The Chromosome End in Yeast:
Its Mosaic Nature and Influence on Recombinational Dynamics

Edward J. Louis,* Elena S. Naumova,† Alice Lee,‡ Gennadi Naumov† and James E. Haber‡

*Yeast Genetics, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, England, †Scientific Research Institute for Genetics and Selection of Industrial Microorganisms, Moscow 113545, Russia, and ‡Rosenstiel Center, Brandeis University, Waltham, Massachusetts 02254

ABSTRACT

Yeast chromosome ends are composed of several different repeated elements. Among six clones of chromosome ends from two strains of Saccharomyces cerevisiae, at least seven different repeated sequence families were found. These included the previously identified Y' and X elements. Some families are highly variable in copy number and location between strains of S. cerevisiae, while other elements appear constant in copy number and location. Three repeated sequence elements are specific to S. cerevisiae and are not found in its evolutionarily close relative, Saccharomyces paradoxus. Two other repeated sequences are found in both S. cerevisiae and S. paradoxus. None of those described here is found (by low stringency DNA hybridization) in the next closest species, Saccharomyces bayanus. The loosely characterized X element is now more precisely defined. X is a composite of at least four small (ca. 45–140 bp) sequences found at some, but not all, ends. There is also a potential "core" X element of approximately 560 bp which may be found at all ends. Distal to X, only one of six clones had $(TG_{1-3})_n$ telomere sequence at the junction between X and Y'. The presence of these internal $(TG_{1-3})_n$ sequences correlates with the ability of a single Y' to expand into a tandem array of Y's by unequal sister chromatid exchange. The presence of shared repeated elements proximal to the X region can override the strong preference of Y's to recombine ectopically with other Y's of the same size class. The chromosome ends in yeast are evolutionarily dynamic in terms of subtelomeric repeat structure and variability.

THE chromosomes of many organisms contain subtelomeric regions composed of repeated sequences: Chironomus (SAIGA and EDSTRÖM 1985); Drosophila (RUBIN 1978; VALGEIRSDÓTTIR, TRAVERSE and PARDUE 1990; KARPEN and SPRADLING 1992); humans (BROWN et al. 1990; CHENG et al. 1991; COOKE et al. 1985; CROSS et al. 1990; DE LANGE et al., 1990; INGLEHEARN and COOKE 1990; RIETHMAN et al. 1989); Plasmodium (CORCORAN et al. 1988; DORE et al. 1990; FOOTE and KEMP 1989; PACE et al. 1990); Saccharomyces (CHAN and TYE 1983a, b; SZOSTAK and BLACKBURN 1982; WALMSLEY et al. 1984); Secale (BEDBROOK et al. 1980); Trypanosome (PAYS and STEINERT 1988); and tomato (BROUN et al. 1992). In many cases there is a mosaic of several different elements that vary in copy number and location between strains. This variation is often sufficient to provide characteristic DNA fingerprints for different strains or lines of organisms (BROUN et al. 1992). Despite their prevalence, very little is known about the origin or function of these elements (ZAKIAN 1989).

The yeast chromosome end is also composed of different subtelomeric repeated elements. The well characterized Y' element is immediately adjacent to the $(TG_{1-3})_n$ telomere sequences on some chromosomes. Y's vary a great deal in copy number and in location between strains, such that a single probe and a single restriction digest yields unique fingerprints for different

strains (HOROWITZ et al. 1984; LOUIS and HABER 1990a). The variation in copy number and location of Y's has been studied extensively in several strains of Saccharomyces cerevisiae (CHAN and TYE 1983b; HOROWITZ et al. 1984; JÄGER and PHILIPPSEN 1989; LINK and OLSON 1991; LOUIS and HABER 1990a; ZAKIAN and BLANTON 1988) as well as in strains of the closely related species Saccharomyces paradoxus (NAUMOV et al. 1992).

Y's fall into two major size classes, 6.7-kb Y'-long (Y'-L) and 5.2-kb Y'-short (Y'-S), that differ by a series of small insertions/deletions (CHAN and TYE 1983a; LOUIS and HABER 1992). In addition there are other types of Y's found in some strains but not others. Two strains out of 24 studied have a few copies of a degenerate form of Y' consisting of the internal half but not the telomere half of the element (LOUIS and HABER 1990a; NAUMOV et al. 1992). One strain has two 7.0-kb Y'-extra longs (Y'-E) (LOUIS and HABER 1992) which have an insertion near the telomere end. Sequence analysis of several Y's revealed a concerted nature to their evolution (LOUIS and HABER 1992) with about 1% divergence within a strain and up to 2% divergence between strains. Recombination among Y's can account for this homogeneity within strains; however, the maintenance of more than one type of Y' within a strain cannot be explained by simple recombination. One possible explanation of the simultaneous maintenance of several forms and homogenization may come from the observation that Y's

tend to recombine preferentially with Y's that are of the same size class (Louis and Haber 1990b).

Another subtelomeric repeat adjacent to Y's or the telomere is X. This element is less well characterized and is defined by restriction mapping and hybridization (Chan and Tye 1983b). This element may exist at all chromosome ends (Chan and Tye 1983b; Walmsley *et al.* 1984). The X sequences that have been sequenced are much less conserved than Y' sequences (Louis and Haber 1991) with only 82–92% identity between various Xs. As recombination among Y's is known to occur and accounts for the high identity between Y's, it is paradoxical that the adjacent sequence elements are not equally homogenized.

Between some X and Y' elements one finds stretches of $(TG_{1-3})_n$ sequences. However, this is not always true. We have previously reported two X-Y' junctions in which the first 49 bp of Y', the $(TG_{1-3})_n$ sequence and part of the X sequence are missing and replaced with an insertion of a 292-bp sequence. This insertion is in part composed of a segment of the mitochondrial intron, bi4 from cytochrome *b* (Louis and Haber 1991). We report below that other X-Y' junctions also lack $(TG_{1-3})_n$ sequences.

Subtelomeric regions have been found to harbor multicopy genes involved in carbon source metabolism. These include the *MAL* (Charron and Michels 1988; Charron *et al.* 1989; Michels *et al.* 1992), *SUC* (Carlson *et al.* 1985) and potentially the *MEL* (Naumov *et al.* 1990, 1991) genes. In addition a 20-kb duplication of the right end of chromosome *I*, containing *PHO11*, is found on chromosome *VIII* (de Steensma *et al.* 1989).

The ubiquity of mosaicism in subtelomeric regions among organisms from many taxa must be due to some common underlying process. Either there is a selective advantage to having mosaic repeats at chromosome ends or a common set of processes leads to the accumulation and maintenance of these repeats. Any advantage to having repeats must be sequence independent as there is no relationship among repeats between organisms (Zakian 1989). This is in contrast to the obvious evolutionary relationship between the functional telomere sequences of different organisms.

Here we present the characterization of several subclones from different chromosome ends: in particular, the variation in copy number and location among strains and between species. The sequence of the elements immediately adjacent to Y's is also presented. In addition, we have explored the types of recombination that occur between a particular pair of Y's and the influence of repeated elements adjacent to them.

MATERIALS AND METHODS

Strains: The strains used in this study are shown in Table 1. Eight strains of *S. cerevisiae* were used. YP1 is congenic with S288C. YP1, S288C and A364a are common laboratory strains

TABLE 1

Strains used in this study

Strain	Source or reference
S. cerevisiae	
YP1	Hieter *et al.* (1985)
S288C	Genetic Stock Center, Berkeley
A364a	Genetic Stock Center, Berkeley
SK1	Kane and Roth (1974)
Y55	McCusker and Haber (1988)
VKMY 502	Naumov (1977)
SBY 2576	Naumov (1977)
YNN 295	Genetic Stock Center, Berkeley
S. paradoxus	
CBS 432	Naumov (1987)
CBS 5829	Naumov (1987)
N7	Naumov (1987)
N8	Naumov (1987)
N9	Naumov (1987)
N11	Naumov (1987)
N12	Naumov (1987)
N15	Naumov (1987)
N17	Naumov (1987)
N18	Naumov (1987)
N25	Naumov (1987)
N34	Naumov (1987)
N36	Naumov (1987)
S. bayanus	
NRRLY 969S	Naumov (1987)
MCYC 623–6C	Naumov (1987)
VKMY 1146–6B	Naumov (1987)

derived by mutant isolation and/or genetic crosses while Y55 and SK1 are wild isolates that have not been outcrossed. SBY 2576 was isolated from wine while the origin of VKMY 502 is unknown. YNN 295 is a strain commonly used as a chromosome marker on CHEF gels. Thirteen strains of *S. paradoxus* were used as well as three strains of *Saccharomyces bayanus*. These have been described elsewhere (Naumov 1987; Naumov *et al.* 1992). Nearly all of these were originally wild isolates without outcrossing. The Y' families of YP1 and Y55 have been extensively studied and are the strains from which clones were obtained (Louis and Haber 1990a, 1992).

Clones: The clones shown in Figure 1 were obtained as previously described (Louis and Haber 1990a; Louis and Haber 1992). Fragments marked F1 through F5 have been subcloned in pEL50, pMD9, pMD3, pXRIRV and pEL59, respectively, for subsequent use as probes.

Sequence analysis: The internal regions of all of the Y' elements have been sequenced and reported on elsewhere (Louis and Haber 1992). The X-Y' junctions of clones 1–3 (from one end of *XV*, *IX* left and *X* left, respectively) have been sequenced and reported elsewhere (Louis and Haber 1991). The chromosome and end determination was made by comparison with the known electrophoretic karyotype and physical map of *S. cerevisiae* (Link and Olsen 1991). The chromosome determination in *S. paradoxus* is based on the electrophoretic karyotype and hybridization to at least eight chromosome-specific probes (Naumov *et al.* 1992). The X-Y' junctions of clones 4, 5 and 6 (from one end each of *XIII*, *XIV* and *VI*, respectively) were sequenced in one direction from the end of the Y' toward the centromere in order to determine whether there was any telomere sequence present at the junctions. This was accomplished by dideoxy-sequencing using Sequenase version 2.0™ as previously described (Louis and Haber 1992). These sequences were then compared to the previously determined X-Y' junction sequences as well to as other sequences derived from chromosome ends. Figures 2 and 3 compare the structures and sequences of the known junctions.

