Expansions and Contractions in 36-bp Minisatellites by 
Gene Conversion in Yeast

Frédéric Pâques,*† Guy-Franck Richard*,‡ and James E. Haber*

*Rosenstiel Center and Department of Biology, Brandeis University, Waltham, Massachusetts 02454-9110, †Institut Jacques Monod, UMR 7592, Université Paris 6 et Paris 7, 75251 Paris Cedex 05, France and ‡Unité de Génétique Moléculaire des Levures, UMR 1300 CNRS and UFR 927 Université Pierre et Marie Curie, Paris, Institut Pasteur, 75724 Paris Cedex 15, France

Manuscript received December 12, 2000
Accepted for publication February 10, 2001

ABSTRACT

The instability of simple tandem repeats, such as human minisatellite loci, has been suggested to arise by gene conversions. In Saccharomyces cerevisiae, a double-strand break (DSB) was created by the HO endonuclease so that DNA polymerases associated with gap repair must traverse an artificial minisatellite of perfect 36-bp repeats or a yeast Y′ minisatellite containing diverged 36-bp repeats. Gene conversions are frequently accompanied by changes in repeat number when the template contains perfect repeats. When the ends of the DSB have nonhomologous tails of 47 and 70 nucleotides that must be removed before repair DNA synthesis can begin, 16% of gene conversions had rearrangements, most of which were contractions, almost always in the recipient locus. When efficient removal of nonhomologous tails was prevented in rad1 and msh2 strains, repair was reduced 10-fold, but among survivors there was a 10-fold reduction in contractions. Half the remaining events were expansions. A similar decrease in the contraction rate was observed when the template was modified so that DSB ends were homologous to the template; and here, too, half of the remaining rearrangements were expansions. In this case, efficient repair does not require RAD1 and MSH2, consistent with our previous observations. In addition, without nonhomologous DSB ends, msh2 and rad1 mutations did not affect the frequency or the distribution of rearrangements.

We conclude that the presence of nonhomologous ends alters the mechanism of DSB repair, likely through early recruitment of repair proteins including Msh2p and Rad1p, resulting in more frequent contractions of repeated sequences.

MINISATELLITES are tandem repeats of a few dozen nucleotides that display an unusually high rate of instability, manifested by changes in number of tandem repeats. In humans, these changes arise during germline formation (Armour and Jeffreys 1992; Jeffreys et al. 1994; Richards and Sutherland 1997) and some alleles are found to change in as many as 13% of the gametes (Vergnaud et al. 1991). As with expansions of microsatellite sequences, expansions of minisatellites have been associated with human disease (Lalioti et al. 1997; Virtaneva et al. 1997) and with chromosome fragility (Yu et al. 1997; Hewett et al. 1998).

Human minisatellite instability apparently arises through gene conversion events during or shortly after meiosis, many of which involve interallelic transfers of information (Baard and Vergnaud 1994; Jeffreys et al. 1994; May et al. 1996; Baard et al. 1998). Most likely these events result from the gene conversion repair of double-strand breaks (DSBs), as recent evidence suggests that meiotic recombination in mammals as well as yeast is initiated by the Spo11p endonuclease (Bergerat et al. 1997; Keeney et al. 1997; Baudat et al. 2000; Romanienko and Camerini-Otero 2000). Similar intranucleotide recombination events are found in MS32 and CEB1 minisatellite sequences, when they are placed close to a meiotic hotspot in Saccharomyces cerevisiae (Appelgren et al. 1997, 1999; Debrauwère et al. 1999). Recently, a natural yeast minisatellite was also shown to be unstable during meiosis (Bishop et al. 2000).

We previously studied frequent rearrangements of repeated sequences associated with mitotic gene conversion in budding yeast, using both 375-bp repeats and trinucleotide repeats (Pâques et al. 1998; Richard et al. 1999, 2000). The experimental system was similar to the one shown in Figure 1. The HO-cleaved ends of the DSB can invade homologous regions on a plasmid and then DNA polymerases can traverse the intervening region on the template that may contain various repeated or unique sequences. We consistently found a much higher rate of both expansions and contractions of the repeated sequences during gene conversion than during replication. Moreover, the rearranged array was nearly always found in the recipient (repaired) molecule, suggesting that the gene conversion events occur by a synthesis-dependent strand annealing (SDSA) pathway, where nearly all newly synthesized DNA sequences are predicted to be in the recipient (Figure 2, E–H and I–M).

In this study, we provide direct evidence for DSB-
induced rearrangements of minisatellite sequences in yeast, using arrays of either perfect or imperfect 36-bp repeats. Both expansions and contractions of minisatel-
lites were induced by recombination, but their ratio, as well as the overall rearrangement frequency, is affected by the presence of nonhomologous sequences sur-
rounding the DSB, and by mutations in the RAD1 and MSH2 genes affecting the removal of such sequences. We suggest that the early recruitment of Rad1p and Msh2p and associated repair proteins needed to trim off the nonhomologous DSB ends affects the sub-
sequent steps of DNA repair, in a way that favors the contraction pathway.

