Sequence organization within and flanking clusters of 5S ribosomal RNA genes in *Tetrahymena*

David S.Pederson*, Meng-Chao Yao+, Alan R.Kimmel§ and Martin A.Gorovsky*

*Department of Biology, University of Rochester, Rochester, NY 14627, +Department of Biology, Washington University, St. Louis, MO 63130, and §Laboratory of Cellular and Developmental Biology, National Institutes of Health, Bethesda, MD 20205, USA

ABSTRACT

Macro- and micronuclei of Tetrahymena thermophila each contain approximately 30 clusters of 5S genes per haploid genome. Structural changes in DNA sequences associated with some of these clusters occur during the development of the transcriptionally active macronucleus from the transcriptionally inert micronucleus. Exonuclease digestion indicates that DNA fragmentation is not responsible for these changes, making it likely that sequence rearrangements occur near some 5S genes during macronuclear development. These rearrangements appear to be random in location with respect to the 5S genes and do not alter either the tandem repeat organization of the genes, the average number (five) or the range in number (one to about 16) of genes per cluster. The 5S gene clusters are not closely linked and are not flanked by common repeating elements large enough to cross-hybridize. Sequence analysis of one tandem repeat suggests that Tetrahymena 5S genes have intragenic promoters. These observations indicate that features other than DNA rearrangements or DNA sequence per se are responsible for the transcriptional activation of 5S genes during macronuclear development.

INTRODUCTION

The ciliated protozoan, Tetrahymena thermophila, has a diploid

micronucleus and an endoreplicated macronucleus. During conjugation, cells

pair and micronuclei undergo meiosis to produce 4 haploid nuclei, one of which

divides to produce 2 gametic nuclei. One gametic nucleus from each cell

migrates to the other cell and fuses with the stationary gametic nucleus.

This zygotic micronucleus then undergoes mitotic divisions to yield daughter

products in the posterior cytoplasm which remain micronuclei, and daughter

products in the anterior cytoplasm which differentiate to form new

macronuclei. During this time, the old macronucleus degenerates. While the

macronucleus is transcriptionally active, little or no RNA is synthesized by

micronuclei (reviewed in 1). In addition, genetic studies with heterokaryons

indicate that dominant genes in the micronucleus are not expressed (2). These

studies, coupled with the recent production of viable nullisomic and

amicronucleate strains (3,4) suggest that most, if not all, genes in the

micronucleus are transcriptionally inactive.

The structural changes that occur during the formation of new transcriptionally active macronuclei include DNA breakage, rearrangement, elimination, adenine methylation, and amplification of the large ribosomal RNA genes (6-15). In an effort to understand how these changes relate to the regulation of a specific gene, we have examined the structure of the 5S ribosomal genes in macro- and micronuclei.

Macro- and micronuclei both contain similar numbers of 5S genes per haploid genome distributed among several clusters. Within each cluster the 120 base pair genes are tandemly repeated in a head to tail fashion, separated by AT rich spacers of about 160 base pairs (16,17). Because the Tetrahymena 5S gene can be transcribed by an oocyte extract from Xenopus and because both the Tetrahymena and Xenopus 5S genes can be transcribed by Tetrahymena nuclear extracts (R. Hallberg, personal communication), it is likely that the Tetrahymena gene is transcribed from an internal promoter, as are the 5S genes in Xenopus and other systems (reviewed in 18, 19). However, this does not necessarily account for all aspects of 5S gene regulation in Tetrahymena. Evidence summarized above strongly suggests that micronuclear 5S genes are transcriptionally inactive while the macronuclear 5S genes are active. Since micro- and macronuclear 5S gene sequences are both derived from the zygotic micronucleus, it is not clear how these different transcriptional states can be mediated solely through the gene's internal promoter. Therefore, we felt it important to examine the sequences which surround different clusters of 5S genes in Tetrahymena to see whether there are any sequences common to all clusters, and determine if these regions might show a consistent pattern of rearrangement during macronuclear development.

This paper describes the sequence of a single 5S gene, the isolation and analysis of 6 clones containing 5S gene clusters, and experiments designed to analyze sequences and DNA rearrangements in and near 5S gene clusters. These studies represent the first detailed characterization of a chromosomal gene family in the macro- and micronuclei of Tetrahymena, and suggest that reorganization of the 5S gene clusters during macronuclear development is not responsible for gene activation.

