Inverted repeated sequences in yeast nuclear DNA

Hannah L.Klein* and Susan K.Welch†

Department of Genetics, University of Washington, Seattle, WA 98195, USA

Received 15 August 1980

ABSTRACT

The inverted repeated sequences (foldback DNA) of yeast nuclear DNA have been examined by electron microscopy and hydroxyapatite chromatography. Of the inverted repeat structures seen in the electron microscope, 34% were hairpins and 66% had a single stranded loop at the end of a duplex stem. The number average length of the repeat was 0.3 kb and the single stranded loop was 1.6 kb. It is estimated that there are approximately 250 inverted repeats per haploid genome. A statistical analysis of the frequency of molecules containing multiple inverted repeats showed that these sequences are non-randomly distributed. The distribution of inverted repeats was also examined by measuring the fraction of total DNA in the foldback fraction that bound to hydroxyapatite as a function of single strand fragment size. This analysis also indicated that the inverted repeats are clustered. Renaturation kinetic analysis of isolated foldback and inverted repeat stem sequence DNA showed that these sequences are enriched for repetitive DNA.

INTRODUCTION

Inverted repeated sequences have been described in a number of eukaryotic genomes (1,2,3,4,5). These sequences are characterized by reannealing with rapid, concentration independent kinetics following denaturation and incubation in renaturing conditions. This fraction of DNA can be observed in the electron microscope as single stranded molecules with an intramolecular duplex segment, or can be isolated by hydroxyapatite chromatography. It is often referred to as foldback DNA (6). In mouse and *Physarum* these sequences have been found to be non-randomly distributed (3,4), while in *Drosophila* the foldback sequences are randomly distributed (2). Inverted repeated sequences have also been identified in the *E. coli* chromosomal DNA (7,8), bacterial plasmids and episomes (for a review see 9), and animal viruses (10,11,12).

One class of inverted repeated sequences, the insertion sequences or IS sequences of *E. coli* have been shown to be involved in the transfer of antibiotic resistance factors and in the integration and excision of F, F' and R episomes (13,14,15,16). The IS sequences have also been found to be hotspots for recA independent recombination in F and R plasmids (17,18,19,20,21).

The role of inverted repeated sequences in the eukaryotic genomes is unknown. Two types of inverted repeated sequences have already been described in yeast. The first is an inverted repeated sequence present in the 2 μ m DNA plasmid (22,23,24,25). This sequence appears to be the site for an intramolecular recombination event, resulting in two different orientations of the plasmid. The second is a series of sequences called delta (δ) sequences (26). These sequences occur in both direct and inverted repetition. There are some suggestions that the δ sequences are involved in the deletion and translocation of genetic material in yeast (26,27,28).

This paper reports the description of the inverted repeated sequences in yeast nuclear DNA by electron microscopy and hydroxyapatite chromatography. These sequences on the average are small, about 0.3 kilobases (kb) in length. Both types of analysis show that these sequences are non-randomly distributed throughout the genome. Renaturation kinetic studies show that the inverted repeats and adjacent sequences are enriched in repetitive DNA. The possible identity of these inverted repeats with the δ and Ty1 sequences of yeast (26) is considered.

MATERIALS AND METHODS

Strains and cell growth

S. cerevisiae strain X322, provided by G. Simchen, was used for electron microscope experiments. The cells were grown at 30°C to stationary phase in YM-1 medium (29) supplemented with 40 μ g/ml adenine and 10 μ Ci/ml [6-³H] uracil (New England Nuclear).

For the hydroxyapatite experiments, *S. cerevisiae* strain 1L, a petite lacking mitochondrial DNA (Whitney, personal communication) was used. Unlabeled DNA was isolated from cells grown at 23°C to stationary phase in YEP supplemented with 40 μ g/ml each of adenine and uracil. ³²P-labeled DNA was isolated from 1L cells grown at 23°C to stationary phase in phosphate-free medium (30) supplemented with 40 μ g/ml each adenine and uracil. ³H-labeled cells were grown to stationary phase in YM-5 (29) supplemented with 40 μ g/ml uracil, 6 μ g/ml adenine, and 1 μ Ci/ml [6-³H]-adenine (specific activity 824 Ci/mmol; New England Nuclear).

DNA preparation for electron microscopy

Cells were converted to spheroplasts with glucosylase according to Klein and Byers (31). Spheroplasts in 1 M sorbitol-25 mM sodium citrate-phosphate-10 mM EDTA (pH 5.9) were lysed over a 0.1 ml layer of 5% Sarkosyl on top of a 4.8 ml linear 15 to 30% sucrose gradient containing 1 M NaCl-10 mM EDTA-10 mM Tris (pH 8.0). Centrifugation was performed at 5°C in a Spinco 50L rotor at 35,000 rpm for 3.5 hr. Fractions were collected and analyzed for DNA content as described (31).

