteroduplex in recombination intermediates, then some position-related bias in the conversion of markers would be expected. This type of model also predicts that most crossover events should occur close to the homology borders, a prediction that the current system was not designed to test. Judd and Petes (13) have actually measured mitotic conversion tract lengths at the *URA3* locus and reported that 90% were 650 bp or longer; almost all of the duplications used in the present study are smaller than the inferred minimal average length of mitotic heteroduplex at *URA3*. Close proximity to a homology border may, of course, eventually preclude either inclusion of a marker in heteroduplex or the correction of the corresponding mismatch. That such a limit may exist comes from the observation that a marker 13 bp from a homology border can convert whereas a marker 5 bp from a homology border apparently cannot (2).

In addition to the effect of substrate length on the total recombination rate, the heterochromosomal data in Fig. 2 indicate that within a given subgroup of strains, the proportion of events resolved as crossovers decreases as the substrate size decreases. If one takes a very simplistic view of crossing over and does not factor in associated conversion events, then one would predict that all exchanges should occur on the 3′ side of the point mutation in order to yield a full-length, wild-type gene. It follows that as the deletion endpoint encroaches on the point mutation from the 3′ side, the exchange frequency should decrease. Although this is exactly the result obtained if each subgroup is examined individually, it should be noted that there is no strict relationship between the distance of a mutation from the 3′ homology border and the frequency of crossing over. The frameshift mutation in strain SJR236 is only 27 bp from the 3′ homology border and yet almost 40% of the recombinants represent crossover events. This proportion of crossing over is as high as that seen in some strains with a centrally located mutation that is 400 bp from either homology border (e.g., SJR203). It seems likely from considerations of this sort that a substantial proportion of crossing over must be occurring on the 5′ side of the mutations used here and therefore must be accompanied by a conversion of the frameshift mutation to wild-type sequence.

The fate of frameshift mutations in the presumed heteroduplex intermediate is of obvious importance: whether the resulting mismatch is efficiently recognized by the mismatch repair machinery and whether there is a bias in the direction of mismatch repair. At best, the data in Fig. 2 indicate only that heteroduplex involving *ura3ΔNco* is more efficiently repaired to wild type than is heteroduplex involving *ura3-Nhe* (e.g., compare the conversion frequencies resulting in a Ura^+ phenotype in strains SJR206 and SJR233). The observed discrepancy might result from differential repair patterns of heteroduplex recombination intermediates containing a four-base as opposed to a one-base loop-out (5, 21). It should be noted that no difference in the conversion rates of the *ura3ΔNco* and *ura3-Nhe* alleles would be expected if most conversions resulted from gap repair.

Since the type of repeat system used for this study sets an artificial upper limit on the length of heteroduplex that can form, our data suggest that the resolution of recombination events may be influenced by the length of heteroduplex. Consistent with results described here, it has been reported that there is a positive correlation between mitotic coconversion tract lengths and the resolution of events as crossovers in yeast cells (1, 3). In principle, gene conversion between repeated sequences requires only that the heteroduplex intermediate persist long enough to be acted upon by the mismatch repair machinery. If one assumes that interactions become progressively more unstable as the length of heteroduplex is limited and that only very stable interactions have the potential to be resolved as crossover events, then one would predict that crossing over should be more sensitive to substrate size than should gene conversion. Alternatively, if recombination initiates with a double-strand break and the intermediate contains two Holliday junctions, then the distance between the junctions might affect the stability and/or resolution of the intermediate.

On the basis of substrate size, the homologous *pGAL* strains (SJR250 and SJR254 in Fig. 2) exhibited unexpectedly low rates of recombination, and very little crossing over between the *ura3* alleles was observed. Given the reports that transcription increases mitotic recombination rates in *S. cerevisiae* (39, 44), it was possible that growth in noninducing YEP/GE medium prior to selective plating might have influenced the recombination rates in SJR250 and SJR254. Pregrowth in galactose-containing medium, however, had very little effect on the total rate of Ura^+ prototroph production (less than 50%; data not shown). The position of the 5′ border of homology as well as the particular sequences duplicated might be expected to influence the resolution of interactions. We speculate that the resolution of Holliday junctions in the *pGAL* region might be strongly biased to yield noncrossover products (crossing over in this region is not possible with the heterologous promoter constructs) and that the loss of recombinants in the crossover class could account for the relatively low recombination rates.

