Genome rearrangements caused by interstitial telomeric sequences in yeast

Anna Y. Aksenova, Patricia W. Greenwell, Margaret Dominska, Alexander A. Shishkin, Jane C. Kim, Thomas D. Petes, and Sergei M. Mirkin

Department of Biology, Tufts University, Medford, MA 02155; and Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710

Contributed by Thomas D. Petes, October 11, 2013 (sent for review September 10, 2013)

Interstitial telomeric sequences (ITSs) are present in many eukaryotic genomes and are linked to genome instabilities and disease in humans. The mechanisms responsible for ITS-mediated genome instability are not understood in molecular detail. Here, we use a model Saccharomyces cerevisiae system to characterize genome instability mediated by yeast telomeric (Ytel) repeats embedded within an intron of a reporter gene inside a yeast chromosome. We observed a very high rate of small insertions and deletions within the repeats. We also found frequent gross chromosome rearrangements, including deletions, duplications, inversions, translocations, and formation of acenitic minichromosomes. The inversions are a unique class of chromosome rearrangement involving an interaction between the ITS and the true telomere of the chromosome. Because we previously found that Ytel repeats cause strong replication fork stalling, we suggest that formation of double-stranded DNA breaks within the Ytel sequences might be responsible for these gross chromosome rearrangements.

Results

System for the Detection of Genome Instabilities Caused by the Ytel Repeats. The reporter gene used to detect changes within the interstitial Ytel sequences is closely related to the one used previously to detect expansions of the (GAA)n repeats (11), which were inserted into an intron of the artificially split URA3 gene (construction details are in SI Methods, Tables S1 and S2). Because long (>1-kb) introns are inefficiently removed from pre-mRNA in yeast (12), large-scale expansions of the (GAA)n tracts produce Ura– clones that might be selected on the medium containing 5-fluoroorotic acid (5-FOA) (11). 5-FOA–resistant clones also resulted from large deletions around those repeats (11, 13), as well as from the repeat-induced mutagenesis in the body of the URA3 gene (13–15).

To determine whether interstitial telomeric repeats also are subject to expansions and other types of instabilities, a similar URA3 reporter with these repeats in its intron was inserted into the chromosome III near the active replication origin ARS306 (Fig. 1 and Fig. S1). The endogenous URA3 locus on chromosome V contained the mutant ura3-52 allele. Although the natural yeast telomeres do not have identical copies of a small repeat, we constructed reporter genes that had either 8 or 15 copies of the perfect octameric repeat (TGTGTGGG), which is complementary to the telomeric RNA template within telomerase (16). The (TGTGTGGG)n sequence was in the nontranscribed strand of URA3 placing it onto the lagging strand template for replication forks emanating from ARS306 responsible for replication of our reporter (11). Yeast strains with

Significance

Telomeres are composed of simple repetitive DNA sequences that normally are located at the ends of the chromosomes. Occasionally, however, they also are found inside chromosomes. Some of these internal or interstitial telomeric sequences colocalize with chromosomal fragile sites, preferred sites of breakage in some cancers and hereditary human diseases. The mechanisms responsible for genome instability at interstitial telomeric sequences are unclear. We developed a system to study genetic instabilities caused by these sequences in a model organism (baker’s yeast) that allowed us to characterize various chromosomal rearrangements and to measure the likelihood of their formation. We found that interstitial telomeric sequences promote the formation of deletions, duplications, inversions, and translocations, and we proposed molecular mechanisms responsible for these events.


The authors declare no conflict of interest.

See Commentary on page 19664.

1To whom correspondence may be addressed. E-mail: tom.petes@duke.edu or sergei.mirkin@tufts.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1319313110/-/DCSupplemental.
either the (TGTGTGGG)$_8$ or the (TGTGTGGG)$_{15}$ insertions (strains SMY752 and SMY749, respectively) were Ura$^+$ and 5-FOA$^+$, although a quantitative measurement of gene expression (Fig. S2) indicated that the (TGTGTGGG)$_{15}$ Ytel tract reduced expression of the spliced UR43 transcript. Because telomeric tracts in the opposite orientation resulted in an even stronger suppression of splicing, we did not attempt a direct comparison of the effects of tract orientation on genome stability.