Electron microscopy

Foldback molecules were prepared according to Cech and Hearst (3). DNA samples in 50 mM sodium phosphate buffer (pH 6.9) were denatured by incubation in 0.11 M NaOH for 5 min at 37°C followed by an additional 5 min at room temperature. The solution was then adjusted to pH 8.5 and made 100 mM Tris, 10 mM EDTA, and adjusted to 0.18 ionic strength (μ) with a 430 mM NaCl–100 mM EDTA–420 mM Tris–HCl pH 8.5 solution. The renaturation temperature was made equivalent to 55°C by the addition of an equal volume of formamide (recrystallized from Matheson, Coleman and Bell, 99% pure). Renaturation was carried out to a Cot of 5×10^{-4} . After renaturation, the hyperphase was diluted in 50% formamide–50 mM Tris (pH 8.5) and cytochrome C (horse heart type VI, Sigma) was added to 80 μ g/ml. The hyperphase was spread onto a hypophase of 20% formamide–5 mM Tris (pH 8.5). Grids were stained with uranyl acetate and then rotary shadowed with Pt–Pd (80:20). Grids were examined in a Philips 300 electron microscope. Contour lengths were measured with a Graf/Pen model GP-3 digitizer connected to a Hewlett-Packard calculator. The measured lengths were calibrated with ϕ X174 single strand circles of molecular length of 5.25 kilobases (kb) and T4 linear duplexes of molecular length 37.9 kb pairs.

Isolation of DNA for kinetic analysis

DNA was isolated by a modification of the method of Bernardi (32). Cells were disrupted with glass beads in a Braun homogenizer in 0.2 M Tris–HCl (pH 8)–0.01 M EDTA–15% sucrose. The lysate was made 1 M in NaCl and 0.1% in SDS, and extracted twice with an equal volume of chloroform:isoamyl alcohol (50:1). The final aqueous phase was ethanol precipitated, the nucleic acid pelleted by centrifugation. Following resuspension in 0.01 M Tris–HCl (pH 8)–2 mM EDTA, the DNA was treated with RNase, pronase, and amylase. The enzyme-treated solution was extracted twice with chloroform phenol:isoamyl alcohol (50:50:1) and the DNA ethanol precipitated from the final aqueous phase. Following centrifugation and resuspension of the nucleic acid, the DNA was sheared to a single strand fragment length of 5–6 kb by passage through a 27 gauge needle. Contaminating RNA was removed by hydroxyapatite chromatography. The DNA containing fractions were pooled, dialyzed against 0.01 M Tris–HCl (pH 8)–0.15 M NaCl–5 mM EDTA, and sedimented to equilibrium in CsCl. Fractions containing DNA were pooled and dialyzed against several changes of 0.12 M phosphate buffer (PB).

Renaturation kinetics

Aliquots of labeled yeast DNA in 0.12 M PB were sealed in sterile disposable glass microsampling pipettes (Corning) by melting the ends of the pipettes. The sealed pipettes were placed in a boiling water bath for 7–8 min to denature the DNA, then transferred to a 60°C bath for renaturation. At appropriate intervals, samples were diluted into 1 ml ice cold 0.12 M PB. Samples taken from the bath and immediately

diluted were considered zero time points for renaturations.

The percent double stranded DNA in each sample was assayed by hydroxyapatite chromatography according to Britten *et al.* (6). Fractionations were carried out in 0.15 ml bed volume columns equilibrated in 0.12 M PB at 60°C. Renaturation samples were quickly heated to 60°C and loaded onto the columns. The single stranded flow-through and 0.12 M PB washes were collected into a scintillation vial, and the double stranded fraction eluted into a second scintillation vial. Samples were then counted in Aquasol.

Isolation of foldback DNA. ³²P-labeled DNA of single strand fragment length 5-6 kb was diluted into 0.12 M PB at a concentration of less than 1 µg/ml. The DNA was denatured by immersion in a boiling water bath, chilled to 0°C, reheated to 60°C, and passed through a number of 0.3 ml bed volume hydroxyapatite columns equilibrated at 60°C in 0.12 M PB. Single stranded DNA was eluted with 3 x 2 ml washes of 0.12 M PB, and the foldback fraction eluted with 1 x 2 ml and 2 x 1 ml washes of 0.5 M PB. Aliquots of the load, 0.12 M, and 0.5 M PB washes were spotted on Whatmann 3 MM filter paper discs for counting. The effective Cot values for this fractionation were 1-1.5 x 10⁻³. Time for reheating the residence time on columns was included in renaturation time in calculating the Cot value. To be certain that the foldback DNA was not contaminated with rapidly renaturing (non-inverted repeat) DNA sequences, a second cycle of purification was carried out. The foldback DNA isolated in the first fractionation was diluted to 0.12 M PB by addition of water, heat denatured, quick-chilled, and carrier DNA added. The DNA was heated to 60°C and passed over hydroxyapatite columns as above. The effective Cot values for this second cycle of purification were at least ten-fold lower than in the first cycle. The foldback DNA isolated in this way made up slightly more than 2% of the total nuclear DNA.

Shearing of foldback and inverted repeat stem DNA. Foldback and inverted repeat stem DNA isolated from fragments of yeast DNA with an average single strand fragment length of 5-6 kb was sheared to smaller fragment lengths for use in renaturation experiments. DNA fragments of 0.4-0.7 kb were obtained by passing the DNA solution in 0.12 M PB three times through a French pressure cell at 12000 PSI. Fragments of single strand length 0.35 kb were prepared by sonication. Each sample was made 0.2 M in sodium acetate and the volume adjusted to 1 ml. The solutions were sonicated for 5 min at 120 watts. Following sonication, the DNA was passed over a Chelex-100 column and in 0.2 M sodium acetate and ethanol precipitated, and resuspended in 0.01 M Tris-HCl (pH 8)- 5 mM EDTA.