MATERIALS AND METHODS

Cell culture, DNA purifications

Tetrahymena thermophila (strains B VII, B IV or CU 357) were grown axenically as described (5) in 1% proteose peptone, 0.2% dextrose, and 0.003%

sequestrine to densities of 2 to 3 x 10^5 cells/ml. Nuclei were isolated as described (5) or with minor modifications (Pederson and Gorovsky, manuscript in preparation). DNA was purified by treatments with proteinase K (Beckman Chemicals), RNAses A and T1 (Sigma), and by extraction with phenol and chloroform (20).

Plasmids in E. coli K802 cells (gal$^-$, met$^-$, supE, hsdR$^+$, hsdM$^+$) grown in LB were isolated from high salt-SDS lysates (20) and purified by extraction with acid phenol (22).

Phage grown in Q358 cells (hsdR$^-$, hsdM$^+$, supF, ϕ80^r) were treated with DNAse and RNAses, and precipitated with polyethylene glycol. Phage DNA was purified by treatment with proteinase K and extraction with phenol and chloroform (20).

Cloning, DNA manipulations

Tetrahymena macronuclear DNA was digested with Bam HI, and fractionated to yield small (280 base pair average) fragments highly enriched for 5S ribosomal genes (17). These were cloned into the plasmid pMK16 and transformants were selected by hybridization with kinase labeled 5S RNA (17), isolated as described (16). One clone, designated pTt16, was used for sequence determination by chemical degradation (23), primary screening of the λ library (see below), and for the experiments described in Figures 3 and 7. The insert in pTt5S was subcloned into the replicative form of phage MP9 for sequence determination by primed synthesis using dideoxynucleotides (24), and into the Bam H1 site of pBR322. The pBR322 transformant, designated pDP5, was used in all other hybridizations requiring a 5S gene probe.

Clones containing 5S gene clusters were isolated from a library of EcoRI digested micronuclear DNA in Charon 4A (12). The construction and screening of a library of Eco RI digested macronuclear DNA in pBR325, used in the isolation of clone pDP6, is described in detail elsewhere (20). All subcloning was done into pBR322 by standard methods (26), using gel purified restriction fragments (21).

Bal 31 digestions were done as described(13). DNA was digested by restriction enzymes (from New England Biolabs or BRL) in 0.1 M NaCl, 10 mM MgSO$_4$, 1 mM β-mercaptoethanol in a 37°C oven, using a combination of enzyme and time calculated to produce a 5 to 10 fold overdigestion on an equivalent amount of λ DNA. In some cases, an aliquot of the restriction digest mixture already containing genomic DNA was mixed with λ DNA assay for complete digestion. Restriction maps were constructed using data obtained from single, double, and (for Bam HI) partial digests of DNA which had been

fractionated and probed, first with the 5S gene, and then with vector DNA.

Nick translations essentially by standard methods (26) are described elsewhere (20). DNA fractionated in horizontal submersible agarose gels was partially hydrolyzed, denatured, and blotted to nitrocellulose (Schleicher and Schuell, Ba 85, 0.45 micron pores) essentially as described (25). Filters were hybridized in sealable plastic bags at 65°C, as described (20), washed twice at room temperature with 2 x SSSPE (SSSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, 8 mM NaOH, 1 mM Na_2EDTA, 5 mM SDS), once at room temperature with 1/2 x SSSPE, and once at 65°C with 1/2 x SSSPE. This last wash is about 9°C below the T_m for the Tetrahymena genome (27). Filters were air dried and exposed to preflashed X-ray film (Kodak XAR5) using a single Dupont Cronex intensifying screen at -80°C. The T_m for the probe pDP7 was estimated by quantitative densitometry of autoradiograms (exposed for specific lengths of time) of filters washed at 60°C first in 0.1 M Na^+, then in 25 mM Na^+, and finally in 6.3 mM Na^+, each step representing a stringency increase of about 10°C. When filters were to be reused with a new probe, the old probe was removed by washing the filters for 15 to 30 min at 75°C in 0.75 mM $Na_{1.5}PO_4$, 0.25 mM SDS, and 0.06 mM Na_3EDTA.

