Volume 12 Number 21 1984 Nucleic Acids Research

Tandemly arranged variant 5S ribosomal RNA genes in the yeast *Saccharomyces cerevisiae*

Michael E.McMahon*, Dorothy Stamenkovich and Thomas D.Petes

Department of Molecular Genetics and Cell Biology, University of Chicago, 920 E.58th Street, Chicago, IL 60637, USA

Received 17 August 1984; Accepted 5 October 1984

ABSTRACT

Most of the ribosomal RNA genes of the yeast Saccharomyces cerevisiae are about 9 kilobases (kb) in size and encode both the 35S rRNA (processed to produce the 25S, 18S, and 5.8S species) and 5S rRNA. These genes are arranged in a single tandem array of 100 repeats. Below, we present evidence that at the centromere-distal end of this array is a tandem arrangement of a different type of rRNA gene. Each of these repeats is 3.6 kb in length and encodes a single 5S rRNA. The coding sequence of this gene is different from that of the "normal" 5S gene in three positions located at the 3' end of the gene.

INTRODUCTION

In most eukaryotes, the genes encoding the large ribosomal RNA's are organized in clusters that are separated from the genes encoding the 5S rRNA's (summarized in ref. 1). In the yeast Saccharomyces cerevisiae and the slime mold Dictyostelium, however, the coding sequences for the 5S genes alternate with those encoding the larger rRNA's. The yeast rRNA gene repeat is about 9 kb in size and has a relatively small amount of non-transcribed spacer (about 2 kb). The repeat has been mapped with restriction enzymes and the position and orientation of the rRNA transcripts within the repeat has been determined (2-6). Most lab strains of yeast have rRNA genes that contain seven EcoRI fragments, labeled A through G in order of size (Figure 1a).

There are about 100 rRNA genes per haploid yeast genome (7). Most (and possibly all) of these genes are located in a single tandem array on yeast chromosome XII (8). In an effort to identify and characterize the junctions between the rRNA gene cluster and other chromosomal sequences in the yeast strain A364a, we analyzed a large collection of recombinant plasmids, looking for EcoRI restriction fragments composed partly of rRNA gene sequences and partly of non-ribosomal RNA gene sequences (9). One of the plasmids isolated in this study, pY1rG11, contained two EcoRI fragments characteristic of the 9 kb repeat (fragments C and G) as well as a novel EcoRI fragment J2 that represented a fusion of rDNA fragment B with other chromosomal sequences. The structure of this plasmid and data obtained from Southern analyses of genomic DNA

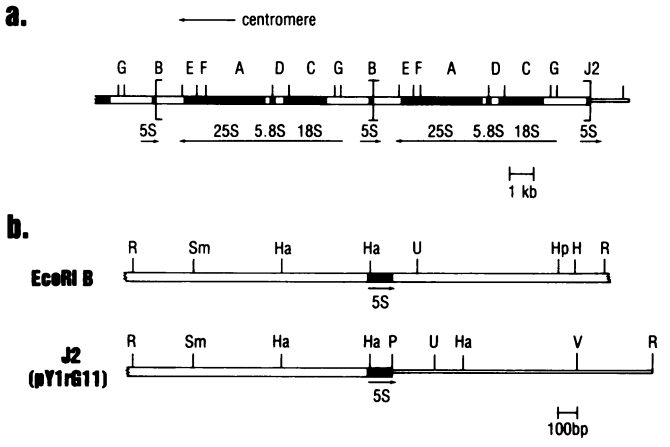

Figure 1. Restriction maps of the centromere-distal junction between the yeast ribosomal RNA genes and other chromosomal sequences. (a) EcoRI restriction map of two rRNA genes (EcoRI fragments B, E, F, A, D, C and G) and the junction EcoRI fragment (J2). The position of the EcoRI sites are indicated by vertical lines; brackets delimit the 9 kb repeat unit. The positions of the transcripts relative to the EcoRI sites are based on the results of Philippsen et al. (6) and the position of J2 relative to the EcoRI fragments is described in reference 9. (b) Detailed restriction maps of the EcoRI B fragment of the rRNA gene and the EcoRI J2 junction fragment (based on data in ref. 9). The code for the restriction sites is: H (HindIII), Ha (HaeIII), P (PstI), R (EcoRI), Hp (HpaII), Sm (SmaI), U (PvuII), and V (PvuI).

allowed construction of the EcoRI restriction map of the centromere-distal junction shown in Figure 1a (9).

Since the J2 fragment from pY1rG11 appeared to contain the junction between rDNA and other chromosomal sequences, a detailed restriction map of this fragment was made (9). As indicated in Figure 1b, a PstI restriction site divides the J2 fragment into two equal-sized segments. Restriction sites to the left of the PstI site are conserved between J2 and EcoRI B. Sites to the right of the PstI site are not conserved between the two fragments. In addition, sub-fragments prepared from the J2 fragment to the right of the PstI site fail to hybridize to any yeast rDNA fragment (9). Since J2 hybridized to 5S rRNA, we suggested that the junction between the rDNA and other chromosomal sequences occurred near the 3' end of the 5S gene.