Hybridization of subtelomeric fragments: Genomic DNA from each strain was prepared for pulsed field gel analysis as previously described (LOUIS and HABER 1990a). The chromosomes were separated on a CHEF DRII apparatus (Bio-Rad) using standard conditions. These gels were blotted onto Hybond-N™ for probing. Probes of fragments F1–F5 were prepared using isolated fragment DNA from the original clones or as linearized subclones. Nonradioactive probes were produced using either digoxigenin-11-dUTP™ incorporation and AMPPD™ for detection (Boehringer Mannheim) or by using direct bonding of horseradish peroxidase to the probe DNA using the ECL-direct™ kit (Amersham). A probe for the small sequences immediately adjacent to the telomere or Y' elements was made using synthetic oligomers and PCR on pYP1-L1 (SAIKI *et al.* 1988). A 340-bp fragment was isolated from an agarose gel and labeled with digoxigenin-11-dUTP™ as described above. Two stringencies of washes were used after hybridization. A normal stringency wash was done with 0.1 × SSC, 1% SDS at 65°. A lower stringency wash used the same buffer at 42°. These data are presented in Figures 4 and 5.

Restriction mapping of a large triplication: Three ends containing homology to fragment F1 were mapped in three strains of yeast in order to determine the level of shared homology. Insertion of markers into F1-homologous locations was accomplished using pAL1 and one step transplacement or pEL63 and integration (ROTHSTEIN 1983). pAL1 contains an insertion of *URA3* and 692 bp of pBR322 inserted into the *ClaI* site of fragment F1 while pEL63 has *URA3* inserted into the vector sequence of pEL50 adjacent to the F1 fragment (see Figure 6). These transformants were analyzed genetically and by Southern blots of restriction fragments and CHEF gels. All three F1-homologous locations in each strain (YP1, S288C and Y55) were transformed with equal frequency. pBR322 was used as an insert specific probe. Restriction sites at all three locations were assayed.

Recombination at Y'-E containing ends: Mitotic recombination between chromosome ends was monitored using a dosage-dependent marker inserted into individual Y's as previously described (LOUIS and HABER 1990b). YP1 diploids with a single *SUP11* marked Y' are adenine requiring while those with two copies are adenine prototrophic. Y's donors for recombination between Y's were created either by transplacing in the *SUP11* gene as previously described (LOUIS and HABER 1990a) or by mitotic segregation of new marked Y's from their donors. Marked Y'-extra longs (Y'-E) were used as donors to add to the previous data base generated using marked Y'-longs (Y'-L) and Y'-shorts (Y'-S) (LOUIS and HABER 1990b). Ade⁺ events were analyzed by CHEF and Southern analysis in order to determine the structure of the second copy of the marked Y' as previously described (LOUIS and HABER 1990b). The data from marked Y'-Ls at one end each of *XV*, *XVI*, *V* or *VIII*, a marked Y'-S at one end of *XIII*, and Y'-Es at one end each of *IX* and *X* were combined for analysis.

The creation of a Y'-S at a location normally occupied by a Y'-E was accomplished by analyzing gene conversions between a strain containing a *SUP11*-marked Y'-S donor at one end of *XIII* and two recipient Y'-Es on chromosomes *IX* and *X*. In one out of four cases the recipient changed size class to a marked Y'-S creating a novel situation with a marked Y'-S at one of the two Y'-E locations. This novel marked Y' was segregated away from its donor for subsequent recombinational studies.

RESULTS

The six clones from YP1 and Y55 (see Figure 1) contain the junctions between the internal end of the Y' and the adjacent sequences. The sequence of three of these

junctions (YP1-L1, YP1-L3 and YP1-L4) as well as three Y'-Y' junctions from strain YP1 have been reported elsewhere (LOUIS and HABER 1991; LOUIS and HABER 1992). Two additional X-Y' junction sequences (pCD2 and YRp120, see Figure 2) have also been reported elsewhere (WALMSLEY *et al.* 1984). Two of the X-Y' junctions (YP1-L3 and YP1-L4) contain no $(TG_{1-3})_n$ sequence but did contain a 292 segment that includes part of the mitochondrial cytochrome *b* intron bi4 (LOUIS and HABER 1991). Clone YP1-L1, as well as pCD2 and YRp120, do contain $(TG_{1-3})_n$ sequence. The three clones from strain Y55 (Y55-L1, Y55-L2 and Y55-S1), were sequenced on one strand from the Y' toward the centromere in order to determine whether there was any $(TG_{1-3})_n$ sequence on the internal end of the Y's. These are shown in Figure 2, in comparison with the previously reported X-Y' junctions and X sequences. In none of these three new sequences is there any significant amount of telomere sequence.

X elements are composed of a mosaic of at least five different short repeat sequences: Immediately adjacent to Y' sequences should be the X element (CHAN and TYE 1983a). However, inspection of these sequences shows that they are all different from each other, though each is partially homologous to different regions of previously defined X elements. Comparison of these sequences to the previous three X-Y' junctions (LOUIS and HABER 1991) as well as chromosome *III* left and *III* right (BUTTON and ASTELL 1986; LOUIS 1994; OLIVER *et al.* 1992) and two other clones of X-Y' junctions (pCD2 and YPp120 described in WALMSLEY *et al.* 1984) reveal the presence of several short sequence elements found at some ends but not others (Figure 2). Four of these subtelomeric repeated elements (STRs), designated STR-A, STR-B, STR-C and STR-D, are present in different arrangements adjacent to the telomere or Y'. In addition, more centromere-proximal is another sequence that may be a "core" X sequence, in that all clones that that have been sequenced far enough share this sequence. The sequences of the "core" X, STR-A, STR-B, STR-C and STR-D are presented in Figure 3, A and B.

STR-A varies in length from 35 to 50 bp. The sequences contain what appear to be repeats of TTAGGG and several degenerate versions of this sequence. Each example of the STR-A element has at least one copy of TTAGGG among the degenerate versions. TTAGGG is the functional telomere repeat of many organisms ranging from fungi to humans.

STR-B is a more conserved element of 56/57 bp (as defined by the eight consecutive repeats of this element in clone Y55-L1) with 40 bases identical in all six examples. These elements are 81–93% identical in pairwise comparison.

STR-C (42–46 bp) is also conserved among four examples. These share 38 identical positions and are 90–
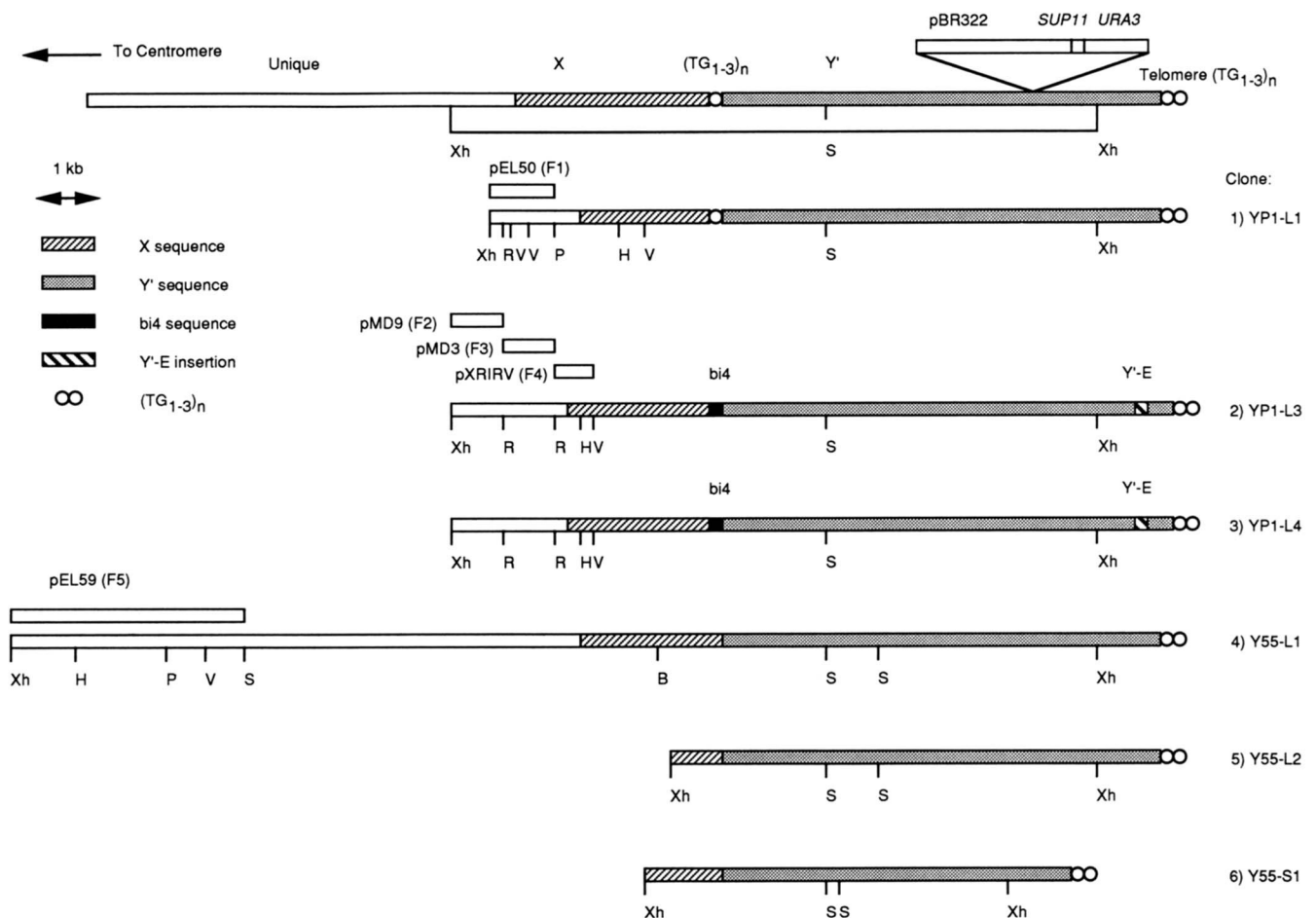
FIGURE 1.—Six clones from two strains containing Y', X and internal sequences. Clones were obtained by integrating selectable markers as indicated into individual Y' elements in both YP1 and Y55. Genomic DNA from these transformants was digested with *Xho*I, ligated and transformed into *E. coli*. Subclones were made in pGEM3Zf(−) as indicated (fragment F1 is in pEL50, F2 in pMD9, F3 in pMD3, F4 in pXRIRV, and F5 in pEL59). Restriction sites are as follows: Xh, *Xho*I; S, *Sal*I; R, *Eco*RI; V, *Eco*RV; P, *Pst*I; and H, *Hin*dIII.

97% identical in pairwise comparison. All three elements, STR-A, STR-B and STR-C contain a G-rich strand. The STR-C element contains the *S. cerevisiae* telomere-like sequence TGGTGGT. However reiterations of TGGTGGT are in fact not found among normal telomeres (SHAMPAY *et al.* 1984; KRAMER and HABER 1993).

The longest of the newly defined elements, STR-D (138–150 bp), is also well-conserved (86–95%) over most of its length among six examples, one of which is truncated by the insertion of the cytochrome *b* bi4 sequences.