MATERIALS AND METHODS

Strains: The S. cerevisiae strains studied in this work all derive from YFP17 (Pâques et al. 1998), which contains a GAL::HO fusion inserted into the chromosomal ADE3 locus (Sandell and Zakian 1993), a deletion of the HO endonuclease cleavage site in the MAT locus, and a 117-bp HO cut site in the KpnI site of the LEU2 gene on chromosome III. We used a one-step disruption method described by Wach et al. (1994) to knock out the MSH2, MSH6, PMS1, and RAD1 genes in YFP17 with a KAN gene. All transformations were performed with the one-step method described by Chen et al. (1992).

Plasmids: A series of five plasmids, described in Figure 1, were derived from Ted, a centromeric plasmid marked by the URA3 gene (provided by W. Kramer). In pFP14 (Figure 3C), a genomic XbaI-SalI fragment including the LEU2 gene was inserted into the polylinker of a URA3-marked centromeric plasmid, as described by Pâques et al. (1998). Part of the Bluescript (Stratagene, La Jolla, CA) polylinker, including XbaI, BamHI, and XhoI, was inserted into the KpnI site of LEU2, resulting in plasmid pFP36, to allow the subsequent insertion of the repeated arrays shown in Figure 1. Insertion of a XbaI-
XhoI fragment from plasmid pPS8.4 (Robine et al. 1996), containing eight copies of the Escherichia coli lac operator, resulted in pFP59 (Figure 1D). Insertion of a dimer of the former insert resulted in pFP58 (Figure 1E). To obtain pFP46 (Figure 1F), a 530-bp fragment containing the natural yeast minisatellite arrays found in Y′ subtelomeric regions (Horowitz and Haber 1984) was amplified by PCR and cloned into pFP36. In pFP225 (Figure 1G), two half-HO cut sites (solid boxes) on each side of the repeated array restore perfect homology to the template.

DSB induction and characterization of recombinants: YEPD and synthetic dropout media used for the growth of S. cerevisiae were made according to Sherman et al. (1986). YEP-glycerol contains 2% galactose (wt/vol) instead of glucose as a carbon source. YEP-glycerol contains 2% glycerol (wt/vol) instead of glucose. Yeast were grown for 24 hr in YEPD, or in synthetic medium lacking uracil or tryptophan if plasmid selection was required. This culture was then used to inoculate 50 ml of YEP-glycerol, at an initial concentration of 10⁶ cells/ml. The YEP-glycerol culture was grown overnight, to a final concentration of 1–5 × 10⁷ cells/ml, to prepare the cells for galactose

![Figure 1](https://example.com/figure1.png)

**Figure 1.**—DSB-induced gene conversions involving repeated sequences. (A) Using the HO endonuclease, one double-strand break is delivered per cell, in the **leu2** gene, which has been modified to contain an HO cut site. (B) Repair by nonhomologous recombination was assessed with cells containing no sequence homologous to **LEU2**. (C) Repair by gene conversion when the homologous donor template carries **LEU2**. (D–F) Three other homologous templates were designed, each one with a repeated array inserted within the **leu2** copy, at the exact site where the DSB is formed in the **leu2** chromosomal copy. The repeated arrays are as follows: eight identical 36-bp repeats in pFP59 (D); two copies of the same repeat in the same orientation, surrounding a 34-bp-long piece of polylinker in pFP58 (E); and a natural yeast minisatellite locus corresponding to 12 diverged copies of a 36-bp repeat in pFP46 (F). In pFP225 (G), two half-HO cut sites (solid boxes) on each side of the repeated array restore perfect homology to the template. The DSB repair efficiencies are listed next to each construction. Oligonucleotides used to characterize the recombinants are shown.
induction. Then, cells were plated on YEPD and YEP-galactose plates, at a concentration of ~200 cells/plate. In the absence of any DSB, colonies appear on YEPD and YEP-galactose with the same efficiency (not shown). For strains with an HO cut site in the chromosomal LEU2 gene, DSB repair efficiency was scored as the ratio of the number of colonies on YEP-galactose to that on YEPD. Independent colonies were patched, and the patch was used to inoculate 2-ml cultures. PCR was performed directly on cells: about one-tenth of a 3-day-old colony was boiled 5 min in the PCR mixture, and then 5 units of Taq polymerase were added; PCR involved 35 cycles, including 1 min at 94°C, 2 min at 42°C, and 4 min at 65°C for elongation. For precise rearrangement mapping (Figure 3B), PCRs were done on DNA and not cells. The DNA sequences of the oligonucleotides shown in Figure 1 are: TCATTTAATTGGTGCTGCTATC (oligo 1), GATAAGTCTAAAGAGAGTCGGATGC (oligo 2), TTGCAGATTCCCTTTTATGGATTCC (oligo 3), and GCTGCTTCTAATGCAAGGATCG (oligo 4).

For statistical analysis, we used Fisher’s exact test. With pFP59 and pFP225 in the wild-type, rad1, and msh2 backgrounds, two independent experiments were performed. We first tested the homogeneity of two sets of results, and then pooled all the events in Tables 1 and 3, to compare them with other substrates and/or genetic backgrounds.