In addition to the plasmid pY1rG11, we previously described a recombinant bacteriophage λZ1 that contained junction-related sequences (9). The λZ1 phage contained four EcoRI restriction fragments of yeast DNA. One of these fragments was approximately the same size as J2 from pY1rG11 and hybridized strongly to J2 from pY1rG11. We suggested that this fragment was J2 cloned in a different chromosomal context and that the other EcoRI restriction fragments (K2, L2, and M2) were located

centramere-distal to J2. We show below that although J2 from pY11rG11 contains regions of homology with J2 from λZ1, these EcoRI fragments are not identical. Henceforth, we will refer to the J2 fragment of λZ1 as "J2*". In addition, our realization of the difference between J2 and J2* has led to a different map of the junction between the rRNA genes and other chromosomal sequence from that described previously (9). Below, we show that J2 represents the beginning of a cluster of four 3.6 kb repeats, each of which contains a single 5S rRNA gene. The 5S rRNA gene in the 3.6 kb repeats has three base pair changes relative to the "normal" 5S genes. Variant yeast 5S rRNA genes with the same three substitutions have also been observed in experiments done independently by Piper et al. (10).

MATERIALS AND METHODS

Yeast and Bacterial Strains

Two haploid yeast strains were used in these studies, A364a (a ade1 ura1 lys2 tyr1 his7 ade2 gal1) and 2262 (α leu2 his5 lys11 ade1 ura1 gal1). The diploid strain +D4 was constructed by mating A364a to 2262. The E. coli host for the recombinant plasmids was HB101 (11). The E. coli host for the recombinant bacteriophage library was LE392 (12).

Recombinant Plasmid and Bacteriophage Banks

Three different clone banks were used in these studies. One recombinant plasmid bank was constructed by inserting shear fragments of DNA from the yeast strain +D4 into the EcoRI site of pMB9 (13). The second plasmid bank was made by inserting HindIII-SalI fragments of DNA from the yeast strain A364a into a pBR322 vector treated with HindIII and SalI. The ligated DNA molecules were then transformed into the E coli strain HB101. The recombinant bacteriophage bank (obtained from J. Woolford, Carnegie Mellon University) was constructed by inserting partial EcoRI fragments of A364a DNA into a λCharon 4A vector.

Isolation and Sub-cloning of Yeast DNA Fragments Derived from the Junction of the rRNA Gene Tandem Array.

The isolation of recombinant plasmids λZ1 and pY1rG11 has been described previously (9). The bacteriophage λZ1 contains an EcoRI restriction fragment, J2*, of 2.65 kb that has a single PstI restriction site. We sub-cloned the two PstI-EcoRI fragments derived from J2* into the PstI-EcoRI site of pBR322. These plasmids were called "pMM1" and "pMM2." As described below, the pMM1 plasmid contains sequences that are unique to the junction of the rRNA genes with other chromosomal sequences.

The recombinant bacteriophage λDS16-14 was identified as a plaque that hybridized (14) to ^{32}P-labeled pMM1 DNA. The recombinant plasmids pTP1-4 and pTP 12-32 were identified in the HindIII-SalI plasmid bank by detecting bacterial colonies that hybridized (15) to ^{32}P-labeled pMM1 DNA.

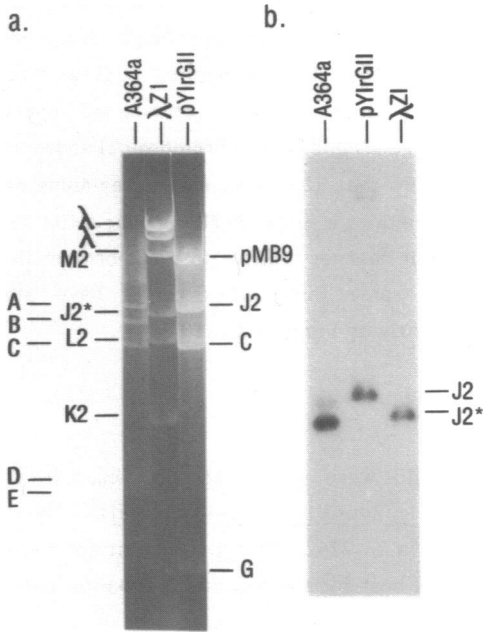

Figure 2. Comparison of the electrophoretic mobilities of EcoRI fragments J2 (derived from the plasmid pY1rG11) and J2* (derived from the bacteriophage λZ1). (a) Gel electrophoresis of EcoRI digests of genomic DNA (A364a), λZ1 and pY1rG11. The bands appearing in the lane containing genomic DNA represent ribosomal DNA fragments A through E (in descending order of size). The J2 fragment is the second largest band in the lane containing pY1rG11 DNA and J2* is the fourth largest band in the lane containing λZ1 DNA. (b) Southern analysis of EcoRI-treated genomic DNA, λZ1 and pY1rG11. After electrophoresis, the DNA fragments were transferred to nitrocellulose (17) and hybridized to ^{32}P-labeled pMM1 DNA; the pMM1 plasmid contains sequences homologous to both J2 and J2*.

Isolation and Southern Analysis of DNA

Plasmid and bacteriophage DNA were isolated by standard techniques (16). Yeast DNA was purified using CsCl gradients containing the fluorescent dye Hoechst 33258 (5).

For Southern analysis, the yeast DNA was treated with restriction enzymes (purchased from New England Biolabs and used with the buffers recommended by the manufacturers) and the resulting fragments separated by agarose gel electrophoresis. The concentration of agarose was varied in different experiments between 0.6% and 1%. Southern transfer procedures were standard (17). The hybridization conditions have been described previously (9).