The "core" X sequence is more variable. These sequences are 57–92% identical in pairwise comparisons. This element includes the ARS element found in X sequences (CHAN and TYE 1983a, b) which is completely conserved in 3 out of 4 of the sequences (Figure 3). The core X is defined as the sequences shared among all ends that have been sequenced so far. The maximal size of the element is bounded at the internal end by shared homology between the ends of chromosome *III*. Whether there is more sub-structure of the core X region awaits additional cloning and sequencing of these regions from more telomeres. There is no significant

homology between these five elements and no internal inverted or direct repeats within each element other than the TTAGGGs of STR-A. The level of identity among the STR-A, STR-B, STR-C, STR-D and core X sequences contrasts with the 98–99% identity found in Y' elements.

The order of the short STR elements is conserved, however, their presence/absence is variable. Four of the ends compared have all four of these STR regions and $(TG_{1-3})_n$ telomere sequence in the order $DCBA(TG_{1-3})_n$ prior to the end or the adjacent Y'. One end has the structure $DBA(TG_{1-3})_n Y'$. Two have the structure D(bi4)Y' where bi4 is a 292-bp element that in part contains the 5' end of the fourth intron of the mitochondrial gene cytochrome *b* (LOUIS and HABER 1991). Two of the three ends from Y55 have the potential "core" X sequence immediately adjacent to the Y' while the third has at least 8 tandem copies of STR-B immediately adjacent to the Y'.

Regions centromere-proximal to X sequences are also composed of repeated sequences: Five different subclones from centromere-proximal regions of clones
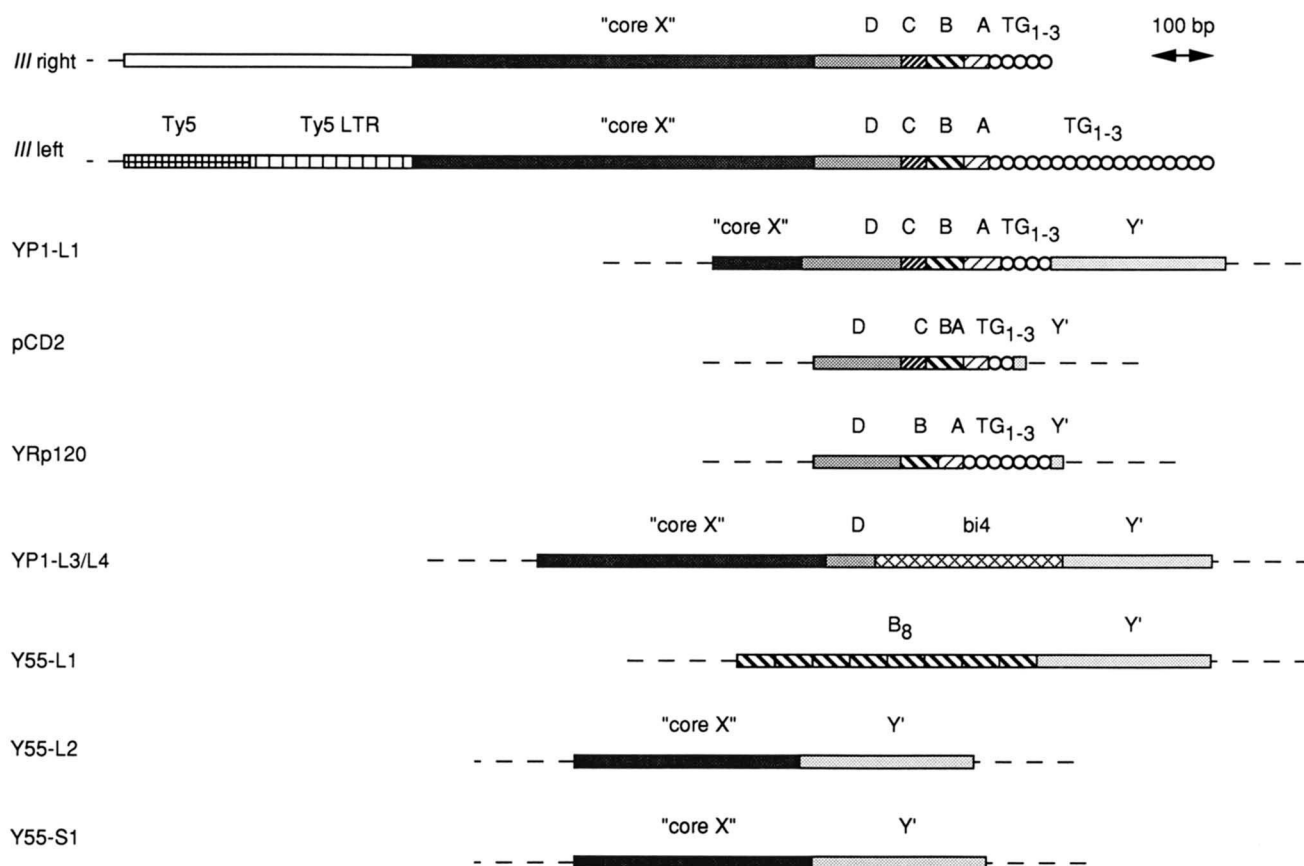
FIGURE 2.—The sequences of the Y′ adjacent regions compared to known chromosome ends reveal a mosaic of several small subtelomeric repeated elements. The new sequences represented here are all in GenBank with accession numbers: M63936, M63937 and M63938 for the X-Y′ junctions of pY55-L1, pY55-L2 and pY55-S1, respectively. The other sequences are previously published: *III* right and *III* left (BUTTON and ASTELL 1986; LOUIS 1993; OLIVER *et al.* 1992); YP1-L1 (LOUIS and HABER 1992); pCD2 and YRp120 (WALMSLEY *et al.* 1984); YP1-L3 and YP1-L4 (LOUIS and HABER 1991, 1992). At least four small elements, STR-A, STR-B, STR-C and STR-D, in addition to a potential "core" X element are defined by these comparisons.

YP1-L1 YP1-L2 and Y55-L1 were isolated (see Figure 1). These probes, designated F1–F5, were each used to probe CHEF blots of several strains of *S. cerevisiae*, *S. paradoxus* and *S. bayanus* (see Table 1 and Figure 4). In every case the probe was homologous to more than one location in the genome. Thus, none of these clones from chromosome ends are entirely composed of unique sequences, even in one that includes sequences 20 kb internal from the end. A summary of the data from these hybridizations is shown in Figure 5.

Three of these new subtelomeric repeats (F1, F2 and F3) are specific to *S. cerevisiae* as no homology was found in any of the *S. paradoxus* or *S. bayanus* strains. The F1 repeat is located at the same three chromosome ends (in chromosomes *XV*, *XIII* and *XVI*) in all seven strains of *S. cerevisiae* (Figure 4B, see below for further characterization).

The second and third *S. cerevisiae*-specific repeats, F2 and F3, are variable in both copy number and location among strains (Figure 4, C and D). Fragment F2 shows strong homologies with chromosome *IV* in SK1 and Y55; chromosome *VII* or *XV* in Y55; chromosome *X* in YP1, S288C, A364a, SK1 and VKMY 502; and chromosome *IX* in YP1, S288c and SK1. All of the *S. cerevisiae* strains have weaker

homology with F2 on chromosome *III*. Homologies to fragment F3 are found on chromosomes *IV*, *XIV*, *X*, *XI*, *IX* and *III* in all seven *S. cerevisiae* strains (except for lack of *X* homology in Y55 and SBY 2576). As with F2, the strongest homologies are at *IX* and *X* in some of the strains.

Fragment F4 is immediately adjacent to F2 and F3 on the telomere side in the genome but has homologies with nearly all chromosomes in all strains of *S. cerevisiae* as well as *S. paradoxus* (Figure 4F). This fragment is only about 500 bp and may overlap part of a "core" X sequence though it does not hybridize to every chromosome in every strain. As can be seen in Figure 4F, the *S. paradoxus* chromosome that comigrates with *S. cerevisiae*'s chromosome *X* does not have F4 homology. F4 homology is weak for some of the chromosomes in *S. cerevisiae* strains YNN 295 and VKMY 502.

The last fragment, F5 in pEL59, is found in all strains of *S. cerevisiae* as well as *S. paradoxus*. It is intermediate in variability in copy number and location. There are several locations that appear constant and other locations that vary in some strains (Figure 4, G and H). There are locations with strong signal and others with weaker signal indicating either homology to only part of the element, more divergence over the fragment, or variable copy number of the element.