RESULTS
Sequence of the 5S ribosomal genes
The sequence of the gene and spacer is shown in Figure 1. Bam HI cuts the gene into a 30 base pair 5' end and a 90 base pair 3' end. The spacer is highly (86%) A+T rich as predicted (17) and contains 2 perfect palindromic sequences of 12 and 16 base pairs, several shorter palindromes, and tracts of oligo dT and oligo dA. The gene sequence agrees with one of the Tetrahymena thermophila 5S RNA sequence variants determined by Luehrsen et al. (28).
Clusters of 5S ribosomal genes in macro- and micronuclei
When DNA digested with enzymes which do not cut within the 5S gene tandem repeats is analyzed, multiple band patterns like those shown in Figure 2 are produced, indicating that 5S genes are distributed among many clusters in both macro- and micronuclei. One of the faintest bands in the profile generated by Eco RI (Figure 2) corresponds in size to a restriction fragment from clone pDP6, which is known to contain just one 5S gene (see Figure 5). The densitometric area of this band was used as an internal standard to calibrate the profiles shown in Figure 2. From an analysis of autoradiograms exposed for different lengths of time, we calculate that the haploid number of 5S genes in both macro- and micronuclei is about 150. In spite of
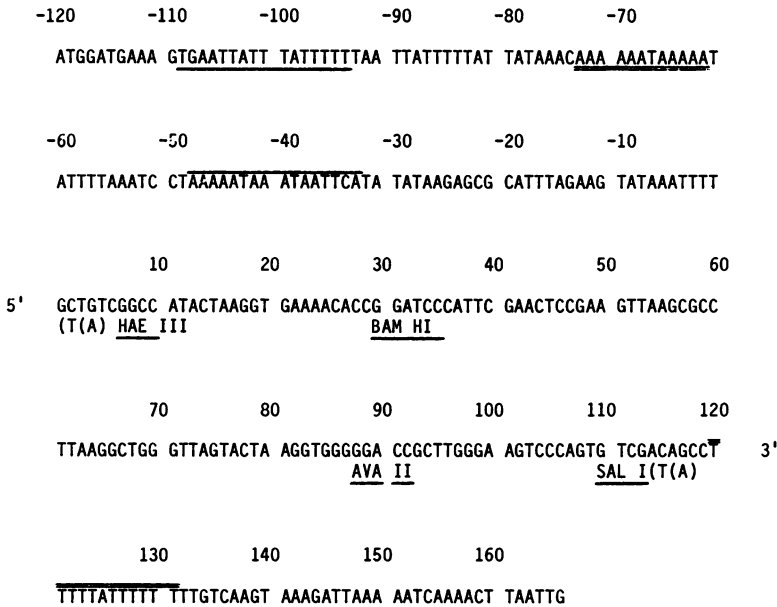
```
    -120        -110        -100         -90         -80         -70

    ATGGATGAAA GTGAATTATT TATTTTTTAA TTATTTTTAT TATAAACAAA AAATAAAAAT


     -60         -50         -40         -30         -20         -10

    ATTTTAAATC CTAAAAATAA ATAATTCATA TATAAGAGCG CATTTAGAAG TATAAATTTT


                 10          20          30          40          50          60

5'  GCTGTCGGCC ATACTAAGGT GAAAACACCG GATCCCATTC GAACTCCGAA GTTAAGCGCC
    (T(A) HAE III                    BAM HI


                 70          80          90         100         110         120

    TTAAGGCTGG GTTAGTACTA AGGTGGGGGA CCGCTTGGGA AGTCCCAGTG TCGACAGCCT      3'
                                     AVA II                  SAL I(T(A)


                130         140         150         160

    TTTTATTTTT TTTGTCAAGT AAAGATTAAA AATCAAAACT TAATTG
```