**Interstitial Ytel Repeats Are Naturally Highly Unstable.** In the 5-FOA$^+$ strains SMY752 and SMY749, we looked for tract expansions or other gene-inactivating changes by selecting 5-FOA$^+$ derivatives. In the control strain (SMY803) with no Ytel sequences in the reporter, the rate of 5-FOA$^+$ derivatives was about $0.4 \times 10^{-7}$ per division. This rate was elevated in strains with (TGTGTGGG)$_8$ and (TGTGTGGG)$_{15}$ Ytel tracts 20-fold and 125-fold, respectively (Table 1). The 5-FOA$^+$ derivatives of all three strains initially were examined by PCR analysis with the primers (UIRL1/UIRL2) that flanked the Ytel tracts (Fig. 1A). In the control strain, most of the independent derivatives (40 of 42) generated a PCR fragment with these primers. We sequenced five of such derivatives, and all had mutations in the *URA3* coding sequence (Table S3).

For the SMY752 strain (Fig. 1C), 24 of 96 derivatives obtained from independent cultures (four derivatives per culture) failed to produce a PCR fragment with primer pairs closely (within 100

---

**Table 1. Rates of genetic alterations observed in 5-FOA$^+$ strains derived from SMY803, SMY752, and SMY749**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Repeat no.</th>
<th>5-FOA$^+$ rate ($\mu \times E^{-07}$)</th>
<th>Tract alteration rate ($\mu \times E^{-08}$)</th>
<th>ura3 point mutation rate ($\mu \times E^{-07}$)</th>
<th>Rate of potential GCR events ($\mu \times E^{-07}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMY803</td>
<td>0 repeats</td>
<td>0.4 ($0.2–0.6$) [1]</td>
<td>Not applicable</td>
<td>0.3 ($0.1–0.5$) [1]</td>
<td>0.03 (0.002–0.06) [1]</td>
</tr>
<tr>
<td>SMY752</td>
<td>(TGTGTGGG)$_8$</td>
<td>8 ($6–10$) [20]</td>
<td>5 ($1–11$) [8]</td>
<td>8 ($4–14$) [27]</td>
<td>4 ($2–6$) [133]</td>
</tr>
<tr>
<td>SMY749</td>
<td>(TGTGTGGG)$_{15}$</td>
<td>50 ($38–64$) [125]</td>
<td>15 ($4–29$) [9]</td>
<td>23 ($12–35$) [77]</td>
<td>36 ($21–53$) [1,200]</td>
</tr>
</tbody>
</table>

*The rates of 5-FOA$^+$ derivatives were determined by measuring the frequencies of 5-FOA$^+$ colonies per culture. Rates and 95% confidence intervals (in parentheses) were calculated using the Ma–Sandri-Sarkar maximum likelihood estimator with a correction for sampling efficiency. Numbers in brackets are fold increases relative to rates in SMY803.

*This category includes 5-FOA$^+$ derivatives that had no change in the number of telomeric repeats, as determined by PCR. DNA sequence analysis confirmed that these strains had mutations in the *URA3* sequences flanking the intron.

*This category includes 5-FOA$^+$ isolates that failed to produce a PCR fragment using primers flanking the telomeric tract.

*Using primers flanking the ITs, we looked for tract alterations in multiple 5-FOA$^+$ derivatives in independent cultures. Eight small-scale alterations were observed among 96 isolates examined, and all cases represented double events, in which the change in the repeat length was accompanied by a mutation in *URA3*. Because the tract alterations are not responsible for the 5-FOA$^+$ phenotype, the true rate of tract alterations is much higher than indicated for SMY752 and SMY749.

*Only two single-repeat expansions were observed among about 50 5-FOA$^+$ isolates. Sequencing of one of them showed that the change in the repeat length was accompanied by a mutation in *URA3*.
bp) flanking the tract. As discussed below, these isolates likely contain chromosome rearrangements. Seventy-two derivatives did produce the PCR fragment with the adjacent primers, of which 64 isolates showed no alterations in the repeat length. We sequenced the \textit{URA3} gene from 33 such isolates (from 11 independent cultures) and found that most of them (26 of 33) had mutations in the \textit{URA3} coding sequence (Table S3). In the eight isolates with altered Ytel tracts, four had an addition and four had a deletion of just one repeat. These tract alterations probably were not responsible for the 5-FOA\(^{-}\) phenotype, because the two sequenced isolates of this class had additional mutations in the \textit{URA3} coding sequence in both cases.