A

```
              1                                                                                            100
Y55-S1
III right 314507 TGAAGTGGCATCCAGTCAAGCAAATGGTAATATTAAGGAACTTTTAAGTTAATGACGTCATGGTAGTGCTCGTACTTCAAGTCAAAGTGTTTATGTATTA
III left  1171   TGAAATGGCATCCGTTCAAGCAATTTGCAATATTAAGGAACTTTTACGTTAATGACGTCATGGTGGTATAGATACGCAAT TGAGTGTGTTTATGTATTA
YP1-L3/L4
Y55-L2
YP1-L1                                                                                  CTTCAAGCCAATGTGTTTATGTATTA
```

```
              101              ARS element                                                                 200
Y55-S1                TTTT CGTTTAGCTGAGTTTAACGGTGATTATTAGGTGGCATTTTATATTAGTCTACATAAAAATAAGTGG T GGATAT
III right     TGGTTGAAGAATAGAATATTTTTATGTTTAGGTGATTTTAGTGGTGATTTTTCTGTAATA TTGACATAAGTGTATAT AAATTAAGTGG TTAGTATAC
III left      TTGTTGAAGAGTAGAATATTTTTATGTTTAGGTGATTTTAGTGGTGATTTTGTGGTTATA TTGACATAAGTGTGTAT AAATTGAGTGG TTAGTATAT
YP1-L3/L4     T GTTGAAGAATAGAAATATTTTTATGTTTAGGTGATTTTGGTGGTGATTTTTATGTAATA TTGGCATAAGTGTATAT AAATTAAGTGG TTAGTATAT
Y55-L2                                                                           AAGTGTATAT AAATTAAGTGGCTTAGTATAT
YP1-L1
```

```
              201                                                                                          300
Y55-S1        CTACATAAAATTGTCATAACGCGTAAACTAAAAATTATTTTTATGATCATTGAGGATCTATAATCAACTATAGACATTAATGTATGGATAATCATGAGGA
III right     GGTGA AAAAGAGGTATAACGTATGTA TTAAGGGAATTTATACGATATTTGGGCCCGCCGAATCAGATATAGATATTAAAATGTGGATAATCGTGGGCT
III left      GGTGT AAAAGTGGTATAACGTATGTA TTAAGGGCAGTTATACAATATTTGTGTCGCCGAATCAGATATAGATATTAAAATGTGGATAATGATGGGCG
YP1-L3/L4     GGTGA AAAAGTGGTATAACGTATGTA TTAAGGGCAGTTATACAATATTTGGGCCGCCGAATCAGATATAGATATTAAAATGTGGATAATCGTGGGCT
Y55-L2        GGTGC AAAAGTGGTATAACGTATGTA TTAAGGGCATTTATACAATATTTGGGCCGCCGAATGAGATATAGATATTAAAATGTGGATAATCGTGGGAG
YP1-L1
```

```
              301                                                                                          400
Y55-S1        TTATAGGTAAATGGGA AGGGTAAAAACCAGTGAGGCCATTTCCGTGTGTAGTGATCCGAACTCAGT  TACT ATTGA TGGAA ATGAGGACTGG G
III right     TTATGGGTAAATGGCACAGGGTATAGACCGCTGAGGCAAGTGCCGTGCATAATGATGTGGGTGCATTTGGTACTGATTTAGTGAGA ATGGGCCATGGAT
III left      TTATGCGTAAATGGCACAGGGTATAGACCGCTGAGGCAA GCCGTGCATAATGATGCGAGTGCATT GGTACTGAT GGTGAGAGGGGTGATG AT
YP1-L3/L4     TTATGGGTAAATGGCACAGGGTATAGACCGCTGAGGCAAGTGCCGTGCATAATGATGTGAGTGCATTT ATACTGATTTAGTGAGA ATGGGCCGTGGAG
Y55-L2        TTATAGGTAAATGGCACAGGGTATAGACCGCTGAGGCAAGTACCGTGCATGATGATGTGGGTGCATTTGGTACTGATTTAGTGAGA ATGGGCCTAGGAG
YP1-L1
```

```
              401                                                                                          500
Y55-S1        TCATGGGGCGCAATGGAGTGAAGTAA           TATATACTTTAGCATACGTGTGCGTATGCCATATCAATATGCTAGTG TGGTG    TGGTG
III right     TG    GAGTGTGAGAGTAGGGTAAC TTGAGAGTGGTATATACTGTAGCATCCGTGTGCGTATGCCCCATCAATAT AAGTGAAGGTGAGTATGGCA
III left            GAATATGAAAGTAGG TAAGTTTGAGA TGGTATATACTGTAGCATCCGTGTGCGTATGACCGATCAGAATACAAGTGAAGGTGAGTATGGCA
YP1-L3/L4     TGGAGTGGAATATGAGAGTAGGGTAAC TTGAGA T GTATATACTGTAGCATCCGTGTGCGTATGGCATATCAGTATACAAGTGAGGGTGAGTGTGGCA
Y55-L2        TG    GAATGTGAGAGTAGGGTAAC TTGAGA TGGTATATACTGTAGCATCCGTGTGCGTACGTAAAATCAGTATACAAGTGAGGGTGAGTATGGCA
YP1-L1                 AGG TAAG TTGAGAGTGGTATATACTGTA CATCCGTGTCG TATCGCATATCAGTATACAAGTGAAGGTGAGTATGGCA
```

```
              501                                          559
Y55-S1        TG
III right     TGTGGTGGTGG    TATAGAGTGGTAGGGTAAGTATGTATGTATTATTTACGATC
III left      TGTGGTAGTGGGAT    TAGAGTGGTAGGATAAGTATATGTGTATTATTTACGATC
YP1-L3/L4     TGTGGTGGTGGGAT    TAGAGTGGTAGGGTAAGTATGTGTGTATTATTTACGATC
Y55-L2        TGTGGTGGTGGTGG    TATAGAGTGGTAGGGTAAGTGTATGTGTATTATTTACGATCA
YP1-L1        AGTGGTGGTGGGATTGGTATAAAGTGGTAGGGTAAGTATGTGTGTATTATTTACGATC
```

B

```
Sequence      D element

III right     ATTTGTTAACGTTTCAATATGGTGGGT GAACAACAGTACAGTGAGTAGGACATGGTGGATGGTAGGGTAATAGTAGGGTAAGTGGTGGTGGAGTTGGAT
III left      ATTTGTAGACGTTTCAATATGGTGGGT GAACAACAGTACAGTGAGTAGGACATGGTCGATGGTA           GGGTAAGCGGTAGTGGAGTTGGAT
YP1-L1        ATTTGTTAACGTTTCAATATGGTAGGTAGAAGAACAGTATGGTGAGTAGCAGATGGTGGATGGTAGAGGAATAGCAGGGTAAGTGGTAGTGGAGTTGGAT
YRp120                    394 ACGTTTCAATATGGAGGGTAGAACAACAGTACAGTGAGTAGGACATGGTGGATGGTAGGGTAATGGCAGGGTAAGTGGCAGTGAGGTTGGAT
pCD2                       349 ACGTTTCAATATGGAGGGTAGAACAACAGTACAGTGAGTAGGACATGGTGGATGGTAGGGTAATAGTAGGGTAAGTGGTAGTGGAGTTGGAT
YP1-L3/L4     ATTTGTTAACATTTCAATATGTTGGGTAGAACAACAGTATAGTGAGTAACAGATGGGCATGGTAGGGTAATGGCAGGGT  ->bi4
```

```
Sequence      D element continued                                     C element

III right     ATGGGTAATTGGAGGGTAACGGTTATGATGGGCGG TGG ATGGTGGTAGT    AAGTAGAGAGATG    GATGGTGGTTGGGAGTGGTATGGTTGAGTG
III left      ACGGGTAGTTGGAGGGTAACGGTTATGGTGGACGG TGGGTTGGTGGTAGT    AAGTAGAGAGATG    GATGGTGGTTGGGAGTGGTACGGTTGAATC
YP1-L1        ATGGGTAATTGGAGGGTAACGGTTATGGTGCACGA TGGGTTGGTGGTAGC    AAGTAGAGAGATG    GATGGTGGTTGG AGCGGTATGGTTGAAGG
YRp120        ATGGGTAATTGGAGGGTAACTGTTATGGTGGATTTGTGGGTTGGTGGTCGG
pCD2          ATGGGTAATTGGAGGGTAACGGTTATGGTGGACGG TGG TTAGTGGT      AAGTAGAGAGATGATGGATGGTGGTTGGGAGTGGTATGGTTGAATG
```

```
Sequence      B element                                              A element

III right     GGGCAGGGTAACGAGTGGGGAGGTAGGGTAATGTGAGGGTAGGTTTGGAGACAGGTA AAATC   AGGGTTAGAATAGGGTAGTGTTAGGGTAGTG
III left      AGACAGGGTAACGAGTGGAGAGGTAGGGTAATG GAGAGTAAGTTGGGAGACAGGTA AAATC   AGGGTTAGAATAGGGTAGTGTTAGGGTAGTGTTAGGG
YP1-L1        GGACAGGGTAACGAGTGGGGAGGTAGCGTAATG GAGGGTAAGTTAAGAGACATGCT AAATC   AGGGTAAGAATAGGGTAGGGTTAGGGTAGGGTTAGGGTAGTGTTAGGG
YRp120        GAGGCGGGTAACGAGTGGGGAGGCAGGGTAATG GAGGGTAAGTTGAGAGACAGGAT GGAAG   GGTTGGATTAGGGTAGGGTTGAGGTAGTATTAGGG
pCD2          AGACAGGGTAACGAGTGGGGAGGTAGGGTAATG GAGGGTAGGTTTGGAGACAGGTT CATCA   GGGTTAGAATAGGGTACTGTTAGGATTGTGTTAGGG
Y55-L1        (GGACAGGGTAACGAGTGGGGAGGTAGGGTAATG GAGGGTAAGTTCAGAGAGAGGTT)gGGACA
```

FIGURE 3.—The sequences of (A) the "core" X element as well as (B) the STR-A, STR-B, STR-C and STR-D elements are compared. Sequence alignments of the five elements defined as in Figure 2 were obtained using PILEUP in the GCG software package (DEVEREUX et al. 1984). The pairwise identities were determined using DISTANCES of the same software package. Underlined bases are identities shared in all clones. The X type ARS element (TTTTATGTTTA) is indicated in the "core" X sequence. The starting base pair as given in GenBank for *III* right, *III* left, YRp120 and pCD2 are shown. All the others start at base pair 1. *III* left, YRp120 and pCD2 are displayed in reverse orientation.
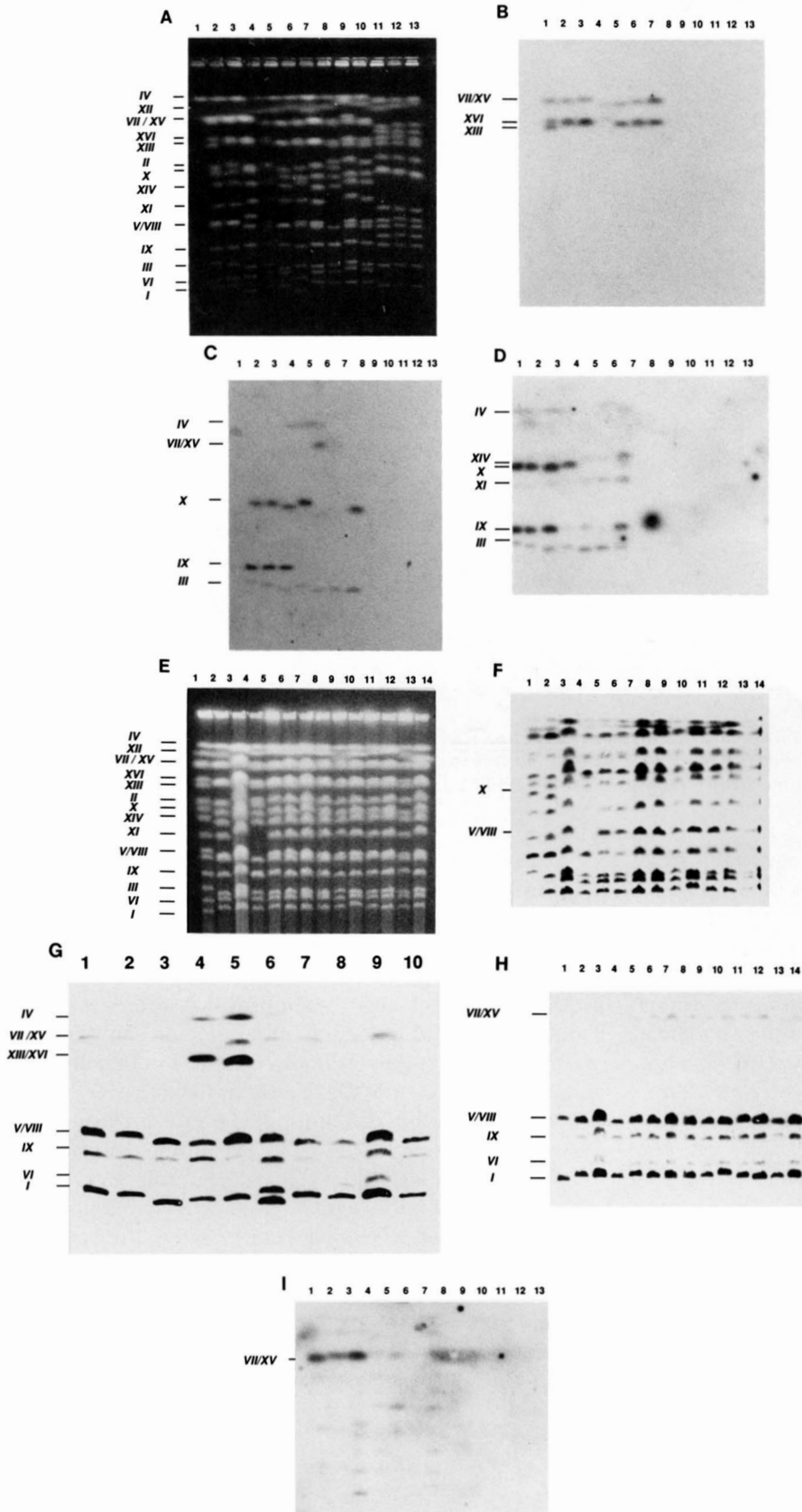
FIGURE 4.—Blots of strains with F1–F5 and DCBA. Panel A shows the ethidium bromide stained CHEF gel for panels B–D, and I. Panel E shows the ethidium stained CHEF gel for panels F and H. The panels B–D and F–I show the hybridization of probes (F1, F2, F3, F4, F5, F5 and DCBA, respectively) to CHEF blots of *S. cerevisiae* strains and *S. paradoxus* strains. Lanes 1–13 in panels A–D and I contain the strains: YP1, S288C, A364a, SK1, Y55, SBY 2576, VKMY 502, CBS 5829, CBS 432, N7, NRRLY 969S, MCYC 623–6C and VKMY1146–6B. Lanes 1–14 in panels E, F and H contain the strains: YNN295, VKMY 502, CBS 432, CBS 5829, N8, N9, N11, N12, N15, N17, N18, N25, N34 and N36. Lanes 1–10 in panel G contain the strains: YP1, S288C, A364a, SK1, Y55, SBY2576, VKMY502, CBS 5829, CBS 432 and N7.