Figure 1: Sequence of the 5S Ribosomal Gene from Tetrahymena thermophila
     Sequences for the 120 base pair 5S gene and a 166 base pair spacer were
obtained from a cloned Bam HI fragment. Chemical sequencing (23) was done in
both directions from both the Bam HI and Sal I sites at positions 30 and 110
in the gene. In addition, the Bam HI fragment was subcloned into the
replicative form of phage MP9 and dideoxy sequencing (24) was performed using
a primer from PL Biochemicals. In the figure, the sequence has been written
so that the coding region is contiguous and is numbered from the 5' end of
the mature transcript, which was identified from the RNA sequence (28).
Since the 5S genes are tandemly repeated, the spacer is both downstream from
one gene, and upstream of another gene. It is thus not yet known what
portion of the spacer functions as an upstream sequence, and the decision to
write the sequence starting at -120 is an arbitrary one. The residues from
-109 to -94 and from -48 to -33 have been underlined and overlined,
respectively, to indicate a perfect 16 base pair palindrome. Similarly, the
sequences from -73 to -62 and from 120 to 131 (double under- and overlines)
form a perfect 12 base pair palindrome. Residues 2, 4, 115, and 117 at which
Luehrsen et al. (28) found RNA sequence heterogeneity are indicated by
showing the alternate bases in parentheses below the main sequence.
Restriction sites referred to in the text and subsequent figures are shown
with the recognition sequences indicated.

differences in the macro- and micronuclear band patterns, macronuclear

development generates no obvious changes in haploid gene number, number of 5S

gene clusters, or in distribution of cluster sizes (Figure 2; center

and bottom tracings). Comparison of profiles from 2 different digests of

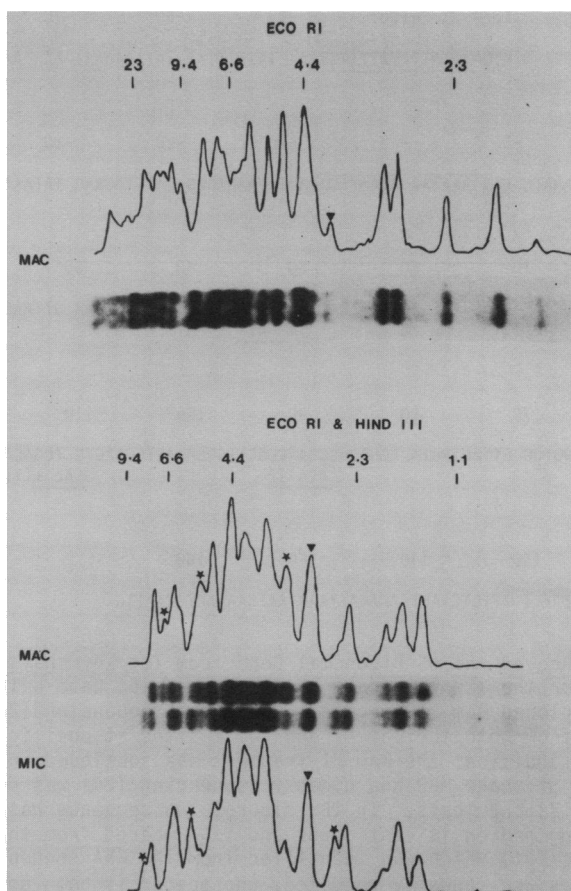macronuclear DNA (Figure 2) shows that in some digests not all the clusters

Figure 2: Clusters of 5S Ribosomal Genes in Macro- and Micronuclei
    DNA was digested with Eco RI, or with Eco RI and Hind III, gel
fractionated, blotted to nitrocellulose, and hybridized with the 5S gene.
Autoradiographs and densitometric profiles of the autoradiographs are shown.
The size (in KB) and position of Hind III cut λ and Hae III cut φX174 marker
fragments relative to bands in the densitometric profiles are indicated by
numbers and lines. Two different digest patterns are shown for the
macronucleus (top and middle tracings) to illustrate the difficulty in
resolving all the 5S gene clusters. The arrows indicate the position of the
fragment corresponding to the clone pDP6 (see Figure 5), which contains a
single 5S gene and was used to calibrate the densitometric profiles. In the
lower photograph which shows the 5S gene hybridization to macro- and
micronuclear DNA digested with EcoRI and Hind III, the fragment corresponding
to clone pDP6 is obscured by a cluster containing several genes. Some peaks
in the micronuclear pattern are much reduced in intensity relative to the
corresponding macronuclear peaks. This is probably due to a low level of
contaminating macronuclear DNA in the micronuclear DNA preparation. Some of
the peaks unique to either macro- or micronuclei have been starred.

are resolved. Therefore, the maximum number of countable clusters in any blot-29-is probably slightly less than the total number in the genome. From an analysis of our highest resolution blots, it appears that the largest cluster contains about 16 genes. Thus, as predicted from previous studies based on buoyant density centrifugation (17), the restriction fragments containing 5S gene clusters are generally larger than the clusters they contain, and are thus not composed entirely of 5S genes.