The same repeats used in the heterochromosomal recombination studies discussed above were also positioned as nontandem direct or inverted repeats on the same chromosome (Fig. 4 to 6). While the general effects of substrate length on heterochromosomal versus intrachromosomal recombination are similar, the absolute rates of heterochromosomal versus intrachromosomal recombination events differ

by at least 2 orders of magnitude. It was expected that mitotic recombination events involving physically linked heteroalleles would be more frequent than those involving the same heteroalleles positioned on separate DNA molecules (17), but the magnitude of the difference in rates was somewhat surprising. It would be interesting to systematically move the intrachromosomal repeats farther apart and to determine how the distance between repeats affects the rate of recombination and whether there is a distance at which physical linkage ceases to be advantageous. One interpretation of the large difference between the intrachromosomal and heterochromosomal recombination rates is that the homology search rather than an initiating event may be the major factor limiting interchromosomal mitotic recombination. If the initiating event occurs during or after chromosome replication, then sister chromatid interactions could be effectively limiting interactions between nonhomologous chromosomes. Finally, although we assume that the initiating event for intrachromosomal versus heterochromosomal recombination is the same, there is no direct evidence for this.

As was seen with the heterochromosomal repeats, the rates of recombination between the direct and inverted repeats, when considered separately, are directly proportional to substrate length. It is particularly noteworthy that the MEPS for the intrachromosomal repeats (271 and 280 bp for the direct and inverted repeats, respectively) are very close to the estimated MEPS (248 bp) for the heterochromosomal repeats, in spite of the fact that the absolute rates of Ura^+ recombinants are very different. Thus, the minimal length of homology necessary for efficient interactions is a constant, regardless of the position of repeats. The rates of intrachromosomal recombination are such that recombination can be detected with substrate lengths less than the MEPS. We could easily detect recombination between 80-bp substrates but were unable to detect recombination between 37-bp substrates (data not shown), an observation which is consistent with the notion that recombination becomes very inefficient with substrate lengths less than the MEPS.

The direct and inverted repeats used here detect intrachromatid gene conversion events, intrachromatid exchange events, and sister chromatid conversion events. In addition, there are two types of interaction possible with the direct repeats that cannot be detected with the inverted repeats: intrachromatid single-strand annealing (a nonreciprocal type of event [18, 37, 41]) and sister chromatid exchange. In the present study, recombination between the direct repeats differs from that involving inverted repeats in two ways. First, for a given substrate size, the rate of recombination between direct repeats is always slightly greater than that between inverted repeats. This may be related to the additional recombination mechanisms available for the direct repeats or could reflect the different distances between the direct versus inverted repeats (2.0 to 2.5 kb and 3.3 to 3.8 kb, respectively). Second, more than 85% of the events involving the direct repeats represent crossing over, irrespective of substrate length. This observation suggests that many of the direct-repeat events may be generated by the single-strand annealing mechanism, a nonreciprocal mechanism that cannot yield viable recombinants with either the heterochromosomal or inverted repeats.

The results obtained with heterochromosomal repeats are relevant to previously reported studies in *S. cerevisiae* that were designed to answer questions similar to those addressed here. Yuan and Keil (45) examined the effect of substrate size on mitotic intrachromosomal recombination events involving nontandem *his4* direct repeats ranging in size from 0.23 to 12.9 kb. They observed that recombination rates leveled off with repeats greater than 1 kb, which is outside the size range of repeats used in the present study, and dropped dramatically when repeat size was decreased from 0.75 to 0.23 kb. The drop in recombination they observed (24-fold) in comparing the smallest repeats is consistent with the *ura3* repeat data in Fig. 2. The effect of substrate length on the rate of mitotic intrachromosomal gene conversion involving plasmid-encoded *his3* repeats has been examined by Ahn et al. (2). In this system, substrate length had an exponential effect on gene conversion, with a 10-fold decrease in substrate size resulting in an astounding 1,000-fold decrease in recombination rate. Although the reason for the discrepancy between our results and those of Ahn et al. is not known, three possible explanations can be entertained. First, since only two of the six substrates used by Ahn et al. were longer than the MEPS deduced from our experiments, it may be that the data of Ahn et al. primarily reflect the relatively inefficient recombination that occurs with substrate lengths less than the MEPS. Second, the parameters governing intramolecular plasmid recombination (2) may be different from those affecting chromosomal recombination (our studies). Third, the differences observed in the two studies may simply be a reflection of the different substrates (*ura3* versus *his3*) used to assay recombination.

In addition to the spontaneous yeast recombination studies cited above, Sugawara and Haber (37) have examined homology dependence in a competition assay in which mitotic recombination events were initiated with an HO-induced double-strand break. In this system, a linear relationship between substrate size and recombination frequency was observed with repeats greater than 89 bp. The relevance of HO-initiated events to the spontaneous events examined here is not clear.