RESULTS

A perfect 36-bp repeat undergoes frequent rearrangements during DSB repair: Using the experimental system described in Figure 1, we examined gene conversion-associated rearrangements of 36-bp repeated
sequences. A DSB was induced by the HO endonuclease in the yeast LEU2 gene (see Materials and Methods). When there is no donor template (Figure 1B), the vast majority of cells die, because homologous recombination is impossible and most cells lose the broken chromosome. Cells were transformed with plasmid pFP59 (Figure 1D) that contains an octamer of directly or indirectly oriented 36-bp repeats of the E. coli Lac operator (Robnett et al. 1996) inserted into a leu2 gene at the KpnI site. The ends of the DSB at the chromosomal leu2 locus begin with 47 or 70 bp of the HO cleavage site that are not homologous to the template, followed by leu2 sequences that can engage in recombination with the donor template on either side of the minisatellite array. The presence of these homologous leu2 sequences on the plasmid allows 15–28% of cells to repair the DSB by homologous recombination, depending on the insertions into the template (Figure 1).

We studied the outcome of the gene conversion events in the survivors by the PCR assay illustrated in Figure 3, which allowed us to identify the structure of both the plasmid donor template and the chromosomal recipient molecule. The DSB was indeed repaired by a gene conversion event wherein the repeated locus was transferred into the broken molecule. Note that this gene conversion is necessarily not associated with crossing over, because such an event would integrate the donor plasmid in the chromosome, resulting in an unstable dicentric chromosome III. As shown in Table 1, 11 rearrangements among 69 products were found only in the recipient molecule (the chromosomal copy). These rearrangements are considered to occur during DSB repair and amount to an average of 15.9% of the repair events. In another case, both the donor and recipient had the same altered number of repeats, suggesting that the donor had rearranged prior to gene conversion, although we cannot exclude that both molecules were concomitantly rearranged during DSB repair. One survivor colony carried on two different donor molecules, and in the recipient, the tandem array had the same size as in the rearranged donor. In this case, prior rearrangement of the donor may have occurred during the S phase that preceded a gene conversion induced in G2. In a control experiment, where HO was not expressed, there was only one rearrangement of plasmid sequences among 176 cells, an expansion of one repeat.

The same experiment was done with pFP58 as a template donor. pFP58 contains an array of 16 repeats,
Minisatellite Expansions

TABLE 1

Expansions and contractions with the pFP59 donor template

<table>
<thead>
<tr>
<th></th>
<th>% rearranged in recipient only</th>
<th>% rearranged in donor only</th>
<th>% rearranged in both</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Expansions</td>
<td>Contractions</td>
</tr>
<tr>
<td>WT</td>
<td>15.9</td>
<td>0</td>
<td>15.9</td>
</tr>
<tr>
<td></td>
<td>(11/69)</td>
<td>(0/69)</td>
<td>(11/69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>NA</td>
<td>0.6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1/176)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rad1</td>
<td>0</td>
<td>0</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>(0/72)</td>
<td>(0/72)</td>
<td>(1/72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>msh2</td>
<td>2.4</td>
<td>2.4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3/126)</td>
<td>(3/126)</td>
<td>(0/126)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>msh6</td>
<td>11.7</td>
<td>0</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>(7/60)</td>
<td>(0/60)</td>
<td>(7/60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pms1</td>
<td>10.0</td>
<td>3.3</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>(6/60)</td>
<td>(2/60)</td>
<td>(4/60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*NR, nonrearranged; R, rearranged only in donor (second column), or only in recipient (third column); BR, both donor and recipient rearranged. The number of events of each kind is followed by the kind of event: for example, 2 (–3) stands for 2 contractions removing two 36-bp units; 1 (+3/NR), in the same survivor colonies, we found two kinds of repeated arrays, one with an expansion (+3) and one that has not been rearranged. NA, there is no recipient in this experiment.

interrupted in the middle by 32 bp of a polylinker. Among 38 recombinants, 10 exhibited rearrangements in the recipient molecules (26.3%), 9 of 10 being contractions (Table 2). One expansion was also found in the donor. Since the recipient molecules of the same cells display a nonrearranged tandem array, this rearrangement in the donor is likely to be a consequence of DSB. However, a spontaneous event occurring immediately after DSB repair cannot be ruled out.

3' nonhomologous ends affect the DSB-induced rearrangement distribution and frequency: With the pFP58 and pFP59 templates, one of the early steps of gene conversion is the efficient removal of the nonhomologous sequences at the ends of the DSB by a MSH2-and RAD1-dependent process (Pâques and Haber 1997; Sugawara et al. 1997). These sequences correspond to the two halves of the 117-bp cleavage site. In plasmid pFP225, full homology with the 3' ends of the DSB was restored, by inserting two half-HO cut sites surrounding the repeated array (Figure 1G). With this plasmid, repair was substantially more efficient (46.5% vs. 28.0%) than with pFP59, consistent with previous results showing that perfect homology between the DSB ends and the template improves significantly the efficiency of the repair process (Pâques and Haber 1997).