Determination of DNA single strand fragment length. Weight average single strand fragment length of the sheared DNA was determined by zone sedimentation through 5-20% alkaline sucrose gradients (0.1 M NaOH, 0.9 M NaCl). Centrifugation was in a Beckman 50 L rotor at 4°C for 6 to 16 hr at 33,000 rpm. Intact T4

bacteriophage or French pressure cell-sheared salmon sperm DNA was used as molecular weight marker. Gradient fractions were neutralized, TCA precipitated, and counted as described in figure legends.

RESULTS

Electron microscopic analysis of inverted repeated sequences

Inverted repeated sequences can be visualized in the electron microscope as duplex branches from a single stranded molecule. A schematic representation of such a molecule is shown in Figure 1. This molecule contains the sequence *bc* and its complement *b'c'* repeated in reverse orientation. The duplex region formed by denaturation and reassociation of the repeated sequences on the same strand is referred to as the stem. When these sequences are separated by an interstitial nonrepeated sequence *d*, the duplex stem has a single stranded loop at one end. Inverted repeated sequences lacking sequence *d* form duplex structures called hairpins.

To examine the inverted repeated or foldback sequences of yeast DNA, high molecular weight nuclear DNA was isolated from vegetatively grown cells. The DNA

Figure 1. Schematic representation of an inverted repeated sequence. The inverted repeated sequence *bc* and *c'b'*, after denaturation, can reanneal to form a fold-back structure consisting of a single stranded molecule with a duplex branch. If the two units of the repeat are separated by the non-repeated interstitial sequence *d*, the duplex branch will have a single stranded loop at one end.

was denatured and renatured as described in MATERIALS AND METHODS to a Cot of 5×10^{-4} to 10^{-3} . Intermolecular reassociation of nuclear DNA begins at Cot 10^{-2} under the conditions used in these experiments (33,34). After renaturation, the DNA was spread and examined in the electron microscope.

The number average length of the single stranded fragments examined was 23 kb. The 2 μ m plasmid DNA, which has an inverted repeat of 0.64 kb (22) and could be recognized by its length, was excluded from this analysis. Of 281 molecules photographed and measured, 33% were found to contain visible inverted repeat sequences. A total of 120 such structures were observed; examples are shown in Figures 2 and 3. The limit of resolution in the electron microscope of regions of convergence

Figure 2. Electron micrographs of inverted repeated sequences in yeast chromosomal DNA. DNA was denatured, briefly renatured in 50% formamide, and spread for electron microscopy. A collection of small inverted repeats, visualized as duplex segments, is shown. Each foldback structure is indicated by an arrow. (c) and (d) each show a double inverted repeat. The bar in (i) represents 1 kb.

Figure 3. Electron micrographs of long inverted repeated sequences. DNA was denatured, renatured in 50% formamide and spread for electron microscopy. The renatured sequences, visualized as duplex stems, are indicated by the arrows. (a) and (b) show foldback molecules with no visible loop. (c) shows a foldback structure with a single stranded loop, formed by a non-repeated interstitial sequence between the repeated sequences. The bar represents 1 kb.

between single stranded DNA to form apparent duplex segments is about 0.075 kb. Apparent duplex regions less than this length were not scored as inverted repeats. Of the foldback structures observed, 34% did not contain a visible single stranded loop. Again, the limit of resolution of single stranded loops in the electron microscope is 0.050 to 0.075 kb, so that turnaround loops smaller than this size were not scored.

The distributions of stem and loop lengths are shown in Figure 4. The number average length of the stems was 0.3 kb and of the loops was 1.6 kb. Although the distribution of the stem lengths was not as broad as the loop lengths, a few very large stem lengths at 3.9 kb and 5.4 kb were observed. Calculations of the frequency of these large repeats suggests that the 3.9 kb inverted repeat is present one to three times per

Figure 4. Observed length distributions of segments of the foldback molecules. A shows the distribution of duplex stem lengths. The mean size is 0.3 kb with a sample size of 120. B shows the distribution of single stranded loop lengths. The mean size is 1.6 kb with a sample size of 78. Foldback structures with no visible loop were excluded from the distribution in B.

genome, while the 5.4 kb inverted repeat is present one to four times per genome. Both the length and the copy number of the 5.4 kb inverted repeat suggest that it is identical with the Ty1 sequence.

Thirty-three percent of the molecules examined were found to contain at least one inverted repeat, and 27% of the molecules that contained an inverted repeat had more than one such structure. The distribution of molecules with multiple inverted repeats is shown in Table 1. These data can be used to calculate the average center to center spacing of inverted repeat pairs (2). If the inverted repeats were randomly distributed in the yeast genome, they would occur at an average spacing of 66 kb. This value is larger than the size of the fragments used for this analysis. Clearly, then, the model of random distribution is not correct. Statistical analysis of the frequency of molecules containing multiple inverted repeats indicate that these structures are clustered in the genome. That is, there are many more examples of molecules containing more than one inverted repeat pair than expected if the inverted repeats

Table 1. Distribution of Molecules with Multiple Inverted Repeats

	Number of Inverted Repeats per Molecule				
	0	1	2	3	>3
Number of Molecules	161	53	9	7	4

were randomly distributed along the genome.