In SMY749, we examined 71 5-FOA\(^{-}\) isolates from 24 independent cultures. Of these isolates, 44 lacked a PCR fragment. As described below, we examined 15 of the isolates that lacked the intron-spanning PCR product by other techniques, and all of them had chromosome rearrangements. The remaining 27 isolates had an unchanged Ytel tract. We sequenced five such derivatives: four had alterations in the \textit{URA3} coding sequence, whereas one had a mutation at the intron–exon junction (Table S3).

Because small changes in the number of repeats did not generate a 5-FOA\(^{-}\) strain, we also looked for small expansions or contractions of the (TGTGTGGG)\(_{12}\) tract in the nonselective conditions. To this end, 20 clones each from 12 independent cultures grown nonspecifically were analyzed by PCR. We detected 13 small-scale tract alterations involving a loss or gain of one to two repeats (Table S4). The calculated rates of alterations (95% confidence limits in parentheses) corresponded to 2.2 (0.7–4.3) \times 10^{-3} \text{per division for expansions and 1.4 (0.4–2.9) \times 10^{-3} per division for contractions.}

In summary, our analysis demonstrates that ITSs are associated with several types of instability. First, the tracts gain or lose small numbers of repeats at high rates, whereas large tract expansions were not observed. Second, ITSs stimulate mutations in flanking sequences. Several studies recently showed that double-strand DNA breaks (DSBs) in long (GAA)_\text{n} tracts (11, 13–15) or inverted repeats (14) elevate rates of mutations in closely linked genes. It is likely that ITSs operate by related mechanisms. Third, ITSs stimulate GCRs as described below.

\textbf{Contour-Length Homogeneous Electric Field Gel and Microarray Analyses of GCR Events Induced by Ytel Sequences.} We examined 15 5-FOA\(^{-}\) derivatives of the SMY749 strain (designated SMY779), in which no PCR fragment was detected with primers flanking the Ytel sequences, by a variety of approaches beginning with contour-length homogeneous electric field (CHEF) gel analysis. Intact chromosomal DNA molecules were separated using CHEF gels, transferred to nylon membranes, and hybridized to probes located either centromere-proximal (\textit{LEU2}) or centromere-distal (\textit{CHA1}) to our \textit{URA3} reporter. Four patterns were observed. In seven isolates (nos. 1, 9, 17, 18, 20, 24, and 27), chromosome III was identical in size to the original chromosome III (about 365 kb), and both probes hybridized to the same chromosome (Fig. 2). We refer to this class as class 1. In the single class 2 strain (no. 4), chromosome III was about 10 kb smaller than the original chromosome III, but still hybridized to both the \textit{LEU2} and \textit{CHA1} probes. Seven isolates had no DNA fragment similar in size to the original chromosome III. In six of them (nos. 7, 10, 11, 19, 21, and 25), the \textit{LEU2} probe hybridized to chromosomes that varied in size from about 435–490 kb, whereas the \textit{CHA1} probe hybridized to a DNA fragment of ~80 kb (Fig. 2C), which is smaller than any of the yeast chromosomes (class 3). In the remaining isolate, designated class 4 (no. 12), the \textit{LEU2} probe hybridized with a chromosome of ~290 kb, and the \textit{CHA1}-hybridizing fragment was ~80 kb long.