DNA Labeling and Sequencing

The plasmid DNA molecules that were used as hybridization probes were labeled

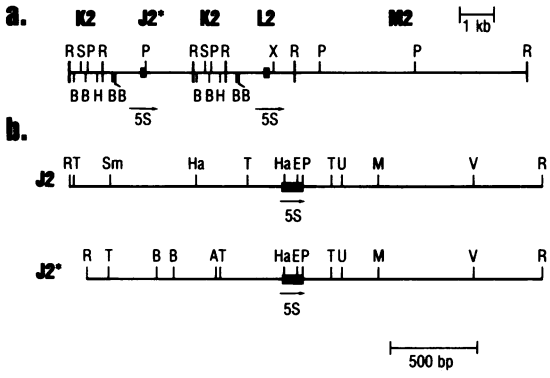

Figure 3. Restriction maps of junctional sequences. (a) Restriction map of the yeast DNA insert of λZ1. Letters K2-M2 above the map indicate EcoRI fragments. (b). Detailed restriction maps of EcoRI fragments J2 and J2*. The code for restriction sites is the same as that in Figure 1 with the following additions: A (AvaII), B (BglII), E (BstEII), M (MspI), S (SalI), T (TaqI) and X (XhoI). There is one additional PstI site in M2 that has not been mapped.

by nick translation (18). For sequencing, the DNA fragments were labeled either at the 5' ends by using T4 polynucleotide kinase or at the 3' ends by using DNA polymerase (19). Methods for strand separation, chemical modification of the DNA and gel electrophoresis of the modified molecules were those described by Maxam and Gilbert (19).

RESULTS

Evidence that J2 of pY1rG11 is not identical to J2* of λZ1.

Three arguments indicate that J2 and J* are not identical. First, there is a slight but reproducible difference in the electrophoretic mobility of these fragments in agarose gels (Figure 2). The estimated size of J2 is 2.75 kb versus 2.65 kb for J2*. Second, a more detailed restriction map of λZ1 (Figure 3) indicated that λZ1 had five EcoRI fragments of yeast DNA, J2*; two copies of K2, and one copy of L2 and M2. Since J2* is flanked by K2 fragments (Figure 3) and J2 in pY1rG11 is flanked on one side by the rDNA EcoRI G fragment, J2 and J2* cannot represent the same fragment. Third, from the maps shown in Figures 1 and 3, it is clear that J2 and J2* do not have identical restriction sites.

To ensure that the observed differences between J2 and J2* were not cloning artifacts, we did two more types of experiments. First, we sub-cloned the J2* sequences as PstI-EcoRI fragments in pBR322. The plasmid with the larger PstI-EcoRI fragment (1.4 kb) of J2* was called pMM1; the plasmid with the smaller fragment (1.25 kb) was called pMM2. The pMM1 plasmid hybridized strongly to J2 from pY1rG11 but

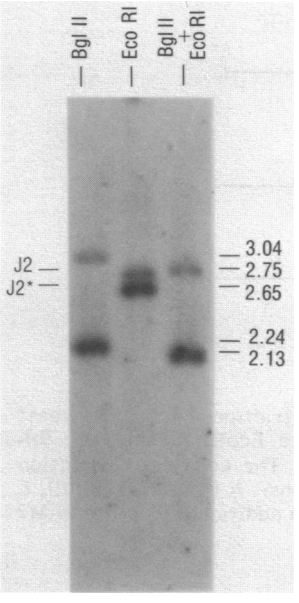

Figure 4. Southern analysis of yeast genomic sequences that hybridize to the junction- specific plasmid pMM1. DNA from the strain A364a was treated with BglII, EcoRI or both enzymes and the resulting fragments separated on a 0.8% agarose gel. The fragments were transferred to nitrocellulose (17) and hybridized to ^{32}P-labeled pMM1 DNA. Fragment sizes were determined by using EcoRI fragments of yeast ribosomal DNA as size standards (6).

pMM2 did not (data not shown). When pMM1 is used as a hybridization probe to EcoRI-digested genomic DNA from the yeast strain A364a, two bands of hybridization are detected, one at the position of J2 and one at the position of J2* (Figures 2 and 4). The fragment at the position of J2* is cut by BglII as expected from the restriction map of J2* (Figure 3). By densitometer tracings of the bands in Figure 4, we determined that hybridization at the J2* position was three times stronger than that observed for J2. This result suggests that either the homology between pMM1 and J2* is considerably better than that between pMM1 and J2 or that there are three times more copies of J2* than J2.

We also demonstrated that J2 and J2* were not likely to be cloning artifacts by re-isolating both types of sequences from a newly constructed recombinant plasmid collection made from A364a DNA. As described in detail in Materials and Methods, we treated A364a DNA with enzymes HindIII and SalI and inserted these fragments into HindIII-SalI treated pBR322. Of approximately 500 transformants examined, nine hybridized to the J2* sequences inserted in pMM1. Of these nine plasmids, seven had EcoRI fragments of the size of J2* and two (pTP1-4 and pTP12-32) had fragments the size of J2. DNA sequencing experiments which will be described below showed that the insert in pTP1-4 was identical to J2 in pY1rG11.

Arrangement of J2 and J2* in the genome.