DNA rearrangements near 5S ribosomal genes during macronuclear development

The macro- and micronuclear band patterns in Figure 2 (center and bottom tracings) show clear differences suggesting that structural alterations occur near some of the 5S gene clusters during macronuclear formation. This difference cannot be due to an inhibition of restriction enzyme cleavages by macronuclear specific methyladenine residues (29), since the differences can be demonstrated with either Eco RI (which cleaves at GAATTC) or Hind III (which cleaves at AAGCTT), and the in vivo methylation site contains the bases 5'-NAT (7). DNA breakage to produce new free ends also takes place during macronuclear development (8, 9, 12, 13). A fragment created by DNA breakage can be identified by partially digesting macronuclear DNA with the exonuclease Bal 31. Fragments derived from free ends are shortened by the exonuclease as has been shown for C_4A_2 containing sequences in macronuclei (13) and for a sequence which flanks the excised rDNA in macronuclei (12). When the same DNA used in those studies was probed with 5S sequences, none of the 5S gene containing fragments in the macronucleus had been shortened (Figure 3). Thus the 5S gene containing fragments unique to macro- or micronuclei reflect true DNA rearrangements which occur during macronuclear development, and the 5S gene clusters are not located at the very tips of macronuclear chromosomes.

The fact that there is no wholesale lengthening or shortening of the tandem repeats during macronuclear development suggests that the rearrangements occur outside the tandem repeats themselves. To see if these rearrangements are unique to either the 5' (upstream) or 3' (downstream) sides of the gene clusters, DNA was digested with enzymes which cut asymmetrically in the 5S gene to probe specifically those fragments lying just upstream or those lying just downstream of the various 5S gene clusters (Figure 4). A similar fraction of the flanking sequences on either side of the clusters is involved in detectable rearrangements (about 1/4). The fraction of 5S gene containing fragments involved in rearrangements also is about 1/4 (Figures 2 and 3). Therefore the number of rearrangements flanking 5S genes clusters is more than enough to account for the number of cluster-containing fragments
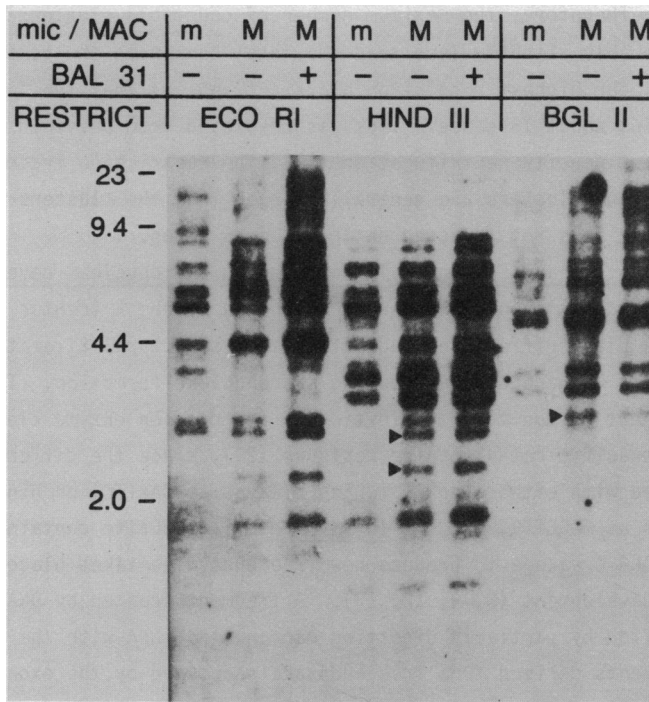
Figure 3: Bal 31 Digestion of the Macronuclear 5S Gene Clusters
    Purified, high molecular weight macronuclear DNA was digested with the exonuclease Bal 31 to remove approximately 2 kilobase pairs of DNA from the free ends of the DNA (13). The DNA was then digested with either Eco RI, Hind III, or Bgl II, and fractionated along with comparable restriction digests of non-Bal 31 treated micro- and macronuclear DNA's. The fractionated DNA's were then blotted and probed with the 5S gene. None of the bands detected in the macronuclear DNA are shortened or eliminated by the Bal 31 treatment. Thus, none of the 5S gene clusters is situated at the end of macronuclear chromosomes, and the fragments unique to macronuclei (examples indicated by arrowheads) reflect the fact that rearrangements rather than simple cleavages have occurred near 5S gene clusters during macronuclear development. By contrast, when the same Bal 31 treated DNA is probed with $(C_4A_2)_n$, these terminally located sequences are found to have been lost (data not shown).