However, the rearrangement rate using pFP225 was four times lower than with pFP59, dropping from 15.9 to 3.8% (Table 3). Nevertheless, these rearrangements still kept the signature of SDSA, for they were mostly found in the recipient molecule. An intriguing feature was that the ratio of expansions and contractions was significantly shifted toward expansions, which now represented one-half (three out of six) of the rearrangements. In contrast, 11 contractions but no expansions were found among 69 DSB repair events with pFP59. By Fisher’s exact test, the contractions are nonhomologous ends affect the DSB-induced rearrangement distribution and frequency: With the pFP58 and pFP59 templates, one of the early steps of gene conversion is the efficient removal of the nonhomologous sequences at the ends of the DSB by a MSH2-and RAD1-dependent process (Pâques and Haber 1997; Sugawara et al. 1997). These sequences correspond to the two halves of the 117-bp cleavage site. In plasmid pFP225, full homology with the 3' ends of the DSB was restored, by inserting two half-HO cut sites surrounding the repeated array (Figure 1G). With this plasmid, repair was substantially more efficient (46.5% vs. 28.0%) than with pFP59, consistent with previous results showing that perfect homology between the DSB ends and the template improves significantly the efficiency of the repair process (Pâques and Haber 1997).

However, the rearrangement rate using pFP225 was

Requirement for the Msh2 and Rad1 proteins in the DSB-induced tandem repeat rearrangements: We have previously shown that the excision endonuclease Rad1p-Rad10p and the mismatch repair proteins Msh2p and Msh3p (Fishman-Lobell and Haber 1992; Ivanov and Haber 1995; Pâques and Haber 1997; Sugawara et al. 1997; Colatácovu et al. 1999) are required for the removal of nonhomologous 3'-ended tails from intermediates of recombination. We proposed that such branched intermediates, shown in Figure 2, F–H and L–M, are responsible for DSB-induced tandem repeat
TABLE 2
Expansions and contractions with the pFP58 donor template

<table>
<thead>
<tr>
<th></th>
<th>% rearranged in recipient only</th>
<th>% rearranged in donor only</th>
<th>% rearranged in both</th>
<th>Donor&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Recipient&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Expansions</td>
<td>Contractions</td>
<td>Total</td>
<td>Expansions</td>
</tr>
<tr>
<td>WT</td>
<td>26.3</td>
<td>2.6</td>
<td>23.7</td>
<td>2.6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(10/38)</td>
<td>(1/38)</td>
<td>(9/38)</td>
<td>(1/38)</td>
<td>(0/38)</td>
</tr>
<tr>
<td>msh2</td>
<td>2.9</td>
<td>0</td>
<td>2.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(1/35)</td>
<td>(0/35)</td>
<td>(1/35)</td>
<td>(0/35)</td>
<td>(0/35)</td>
</tr>
<tr>
<td>msh6</td>
<td>21.1</td>
<td>2.6</td>
<td>18.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(7/38)</td>
<td>(1/38)</td>
<td>(7/38)</td>
<td>(0/38)</td>
<td>(0/38)</td>
</tr>
<tr>
<td>pmsl</td>
<td>13.5</td>
<td>0</td>
<td>13.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(5/37)</td>
<td>(0/37)</td>
<td>(5/37)</td>
<td>(0/37)</td>
<td>(0/37)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Notations are the same as in Table 1. In addition, a letter L, R, or M is appended to each rearrangement event to indicate where the rearrangement took place: it can be confined to the left (L) or right (R) tandem array, or, in a big deletion event, remove the piece of polylinker in the middle together with a number of repeated units on each side (M), leaving only one tandem array. In two cases, rearrangements occurred on both arrays independently (example: 1L and 2R).

rearrangement (Pâques and Wegnez 1993; Pâques et al. 1998). In addition, Rad1p and Msh2p, and presumably Rad10p and Msh3p, are required for the removal of heteroduplex loops formed during gene conversions (Kirkpatrick and Petes 1997; Likeman et al. 2001). Intermediates containing heterologous loops could arise by replication slippage during gene conversion where the polymerase must traverse a set of repeats (Figure 2C).

TABLE 3
Expansions and contractions with the pFP225 donor template

<table>
<thead>
<tr>
<th></th>
<th>% rearranged in recipient only</th>
<th>% rearranged in donor only</th>
<th>% rearranged in both</th>
<th>Donor&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Recipient&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Expansions</td>
<td>Contractions</td>
<td>Total</td>
<td>Expansions</td>
</tr>
<tr>
<td>WT</td>
<td>3.8</td>
<td>1.9</td>
<td>1.9</td>
<td>0</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>(6/157)</td>
<td>(3/157)</td>
<td>(3/157)</td>
<td>(0/157)</td>
<td>(1/157)</td>
</tr>
<tr>
<td>rad1</td>
<td>0.6</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>(1/181)</td>
<td>(0/181)</td>
<td>(1/181)</td>
<td>(0/181)</td>
<td>(2/181)</td>
</tr>
<tr>
<td>msh2</td>
<td>3.8</td>
<td>0.5</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(1/182)</td>
<td>(1/182)</td>
<td>(6/182)</td>
<td>(0/182)</td>
<td>(0.182)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Notations are as in Table 1.
gous to the template; however, msh2 does reduce repair by 2.5-fold. This small drop might depend on the repeat structure, for it was not observed in previous assays measuring gene conversion between nonrepetitive sequences (Paques and Haber 1997).

We then analyzed the survivors. In the rad1 strain, we characterized 71 survivor cells from two independent experiments with pFP59. Six proved to have repaired the DSB by nonhomologous end-joining, a RAD1-independent process (Moore and Haber 1996). Only one of the remaining 72 colonies showed a rearrangement, which was found in both donor and recipient and thus is not necessarily a rearrangement associated with gene conversion. Assuming that this one rearrangement event is recombination associated, it corresponds to an estimated rate of 1.4% instead of 15.9% in wild-type cells. In the msh2 strain, three events were found out of 126 survivors (2.4%). Similarly, with pFP58, msh2 cells had 2.9% rearrangements instead of 28.9% in wild type.