From the sequences in YP1-L1, primers were synthesized that allowed PCR amplification of the STR-A, STR-B, STR-C and STR-D elements as a single 340-bp fragment for probing Southern blots. Most chromosomes of *S. cerevisiae* strains have homology to these elements under low stringency conditions. Under the

Strains:
S. cerevisiae
S. Paradoxus
Chromosomes
YP1
S288C
A364a
SK1
Y55
SBY 2576*
VKMY 502
CBS 5829
CBS 432
N7
IV
XII
VII and XV
XIII and XVI
II
XIV
X
XI
V and VIII
IX
III
VI
I

FIGURE 5.—Summary of chromosomal locations of fragments F1–F5 as well as Y′s and the mitochondrial intron bi4. The locations of F4 are not shown, rather the absence of F4 in the few locations in *S. paradoxus* strains is noted. The bi4 locations come from LOUIS and HABER (1991). The Y′ locations come from LOUIS and HABER (1990b) and NAUMOV *et al.* (1992). *The Y′ locations in strain SBY 2576 were not determined.

normal stringency washes at 65°, homology to this fragment is found at only some of the chromosomes (Figure 4I). In contrast the *S. paradoxus* strains only have homology at one chromosome at the lower stringency conditions (data not shown).

None of the fragments exhibited homology in strains of *S. bayanus*, even though several unique sequence probes from *S. cerevisiae* have been shown to hybridize to DNA from *S. bayanus* strains (NAUMOV *et al.* 1992).

To further characterize the degree of shared DNA sequences at three chromosome ends sharing fragment F1 homology, we inserted *URA3* into each of the F1-homologous regions, by transplacements of an F1 fragment into which *URA3* had been inserted. Insertions of *URA3* have been obtained at all three chromosomes for strains YP1, S288C and Y55. Southern blot analysis of restriction digests of these transformants could then be used to map the restriction sites adjacent to the *URA3* marker. The restriction maps around F1 at all three locations (*XIII*, *XV* and *XVI*) in YP1, S288C and Y55 are shown in Figure 6. Except for a Y′ missing at the *XIII* location in both YP1 and S288C and the absence of Y′s at all three locations in Y55, the maps are nearly identical for up to 12 kb. The chromosome *XIII* location in YP1 is missing the *Bam*HI site. It is likely that the shared restriction maps and telomere-proximal locations for this repeat hold in all strains of *S. cerevisiae*.

The presence of shared subtelomeric regions may account for some of the bias in Y′ × Y′ ectopic recombination: Recombination events involving a Y′ element with a dosage-dependent *SUP11* gene can be readily detected in an adenine-requiring strain as Ade$^+$ cells (LOUIS and HABER 1990b). In previous studies we noted that a marked Y′ tended to undergo gene conversion or reciprocal recombination with other Y′s of the same size class (Y′-S, Y′-L or Y′-E) (LOUIS and HABER 1990b). For example, in strain YP1, among 33 events involving a Y′-short (Y′-S) donor, 23 (70%) recombined with other Y′-Ss, even though they represented only 36% of the total Y′ population (Table 2A). Similarly among 65 events involving a Y′-long (Y′-L) donor, the recipient was also a Y′-L 89% of the time. The few events involving Y′-less chromosome ends presumably occurred by recombination of more centromere-proximal subtelomeric repeats such as the X region.

In strain YP1 the Y′s at the left arms of *IX* and *X* are a novel size class. They are the only two Y′s with an insertion of about 300 bp near the telomere end. When one of these two extra long Y′s (Y′-Es) is marked and used in the same recombination assay, 59% of the events involve the other Y′-E as the recipient. Most of the rest were interactions with Y′-longs (Table 2B). This is consistent with the observation that Y′s interact within their own size class.
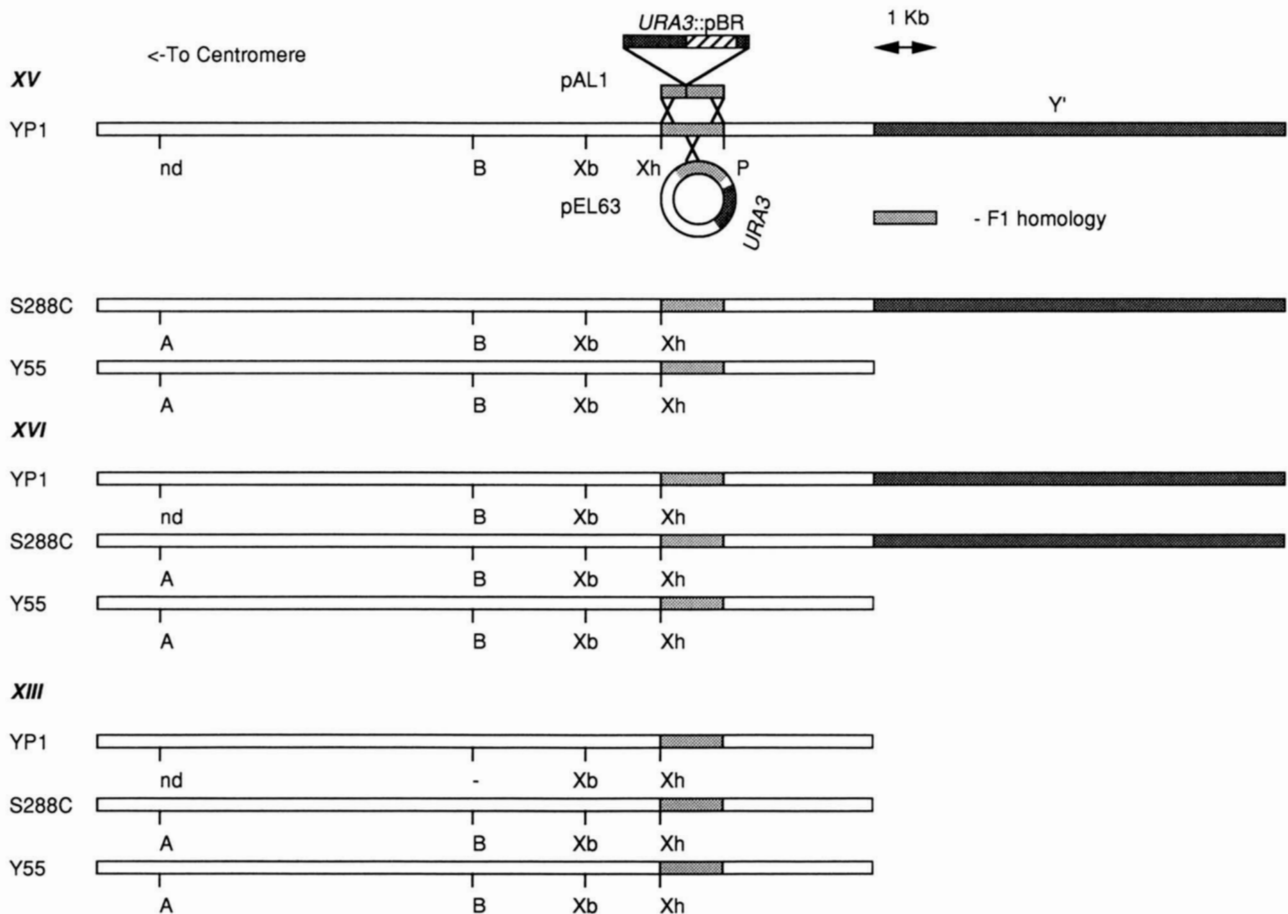
FIGURE 6.—Restriction maps of the ends of *XIII*, *XVI* and *XV* which all have homology to fragment F1. The three locations in YP1 were marked with pAL1 via transplacement. The pBR322 insertion was used to aid in restriction mapping. There is a *Bam*HI site missing in the corresponding location on chromosome *XIII* in YP1. The *Asp*718 sites were not assayed in this strain. S288C and Y55 were marked by integrating pEL63 cut with *Cla*I to direct to F1 homologous locations. Again the pBR322 homology was used to aid in restriction site mapping. All three locations in these two strains share maps up to 12 kb from the telomere or beginning of Y'. nd, not determined; A, *Asp*718; B, *Bam*HI; Xb, *Xba*I; Xh, *Xho*I; and P, *Pst*I.