The same DNA samples then were examined by comparative genome hybridization (CGH) microarray (17). Deletions or duplications of single-copy regions ≥5 kb throughout the genome can be detected readily. Fig. 3 shows CGH data for chromosome III of the four classes described above. In class 1 isolates, no deletions or duplications on chromosome III or any other yeast chromosome were observed. In class 2 isolates, there is a deletion of ~10 kb that removed sequences from the \textit{URA3::Ytel} reporter gene to a Ty1 element of left arm transposition hotspot (LAHS) (Fig. 1B). Two class 3 isolates have the same 10-kb deletion as in class 2, but also have a duplication of sequences derived from the right arm of chromosome III ranging from ~190 kb in no. 11 to ~215 kb in no. 19. The starting point of these duplications is near \textit{YCRCdelta6} in no. 11 and near the Ty/delta elements of fragile site 1 (FS1) in no. 19 (Fig. 1B). In the class 4 isolate (no. 12), chromosome III has the same 10-kb deletion as above plus a duplication of 30-kb DNA segment located at the left telomere of chromosome II starting near \textit{YBLWTy2-1} (Fig. S3B).

Both class 3 and 4 strains have an 80-kb minichromosome that hybridizes to the \textit{CHA1} probe. To identify the sequences within this minichromosome, we purified it from a CHEF gel and hybridized it to the CGH arrays. As shown in Fig. 3, it contains all the chromosome III sequences between the left telomere and \textit{URA3::Ytel} reporter.

\textbf{Inversions Between Interstitial and Terminal Telomeric Repeats in Class 1 Strains.} Unlike the class 2–4 strains, class 1 strains did not have detectable deletions or duplications. Using PCR analysis and Southern analysis (SI Methods, Fig. S4), we found that class 1 strains do have a rearrangement on chromosome III, which is an 80-kb–long inversion of the chromosomal segment between the interstitial Ytel repeat and the natural telomere. Fig. 4 shows that in class 1 strains, a PCR fragment may be generated with a primer located within the 3′ part of the \textit{URA3} reporter and a primer located near the left telomere. The size of the resulting PCR product varies, because the length of telomeric DNA is different in different isolates. Based on the size of this fragment, the approximate lengths in base pairs of the
interstitial telomeric DNA after the inversion in different isolates varies between 315 and 540 bp. We also sequenced part of the PCR product containing the ITS in the class 1 strain SMY779-#27. The first 60 bp of the ITS adjacent to the X repeat (Fig. 4A) were identical to the telomeric sequences adjacent to the X repeat in the *Saccharomyces* Genome Database (SGD), whereas the sequences at the other junction were six perfect repeats of URA3::Ytel. These results are consistent with the model proposed in Fig. 4A.

Class 2–4 strains were Trp⁺ because the 10-kb deletion associated with the chromosome rearrangement removed the TRP1 gene located adjacent to the URA3::Ytel gene. Unexpectedly, however, class 4 strains were Trp⁻ as well. Long yeast telomeric repeats located internally on the chromosome have been shown to repress expression of nearby genes (18, 19), and this Rap1- and Sir2-dependent telomeric silencing (1) is reversed by nicotinamide (20). Because class 1 strains had Ytel repeats that were at least twice as long as the original strain without the chromosome rearrangement (120 bp), we hypothesized that the Trp⁺ phenotype reflected telomeric silencing. Consistent with this possibility, we found that Rap1p is recruited to the Ytel repeats and the flanking URA3 sequences by chromatin immunoprecipitation (Fig. S5). Furthermore, class 1 strains grown in the presence of nicotinamide have Trp⁻ phenotype (Fig. S6). Thus, chromosome rearrangements induced by ITS sequences may affect gene expression as well as chromosome structure.

**Discussion**

Genetic instabilities caused by DNA microsatellites have been studied for several decades in connection with hereditary human diseases, cancer, and chromosomal fragility. The effects of multiple types of microsatellites, particularly trinucleotide repeats, on genome stability have been investigated in detail in *S. cerevisiae*. In yeast, certain microsatellites [such as (CTG)n and (GAA)n] may act as hotspots for mitotic DSBs that are recombinogenic and/or associated with chromosome rearrangements (21, 22).

From our analysis of ITSs, it is clear that some of the genome destabilizing effects of these sequences are shared with other types of microsatellites. In strains with the (TGTGTGGG)ₚ repeat, deletions and additions of one repeat were remarkably frequent, about 10⁻³ per cell division. For comparison, deletions and additions of repeats in a (CAATCGGT)₁₀ tract occurs at a rate of about 10⁻⁵ per division (23). Of the 5-FOA⁺ isolates derived from SMY752 and SMY749, none had large expansions. In contrast, large expansions were the main class leading to 5-FOA resistance for (GAA)n tracts longer than 100 repeats.