which are rearranged, and lends support to our suggestion that rearrangements do not directly involve the 5S gene tandem repeat units. The same conclusions can be drawn when two different restriction enzymes (Bam HI and Ava II) are used to examine, respectively, the downstream and upstream fragments (data not shown).

    Because of sequence polymorphism (reflected in the 5S RNA sequence of reference 28; see Figure 1), not all 5S gene containing clusters are cleaved
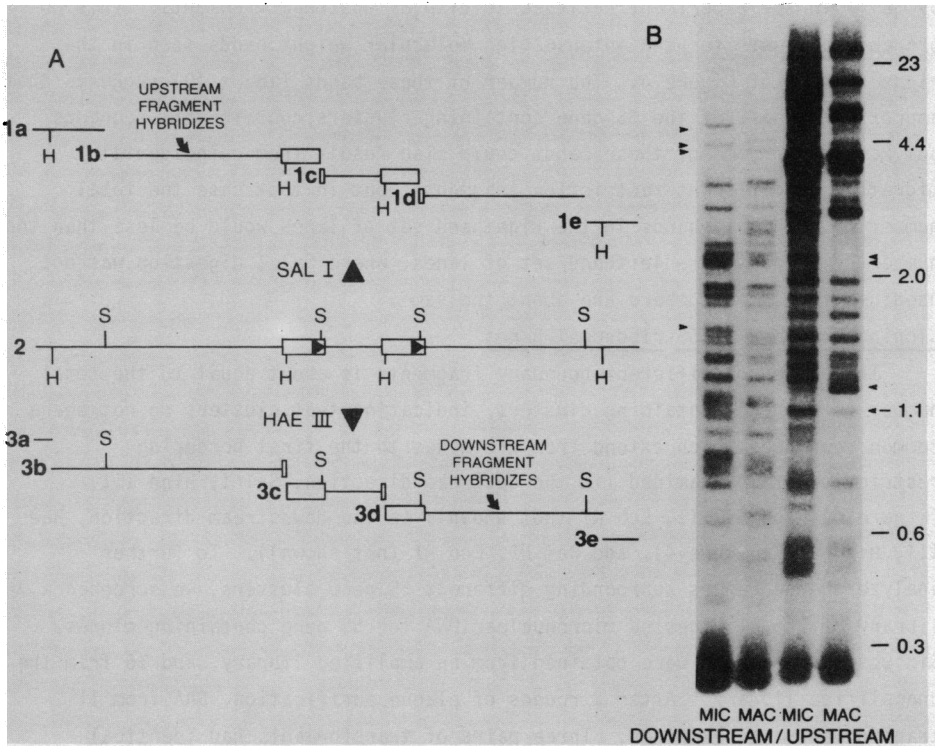
Figure 4: <u>Hybridization to Genomic DNA Upstream and Downstream to 5S Gene</u>
<u>Clusters</u>
      The experimental design is illustrated by the diagram in A, in which a
cluster-containing two 5S genes (open boxes) transcribed from left to right
is shown in the center (line 2). When the DNA is cut with Hae III (H), the
fragments 3a through 3e are generated. Of these, only the 0.28 KB tandem
repeat fragment (3c) and the downstream fragment (3d) contain enough of the
5S gene sequence to hybridize. Conversely, when the DNA is cut with Sal I
(S), fragments 1a through 1e are generated, and among these, only the
upstream fragment (1b) and the 0.28 KB tandem repeat fragment (1c) contain
enough of the gene to hybridize. The photograph in B shows the set of all
fragments lying downstream from the 5S gene clusters, generated by digestion
with Hae III and Hind III (left 2 lanes); and the set of all fragments lying
upstream from the 5S gene clusters, generated by digestion with Sal I and
Hind III (right 2 lanes). Hind III was included in the digests to reduce the
fragments to sizes which could be readily fractionated. Since it does not
cut within a 5S cluster, it's use does not alter the principle of the
experimental design. The arrows in B point to some of the bands which are
unique to either macro- or micronuclei. The probable origin of the intense
high molecular weight bands in the right 2 lanes is discussed in the text.