The three events found in the msh2 strain with pFP59 are all +1 expansions. This contrasts with wild-type cells, where only contractions were seen in 11 cases. The higher rate of contractions in the wild-type strain is clearly significant ($P = 6 \times 10^{-6}$), but the expansion rates are not distinguishable ($P = 0.81$), indicating that—as with the comparison of events in templates that did or did not contain homology to the DSB ends—only the contraction rate is affected by msh2. It is also important to note that in two of the three cases of expansion in msh2, the expansion was found in a mixed colony in which one-half of the cells had an unarranged number of repeats and one-half had an expansion. Sected colonies could be attributed to the lack of Rad1p- and Msh2p-dependent mismatch correction of a heteroduplex containing a 36-bp loop, analogous to postmeiotic segregation that were shown to have a similar dependence on Rad1p and Msh2p (Kirkpatrick and Petes 1997).

We also tested deletions of MSH6 and PMS1. These two genes act with MSH2 in the mismatch repair pathway but do not participate in nonhomologous tail removal. In addition, PMS1 is required together with MSH2 and RAD1 for heteroduplex loop correction during HO-induced mitotic gene conversion (Kirkpatrick and Petes 1997; Clikeman et al. 2001). The rearrangement rate among the gene conversion events also decreased, but less than twofold (see Tables 1 and 2); this effect is much weaker than the effect of the msh2 mutation. As in wild-type cells, most of the rearrangements were contractions, although two expansions were found in pms1. These results show that, although Msh6p and Pms1p are required for full rearrangement efficiency of pFP58 and pFP59, Msh2p and Rad1p have a much greater role in the process.

A completely different situation was observed with plasmid pFP225 (see Tables 1 and 3). Rearrangements were already relatively rare in a wild-type strain, but neither msh2 nor rad1 mutations affected the contraction rate ($P = 0.16$). This is clearly different from the effect in pFP59. There was also no significant change in the expansion rate. A single expansion was observed in msh2 and none in rad1 out of 363 total colonies, vs. 3/157 in wild type; however, the result is statistically not significant ($P = 0.08$).

DSB-induced rearrangements appear with a low frequency in a natural diverged yeast minisatellite: Natural minisatellites generally contain diverged repeats. To determine if base pair differences within the repeats have an effect on minisatellite stability during mitotic recombination, we replaced the artificial perfect repeat by a natural yeast minisatellite locus, normally found in subtelomeric Y’ sequences (Horowitz and Haber 1984). The repeats are 36 bp in length, as in the artificial minisatellite we used, but are very polymorphic. Although the Y’ 36-bp repeats fall into three main categories, there are only two pairs of perfectly identical units; the other ones differ by up to eight substitutions scattered among 16 sites of polymorphism. The repeat copy number at different Y’ elements within a single strain varies from 8 to 20 copies (Horowitz and Haber 1984). The minisatellite locus we cloned in the pFP46 plasmid contains 12 copies.

When we induced DSB repair with pFP46 as a donor template, the repair efficiency did not change compared to the result obtained with a perfect repeat (Figure 1); however, the frequency of rearrangements was greatly reduced. Among 194 gene conversion events, only two rearrangements, one expansion and one contraction, were found within the minisatellite, both in the recipient molecule (Table 4). This 1% rate of rearrangement is 16 times lower than what was observed with a perfect repeat. We conclude that the difference in the rearrangement rate is very likely to be due to the sequence divergence, although we cannot rule out that the difference results from some sequence-specific features.

In yeast and bacteria, recombination between diverged sequences is inhibited by the mismatch repair system (Rayssiguier et al. 1989; Borts et al. 1990; Selva et al. 1995; Chambers et al. 1996; Datta et al. 1996, 1997). In our system, repeat rearrangement is likely to reflect repeat misalignment at some point of the DSB repair process. In the context of a diversified tandem array, such misalignment would include many mismatches. Therefore, we tested if mismatch repair mutants would display a higher rate of expansions and contractions in our assay. Since the divergence in the Y’ minisatellite is only from base substitutions (Horowitz and Haber 1984), recombination between two different repeats might be inhibited by Msh2p-Msh6p, as the Msh6p-Msh2p heterodimer has been implicated in correcting single base mispairs, whereas heteroduplexes involving frameshift mutations are primarily recognized by the Msh2p-Msh3p heterodimer (for review, see Kołodner and Marsischky 1999). Pms1p forms a di-
Expansions and contractions with the pFP46 donor template