TABLE 2

Recombinational interactions among Y's

	Recipients				
Donors	Short	Long	Extra long	Y'-less	Tandem
A. Earlier study (Louis and Haber 1990b):					
Short	11	4[a]	1	0	5
Long	3	45	1	3	16
B. This study:					
Short	12	2	3	0	ND[b]
Long	0	14	1	0	ND
Extra long	0	14	23	2	0
Short @ E	7	3	10	0	0

[a] Underlined data are cross size class interactions analysed for change in the recipient size class (see text).
[b] Not determined.

A question that arises is whether the two Y'-E ends share Y' homology because they tend to interact with each other (resulting in their homogenization) or whether they interact with each other due to their shared Y' homology. One way to address the issue is to replace one of the Y'-Es with a Y' of a different size class. From events involving a marked Y'-S on chromosome

XIII, we had previously obtained several transfers of the *SUP11* and *URA3* marker sequences to either the Y'-E at *IX* or the one at *X*. These were analyzed to identify one in which the recipient Y' was now a marked Y'-S. In one of four cases, the newly marked Y' was short instead of extra long. As with most other interactions between Y's of different size classes, the recipient usually retains its original

A

B
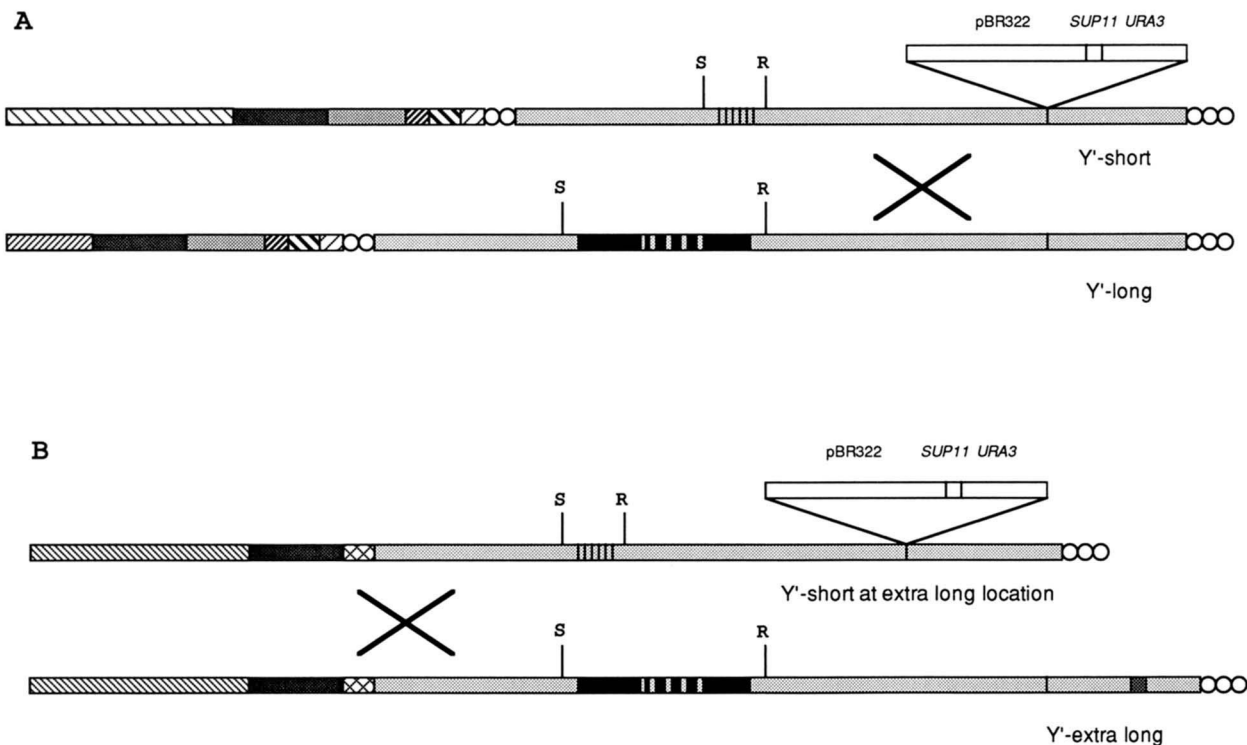
FIGURE 7.—Influence of adjacent repeated sequences on Y′ × Y′ recombination. The shading of subtelomeric elements is as in Figure 2. The black blocks between the *Sal*I and *Eco*RI sites within the Y′ indicate the insertion/deletion differences between Y′-longs and Y′-shorts. The position of the crossover is indicated by the large X. (A) A schematic of Y′-Y′ interactions shows how in the normal case crossovers occur between the inserted marker sequence and the long *vs.* short sequence differences. (B) In contrast, the interactions involving the left ends of *IX* and *X* all have crossover points proximal to the long *vs.* short differences. S, *Sal*I; R, *Eco*RI.

size class. This new marked Y′-S at chromosome *IX* was segregated away from its donor Y′ and then used in subsequent recombinational studies. When this Y′-S is used as a donor, the other Y′-E is still the recipient in 50% of the events [making these events independent of Y′ size class (see Table 2)]. Among the remaining 10 events there still was a preference for Y′s of the same size class as the Y′-S donor.

One way to explain the continued preference for a Y′-E target by the new Y′-S (occupying the location normally held by a Y′-E) is illustrated in Figure 7. Gene conversions and reciprocal exchanges can occur between Y′s without changing the size class of the recipient (Figure 7A). However, reciprocal recombination between more centromere-proximal repeated sequences can also result in the acquisition of donor Y′s size class (Figure 7B). If two chromosomes share extensive amounts of subtelomeric homology, many of the events can be in this region. RILES *et al.* (1993) have found extensive homology of the two ends involved in this experiment (*IX* left and *X* left). In all such cases, the recipient chromosome end will acquire the size class of the Y′ at the donor. When analyzed by restriction mapping, we found that 10/10 acquisitions of a marked Y′ on the other Y′-E chromosome were Y′-S rather than remaining Y′-E as expected from previous events. The 10/10 switches of size class is statistically different from the 2/13 observed in other cross size class interactions (underlined data in Table 2, A and B).

Expansion into a tandem array depends on internal TG$_{1-3n}$ sequences: None of the events with a marked Y′-E involved the tandem duplication of the marked end (Table 2B), which is statistically significantly different from the 25% tandem duplications found with other Y′s (Table 2A) (Louis and Haber 1990b). This difference can be explained by the fact that there are no $(TG_{1-3})_n$ sequences between X and Y′-E, so that unequal exchanges between $(TG_{1-3})_n$ sequences on either side of a Y′ cannot occur.

DISCUSSION

Mosaicism of chromosome ends: The chromosome end in yeast is much more complex than previously thought. The sequences immediately adjacent to the telomere sequences or to Y′s are not the same at all ends. From our hybridization and DNA sequence analysis of six cloned subtelomeric regions plus the previously published results of four other chromosome ends, we can define at least 11 distinct elements. These include Y′s, bi4, a "core" X, four new elements, termed STR-A, STR-B, STR-C and STR-D and four new elements defined by hybridization (F1, F2, F3 and F5). Fragment F4 and the "core" X sequence may overlap and so it is not counted as a new repeated element.

What was previously defined as X is a composite of several repeats, the "core" X and STR-A, STR-B, STR-C and STR-D. Fragment F4 is only 500 bp long and has homology at most chromosomes. F4 may overlap part of the "core" X or contain another subelement of X but even it does not hybridize to all chromosomes in all strains. F4 will need to be sequenced to see if it contains part of the core X. The core X is maximally defined as the 559 bp element shared by *III* left and *III* right. As more telomere regions are sequenced the X region will be more precisely defined. There may be a part of X that is located at all ends, but which has not yet been identified. Alternatively the X region may be a composite of several repeated elements such that every end has at least one of the set of repeats.

The mosaic pattern of repeated elements near yeast telomeres continues centromere proximal to the X element(s). Subclones of more centromere-proximal regions from only three ends contain at least four different repeated elements (F1–F3 and F5) that can vary in copy number and location between strains. These latter elements have so far been defined only by hybridization pattern, but will be given STR designations once they have been analyzed by DNA sequencing.

One of the newly defined elements, STR-A, contains at least one copy of TTAGGG and several degenerate versions of the sequence. Human telomere regions also contain tandem copies of degenerate forms of the TTAGGG repeat (ALLSHIRE *et al.* 1989). This similarity to the vertebrate telomere sequence may be accidental or it may be a reflection of the evolutionary relatedness of telomere sequences. Perhaps these are remnants of the ancestral Saccharomyces telomere. This is possible, as the vertebrate telomere sequence TTAGGG is also used in many distant taxa including mammals, slime molds, filamentous fungi, and kinetoplastid protozoa (BLACKBURN 1991). It is also possible that TTAGGG sequences may still play an important role in yeast. Recently a protein in *S. cerevisiae* that shows high affinity for repeats of TTAGGG as well as the STR-A region of some telomere clones has been described (BRIGATI *et al.* 1993; LIU and TYE 1991). A 10-bp repeat containing the sequence TTAGGG is also found within 100 bp of the 3' end of Y' elements (LOUIS and HABER 1992).

None of the STR-A, STR-B, STR-C or STR-D elements is essential for chromosome integrity, as among the eight informative ends shown in Figure 2, there is at least one example for each of these elements in which it is missing. In *S. paradoxus*, the ends appear to have a "core" X element but only one chromosome has any homology to the STR-A, STR-B, STR-C or STR-D elements at low stringency.

Origin and evolution of STR elements: Like Y's, some of the repeated elements described here are found in both *S. cerevisiae* and *S. paradoxus*. Their presence in the closest relative to *S. cerevisiae* indicates that they originated prior to the *S. cerevisiae-S. paradoxus* split. Within-species variation (post divergence) is evidence of their movement among chromosome ends.

A few of these elements are specific to *S. cerevisiae*. The previously described bi4 element is found in only a small number of locations, and not in all *S. cerevisiae* strains, and is likely to be a very recent acquisition by the chromosome end. Its presence at more than one chromosome end can be explained by subsequent homologous recombination events involving repeated sequences more centromere-proximal to the bi4 insertion. Its absence in other strains can also be explained by recombination and subsequent chromosome segregation.

In contrast the element revealed by fragment F1 is found in three constant locations in all strains of *S. cerevisiae* we examined. The constancy among strains indicates that this triplication was early in the evolution of *S. cerevisiae* as a species and that it has not been changing in location as other repeats have been over that time (*i.e.*, Y's are at least as old as F1 but are highly variable in location). Restriction mapping of this region on the three chromosomes using inserted markers as specific probes indicates a large area of triplication.