Two hundred eighty-one molecules of average length of 23 kb were photographed and measured. A total of 120 inverted repeats were observed in this 1.6×10^6 kb of DNA. Assuming a haploid genome size of 1.35×10^4 kb (35) and that the DNA fragments examined are a random sample of the total nuclear DNA, we calculate that there are approximately 250 inverted repeat sequences per genome.

Characterization of inverted repeats by hydroxyapatite chromatography

Inverted repeated sequences are detected by rapid concentration-independent kinetics of reassociation following denaturation and incubation in renaturing conditions. The fraction of foldback DNA that is retained on hydroxyapatite as a function of DNA chain length has been shown to reflect the distribution of inverted repeated sequences along the DNA (1). High molecular weight ^{32}P -labeled yeast DNA was isolated as described in MATERIALS AND METHODS and sheared by passage through a French press or through needles of various gauges to single strand fragment lengths of 0.5 kb to 40 kb. Single strand fragment lengths were determined by zone sedimentation through alkaline sucrose gradients. The fragments were denatured by heat or alkali and renatured to Cot values of 10^{-5} to 10^{-3} mole l^{-1} sec. Repetitive DNA in yeast does not begin to renature until Cot 10^{-2} mole l^{-1} sec (33,34,36). After renaturation the DNA was passed over hydroxyapatite and the percent DNA in the foldback fraction determined.

The percent DNA in foldback as a function of single strand fragment length and Cot of renaturation is shown in Table 2. The amount of DNA in the foldback fraction increased from about 2% at 0.5 kb to 8% at 45 kb. It can be seen that the percent foldback is independent of Cot over a ten-fold range. This is characteristic of intramolecular renaturation events. The data in Table 2 show that the percent of DNA in foldback does not increase proportionately to the single strand fragment length, suggesting that the inverted repeated sequences are not randomly spaced, but rather are clustered in a few percent of the yeast genome.

The percent DNA in foldback at a given fragment length can be used to estimate the number of inverted repeats per genome. At a fragment length of 0.5 kb about 2% of the DNA contains a foldback duplex that binds hydroxyapatite. The haploid nuclear

Table 2. Percent Foldback in Yeast DNA

Fragment Size	Cot 1.9×10^{-4}	Cot 5.7×10^{-4}	Cot 1.7×10^{-3}	Cot 3.2×10^{-3}
0.5 kb	2.4	2.1	3.1	2.8
2-3 kb	4.0	5.2	5.2	4.2
8-10 kb	4.6	4.6	5.3	4.9
20 kb	2.8	2.9	3.7	3.7
40 kb	3.3	3.5	4.0	4.3
45 kb	5.8	6.2	6.7	8.7

genome is made up of 1.35×10^4 kb (35), or about 2.5×10^4 fragments of 0.5 kb each. If 2% of these fragments bind hydroxyapatite because they contain an inverted repeat, there are about 500 inverted repeats in the genome. Similar calculations for the 20 kb foldback fraction indicate that there are about 30 inverted repeat sequences per genome. This apparent increase in the estimated number of inverted repeats as the DNA fragment size decreases is further evidence for clustering of these structures. Multiple inverted repeats which are closely spaced on a longer fragment are scored as only one inverted repeat. The individual inverted repeat pairs of each cluster are not scored until the fragment size becomes small enough to separate the individual members of a cluster.

The binding of this small fraction of DNA to hydroxyapatite is specific to the formation of duplex regions and does not represent non-specific binding of a random fraction of the DNA. Less than 1% of T4 DNA was retained by the column under the conditions used to isolate yeast foldback sequences. Secondly, the percent foldback of yeast DNA remained constant in the presence or absence of a forty-fold excess of unlabeled calf thymus carrier DNA. Repassage experiments showed that less than 0.6% of a 0.5 kb single stranded fraction bound when passed over a second column while all of the foldback DNA was retained. In addition, RNase treatment in low salt conditions to digest double stranded RNA prior to passage over hydroxyapatite did not affect the yield of yeast DNA in the foldback fraction.

The thermal stability of the foldback DNA was examined for the foldback fractions from single stranded DNA fragment lengths of 0.4-0.7 kb and 40 kb. The foldback fraction was isolated by hydroxyapatite chromatography as described in MATERIALS AND METHODS and reabsorbed to a new hydroxyapatite column. The temperature of the column was raised in 5°C increments, equilibrated, and the DNA melted at that temperature eluted with 0.12 M phosphate buffer. The results are shown in Figure 5. The profile of the 0.4-0.7 kb foldback DNA was sharp and showed little material melting at temperatures below the T_m of yeast DNA-DNA duplexes (83°C)

Figure 5. Thermal denaturation profiles of yeast foldback DNA isolated from DNA fragments of 0.5 and 40 kb in length. Foldback was isolated as described in MATERIALS AND METHODS by alkali denaturation of DNA and renaturation at 60°C for 10 min. Concentration of DNA in renaturation mixes was 0.21 $\mu\text{g/ml}$ and final Cot was 9×10^{-4} . Partially renatured solutions were passed through hydroxyapatite columns equilibrated at 60°C in 0.12 M PB, and the single strand and foldback fractions collected. The foldback fractions were diluted from elution concentration of 0.5 M PB to 0.12 M PB by addition of sterile water, heated to 60°C, and loaded onto a hydroxyapatite column poured in a jacketed glass column. The DNA was eluted at a series of temperatures as described in MATERIALS AND METHODS. A 0.5 M PB was used to elute any DNA bound to the column at 90°C. The percent DNA eluting at each temperature was plotted versus the temperature. Symbols: Δ , 0.5 kb foldback DNA; $\circ$, 40 kb foldback DNA.