We re-examined the recombinant λ bank containing A364a DNA, looking for additional phage that hybridized to pMM1. In addition to λZ1-like phage (containing

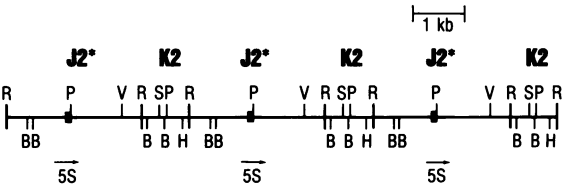

Figure 5. Restriction map of the yeast DNA insert in phage λDS 16-14. The code for restriction sites is the same as that described in Figures 1 and 3.

EcoRI fragments J2*, K 2, L2, and M2), we found a recombinant phage (λDS16-14) that had only J2* and K 2 EcoRI fragments. Since the Charen 4A vector requires an insert of greater than 7 kb to produce a viable phage and since J2* and K 2 represent only 3.6 kb of DNA, we expected that λDS16-14 contained more than one copy of J2* and K 2. Densitometer tracing of the bands representing J2*, K 2 and the λ "arms" showed that λDS16-14 had three copies of both J2* and K 2. The restriction map of λDS16-14 is shown in Figure 5.

A composite restriction map of the EcoRI fragments at the centromere-distal junction of the rRNA gene cluster of A364a is shown in Figure 6. The only portion of the composite that is not formed by overlapping restriction maps derived from recombinant clones is the K 2 fragment shown as the junction between J2 and J2*.

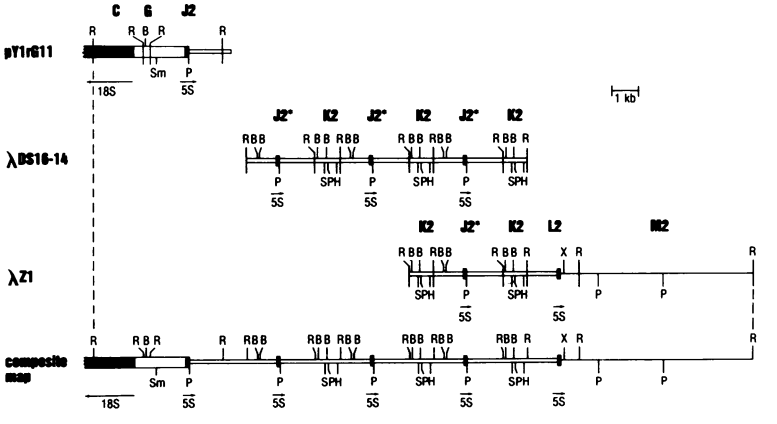

Figure 6. Composite restriction map of the centromere-distal junction between the rRNA genes and other chromosomal sequences. The thick lines represent ribosomal DNA sequences (black lines show rRNA coding sequences and "open" lines represent ribosomal DNA spacer sequences). Medium width lines represent the 3.6 kb repeats and thin lines show the unique chromosomal sequences. The code for restriction sites is given in Figures 1 and 3.

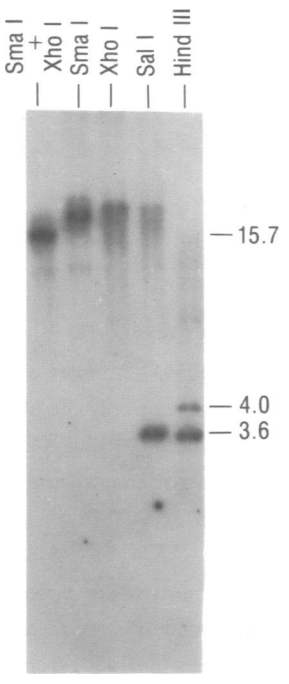

Figure 7. Southern analysis of A364a DNA treated with SmaI, XhoI, SmaI-XhoI, SalI or HindIII. Digested samples were analyzed by gel electrophoresis, transferred to nitrocellulose (17) and hybridized to pMM1 DNA. Sizes of the fragments were determined as described in Figure 4.

Several arguments indicate the presence of K 2 at this position. First, the sequences in pY1rG11 that are centromere-distal to J2 hybridize to K 2. Second, when A364a DNA is treated with either HindIII, SalI, EcoRI, BglII-EcoRI or XhoI-SmaI and hybridized to pMM1 (which hybridizes to both J2 and J2*), bands are found at the expected positions (Figures 4 and 7). In these figures, when there are two bands of hybridization, the dark band represents hybridization of pMM1 to J2* and the light band shows the position of the single-copy J2 sequences. The correlation between the observed size of the DNA fragments and the expected sizes based on the composite map of Figure 6 is shown in Table 1.

It is important to point out that the arrangement of sequences shown in Figure 6 was obtained by analysis of a single yeast strain, A364a. It does not, therefore, necessarily represent an arrangement that will be found in all S. cerevisiae strains. In particular, we previously showed that sequences from J2 centromere-distal to the Pst I site were not present in the yeast strain 2262 (9). In crosses between A364a and 2262, we found that the J2 sequences segregated 2:2 (two spores have the J2 sequences and two do not) and that the J2 sequences co-segregated with the rRNA gene cluster derived from A364a (9). Since the sequences from the centromere-distal portion of J2 also hybridize to J2*, the 2:2 segregation of these sequences indicates that J2 and J2* fragments are tightly linked in a genetic as well as a physical sense.

Table 1. Comparison between the predicted and observed sizes of A364a genomic DNA fragments that hybridize to the pMM1 plasmid. The predicted sizes are based on restriction maps of cloned fragments as summarized in Figure 6. The observed sizes are based on a Southern analysis of A364a DNA. Since the pMM1 plasmid hybridizes to both J2 and J2* sequences, for most digests, two bands of hybridization are expected, a weak band (corresponding to the single-copy J2 sequences) and a strong band (corresponding to the multi-copy J2* sequences).