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Expansions</th>
<th>Contractions</th>
<th>% rearranged in recipient only</th>
<th>% rearranged in donor only</th>
<th>% rearranged in both</th>
<th>Donor*</th>
<th>Recipient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR: 194</td>
<td>NR: 192</td>
</tr>
<tr>
<td></td>
<td>(2/194)</td>
<td>(1/194)</td>
<td>(1/194)</td>
<td>(0/194)</td>
<td>(0/194)</td>
<td></td>
<td>R: 1 (-3); 1 (+1)</td>
<td></td>
</tr>
<tr>
<td>msh6</td>
<td>1.4</td>
<td>0</td>
<td>1.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR: 72</td>
<td>NR: 71</td>
</tr>
<tr>
<td></td>
<td>(1/72)</td>
<td>(0/72)</td>
<td>(1/72)</td>
<td>(0/72)</td>
<td>(2/72)</td>
<td></td>
<td>R: 1 (-1)</td>
<td></td>
</tr>
<tr>
<td>pms1</td>
<td>2.7</td>
<td>1.4</td>
<td>1.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR: 74</td>
<td>NR: 72</td>
</tr>
<tr>
<td></td>
<td>(2/74)</td>
<td>(1/74)</td>
<td>(1/74)</td>
<td>(0/74)</td>
<td>(0/74)</td>
<td></td>
<td>R: 1 (+1)</td>
<td></td>
</tr>
<tr>
<td>msh2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR: 40</td>
<td>NR: 40 NR</td>
</tr>
<tr>
<td></td>
<td>(0/40)</td>
<td>(0/40)</td>
<td>(0/40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Notations are as in Table 1.

Amer with Mlh1p, which interacts with both Msh6p-Msh2p and Msh3p-Msh2p, and is thus involved in the processing of all types of mismatches.

Neither msh6 nor pms1 mutations had much effect on this system; the frequencies of rearrangements were 1.4% (1/72) and 2.7% (2/74), respectively, which are not statistically significantly different from the 1% observed in wild type and still far from the 18.5% contractions observed with a shorter but perfect repeat. We also tested a msh2 mutant. As with the other templates, a low rate of survivors was obtained because of the requirement of Msh2p to remove nonhomologous tails. We tested 40 survivors, and no rearrangement was observed.

**DISCUSSION**

**Gene conversion as a major source of tandem repeat rearrangements:** Expansions and contractions of tandemly repeated sequences, from micro- and minisatelitites to genesized repeats, occur during or around meiosis (Welch et al. 1990, 1991; Fu et al. 1991; Buard and Vergnaud 1994; Jeffreys et al. 1994; Malter et al. 1997). As reviewed in the Introduction, there is growing evidence that, at least for human minisatelitites, the rearrangements occurring during germline formation are the consequence of DSB repair.

We developed a mitotic system in Saccharomyces, where it is possible to examine repeat instability accompanying DSB repair in great detail and provide a paradigm for the study of the mechanism and genetic requirements of this instability. This approach was used to examine rearrangements in 375-bp repeats (Paques et al. 1998), artificial and natural 36-bp minisatelitites (this work), and CTG microsatellites (Richard et al. 1999, 2000). In all three cases, gene conversion is frequently associated with rearrangements of tandem repeated sequences.

An examination of the rates and types of rearrangements with different templates provides some further insight into the origins of expansions and contractions of repeated sequences. First, the rearrangement rate is not strongly a function of the total length of the interval between the ends of the DSB. With 8 375-bp repeats (total length 2900 bp) 36% of genes had either fewer or >8 repeats (Paques et al. 1998). With 8 36-bp repeats (288 bp) 16% of genes conversions produced contractions, whereas 43% contractions and expansions were observed with a CAG repeat of approximately the same length (Richard et al. 2000). However, rearrangements appear to be more frequent when the number of repeats increases, rising from 16% with 8 36-bp repeats to 26% with 16 repeats and from 16% with a CAG repeat to 43% with a CAG repeat.

On the other hand, either the size of the repeats or some sequence-specific feature influences the nature of the rearrangements. With 8 375-bp repeats, more than one-third of the events were expansions, whereas virtually all the changes with 8 36-bp repeats were contractions. A few expansions were seen with pFP58, which has 16 repeats, whereas all events with a template carrying 8 repeats were contractions. Similarly, with a donor template carrying (CAG) repeat, all rearrangements were contractions (Richard et al. 1999), but with (CAG) repeat, 30% were expansions (Richard et al. 2000). In contrast, with a template harboring (CAA) repeat, all of the rearrangements were contractions (Richard et al. 2000).

It appears that those sequences that have high frequencies of recombination-associated expansions are capable of forming stable single-stranded secondary structures. CAG repeats are known to form hairpin structures in vitro and show frequent rearrangements, whereas CAA repeats, which do not form stable single-strands hairpin structures, have low rates of rearrangement (Gacy et al. 1995; Yu and Mitas 1995; Richard et al. 2000). There is growing evidence that large expansions of human simple repeats also depend on the ability of single-stranded tandem repeats, including micro- and...
minisatellites, to form secondary structures (McMurray 1999). Similarly, 5S RNA encoded by the 375-bp repeats has a complex secondary structure involving large hairpins (see Correll et al. 1997, and references therein); and it is possible that single-stranded 5S rDNA also adopts such a conformation, which could account for the frequent expansions we observed.

**Ratio of expansions vs. contractions:** In many previous studies on tandem repeat instability in yeast, contractions were generally found to be the major class of events (Richard and Pâques 2000). Yet we seek a model in which expansions are as frequent as contractions, or even more frequent, as is seen in the changes of unstable human minisatellites (Buard and Vergnaud 1994; Jeffrey et al. 1994; May et al. 1996). In this study, we found two situations where expansions were as frequent as contractions, or even in the majority.