(34,35). In contrast, about 20% of the foldback fraction from the 40 kb fragments did not rebind. In addition, this foldback fraction contained a substantial amount of material that eluted at low temperatures. This behavior of foldback DNA isolated from large single stranded DNA fragments has also been observed with *Drosophila* DNA (2).

Kinetic components in and around the inverted repeats

Renaturation analysis was used to determine whether the inverted repeat sequences and the adjacent DNA included single copy or repetitive sequences. Foldback DNA, which includes both the inverted repeat stem and the flanking sequences, was

isolated from 5-6 kb long single stranded DNA fragments by 2 cycles of hydroxyapatite chromatography as described in MATERIALS AND METHODS. The foldback DNA was then sheared to smaller fragment sizes (0.3-0.4 kb) for use in reassociation experiments. This foldback fraction consistently made up a little over 2% of the total nuclear DNA. The inverted repeat stem sequences themselves were isolated by renaturing total yeast DNA (5-6 kb single strand fragment length) to low Cot, then digesting the single stranded regions with S1 nuclease. Double stranded DNA was separated from the digestion products by hydroxyapatite and Sephadex chromatography. The purified stem sequence DNA, which made up 4.2% of the total DNA, was then sheared by sonication to a single strand fragment length of 0.33 kb for use in renaturation experiments.

Two types of renaturation experiments were carried out. First, the foldback and stem sequence DNA were renatured with an excess of total yeast DNA driver to determine the copy number of these DNA fractions. Second, the foldback and stem DNA preparations were renatured without total DNA added to determine their sequence complexities.

Frequency classes of foldback and stem DNA

The kinetics of renaturation of foldback and stem DNA with excess total DNA driver are shown in Figure 6. The lines drawn through the data are theoretical curves, derived by least squares fit, for the second order reassociation of DNA of the sequence composition shown in Table 3. These data indicate that total yeast DNA contains about 11% repetitive sequences which are present in about 250 copies per genome. Schweizer *et al.* (37) estimate from hybridization experiments that there are about 140 ribosomal cistrons in yeast, making up about 10% of the total DNA. Yeast cells also contain a 2 μ m plasmid which makes up about 2% of the total DNA and is present in 35 copies per haploid genome (23). These repetitive DNA sequences probably account for most of the rapidly renaturing component in the total DNA in Figure 6, although other repeated sequences which comprise only a few percent of the DNA cannot be ruled out.

Figure 6 also shows that foldback DNA renatures much faster than total DNA, such that over 50% was renatured by a Cot of 1.0. The sequences in foldback DNA occur in a broad range of copy numbers, from only a few copies to several hundred copies per haploid genome. The Cot $\frac{1}{2}$ values, and copy number for the three frequency classes of foldback DNA, calculated from a least squares fit of the renaturation data, are shown in Table 3. Foldback DNA is clearly enriched for repetitive DNA sequences. About 15% of this DNA is made up of sequences which are present in about 500 copies per haploid genome. Another 50% comprises sequences which are present in 14 or more copies. Hybridization data (34) show that none of this foldback is ribosomal DNA. However, the 2 μ m plasmid, which contains an inverted repeat sequence (22,23,24,25), probably contributes to this foldback fraction, but cannot account for all of it. Plasmid

Figure 6. Renaturation kinetics of tracer amounts of ^{32}P -labeled total, foldback, and stem DNA with 2000x excess of ^3H -labeled total DNA. Reaction mixtures were prepared at two driver DNA concentrations to allow accurate determination of renaturation kinetics at both low and high Cot values. Mixes contained the following DNAs in 0.12 M PB: 1) Total DNA, 0.069 $\mu\text{g/ml}$ ^{32}P -labeled total DNA, and either 30.5 or 204 $\mu\text{g/ml}$ ^3H -labeled total DNA; 2) Foldback DNA, 0.059 $\mu\text{g/ml}$ ^{32}P -labeled foldback DNA, and either 30.5 or 143 $\mu\text{g/ml}$ ^3H -total DNA; 3) Stem DNA, 0.016 $\mu\text{g/ml}$ ^{32}P -labeled stem DNA, and either 30.5 or 143 $\mu\text{g/ml}$ ^3H -total DNA. ^{32}P -labeled DNA, specific activity 1.2×10^6 cpm/ μg , and ^3H -total DNA was prepared by the method of Bernardi (32) and had a specific activity of 2000 cpm/ μg . Each renaturation mix was divided into 8 aliquots, each of which was sealed into a glass capillary. DNA was denatured by placing the capillaries in a boiling water bath, and samples were moved to a 60°C bath for renaturation. Following renaturation for 3–2500 min, samples were diluted into 1 ml cold 0.12 M PB, and the percent DNA in hybrid assayed by HAP chromatography as described in MATERIALS AND METHODS. Shear size of DNAs was 0.75 kb for ^3H -total DNA, 0.11 kb for ^{32}P -total DNA, 0.36 kb for foldback DNA, and 0.33 kb for stem DNA. Kinetics were corrected for differences in rate due to differences in fragment length by applying the relationship $Cot \frac{1}{2} \propto 1/l^{1/2}$, and using the $Cot \frac{1}{2}$ of total yeast DNA at 0.75 kb as 6. Percent DNA in hybrid at each Cot value was corrected for zero time binding by subtracting the percent DNA in duplex at 0 min of renaturation from each duplex fraction. Symbols: $\blacksquare$, $\bullet$, ^3H -total DNA; $\circ$, ^{32}P -total DNA; Δ , ^{32}P -foldback DNA; $\square$, ^{32}P -stem DNA.