Mutations in the coding sequence of the *Ura3::Ytel* reporter were substantially (>20-fold) induced by both the 8- and 15-copies of Ytel tracts (Table 1). Similar induction of mutagenesis was observed for (GAA)n and inverted repeats [repeat-induced mutagenesis (RIM)] (11, 13–15). All these repeats stall the replication fork progression (24–26), and the (TGTGTGGG)ₚ run is a particularly potent replication block. A connection between stalled replication forks and DSBs in yeast also has been observed for CTG repeats (22) and in mecl cells treated with hydroxyzene (27). The recombinogenic repair of DSBs is known to be error prone (28), and single-stranded DNA recombination intermediates are particularly susceptible to mutations (29). We believe, therefore, that error-prone DSB repair might be responsible for RIM.

In about half the 5-FOA⁺ derivatives of the *Ura3::Ytel* strains, we detected no PCR fragments using primers flanking the telomeric tract. The rate of formation of these strains was elevated about 100-fold by the (TGTGTGGG)ₚ tract and about 1,000-fold by the (TGTGTGGG)ₚ tract. All these strains contained chromosome rearrangements that were consistent with a DSB within the ITS as their initiating event; these results, however, have not been demonstrated by direct physical methods.

In class 1 strains, we suggest that the broken ends generated by the DSB within the ITS are processed by exonucleases (30) to yield a protruding G-rich single strand on the 80-kb acentric chromosome fragment, and a protruding C-rich single strand on the centromere-containing fragment (Figs. 4Aand S4). The protruding C-rich single strand of the centromere-containing fragment might anneal with the left telomere (which has a protruding G-rich strand) to produce the observed inversion. This event would split the *Ura3* cassette into two noncontiguous fragments, resulting in the 5-FOA⁻ phenotype. Although a similar structure might be produced by a mechanism in which two broken chromatids anneal to produce one chromosome with an inversion, similar to a half-crossover, we prefer the model shown in Fig. 4A, because half-crossovers usually are observed only in strains defective in homologous recombination (31).

By the model shown in Fig. 4A, the telomeres derived from the ITS sequence would have a maximum length of 120 bp. From Southern analysis of class 1 strains (Fig. S4C), we calculated that the newly formed telomeres in class 1 strains are about 360 bp, similar in length to telomeres in wild-type strains (1). These results suggest that telomerase extended the telomeric repeats associated with the 5' end of the reporter gene.

The single class 2 strain had a deletion on chromosome III that removes the intron with Ytel sequences and the 3' end of the *Ura3* reporter, as well as all sequences between the *Ura3* reporter up to the Watson-oriented Ty element on the left arm; the 5' portion of the reporter gene was retained. By PCR analysis, followed by DNA sequencing, we identified the junction between the 5' portion of *Ura3* and the Ty element on chromosome III as identical to the junction between the *Ura3* coding sequence and Ty sequences of the *ura3*-52 allele on chromosome V (32). This observation is consistent with the model, in which one end of the DSB within the ITS sequence is degraded into the 5' part of the *Ura3::Ytel* reporter, whereas its other end is degraded into the Ty element of LAHS. The two ends then invade the *ura3*-52 allele, and the subsequent gene conversion event would yield the class 2 rearrangement (Fig. 5B).

Class 3 and 4 events reflect similar mechanisms (Fig. 5 C and D) likely initiated by a DSB within the ITS. Because classes 3 and 4 have 80-kb minichromosomes, this acentric fragment does not recombine with other yeast sequences. Note that the broken end of this acentric fragment has the structure required for a
functional telomere: the 3' end has G\textsubscript{1}–G3 T sequences. In the centromere-containing fragment, however, the orientation of the telomeric repeats (3' end with C\textsubscript{1}–C3 A sequences) is in an orientation precluding telomere formation. We term such sequences "reverse-telomeres." For the centromere-containing fragment to form a stable chromosome, this reverse-telomere sequence must be removed and replaced by a functional telomere. Thus, the reverse-telomeres are degraded until the Ty sequences in the LAHS are rendered single stranded, the resulting end invades Ty or delta elements on the right arm of the chromosome, and sequences distal to the point of invasion are then copied by break-induced replication (BIR).