Restriction enzyme(s)	Predicted size (kb)		Observed size (kb)	
	Weak band	Strong band	Weak band	Strong band
EcoRI	2.75	2.65	2.75	2.65
BglII	3.04	2.24	3.0	2.1
BglII-EcoRI	2.75	2.13	2.75	2.0
Sal I	>900	3.60	>20	3.6
Hind III	4.05	3.60	4.2	3.6
XhoI	>900	>900	>20	>20
SmaI	>23	>23	>20	>20
XhoI-SmaI	15.7	15.7	15.8	15.8

Thus, at the end of the cluster of 9 kb ribosomal RNA genes, there is a small cluster of 3.6 kb repeats. We discuss below DNA sequence analysis that indicates each of these repeats contains a variant 5S rRNA gene. We call the 3.6 kb repeats, the $5S^{VAR}$ repeats (VAR denoting variant).

Three regions of DNA were sequenced in these studies, the 5S rRNA coding sequences derived from J2, J2*, and L2. Data for each class of sequence will be discussed separately below.

Sequence Analysis of J2.

We found that a single 496 bp TaqI fragment derived from J2 hybridized to ^{32}P-labeled 5S rRNA (data not shown). This fragment was sequenced in both strands (Figure 8). The DNA sequences upstream of the 5S coding sequences in J2 are very similar to those upstream of the 5S gene in the 9 kb repeat (20). Only 7 differences were detected in 216 bp. These differences are not likely to be specific to J2 since we have found the same differences between a 9 kb repeat derived from the yeast strain 2262 and the strain used by Valenzuela et al. (unpublished data). We also sequenced the DNA from the region upstream of the 5S gene of the J2 fragment of pTP1-4. This sequence was identical to that shown in Figure 8.

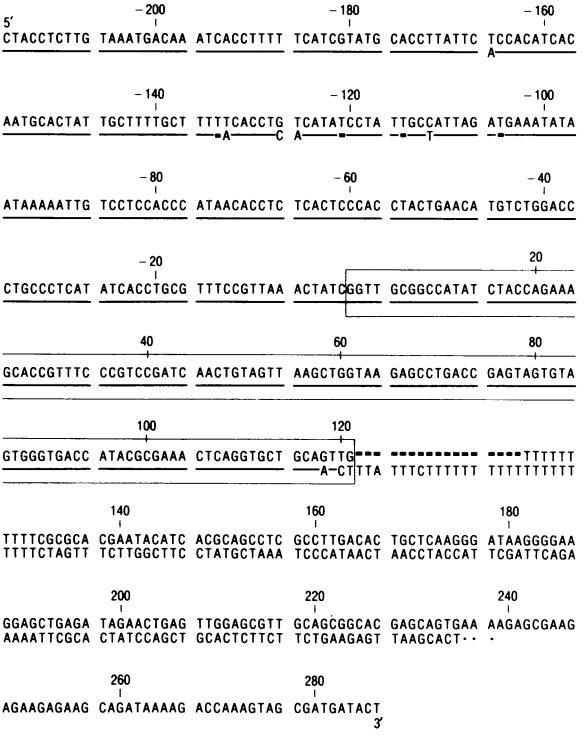

Figure 8. Nucleotide sequence of the noncoding strand of a 496 bp TaqI fragment derived from J2. This fragment includes the 5S rRNA coding sequences as well as flanking sequences. The top lines in the figure represent J2 sequences; the bottom lines represent the 5S rRNA coding sequences and flanking sequences from the normal 9 kb repeat (20). Identity between the sequences is indicated by a continuous thin line, divergence is indicated by depicting both sequences and a deletion is indicated by a short thick line. The 5S rRNA coding sequences are enclosed in a box. Both strands were sequenced.

The 5S coding sequences in J2 are the same as those in the 9 kb repeat for the first 117 bp. Of the last four bp, three are different. These differences produce a PstI restriction site that is a convenient indication of the $5S^{VAR}$ gene. Downstream of the 5S coding sequences, there is a tract of 10 T's rather than the usual tract of 22 T's (ref. 20). Beyond the poly T tract, there is no obvious homology between J2 and the 9 kb repeat. Thus, the breakpoint between the 9 kb repeats and the 3.6 kb $5S^{VAR}$ repeats is at the end of the 5S rRNA gene.

Sequence Analysis of 5S Gene in J2*.

In order to sequence the 5S gene in J2*, we analyzed the sequences of the plasmids pMM1 and pMM2. In Figure 9, we show the coding sequences of the 5S gene