DNA is most likely represented by the kinetic component that renatures with $Cot \frac{1}{2} = 3.2 \times 10^{-1}$, the expected value for renaturation of plasmid in the presence of total DNA driver. The remaining foldback DNA is made up of single copy sequences and sequences which are more repetitive than the plasmid DNA. Thus, all frequency components of the yeast genome are included in the inverted repeat stems, loops, and the flanking

Table 3. Kinetic Components of Total, Foldback and Stem Sequence DNA

	<u>Percent</u>	<u>Cot ½</u>	<u>Copy Number</u>
Total DNA	11%	3.7×10^{-3}	250
	89%	8.8	1
Foldback DNA	15%	2.0×10^{-3}	500
	57%	3.2×10^{-1}	14
	34%	8.5	1
Stem DNA	44%	5.3×10^{-3}	700
	30%	6.7×10^{-2}	40
	26%	8.5	1

DNA sequences. Kinetic analysis of *Drosophila* foldback DNA has shown similar results (2).

Figure 6 and Table 3 describe the kinetics of renaturation of the S1-nuclease resistant stem sequences with total DNA driver. This DNA contains sequences which are present in a broad range of copy numbers, from only a few to several hundred per haploid genome, and is even more enriched for repetitive sequences than foldback DNA. Almost half of this DNA is made up of sequences which are present in about 700 copies per haploid genome (Table 3). It is likely that these repetitive sequences are the same ones which make up the most rapidly renaturing component of foldback DNA. Hybridization data (34) show that approximately 25% of the S1 nuclease resistant DNA preparation used in these experiments was ribosomal DNA, present due to intermolecular reassociation during preparation. In addition, we expect the 2 μ m plasmid to make up most of the kinetic component with $\text{Cot } \frac{1}{2} = 6.7 \times 10^{-2}$. The remaining DNA renatures too fast to be either ribosomal or 2 μ m DNA, and must represent some other repetitive sequence.

Sequence complexity of foldback and stem DNA

Renaturation kinetics of the foldback and stem sequences were measured in the absence of excess total driver DNA to determine the sequence complexity of these DNA fractions. Figure 7 shows that at least half the foldback DNA renatures much faster than total DNA, and is therefore made up of a subset of the total yeast genome. The $\text{Cot } \frac{1}{2}$ values and sequence complexity of each kinetic component are indicated in Table 4. The middle component ($\text{Cot } \frac{1}{2} = 2.3 \times 10^{-2}$) is probably 2 μ m plasmid DNA, as discussed above. About 40% of the foldback DNA renatures with a $\text{Cot } \frac{1}{2} = 1.8 \times 10^{-4}$ and has a sequence complexity of 1.3×10^5 daltons. This value is about 1000 times less than the sequence complexity expected if this DNA fraction were all single copy sequences. Thus, some sequences in foldback DNA are present in about 1000 copies per genome. If most of the approximately 500 small inverted repeats are included in this

Figure 7. Renaturation kinetics of foldback DNA and inverted repeat stem sequence DNA with added total DNA driver. Renaturation mixes containing 0.039 $\mu\text{g/ml}$ ^{32}P -labeled foldback DNA (shear size 0.36 kb) or 0.053 $\mu\text{g/ml}$ ^{32}P -labeled inverted repeat stem DNA were prepared in 0.12 M PB. The mixes were divided into aliquots, sealed in glass capillaries, and the renaturation kinetics measured at 60°C as described in MATERIALS AND METHODS. Approximately 80 μg sheared salmon sperm carrier DNA was added to each sample before the percent duplex DNA was determined by hydroxyapatite chromatography. Symbols: Δ , foldback DNA; $\square$, inverted repeat stem DNA; O, total DNA.

foldback DNA fraction, then there are at most 2-3 different sequences included in the inverted repeats.

The renaturation kinetics of stem DNA without total DNA driver are also shown in Figure 7. The $\text{Cot } 1/2$ values and sequence complexities of each theoretical component are shown in Table 4. The stem DNA preparation renatures much faster than either total or foldback DNA preparations, and therefore has a lower sequence complexity.