In addition to the yeast studies, there is some information available concerning the effect of substrate length on homologous recombination in prokaryotes and in mammalian cells. Studies done in *E. coli* (32, 34, 43) indicate that the MEPS in bacterial cells is approximately 50 bp. With substrate lengths greater than the MEPS, a linear relationship between size and recombination rate is observed; substrates smaller than the MEPS can recombine but do so relatively inefficiently. In studies done using mammalian cells (19, 29), the MEPS appears to be in the 250-bp range; below 250 bp, recombination is inefficient but can occur with as little as 14 bp of homology. The MEPS for yeast recombination determined in this study (250 bp) is larger than that determined in bacterial studies but is similar to that found in mammalian cells. Although it has been suggested that the MEPS might increase as genome complexity increases in order to prevent ectopic interactions (40), the available data concerning yeast and mammalian cells indicate that this probably is not the case. The data are consistent, however, with there being a highly conserved mechanism of recombination in eukaryotes that has different substrate requirements from that found in prokaryotes. It should be noted that all of the substrate length experiments done to date have examined mitotic recombination events. Use of *S. cerevisiae* as an experimental system, however, will allow an examination of the comparable meiotic events, both with respect to the general relationship between substrate size and recombination and with respect to the relationship between substrate size and the resolution of recombination intermediates.

The data reported here are relevant to considerations of the rate of ectopic interactions between naturally occurring

repeated sequences as well as to considerations of general features of recombination. First, we have precisely defined the relationship between substrate size and recombination rate for intrachromosomal and heterochromosomal repeats in the 80- to 900-bp range. The relationship is linear for yeast repeats greater than approximately 250 bp in size; the rate decreases rapidly for repeats smaller than the MEPS. Second, crossing over appears to be more sensitive to substrate size than does gene conversion. If the same is true of naturally occurring repeats, then interactions between small repeats would be less likely to be detrimental (i.e., result in chromosomal rearrangements) than those involving very large repeats. Finally, intrachromosomal repeats can recombine at rates 2 orders of magnitude greater than those for repeats on nonhomologous chromosomes. On the basis of this observation and those of Lichten and Haber (17), one would predict that members of a repetitive gene family would be more similar in sequence when they are located close together on the same chromosome than when they are on nonhomologous chromosomes, or perhaps even far apart on the same chromosome. Despite the great difference in intrachromosomal versus heterochromosomal recombination rates, it is particularly noteworthy that the experimentally determined MEPS is essentially the same and is not influenced by the position of repeats.

ACKNOWLEDGMENTS

We acknowledge the excellent technical assistance of Tamara Murphy and contributions made by Qinghua Feng during a laboratory rotation. We thank Abhijit Datta for helpful comments on the manuscript and David Steele for discussion during the course of this project.

This work was supported by NIH grant GM38464 to S.J.-R.