The first case concerns plasmid pFP225. When an HO-cleaved chromosomal site is repaired using plasmids pFP59, pFP58, and pFP46 as the template, there is no homology with the HO cut site, and the two nonhomologous sequences must be excised. In pFP225, two half-HO cut sites restore perfect homology with the DSB extremities. pFP59 and pFP225 are identical in all other respects, and yet yield strikingly different results. The overall rearrangement frequency is fourfold higher with pFP59, but this difference is due entirely to a very high rate of contractions. These contractions are suppressed in pFP225, but expansions are now found and amount to one-half of the events. Thus the presence of nonhomologous tails profoundly influences what kind of repair events are seen.

A second case where expansions are also prevalent is when there are nonhomologous tails, using pFP59 as the template, but in msh2 and rad1 derivatives, where efficient removal of such tails is prevented. In fact, only expansions were found, whereas the frequent contractions observed in wild type were completely suppressed. Why should nonhomologous ends result in a substantially higher rate of contractions? One possibility is that such 3' overhangs, which have to be removed by a Msh2p-Rad1p complex, would channel DSB repair in a specific, contraction prone, recombination process. But curiously, in the few survivors we recover with pFP59 in the msh2 and rad1 mutant strains, contractions are infrequent. We previously showed that there is a relatively inefficient Rad1p-, Msh2p-independent pathway to remove nonhomologous DNA tails (Colaiácovo et al. 1999; Holmes and Haber 1999). Although this backup excision process seems especially effective when one end has homology, it probably accounts for the remaining gene conversions we observed here with two nonhomologous ends. Thus it appears that a high contraction rate is observed only when 3' overhangs are cut by Msh2p and Rad1p early in the recombination process. We propose that the early recruitment of Msh2p-Rad1p influences the contraction pathway in a subsequent step.

**Mechanisms of expansions and contractions:** Recently, two groups tested the impact of msh2 on minisatellite rearrangement in yeast meiosis (Debrauwère et al. 1999; Bishop et al. 2000). Bishop et al. (2000) did not find any significant changes in the overall rearrangement rates, and Debrauwère et al. (1999) found a small increase (threefold) in one case and no effect in two others. In meiosis, the DSB ends generally match their homologous template; hence these results are mostly in agreement with what we find with pFP225.

In a yeast strain defective for both msh2 and pms1, Debrauwère et al. (1999) observed a new class of meiotic rearrangements, likely resulting from unrepaired heteroduplex loops. Heteroduplex loops appear in most replicative models of simple repeat rearrangements, because they can result from polymerase slippage (see Figure 2, C and D). However, they might also arise during annealing, as pointed out by Debrauwère et al. (1999). With minisatellite repeats, these loops should be at least 36 bp long and therefore are likely to be processed by the Rad1p/Msh2p complex (Kirkpatrick and Petes 1997), which, for HO-induced gene conversion, also involves Pms1p (Glikeman et al. 2001). In our system, such loops would result in an increased number of mixed colonies containing both a rearranged minisatellite locus and an unarranged one, in rad1, msh2, or pms1 strains. We found a few sectored colonies indicative of such heteroduplex loops in the mutant strains: three events with pFP59 in msh2 and pms1, one event with pFP58 in pms1, and three events with pFP225 in msh2 and rad1 (scoring donor as well as recipient molecules). Hence, we think such events are not the majority. In addition, they do not appear to be favored by mutations in the MSH2, RAD1, or PMS1 genes, for they were found in the wild-type strain with a comparable, if not higher, rate (one event out of 69 for pFP59, two events out of 157 for pFP225).

Instead, we interpret most of the observed rearrangements in terms of out-of-frame annealing or out-of-frame reinvasion during SDSA. SDSA models (reviewed in Pâques and Haber 1999) were proposed to account for DSB-induced rearrangements of tandem repeats (Pâques and Wegnez 1993; Buard and Jeffrey 1997; Pâques et al. 1998). Several different versions of this class of model were proposed, differing in whether DNA synthesis initiates at one or both ends of the DSB (Ferguson and Holloman 1996; Pâques et al. 1998; Holmes and Haber 1999; Pâques and Haber 1999). One class of SDSA models is shown in Figure 2, E–H, in which the two ends of the DSB initiate new DNA synthesis and then the two newly synthesized single strands anneal to repair the recipient. Out-of-frame annealing would result in expansions (Figure 2H) and contractions (Figure 2, F and G). An alternative repair mechanism (Figure 2, I–M) suggests the invasion of only
one end, leading to the establishment of a complete replication fork involving leading and lagging strand synthesis, as is seen in HO-induced gene conversion of the MAT locus, where one end of the DSB has a long nonhomologous tail (Holmes and Haber 1999). In this case, the second end may be a more passive partner, needed to anneal to the migrating replication D-loop to terminate the process. Here, the source of rearrangements must be from replication slippage-like events rather than from misaligned strand annealings. Such events can result from branch migration events that would dissociate the newly synthesized strands from their template or from dissociation of the DNA polymerases from the template (Pâques et al. 1998; Figure 2, I and J), followed by out-of-frame reinvansion of the template (Figure 2, L or M). Although formally equivalent to classical replication slippage (as featured in Figure 2C), this process does not involve any heteroduplex loop.