Table 4. Sequence Complexity of Foldback and Stem Sequence DNA

	Percent	$\text{Cot } 1/2$	Sequence Complexity
Foldback DNA	37%	1.8×10^{-4}	1.3×10^5
	30%	5.2×10^{-2}	3.7×10^7
	33%	4	—
Stem DNA	55%	2×10^{-5}	1.4×10^4
	20%	1.2×10^{-3}	8.5×10^5
	25%	$> 2 \times 10^{-3}$	—

Hybridization data show that 25% of this DNA preparation is ribosomal DNA, which is expected to renature with a $Cot \frac{1}{2} = 5.6 \times 10^{-3}$ in the presence of 75% other sequences. The 2 μ m plasmid inverted repeats would renature with a $Cot \frac{1}{2} = 8 \times 10^{-4}$ if they made up all of the remaining 75% of the preparation, and more slowly if other sequences were present. Figure 7 shows that at least half of the stem DNA renatures faster than either of these two DNAs, and must represent sequences with a lower complexity.

The observed complexity of the rapidly renaturing component of the stem DNA is about 1.4×10^4 daltons. Electron microscopy and hydroxyapatite binding studies indicate that yeast nuclear DNA contains approximately 350 inverted repeat pairs of average length 0.075 kb, and 25 pairs of average length 0.3 kb or longer. The complexity of these inverted repeats would be 2.3×10^7 daltons if each inverted repeat pair were different. The complexity of the inverted repeat stem sequences calculated from the renaturation data is about 1000 times less than this, indicating that the inverted repeat sequences are not unique, but occur in 1000 or more copies per haploid genome.

DISCUSSION

The inverted repeats of the yeast genome were examined by electron microscopy and hydroxyapatite chromatography for size, distribution in the genome and sequence composition. Electron microscopic examination of foldback DNA showed that the number average size of the duplex repeat length was 0.3 kb. It is estimated that there are 250 inverted repeats per genome that can be detected by electron microscopy. Applying the Poisson distribution to the frequency of molecules containing multiple inverted repeats showed that these sequences are non-randomly spaced in the genome.

The distribution of inverted repeats was also examined by measuring the fraction of total DNA in the foldback fraction as a function of single strand fragment size. At 0.4-0.7 kb 2% of the DNA was observed to bind to hydroxyapatite (Table 2). This fraction increased to 6-8% at 20-40 kb (Table 2). The fact that the percent DNA foldback did not increase proportionately to the single strand fragment length is further evidence that the inverted repeats are clustered.

Estimates of the number of inverted repeats in the yeast genome, calculated from the percent DNA that binds hydroxyapatite as foldback DNA at different fragment lengths, vary from 500 inverted repeats per genome at a single strand fragment length of 0.5 kb to 30 inverted repeats per genome at 20 kb. The apparent increase in the number of total inverted repeats as the DNA is sheared to smaller fragment lengths can be partially explained by clustering of the inverted repeats. Each single stranded fragment which binds hydroxyapatite is assumed to contain only one inverted repeat. The individual members of each cluster of inverted repeats are not detected until the

fragment length becomes small enough to separate the pairs in the cluster.

There is some discrepancy between the number of inverted repeats per genome calculated from data obtained by electron microscopy and that obtained from the hydroxyapatite binding studies. At small fragment sizes (0.5 kb) the number of inverted repeats detected by hydroxyapatite binding was 500, twice the number seen in the electron microscope. This is probably due to the presence of many inverted repeats which are too small to be resolved by the electron microscope (0.075 kb) but which bind to hydroxyapatite and are included in the foldback fraction.

The sequence composition of foldback DNA was determined by renaturation kinetic analysis of isolated foldback and inverted repeat stem sequence DNA in the presence and absence of excess total nuclear DNA. These experiments indicate that the inverted repeat sequences are themselves repetitive sequence elements that are interspersed with single copy DNA. Even when the data are corrected for the presence of 2 μ m DNA, known to contain an inverted repeat (22,23,24,25) and the repetitive rDNA sequences, it is clear that the foldback and stem fractions are enriched for repetitive sequence DNA. The stem sequences are in the most rapidly renaturing fraction, suggesting that these sequences are the repetitive element in the foldback DNA fraction.

The complexity of the foldback and stem sequences was estimated by reassocia-tion kinetics in the absence of excess total nuclear DNA. The complexity of the most rapidly renaturing component of foldback DNA is 1000 times less than the complexity expected if this fraction were composed of single copy DNA. This result suggests that some sequences in foldback DNA are present in multiple copies. A similar analysis with the stem sequences showed that an even higher proportion of this fraction had a complexity 1000 times less than single copy DNA. Comparison of the copy numbers derived from renaturation kinetics with total numbers estimated by electron microscopy and hydroxyapatite binding suggests that there are at most 2-3 different sequences included in the inverted repeated sequences present at the highest copy numbers.

The repeated 0.3 kb δ sequences of yeast are present at about 100 copies per haploid genome and sometimes occur in an inverted repeat orientation (26). The relationship of the inverted repeated sequences observed in this study to the δ repeated sequences is not clear. However, the average size of the inverted repeats observed in the electron microscope in this study of 0.3 kb is remarkably similar to the size of the δ sequences. It is also of interest to note that both electron microscopy and hydroxyapatite analyses show that the inverted repeats are non-randomly spaced. The δ sequences of yeast occur both as single sequences and in clusters with several units in inverted orientation (26). Recent work has suggested that the inverted repeated δ sequences are involved in the rearrangement of DNA sequences in yeast (26,27,28),

when the δ sequences flank a repeated element called Tyl. The Tyl element is a 5.6 kb repeated sequence that occurs at approximately 35 copies per genome (26) and behaves as a transposable element (28). Both the δ and Tyl sequences may represent some of the repeated components of the yeast genome observed in this study.