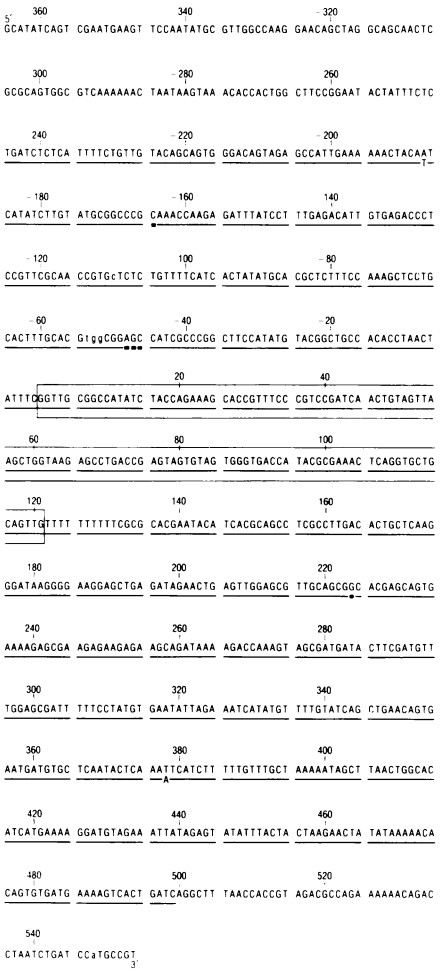

Figure 9. Nucleotide sequence of the portion of the noncoding strand of J2* that includes the 5S rRNA gene and flanking sequences. Most of the 5S gene and upstream sequences were identified by sequence analysis of an AvaII fragment of the pMM2 plasmid. The remaining portion of the 5S gene and the downstream sequences were obtained by sequence analysis of an MspI fragment of the plasmid pMM1. The J2* sequence (top lines) is compared to the corresponding region of DNA of λgtE83 (10). Sequence identities are shown by thin horizontal lines, deletions are represented by short horizontal lines and differences are indicated by showing both sequences. The 5S rRNA genes are boxed. Both complementary strands were sequenced; lower case letters represent remaining sequence ambiguities.

and about 400 bp upstream and downstream of the coding sequence. The 5S coding sequences as well as the sequences downstream from the 5S rRNA gene are completely conserved between J2 and J2* (Figure 8). The upstream sequences, however, show essentially no homology between J2 and J2*. In addition, the sequences flanking the 5S rRNA gene of J2* have no homology to those that flank the 5S gene in the 9 kb repeats.

Piper et al. (10) recently isolated recombinant bacteriophage containing insertions of yeast DNA that hybridized to 5S rRNA but did not have the restriction fragments typical of the 9 kb repeat. One of these phage (λgtE83) had a restriction fragment very similar to J2*. The region of λgtE83 containing the 5S gene was sequenced by Piper et al. If we compare sequences extending from 240 bp upstream of the 5S gene to 500 bp downstream of the gene, only 5 changes are detected between the sequences of

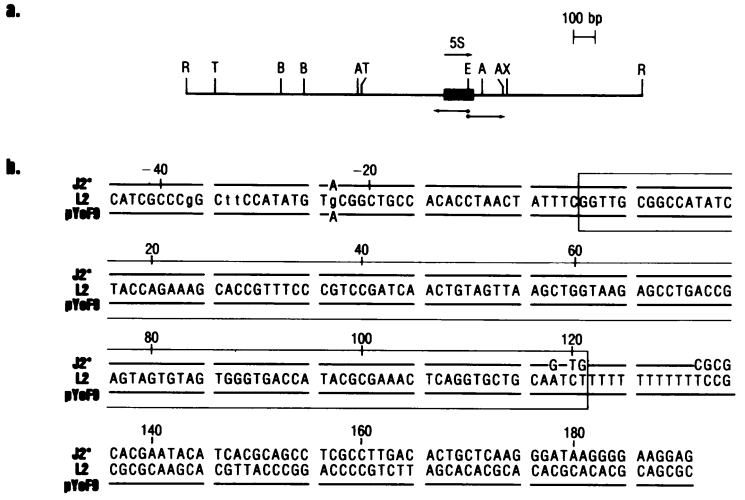

Figure 10. Restriction map and DNA sequence analysis of the L2 fragment. (a) Restriction map of L2. The code for the restriction sites is described in Figures 1 and 3. The arrows below the map indicate the region sequenced. (b) Nucleotide sequence of a portion of the noncoding strand of L2. Both complementary strands were sequenced by 5' and 3' labeling of DNA at the BstEII site. The L2 sequence is compared with the coresponding regions of J2* and λgtF9 (10). Sequence homologies, differences and ambiguities are indicated as in Figures 8 and 9.

J2* and those of λgtE83. The 5S coding sequences were identical in J2* and λgtE83. The small number of sequence differences located outside the coding sequences could reflect either strain-specific differences (since different yeast strains were used as a source of DNA for the recombinant clones) or heterogeneity between different J2* repeats within a strain.

Sequence Analysis of the 5S rRNA Gene of L2.

The L2 EcoRI fragment of λZ1 hybridized to 5S rRNA but did not contain a PstI site. It was clear, therefore, that this fragment did not contain a 5S gene identical to those in J2 and J2*. Sequence analysis of L2 (summarized in Figure 10) indicated that the 5S coding sequence was identical to that of the 9 kb repeat. The upstream sequences were very similar to those of J2* (39 of 40 bp were identical, one potential difference was an ambiguous base in L2). The sequences found downstream of the 5S gene in L2 had no obvious homology to the downstream sequences of J2, J2* or the 9 kb repeat.