by Sal I in the experiment designed to display upstream fragments. This
probably accounts for the intense high molecular weight bands seen in the
right 2 lanes of Figure 4. The number of these bands (about 10) suggests that
approximately 1/3 of the 5S gene containing clusters contain this sequence
polymorphism. Some of these bands could also result from coincidental
migration of different restriction fragments, but in that case the total
number of countable bands in the righthand set of lanes would be less than the
number of bands in the lefthand set of lanes, where Sal I digestion was not
used; in fact, the numbers are about the same.

Cloning clusters of 5S ribosomal genes

The number of different boundary fragments is about equal to the total
number of 5S gene containing clusters, indicating that clusters do not share
common sequences which extend from the genes to the first bordering
restriction sites examined (in the upstream direction, Sal I, Hind III
(Figure 4), and Ava II, Eco RI (not shown); in the downstream direction, Hae
III, Hind III (Figure 4), and Bam HI, Eco RI (not shown)). To further
analyze the sequences surrounding different 5S gene clusters, we screened a λ
library of Eco RI digested micronuclear DNA for 5S gene containing clones.
Six positive plaques were obtained from an amplified library, and 16 from the
unamplified library. After 2 rounds of plaque purification, DNA from 11
transformants was examined. Three pairs of transformants had identical
length inserts (data not shown). One pair may have derived from
cross-contaminated 5S gene positive plaques. Only one of the pairs derived
from the amplified library indicating that there was no problem with
preferential amplification of clones. Poisson analysis based on 5 clones
recovered as single isolates and 2 clones recovered as genuine pairs,
suggests that the library contains at least 14 unique 5S gene containing
clones. None of the clones examined contains more than a single Eco RI
fragment (Figure 5 and data not shown), suggesting that the library is
dominated by fragments from complete Eco RI digestion. This being the case,
only about half of the 5S gene clusters are of a size (7 to 21 kb) which
could have been packaged successfully into Charon 4A phage (see Figure 2).
Thus, the cloning results are consistent with the estimate from genomic blots
that there are approximately thirty 5S gene containing clusters.

Six clones were chosen for more complete analysis, 3 from the amplified
λ library, 2 from the unamplified λ library and one, pDP6, from a library of
macronuclear Eco RI fragments in pBR325. Restriction maps for all these
clones and the subclones derived from them are shown in Figure 5. Together,
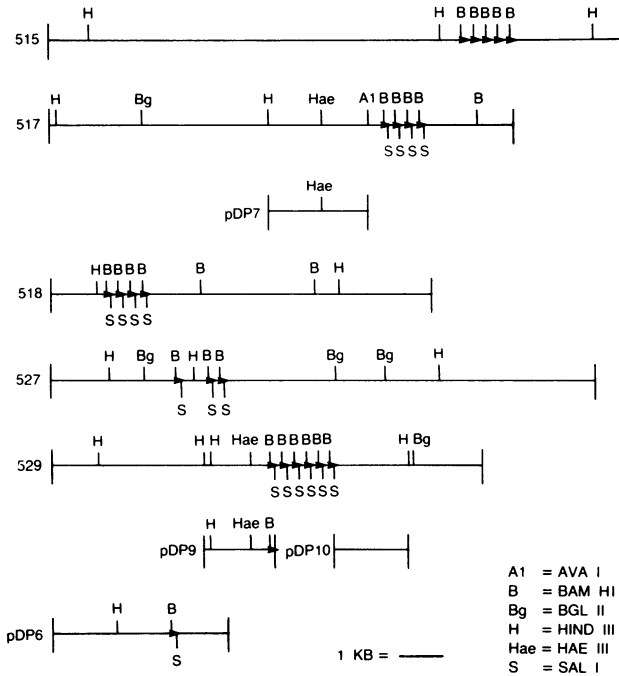
Figure 5: Restriction Maps of 5S Gene Containing Clones
     Restriction maps of the 5S gene containing clones selected from a
library of micronuclear Eco RI fragments in Charon 4A (λ 515, 517, 518, 527,
and 529), and from a library of macronuclear Eco RI fragments in pBR325
(pDP6) are shown. The maps were constructed as described in the Experimental
Procedures and in all cases, have been oriented so that the 5S gene
transcription unit (indicated by arrowheads) runs from left to right. Both λ
527 and pDP6 contain single 5S genes not linked in tandem arrays to other 5S
genes. The figure also shows subclones pDP7, pDP9, and pDP10, used in the
characterization of sequences which neighbor 5S gene clusters (Figure 6).
Mapping for Hae III and Ava I sites was done only on subclones, and therefore
may not be complete for the parent clones.