ACKNOWLEDGMENTS

We thank Drs. Petes, Byers, Hall and Fangman for comments on the manuscript. This research was supported by Public Health Service grant GM 18541 to Breck Byers and Public Health Service grant GM 11895 to Ben Hall.

Current address: *Department of Microbiology, The University of Chicago, Chicago, IL 60637, and [†]Department of Zoology, The University of Washington, Seattle, WA 98195, USA

REFERENCES

1. Wilson, D.A. and Thomas, C.A., Jr. (1974) *J. Mol. Biol.* 84, 115-138.
2. Schmid, C.W., Manning, J.E. and Davidson, N. (1975) *Cell* 5, 159-172.
3. Cech, T.R. and Hearst, J.E. (1975) *Cell* 5, 429-446.
4. Hardman, N. and Jack, P.L. (1977) *Eur. J. Biochem.* 74, 275-283.
5. Krumlauf, R. and Marzluff, G.A. (1980) *J. Biol. Chem.* 255, 1138-1145.
6. Britten, R.J. and Smith, J. (1970) *Carnegie Inst. of Wash. Year Book* 68, 378-386.
7. Chow, L.T. (1977) *J. Mol. Biol.* 113, 611-621.
8. Ohtsubo, H. and Ohtsubo, E. (1977) in *DNA Insertions, Plasmids, and Episomes*, Bukhari, A., Shapiro, J.A. and Adhya, S.L., Eds., pp. 49-65, Cold Spring Harbor, New York.
9. Cohen, S.N. (1976) *Nature* 263, 731-738.
10. Hsu, M. and Jelinek, W.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1631-1634.
11. Shine, J., Czernilofsky, A.P., Friedrich, R., Bishop, J.O. and Goodman, H.M. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1473-1477.
12. Strauss, S.E., Sebring, E.D. and Rose, J.A. (1976) *Proc. Natl. Acad. Sci. USA* 73, 742-746.
13. Kleckner, N., Chan, R.K., Tye, B.-K. and Botstein, D. (1975) *J. Mol. Biol.* 97, 561-575.
14. Berg, D., Davies, J., Allet, B. and Rochaix, J.-D. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3628-3632.
15. Chow, L.T. and Bukhari, A. (1977) in *DNA Insertions, Plasmids, and Episomes*, Bukhari, A., Shapiro, J.A. and Adhya, S.L., Eds., pp. 295-307, Cold Spring Harbor, New York.
16. MacHattie, L.A. and Jackowski, J.B. (1977) in *DNA Insertions, Plasmids, and Episomes*, Bukhari, A., Shapiro, J.A. and Adhya, S.L., Eds., pp. 219-229, Cold Spring Harbor, New York.
17. Davidson, E.H., Hough, B.R., Klein, W.H. and Britten, R.J. (1975) *Cell* 4, 217-238.
18. Hu, S., Ohtsubo, E., Davidson, N. and Saedler, H. (1975) *J. Bacteriol.* 122, 764-775.
19. Mickel, S. and Bauer, W. (1976) *J. Bacteriol.* 127, 644-655.
20. Ptashne, K. and Cohen, S.N. (1975) *J. Bacteriol.* 122, 776-781.
21. Kopecko, D.J. and Cohen, S.N. (1975) *Proc. Natl. Acad. Sci. USA* 72, 1373-1377.
22. Livingston, D.M. and Klein, H.L. (1977) *J. Bacteriol.* 129, 472-481.
23. Gubbins, E.J., Newlon, C.S., Kann, M.D. and Donelson, J.E. (1977) *Gene* 1, 185-207.

24. Guerineau, M., Grandchamp, C. and Slonimski, P.P. (1976) *Proc. Natl. Acad. Sci. USA* 73, 3030-3034.
25. Hollenberg, C.P., Degelmann, A., Kustermann-Kuhn, B. and Roger, H.-D. (1976) *Proc. Natl. Acad. Sci. USA* 73, 2072-2076.
26. Cameron, J.R., Loh, E.Y. and Davis, R.W. (1979) *Cell* 16, 739-751.
27. Rothstein, R. (1979) *Cell* 17, 185-190.
28. Roeder, G.S. and Fink, G.R. (1980) *Cell*, in press.
29. Hartwell, L.H. (1967) *J. Bacteriol.* 93, 1662-1670.
30. Rubin, G.M. (1974) *Eur. J. Biochem.* 41, 197-202.
31. Klein, H.L. and Byers, B. (1978) *J. Bacteriol.* 134, 629-635.
32. Bernardi, G. (1971) in *Methods in Enzymology*, Grossman, L. and Moldave, K. Eds., Vol. 21D, pp. 95-102, Academic Press, New York.
33. Lauer, G.D., Roberts, T.M. and Klotz, L.C. (1977) *J. Mol. Biol.* 114, 507-526.
34. Welch, S.K. (1978) Ph.D. Thesis, the University of Washington.
35. Bicknell, J.N. and Douglas, H.C. (1970) *J. Bacteriol.* 101, 505-512.
36. Whitney, P. and Hall, B.D. (1974) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* 33, 1282a.
37. Schweizer, E., MacKechnie, C. and Halvorson, H.O. (1969) *J. Mol. Biol.* 40, 261-267.