Piper et al (10) did a sequence analysis of a recombinant bacteriophage (λgtF9) that hybridized to yeast 5S rRNA and had an EcoRI restriction fragment similar to L2. In comparing the sequences extending from 40 bp upstream of the coding sequence to 240 bp downstream of the coding sequence, at most a single difference exists between

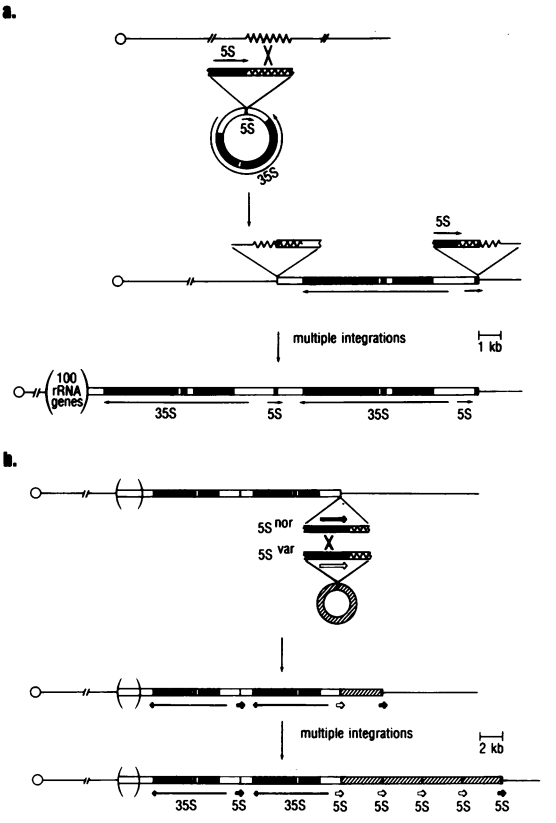

Figure 11. Model for the origin of the 9 kb rRNA gene tandem array and the 3.6 kb tandem array. (a) Origin of the 9 kb repeat array. A recombination event between a polyA tract on an extrachromosomal 9 kb circular repeat and a polyA tract in the chromosome generates a single integrated rRNA gene. Subsequent integration events generate the array. In the diagram, the portion of the DNA encoding the 5S gene and the polyA tract is shown in expanded form (10-fold) as well as to scale. (b) Possible origin of the 3.6 kb repeat tandem array. A 3.6 kb extrachromosomal circle recombines with the terminal 9 kb repeat by a homologous exchange between the 5S genes. Multiple integrations generate the 3.6 kb tandem array. In these diagrams, the thin straight line represents chromosomal sequences other than ribosomal DNA, the thick black lines represent rRNA coding sequences and the open lines represent ribosomal DNA spacer sequences. Poly A tracts are shown by zig-zag lines and the 3.6 kb repeats are cross-hatched.

the sequences of L2 and λgtF9 (Figure 10); this single difference involves the ambiguous base at position -24 of the L2 sequence.

DISCUSSION

In most eukaryotes, the 5S rRNA genes are encoded in a tandem array of repeats that is unlinked to the tandem array of repeats encoding the larger rRNA's (reviewed in

ref. 1). In contrast, in S. cerevisiae and D. discoideum, most of the 5S rRNA genes are encoded in the same repeat that encodes the large rRNA's. The results described above, however, indicate that yeast may actually represent an intermediate in the evolution of separate clusters for 5S rRNA's and large rRNA's.

In examining the junction between the 9 kb tandem array with the 3.6 kb tandem array and the junction between the 3.6 kb tandem array and other chromosomal sequences, we observed that these junctions occur very close to the end of the 5S rRNA coding sequences. Before discussing mechanisms for the evolution of the clusters that are consistent with these junctions, we will summarize the important features of the organization of these DNA sequences: 1) the junction between the 9 kb repeats and the 3.6 kb repeats (represented by the J2 fragment) occurs within the 5S gene; the 5S rRNA gene at this junction is a $5S^{VAR}$ gene, 2) the end of the 9 kb repeat tandem array represents the beginning of an array of four 3.6 kb repeats, 3) all of the 3.6 kb repeats contain a $5S^{VAR}$ gene except the last repeat in the array; the last repeat has a 5S rRNA gene identical to that found in the 9 kb repeats and 4) the breakpoint between the end of the 3.6 kb repeat tandem array and other chromosomal sequences occurs immediately after the poly A tract found at the end of the 5S rRNA gene.

This particular arrangement of sequences can be explained as a two step process; the first creating the 9 kb rRNA gene cluster, the second forming the 3.6 kb gene cluster (Figure 11). We suggest that the 9 kb gene cluster may have originated as the result of integrating an extrachromosomal 9 kb rDNA repeat; extrachromosomal ribosomal RNA genes are fairly common in lower eukaryotes (21). To explain the position of the junction, we suggest that the integration event involved homologous recombination between the poly A tract at the end of the 5S gene of the extrachromosomal 9 kb plasmid and a pre-existing chromosomal poly A tract. After insertion of the first 9 kb repeat, repeated integrations of other plasmids would expand the array without changing the original junctions. Unequal recombination between 9 kb tandem arrays could also lead to expansion.

We suggest that the creation of the 3.6 kb repeat cluster may be the result of a second similar type of integration (Figure 11b). In this case, the extrachromosomal plasmid is 3.6 kb in size and contains all sequences that are present in J2* and K2, including the $5S^{VAR}$ gene. The integration event is a homologous recombination between the $5S^{VAR}$ gene on the plasmid and the terminal 5S gene of the 9 kb tandem array. If this recombination event occurs centromere-proximal to the $5S^{VAR}$-specific changes at the 3' end of the gene, then a variant $5S^{VAR}$ gene would be the junction between the 9 kb and 3.6 kb repeats and a normal 5S gene would be the junction between the 3.6 kb repeats and other chromosomal sequences. Repeated subsequent integration events of 3.6 kb plasmids would expand the array without changing these

junctions. Thus, the mechanisms shown in Figure 11 are consistent with the sequence arrangements summarized in Figure 6. One obvious question concerning the events described in Figure 11 is the source of the 3.6 kb extrachromosomal repeats. One possibility is that such plasmids were introduced in the lab strains from non-S. cerevisiae species during crosses done in the early days of yeast genetics (22). In this connection, it should be stressed that although the 3.6 kb sequences are found in some lab strains, such as A364a and X2180-1B (10), these sequences are not present in all lab strains (9).