A number of SDSA intermediates leading to repeat rearrangement require the processing of 3′-overhangs by Msh2p-Rad1p (Figure 2, F and L). One possibility is that the early recruitment of Msh2p and Rad1p favors the later processing of these intermediates, maybe because Rad1p and Msh2p remain somewhat associated with the replication proteins. Without this early interaction, these proteins would have to be recruited de novo, and the potentially unstable intermediates shown in Figure 2, F and L, would dissociate most of the time before Rad1p and Msh2p have a chance to process them. In this case, only contractions would depend on Rad1p and Msh2p, because most of the intermediates leading to expansions (Figures 2, H and M) can right away initiate new DNA synthesis.

Another possibility is that the recruitment of Msh2p and Rad1p to the end of the DSB, when nonhomologous tails must be removed, facilitates the loading or retention of proteins that decrease the processivity of repair DNA synthesis. When there are no nonhomologous tails, repair synthesis proceeds with only occasional dissociation. Then most gene conversions should be accurate, provided one-ended strand invasion (Figure 2, I–L) is the major pathway. When Rad1p-Rad10p and Msh2p-Msh3p are recruited (along with the participation of both Rad59p and Srs2p; Sugawara et al. 1997, 2000; Evans et al. 2000) to enable the excision of the nonhomologous tail, the polymerase falls off the template more often, leading to an increased rate of rearrangement. The preference for contractions over expansions might be explained if strand invasion occurs preferentially at one of the two ends of the DSB, thereby establishing a biased direction of the repair synthesis in traversing the repeat sequences. Such a bias was observed in the formation of contractions in CTG/CAG microsatellites, depending on the orientation of the array relative to the direction of replication (Maurer et al. 1996; Freudenreich et al. 1997; Balakumaran et al. 2000; Ireland et al. 2000). Indeed we found such a bias when the HO-cleaved LEU2 locus used in our experiments copied a template containing CAG vs. CTG repeats (Richard et al. 2000).

**DSB-induced repeat rearrangements are less frequent in a heterogeneous repeat than in a perfect one:** We also examined the effect of sequence divergence on repeat rearrangement. Although we observed nearly 16% of DSB-induced rearrangements in a perfect 36-bp repeat, the rearrangement frequency was only 1% with a heterogeneous 36-bp repeat, encompassing even more repeat units (12 instead of 8). This result is not surprising, for homeologous recombination is usually impaired by the mismatch repair machinery. In the human CEB1 locus, the very heterogeneous alleles are more stable than the rather (but never fully) homogeneous ones (Buard et al. 1998).

Mismatch repair proteins discourage recombination between mismatched substrates. However, msh2, msh6, and pms1 mutations did not restore a level of recombination-induced rearrangements similar to that observed with perfect repeats. Previous studies suggested that sequences diverged by 10% would be outside the range that could be suppressed by mismatch repair mutants during mitosis (Datta et al. 1996, 1997). Indeed, Bishop et al. (2000), who studied the stability of the same Y′ minisatellite during meiosis, did not see any increased instability in msh2, msh3, mhl1, and pms1 mutants. A human CEB1 minisatellite does not display any increased meiotic instability in a yeast msh2 pms1 mutant either, when the minisatellite locus is homozygous (Derraudeau et al. 1999). However, the same mutations induced a threefold increase in the rearrangement rate when the two homologous chromosomes carry two different CEB1 alleles. Nevertheless, mismatch repair proteins seem to contribute only weakly in the inhibition of rearrangement between short diverged tandem repeats.

Nevertheless, we do see an ~1% rate of change in the size of the Y′ sequences during recombination, which is an evolutionarily significant rate. In meiosis, the same minisatellite array is rearranged in 0.5% of the tetrads (Bishop et al. 2000). It is difficult to compare directly the results we obtained in mitotic cells with these meiotic results, because we would have to know the frequency of DSBs that occur in the vicinity of the Y′ sequences in meiosis. We note also that our mitotic system is a gap repair system in which repair DNA polymerases must traverse the entire Y′ array, whereas the events in meiosis are likely to be initiated by DSBs outside the array itself. Buard et al. (1998) observed 0.1–4.4% of rearrangements per gamete for human CEB1 minisatellite alleles ranging from 10 to 14 repeats in length, which would thus correspond to a minimum rate of 1% to a maximum of 44% rearrangements per DSB, assuming a uniform 10% DSB rate. Thus, during DSB repair, our Y′ yeast minisatellite would be relatively stable, but comparable to its human analogs during meiosis.
The results we presented provide evidence that, depending on the exact nature of the DNA ends and of the template sequences, both expansions and contractions can be obtained. Several different mechanisms appear to be important, depending on these different variables. The development of a mitotic recombination system in which virtually all cells can be induced to undergo recombination at the same time, from a defined DSB, now provides us with a way of exploring in greater detail the process of minisatellite repeat rearrangements.

We thank T. Grange for critical reading of the manuscript and D. Higuët for statistical analysis. This work was supported by National Institutes of Health grant GM20956. F.P. was a fellow from the American Cancer Society.

LITERATURE CITED


Pâques, F., and J. E. Haber, 1997 Two pathways for removal of non-


Communicating editor: L. S. Symington