The single class 4 event likely is generated by a mechanism similar to class 3, except the processed centromere-containing fragment invades a Ty element located on chromosome II, which results in a BIR event extending from the point of invasion to the left end of chromosome II (Fig. 5D). Indeed, GCRs caused by recombination between Ty or delta elements were observed in many previous studies in yeast (33, 34).

Because the 80-kb minichromosome is acentric, it would be expected to be very unstable. As this minichromosome contains essential genes, cells that lose it would not be viable, which should reduce the growth rate of the corresponding strains substantially, yet class 3 strains formed colonies that were only slightly smaller than a control strain (Fig. S7). We hypothesize that some sequence on this minichromosome might have a centromere-like activity, or that the segregation of the minichromosome is coupled to the segregation of other chromosomes (35). Paek et al. (36) previously used PCR to demonstrate acentric fragments in yeast generated by template switching during DNA replication, although the stability of these fragments could not be quantitated.

All four classes of genome rearrangements are explained readily as the outcomes of a DSB repair within the ITS. A DSB at this position likely is caused by the very potent replication block observed in strains with the (TGTGTGGG)\textsubscript{18} tract (24), because DSBS were observed previously at stalled replication forks in yeast (22, 27). The latter likely is the result of a tight nucleoprotein complex formed at the interstitial Y\textsubscript{tel} sequences. We show here that the telomere-binding Rap1 protein covers the interstitial Y\textsubscript{tel} sequence (Fig. S5). Furthermore, fork stalling at ITTs decreases dramatically if we knock out Tof1 protein (24), which is known to preserve protein-mediated replication blocks (37).

In summary, we observed that interstitial telomeric repeats are potent sources of genome instability and unraveled molecular mechanisms of those events. An unexpected payback from this...
analysis is that the \textit{URA3::Yel} reporter may be used as a tool to generate chromosomes of desired sizes.

**Methods**

**Yeast Strains.** We used three isogenic haploid strains with an intron-containing \textit{URA3} gene on chromosome III: SMY803 (no telomeric repeats), SMY752 (\textit{URA3::Yel} with eight telomeric repeats), and SMY749 (\textit{URA3::Yel} with 15 repeats). (See \textit{SI Methods} for details.)

**Measurements of the Rates of Gene-Inactivating Events.** Strains with the starting \textit{URA3-Int} cassettes are \textit{Ura}^{-} and, thus, sensitive to the \textit{5-FOA}. For each strain, we determined the frequencies of \textit{5-FOA} derivatives in multiple independent cultures and converted those data into rate estimates (details in \textit{SI Methods}). Genomic DNA samples from individual \textit{5-FOA} isolates were examined by PCR using primers flanking the \textit{Int} (URI71 UIR72 or 1829R/1829R). This analysis indicated whether the number of telomeric repeats had changed in those isolated or whether they contained chromosome rearrangements. Some of the isolates with an unaltered tract were sequenced to detect point mutations in the body of the \textit{URA3} cassette (\textit{SI Methods}).

**Analysis of Chromosome Rearrangements.** The \textit{5-FOA} derivatives of SMY749 that did not produce a PCR fragment containing \textit{Yel} repeats were examined for chromosome rearrangements using CHEF gel electrophoresis and CGH microarrays (\textit{SI Methods}). For some of these derivatives, the rearrangement breakpoints were analyzed by standard Southern analysis and PCR (\textit{Fig. S3} and \textit{Table S5}).

**ACKNOWLEDGMENTS.** We thank K. Bloom, D. Gottschling, J. Rine, and members of the S.M.M. and T.D.P. laboratories for useful discussions; Gil Han for technical support; and Durwood Marshall from Tufts University Informatics Technology, Research and Geospatial Technology Services for statistical consulting. The research was supported by National Institutes of Health Grant GM60987 (to S.M.M.) and Grants GMS2319 and GM24110 (to T.D.P.).