Fragments F2 and F3 are also specific to *S. cerevisiae* but are more variable in location than F1. In contrast to bi4 they are found in all strains. F2 and F3 may represent repeated gene families, similar to the *SUC, MAL* and *MEL* genes that have been previously described (CARLSON *et al.* 1985; CHARRON and MICHELS 1988; CHARRON *et al.* 1989; MICHELS *et al.* 1992; NAUMOV *et al.* 1990, 1991). It is also possible that these regions contain a transposable sequence such as *Ty5*, one copy of which was found to be just proximal to the X element of *III* L in the strain which was used for sequencing chromosome *III* (OLIVER *et al.* 1992; VOYTAS and BOEKE 1992).

Some subtelomeric regions were found to be present not only in *S. cerevisiae* but in its evolutionarily closest relative, *S. paradoxus*, but none were found to be shared by a more distant member of the genus, *S. bayanus*. Fragment F5 is found in all strains of *S. cerevisiae* and *S. paradoxus*. In *S. paradoxus* it appears constant in location but in *S. cerevisiae* there are some variable locations. Y's are also common to these two species.

Unlike Y's and fragment F5 which are found at many chromosomes in all strains of *S. paradoxus*, the STR-A, STR-B, STR-C and STR-D elements found distal to F5 and proximal to Y's are only found in one location in strains of *S. paradoxus*.

Internal repeated elements influence the apparent recombination among Y's: In YP1, the fragments F2 and F3 are found on *IX* left and *X* left. These two ends are also the ends that have the only two copies of the bi4 insertion and the only two copies of Y'-E. Therefore they share at least 4 kb of sequence internal to the Y', most of which is not found elsewhere in the genome of YP1.

Riles *et al.* (1993) noted a large region of shared homology between *IX* left and *X* left in S288C from which YP1 was derived. The Y's at these two ends tend to interact with each other more often than with Y's at other locations. This is consistent with the previously observed preference for recombination among Y's of the same size class. However, when one of the two Y'-extra longs is replaced by a Y'-short, the new Y' interacts with the other extra long end in 10 out of 20 recombination events. Furthermore, in 10/10 cases involving transfer of the *SUP11* marker to Y'-E, it was replaced by the Y'-S sequences of the donor. This places the end point of the recombination event somewhere centromere proximal to the long-short difference between Y's (Figure 7B). This result is unusual, because we have previously shown that most cross size class interactions have recombinational end points within the Y' element (Figure 7A). It seems most likely that the exceptional, preferential interactions between Y'-S on *IX* and the Y'-E end of *X* occur because the shared sequences internal to the Y' are directing these recombination events, rather than the size of the Y' itself. Thus, these events are not Y' × Y' events at all, but recombination involving the more internal shared subtelomeric repeats.

A question concerning this correlation remains. Have the two ends become homogenous due to preferential interactions or is the preferential interaction due to the shared sequences? This question can be answered by continuing the replacement process until the subtelomeric repeats at these two ends are no longer shared. Experiments are underway to address this issue.

The creation of tandemly arrayed Y's depends on internal $(TG_{1-3})_n$ sequences: To study recombination among Y' elements, we have used the duplication of *SUP11* originally inserted into a given Y' as a signal of a recombinational event which produces a cell with two copies of the marker. We have previously shown that, for some Y's, about 25% of the duplication events are actually the creation of tandemly arrayed Y's. When we examine the structure of one of those ends (YP1-L1) we find that the donor Y' element has $(TG_{1-3})_n$ sequences flanking the element, so that expansion into a tandem array can occur simply by unequal sister-chromatid exchange. In contrast, the two Y'-E containing ends do not have any telomere sequence at the X-Y' junction. They also do not yield tandem duplications of marked Y's. In support of this, the three X-Y' clones from strain Y55 do not have substantial amounts of telomere sequence and none of the Y's in Y55 are tandemly arrayed (Louis and Haber 1990a). This lack of expansion into a tandem array in the absence of the internal $(TG_{1-3})_n$ sequences is further support of the recombinational nature of Y' duplications and is not consistent with a transposition model of Y' duplication (Louis and Haber 1992).

Is mosaicism functional or the result of DNA metabolism? The known recombinational dynamics among Y's

can be invoked to account for the spread and variability of subtelomeric repeated elements in general (Louis and Haber 1990b, 1992). However, in addition to experiencing frequent homologous recombination events, subtelomeric regions also seem to be the targets of other, non-homologous recombination events. The finding of a mitochondrial intron in telomere adjacent sequences may hint at the ability of telomere regions to obtain or trap sequences that subsequently become subtelomeric repeats (Louis and Haber 1991). Among the *MAL* genes, there is no evidence for homologous sequences centromere-proximal to the genes which may be indicative of a non-homologous rearrangement process leading to their duplication (Michels *et al.* 1992). A non-homologous recombination process may account for the evolutionarily recent appearance in *S. cerevisiae* and *S. paradoxus* of the Y' element, which has features of a transposable element (Louis and Haber 1992) though the current movement and homogenization of Y's is via homologous recombination. There may be different processes involved in the origin *vs.* the maintenance and spread of subtelomeric repeated elements.

The accumulation of such elements in telomeric locations may not have any detrimental effect on the organism unless telomere function is disrupted. Acquisitions of sequences in telomere locations could occur via random integration, transposition or some other recombinational mechanism. Homologous recombination involving preexisting repeated sequences, such as the $(TG_{1-3})_n$ repeats found at all ends, can spread these newly acquired sequences to other telomere locations. These can then act as repeated sequences involved in spreading future acquisitions to new locations. Continuous integration of new elements followed by recombinational spread may result in the mosaicism seen for subtelomeric regions. The invariant order of the "core" X, STR-D, STR-C, STR-B and STR-A elements may be a result of this type of accumulation and spread.

In addition, having a mosaic of repeats may have advantages for the organism. In yeast, cell senescence caused by the ever-shortening telomere mutation, *est1*, can be rescued by a dramatic proliferation of Y' elements to many chromosome ends (Lundblad and Blackburn 1993). This process most likely occurs by homologous recombination between repeated subtelomeric sequences at a no longer protected telomere-less chromosome end with an end containing a Y' element. Another possible role for subtelomeric repeats is the protection of unique sequences from heterochromatization and transcriptional silencing (Gottschling *et al.* 1990). Any advantage, even if transient, would allow the evolutionary maintenance of mosaic structures to chromosome ends.

The situation in Drosophila may be similar. A complex family of repeated DNA, HeT, is found specifically in heterochromatin (Valgeirsdóttir *et al.* 1990) with a

subset of HeT-A repeats found at telomeres (DANI-LEVSKAYA *et al.* 1993). These elements are specifically involved in the healing of broken chromosome ends via a transposition mechanism (BIESSMANN *et al.* 1990, 1992a, b), but are also present in normal unbroken telomere regions interspersed with other sequences (DANI-LEVSKAYA *et al.* 1993). Within these elements there are regions thought to be important structural components of heterochromatin (DANILEVSKAYA *et al.* 1993). Though the spread and maintenance of HeT-A elements appears to be mostly due to transposition in contrast to the recombination of Y's in yeast, they both can function to maintain telomeres when necessary for survival.

The evolutionary dynamics of chromosome ends in yeast are complex. There are several unresolved issues and paradoxes. Y's interact by recombination leading to their homogenization as expected (LOUIS and HABER 1992). But the immediately adjacent sequences do not exhibit the same level of homogenization even though they are evolutionarily related. There appears to be an abrupt transition from 98–99% identity to only 60–95% at the junction of Y's to the adjacent sequences. How is recombinational homogenization restricted to the Y' sequences? Experiments will be undertaken to measure the level of ectopic recombination among the Y' adjacent sequences to see if the lack of homogeneity corresponds with reduced recombination. If at some ends recombinational homogenization is low enough to allow some accumulation of differences, then a point may be reached where recombination between the duplications is reduced allowing further divergence (WALSH 1987).

Previously found subtelomeric repeats of *SUC, MAL* and possibly *MEL* genes in *S. cerevisiae* strains could be attributed to selective pressure induced by human domestication. Fermentation processes can lead to selection for more efficient carbon source utilization. *S. paradoxus* strains are not used in fermentation processes and are found in nature mainly in exudates from various plants (NAUMOV 1987; NAUMOV *et al.* 1992). Their electrophoretic karyotype is indistinguishable from *S. cerevisiae* (NAUMOV *et al.* 1992). Here we have shown that they also have subtelomeric duplications similar to those of *S. cerevisiae.* If human activity has influenced some of the structural evolution of *S. cerevisiae* it acted on preexisting processes that lead to and maintain subtelomeric duplications.

The structure of chromosome ends themselves influence their recombinational dynamics which in turn affects their structure. The complex nature of chromosome ends is the result of this interaction between recombinational dynamics and structure.

LITERATURE CITED

ALLSHIRE, R. C., M. DEMPSTER and N. D. HASTIE, 1989 Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Res. **17**: 4611–4627.

BEDBROOK, J. R., J. JONES, M. O'DELL, R. D. THOMPSON and R. B. FLAVELL, 1980 A molecular description of telomeric heterochromatin in *Secale* species. Cell **19**: 545–560.

BIESSMANN, H., J. M. MASON, K. FERRY, M. D'HULST, K. VALGEIRSDÓTTIR, K. L. TRAVERSE and M. PARDUE, 1990 Addition of telomere-associated HeT DNA sequences "heals" broken chromosome ends in Drosophila. Cell **61**: 663–673.

BIESSMANN, H., L. E. CHAMPION, M. O'HAIR, K. IKENAGA, B. KASRAVI and J. M. MASON, 1992a Frequent transpositions of *Drosophila melanogaster* HeT-A transposable elements to receding chromosome ends. EMBO J. **11**: 4459–4469.

BIESSMANN, H., K. VALGEIRSDÓTTIR, A. LOFSKY, C. CHIN, B. GINTHER, R. W. LEVIS and M. PARDUE, 1992b HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in *Drosophila melanogaster*. Mol. Cell. Biol. **12**: 3910–3918.

BLACKBURN, E. H., 1991 Telomeres. Trends Biochem. Sci. **16**: 378–381.

BRIGATI, C., S. KURTZ, D. BALDERES, G. VIDALI and D. SHORE, 1993 An essential yeast gene encoding a TTAGGG repeat-binding protein. Mol. Cell. Biol. **13**: 1306–1314.

BROUN, P., M. W. GANAL and S. D. TANKSLEY, 1992 Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. Proc. Natl. Acad. Sci. USA **89**: 1354–1357.

BROWN, W. R. A., P. J. MACKINNON, A. VILLASANTÉ, N. SPURR, V. J. BUCKLE and M. J. DOBSON, 1990 Structure and polymorphism of human telomere-associated DNA. Cell **63**: 119–132.