There are at least two other mechanisms that could generate the tandem array of 3.6 kb repeats. Childs et al. (23) observed that some eukaryotic genomes contain solitary dispersed repeats ("orphons") derived from tandemly arranged families of repeats. They proposed that orphons might arise by a mechanism similar to prokaryotic transposition events. Since the 3.6 kb repeats are contiguous to the 9 kb repeats, we believe that it is unlikely that this type of transposition is involved. A second possible explanation for the origin of the 3.6 kb repeats is that these repeats represent degenerate forms of the 9 kb repeats. Since it is difficult to understand why the 5S rRNA coding sequences are well conserved and the non-coding sequences completely diverged, we do not favor this explanation.

Since some lab strains do not contain the 3.6 kb repeats, these sequences are unlikely to serve some indispensable function in the yeast cell. It appears, however, that the $5S^{VAR}$ genes are expressed. Piper et al. (10) found that about 5% of the 5S rRNA isolated from vegetative cells of the yeast strains A364a and X2180-1B had the ribonuclease T1 fingerprint expected for $5S^{VAR}$ rRNA. In addition, Piper et al. (10) and D. Brow and E. P. Geiduschek (personal communication) have preliminary evidence indicating that some of the $5S^{VAR}$ genes can be efficiently transcribed by RNA polymerase III in vitro. It seems likely that if the $5S^{VAR}$ genes are transcribed, the $5S^{VAR}$ RNA is functional since the observed changes at the 3' end should still allow the formation of a terminal stem (24).

ACKNOWLEDGMENTS

We thank Drs. Vincent, Goebl, Lustig, Walmsley and Robertson for advice concerning these experiments and Drs. Brow and Geiduschek for providing information concerning the in vitro transcription of the $5S^{VAR}$ genes. The research was supported by N.I.H. grants GM24110 and R.C.D.A. 1 KO4 AG0007 (T.D.P.) and a grant from the National Cancer Institute CA 19265 (T.D.P. and M.E.M.). In addition, M.E.M. had support from the Leonard S. Florsheim, Jr. Fund, the Logan Fund and a PHS Training Grant (CA9267).

*Present address: Abbott Laboratories, Abbott Park, IL 60064, USA

REFERENCES

1. Long, E. O. and Dawid, I. B. (1980) Ann. Rev. Biochem. 49, 727-764.
2. Nath, K. and Bollon, A. P. (1977) J. Biol. Chem. 252, 6562-6571.
3. Bell, G. I., DeGennaro, L. J., Gelfand, D. H., Bishop, R. J., Valenzuela, P. and Rutter, W. J. (1977) J. Biol. Chem. 252, 8118-8125.
4. Cramer, J. H., Farrelly, F. W., Barnitz, J. T. and Rownd, R. H. (1977) Mol. Gen. Genet. 151, 229-244.
5. Petes, T. D., Hereford, L. M. and Skryabin, K. G. (1978) J. Bact. 134, 295-305.
6. Philippsen, P., Thomas, M., Kramer, R. A. and Davis, R. W. (1978) J. Mol. Biol. 123, 387-404.
7. Schweizer, E., MacKechnie, C. and Halvorson, H. O. (1972) J. Mol. Biol. 40, 261-277.
8. Petes, T. D. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 410-414.
9. Zamb, T. J. and Petes, T. D. (1982) Cell 28, 355-364.
10. Piper, P. W., Lockheart, A. and Patel, N. (1984) Nucleic Acids Res. 10, 4083-4096.
11. Boyer, H. W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
12. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning, p. 504, Cold Spring Harbor Laboratory, New York.
13. Petes, T., Broach, J. R., Wensink, P. C., Hereford, L. M., Fink, G. R. and Botstein, D. (1978) Gene 4, 37-49.
14. Benton, W. D. and Davis, R. W. (1977) Science 196, 180-182.
15. Grunstein, M. and Hogness, D. S. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3961-3965.
16. Davis, R. W., Botstein, D. and Roth, J. R. (1980) Advanced Bacterial Genetics, pp. 106-125, Cold Spring Harbor Laboratory, New York.
17. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
18. Schatchat, F. W. and Hogness, D. S. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, 371-381.
19. Maxam, A. M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
20. Valenzuela, P., Bell, G. I., Masiarz, F. G., DeGennaro, L. J. and Rutter, W. J. (1977) Nature 267, 641-643.
21. Findley, R. C. and Gall, J. G. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 3312-3316.
22. Roman, H. (1981) in The Molecular Biology of the Yeast Saccharomyces, Strathern, J. N., Jones, E. W. and Broach, J. R., eds., Vol. 1, pp. 1-9, Cold Spring Harbor Laboratory, New York.
23. Childs, G., Maxson, R., Cohn, R. H. and Kedes, L. (1981) Cell 23, 651-663.
24. Erdmann, V. A., Wolters, J., Huysmans, E., Vanderburghe, A. and DeWacter, R. (1984) Nucleic Acids Res. 12, r133-r166.