REFERENCES

1. **Aguilera, A., and H. L. Klein.** 1989. 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.
2. **Ahn, B.-Y., K. J. Dornfeld, T. J. Fagrelius, and D. M. Livingston.** 1988. Effect of limited homology on gene conversion in a *Saccharomyces cerevisiae* plasmid recombination system. Mol. Cell. Biol. **8**:2442–2448.
3. **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.
4. **Baltimore, D.** 1981. Gene conversion: some implications for immunoglobulin genes. Cell **24**:592–594.
5. **Bishop, D. K., and R. D. Kolodner.** 1986. Repair of heteroduplex plasmid DNA after transformation into *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:3401–3409.
6. **Boeke, J. D., J. Trueheart, G. Natsoulis, and G. R. Fink.** 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. **154**:164–175.
7. **Carpenter, A. T. C.** 1987. Gene conversion, recombination nodules, and the initiation of meiotic synapsis. BioEssays **6**:232–236.
8. **Edelman, G. M., and J. A. Gally.** 1970. Arrangement and evolution of eukaryotic genes, p. 962–972. *In* F. O. Schmitt (ed.), The neurosciences: second study program. Rockefeller University Press, New York.
9. **Hoffman, C. S., and F. Winston.** 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene **57**:267–272.
10. **Holliday, R.** 1964. A mechanism for gene conversion in fungi. Genet. Res. **5**:282–304.
11. **Jinks-Robertson, S., and T. D. Petes.** 1986. Chromosomal translocations generated by high-frequency meiotic recombination between repeated yeast genes. Genetics **114**:731–752.
12. **Johnston, M., and R. W. Davis.** 1984. Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:1440–1448.
13. **Judd, S. R., and T. D. Petes.** 1988. Physical lengths of meiotic and mitotic gene conversion tracts in *Saccharomyces cerevisiae*. Genetics **118**:401–410.
14. **Kupiec, M., and T. D. Petes.** 1988. Meiotic recombination between repeated transposable elements in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:2942–2954.
15. **Lea, D. E., and C. A. Coulson.** 1949. The distribution of the numbers of mutants in bacterial populations. J. Genet. **49**:264–284.
16. **Lichten, M., R. H. Borts, and J. E. Haber.** 1987. Meiotic gene conversion and crossing over between dispersed homologous sequences occurs frequently in *Saccharomyces cerevisiae*. Genetics **115**:233–246.
17. **Lichten, M., and J. E. Haber.** 1989. Position effects in ectopic and allelic mitotic recombination in *Saccharomyces cerevisiae*. Genetics **123**:261–268.
18. **Lin, F.-L., K. Sperle, and N. Sternberg.** 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol. Cell. Biol. **4**:1020–1034.
19. **Liskay, R. M., A. Letsou, and J. L. Stachelek.** 1987. Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. Genetics **115**:161–167.
20. **Meselson, M. S., and C. M. Radding.** 1975. A general model for genetic recombination. Proc. Natl. Acad. Sci. USA **72**:358–361.
21. **Muster-Nassal, C., and R. Kolodner.** 1986. Mismatch correction catalyzed by cell-free extracts of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **83**:7618–7622.
22. **Petes, T. D., and C. W. Hill.** 1988. Recombination between repeated genes in microorganisms. Annu. Rev. Genet. **22**:147–168.
23. **Petes, T. D., R. E. Malone, and L. S. Symington.** 1991. Recombination in yeast, p. 407–521. *In* J. R. Broach, J. R. Pringle, and E. W. Jones (ed.), The molecular biology of the yeast *Saccharomyces*: genome dynamics, protein synthesis and energetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Radman, M.** 1991. Avoidance of inter-repeat recombination by sequence divergence and a mechanism of neutral evolution. Biochimie **73**:357–361.
25. **Rose, M., and D. Botstein.** 1983. Structure and function of the yeast *URA3* gene: differentially regulated expression of hybrid β-galactosidase from overlapping coding sequences in yeast. J. Mol. Biol. **170**:883–904.
26. **Rose, M., P. Grisafi, and D. Botstein.** 1984. Structure and function of the yeast *URA3* gene: expression in *Escherichia coli*. Gene **29**:113–124.
27. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics: a laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
28. **Rothstein, R.** 1991. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. **194**:281–301.
29. **Rubnitz, J., and S. Subramani.** 1984. The minimum amount of homology required for homologous recombination in mammalian cells. Mol. Cell. Biol. **4**:2253–2258.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Schiestl, R. H., and R. D. Gietz.** 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. **16**:339–346.
32. **Shen, P., and H. V. Huang.** 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics **112**:441–457.
33. **Sherman, F.** 1991. Getting started with yeast. Methods Enzymol. **194**:3–20.
34. **Singer, B. S., L. Gold, P. Gauss, and D. H. Doherty.** 1982.

Determination of the amount of homology required for recombination in bacteriophage T4. Cell **31**:25–33.
35. **Smithies, O., and P. A. Powers.** 1986. Gene conversions and their relation to homologous chromosome pairing. Philos. Trans. R. Soc. Lond. B **312**:291–302.
36. **Steele, D. F., and S. Jinks-Robertson.** Time-dependent mitotic recombination in *Saccharomyces cerevisiae*. Curr. Genet., in press.
37. **Sugawara, N., and J. E. Haber.** 1992. Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. Mol. Cell. Biol. **12**:563–575.
38. **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.
39. **Thomas, B. J., and R. Rothstein.** 1989. Elevated recombination rates in transcriptionally active DNA. Cell **56**:619–630.
40. **Thomas, C. A., Jr.** 1966. Recombination of DNA molecules. Prog. Nucleic Acid Res. Mol. Biol. **4**:315–348.
41. **Voelkel-Meiman, K., and G. S. Roeder.** 1990. Gene conversion tracts stimulated by *HOT1*-promoted transcription are long and continuous. Genetics **126**:851–867.
42. **Wallis, J. W., G. Chrebet, G. Brodsky, M. Rolfe, and R. Rothstein.** 1989. A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. Cell **58**:409–419.
43. **Watt, V. M., C. J. Ingles, M. S. Urdea, and W. J. Rutter.** 1985. Homology requirements for recombination in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82**:4768–4772.
44. **White, M. A., P. Detloff, M. Strand, and T. D. Petes.** 1992. A promoter deletion reduces the rate of mitotic, but not meiotic, recombination at the *HIS4* locus in yeast. Curr. Genet. **21**:109–116.
45. **Yuan, L.-W., and R. L. Keil.** 1990. Distance-independence of mitotic intrachromosomal recombination in *Saccharomyces cerevisiae*. Genetics **124**:263–273.