BUTTON, L. L., and C. R. ASTELL, 1986 The *Saccharomyces cerevisiae* chromosome *III* left telomere has a type X, but not a type Y', ARS region. Mol. Cell. Biol. **6**: 1352–1356.

CARLSON, M., J. L. CELENZA and F. J. ENG, 1985 Evolution of the dispersed *SUC* gene family of Saccharomyces by rearrangements of chromosome telomeres. Mol. Cell. Biol. **5**: 2894–2902.

CHAN, C. S. M., and B. TYE, 1983a A family of *Saccharomyces cerevisiae* repetitive autonomously replicating sequences that have very similar genomic environments. J. Mol. Biol. **168**: 505–523.

CHAN, C. S. M., and B. TYE, 1983b Organization of DNA sequences and replication origins at yeast telomeres. Cell **33**: 563–573.

CHARRON, M. J., and C. A. MICHELS, 1988 The naturally occurring alleles of *MAL1* in Saccharomyces species evolved by various mutagenic processes including chromosomal rearrangement. Genetics. **120**: 83–93.

CHARRON, M. J., E. READ, S. R. HAUT and C. A. MICHELS, 1989 Molecular evolution of telomere-associated *MAL* loci of Saccharomyces. Genetics **122**: 307–316.

CHENG, J.-F., C. L. SMITH and C. R. CANTOR, 1991 Structural and transcriptional analysis of a human subtelomeric repeat. Nucleic Acids Res. **19**: 149–154.

COOKE, H. J., W. R. A. BROWN and G. A. RAPPOLD, 1985 Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. Nature **317**: 687–697.

CORCORAN, L. M., J. K. THOMPSON, D. WALLIKER and D. J. KEMP, 1988 Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in *P. falciparum*. Cell **53**: 807–813.

CROSS, S., J. LINDSEY, J. FANTES, S. MCKAY, N. MCGILL and H. COOKE, 1990 The structure of a subterminal repeated sequence present on many human chromosomes. Nucleic Acids Res. **18**: 6649–6657.

DANILEVSKAYA, O., A. LOFSKY, E. V. KURENOVA and M. PARDUE, 1993 The Y chromosome of *Drosophila melanogaster* contains a distinctive subclass of HeT-A-related repeats. Genetics **134**: 531–543.

DE LANGE, T., L. SHIUE, R. M. MYERS, D. R. COX, S. L. NAYLOR, A. M. KILLERY and H. E. VARMUS, 1990 Structure and variability of human chromosome ends. Mol. Cell. Biol. **10**: 518–527.

DE STEENSMA, H. Y., P. DE JONGE, A. KAPTEIN and D. B. KABACK, 1989 Molecular cloning of chromosome *I* DNA from *Saccharomyces cerevisiae*: localization of a repeated sequence containing an acid phosphatase gene near a telomere of chromosome *I* and chromosome *VIII*. Curr. Genet. **16**: 131–137.

DEVEREUX, J., P. HAEBERLI and O. SMITHIES, 1984   A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:** 387–395.

DORE, E., T. PACE, M. PONZI, L. PICCI and C. FRONTALI, 1990   Organization of subtelomeric repeats in *Plasmodium berghei*. Mol. Cell. Biol. **10:** 2423–2427.

FOOTE, S. J., and D. J. KEMP, 1989   Chromosomes of malaria parasites. Trends Genet. **5:** 337–342.

GOTTSCHLING, D. E., O. M. APARICIO, B. L. BILLINGTON and V. A. ZAKIAN, 1990   Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. Cell **63:** 751–762.

HOROWITZ, H., P. THORBURN and J. E. HABER, 1984   Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4:** 2509–2517.

INGLEHEARN, C. F., and H. J. COOKE, 1990   A VNTR immediately adjacent to the human pseudoautosomal telomere. Nucleic Acids Res. **18:** 471–476.

JÄGER, D., and P. PHILIPPSEN, 1989   Many yeast chromosomes lack the telomere-specific Y' sequence. Mol. Cell. Biol. **9:** 5754–5757.

KANE, S., and R. ROTH, 1974   Carbohydrate metabolism during ascospore development in yeast. J. Bacteriol. **118:** 8–14.

KARPEN, G. H., and A. C. SPRADLING, 1992   Analysis of subtelomeric heterochromatin in the Drosophila minichromosome Dp1187 by single P element insertional mutagenesis. Genetics **132:** 737–753.

KRAMER, K. M., and J. E. HABER, 1993   New telomeres in yeast are initiated with a highly selected subset of TG_{1-3} repeats. Genes Dev. **7:** 2345–2356.

LINK, A. J., and M. V. OLSON, 1991   Physical map of *Saccharomyces cerevisiae* genome at 110-kilobase resolution. Genetics **127:** 681–698.

LIU, Z., and B. TYE, 1991   A yeast protein that binds to vertebrate telomeres and conserved yeast telomeric junctions. Genes Dev. **5:** 49–59.

LOUIS, E. J., 1994   Corrected sequence of the right telomere of yeast chromosome *III*. Yeast **10:** (in press).

LOUIS, E. J., and J. E. HABER, 1990a   The subtelomeric Y' repeat family in *Saccharomyces cerevisiae*: An experimental system for repeated sequence evolution. Genetics **124:** 533–545.

LOUIS, E. J., and J. E. HABER, 1990b   Mitotic recombination among subtelomeric Y' repeats in *Saccharomyces cerevisiae*. Genetics **124:** 547–559.

LOUIS, E. J., and J. E. HABER, 1991   Evolutionarily recent transfer of a group I mitochondrial intron to telomere regions in *Saccharomyces cerevisiae*. Curr. Genetics **20:** 411–415.

LOUIS, E. J. and J. E. HABER, 1992   The Structure and Evolution of Subtelomeric Y' Repeats in *Saccharomyces cerevisiae*. Genetics **131:** 559–574.

LUNDBLAD, V., and E. H. BLACKBURN, 1993   An alternative pathway for yeast telomere maintenance rescues *est1-* senescence. Cell **73:** 347–360.

McCUSKER, J. H., and J. E. HABER, 1988   Cycloheximide-resistant temperature sensitive lethal mutations of *Saccharomyces cerevisiae*. Genetics **119:** 303–315.

MICHELS, C. A., E. READ, K. NAT and M. J. CHARRON, 1992   The telomere-associated *MAL3* locus of Saccharomyces is a tandem array of repeated genes. Yeast **8:** 655–665.

NAUMOV, G. I., 1987   Genetic basis for classification and identification of the ascomycetous yeasts. Stud. Mycol. **30:** 469–475.

NAUMOV, G. I., H. TURAKAINEN, E. S. NAUMOVA, S. AHO and M. KORHOLA, 1990   A new family of polymeric genes in *Saccharomyces cerevisiae*: α-galactosidase genes *MEL1-MEL7*. Mol. Gen. Genet. **224:** 119–128.

NAUMOV, G. I., E. S. NAUMOVA, H. TURAKAINEN and P. K. SUOMINEN M., 1991   Polymeric genes *MEL8*: *MEL9* and *MEL10*-new members of the α-glucosidase gene family in *Saccharomyces cerevisiae*. Curr. Genetics **20:** 269–276.

NAUMOV, G. I., E. S. NAUMOVA, R. A. LANTTO, E. J. LOUIS and M. KORHOLA, 1992   Genetic homology between *Saccharomyces cerevisiae* with its sibling species *S. paradoxus* and *S. bayanus*: electrophoretic karyotypes. Yeast **8:** 599–612.

OLIVER, S. G., J. M. VAN DER AART, M. L. AGOSTONI-CARBONE, M. AIGLE, L. ALBERGHINA *et al.*, 1992   The complete DNA sequence of yeast chromosome *III*. Nature **357:** 38–46.

PACE, T., M. PONZI, E. DORE, C. JANSE, B. MONS and C. FRONTALI, 1990   Long insertions within telomeres contribute to chromosome size polymorphism in *Plasmodium berghei*. Mol. Cell. Biol. **10:** 67–59–6764.

PAYS, E., and M. STEINERT, 1988   Control of antigen gene expression in african trypanosomes. Annu. Rev. Genet. **22:** 107–126.

RIETHMAN, H. C., R. K. MOYZIS, J. MEYNE, D. T. BURKE and M. V. OLSEN, 1989   Cloning human telomeric DNA fragments into *Saccharomyces cerevisiae* using a yeast-artificial-chromosome vector. Proc. Natl. Acad. Sci. USA **86:** 6240–6244.

RILES, L., J. E. DUTCHIK, A. BAKTHA, B. K. McCAULEY, E. C. THAYER *et al.*, 1993   Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2.6 kilobase pairs. Genetics **134:** 81–150.

ROTHSTEIN, R., 1983   One-step gene disruption in yeast. Methods Enzymol. **101:** 202–211.

RUBIN, G. M., 1978   Isolation of a telomeric DNA sequence from *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. **42:** 1041–1046.

SAIGA, H., and J. E. EDSTRÖM, 1985   Long tandem arrays of complex repeat units in Chironomus telomeres. EMBO J. **4:** 799–804.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI *et al.*, 1988   Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:** 487–491.

SHAMPAY, J., J. W. SZOSTAK and E. H. BLACKBURN, 1984   DNA sequences of telomeres maintained in yeast. Nature **310:** 154–7.

SZOSTAK, J. W., and E. H. BLACKBURN, 1982   Cloning yeast telomeres on linear plasmid vectors. Cell **29:** 245–255.

VALGEIRSDÓTTIR, K., K. L. TRAVERSE and M.-L. PARDUE, 1990   HeT DNA: A family of mosaic repeated sequences specific for heterochromatin in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **87:** 7998–8002.

VOYTAS, D. F., and J. D. BOEKE, 1992   Yeast retrotransposon revealed. Nature **358:** 717.

WALMSLEY, R. W., C. S. M. CHAN, B. TYE and T. D. PETES, 1984   Unusual DNA sequences associated with the ends of yeast chromosomes. Nature **310:** 157–160.

WALSH, J. B., 1987   Sequence-dependent gene conversion: can duplicated genes diverge fast enough to escape conversion? Genetics **117:** 543–557.

ZAKIAN, V. A., 1989   Structure and function of telomeres. Annu. Rev. Genetics **23:** 579–604.

ZAKIAN, V. A., and H. M. BLANTON, 1988   Distribution of telomere-associated sequences on natural chromosomes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:** 2257–2260.

The sequence data presented in this article have been submitted to the EMBL/GenBank Data Libraries under the accession numbers: M63936, M63937 and M63938 for the X-Y' junctions of pY55-L1, pY55-L2 and pY55-S1, respectively.

Communicating editor: S. JINKS-ROBERTSON