these clones contain 23 5S genes in 7 clusters ranging in size from one to 6

genes. Except for λ518, each clone corresponds to identical sized fragments

in both macro- and micronuclei (Figure 6 and data not shown), suggesting that

none contain sequences involved in genome rearrangements, and making it

unlikely that any sequences were lost during cloning. λ518 comigrates with a

fragment found in macro- but not micronuclei. It may therefore derive from a

macronuclear fragment which contaminated the library of micronuclear

fragments, and may represent a rearranged cluster. Except for λ527, all

clones contain just one 5S gene cluster suggesting that the clusters
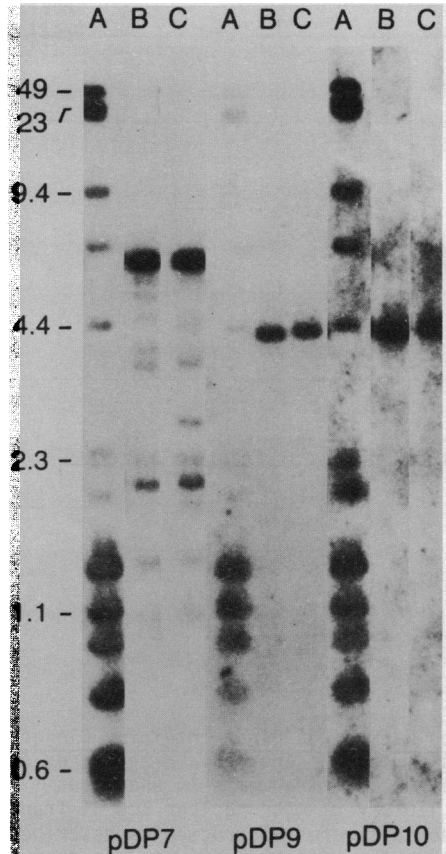
Figure 6: Hybridization to Genomic DNA by Sequences Flanking 5S Gene
Clusters
        Macronuclear (lanes B) and micronuclear (lanes C) DNA's were digested
with Eco RI and Hind III and electrophoresed with marker DNA (lanes A)
consisting of Hind III digested λ DNA and Hae III digested φX174 DNA. The
blot was probed successively with the clones indicated, and with radioactive
marker DNA's. After each hybridization, the old probe was removed before
hybridization to a new probe. The hybridization with clone pDP9 was done
using only the portion of the insert bounded by the Hind III and Hae III
sites (see Figure 5).

themselves are not closely linked in the genome.

5S gene clusters are on more than one chromosome

        The distribution of 5S-gene clusters was compared in micronuclear DNA

isolated from a wild-type strain (B-III) and a nullisomic strain (3) which

was missing both copies of one of the 5 chromosomes (strain CU357, missing

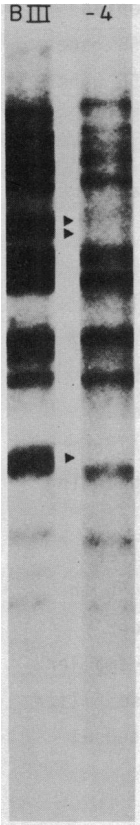chromosome 4) (Fig. 7). The absence of chromosome 4 is correlated with the

Figure 7: Hybridization to 5S Gene Clusters in Micronuclear DNA from a Wild Type and a Nullisomic Strain
Micronuclear DNA was isolated from a wild type strain (B III) and from a nullisomic strain (-4, missing both copies of chromosome 4), digested with Eco·RI and analyzed by Southern blotting using pTt5S as probe.

absence of at least 3 major clusters. Thus, there must be 5S gene clusters on at least 2 different chromosomes. A detailed analysis of the arrangement of 5S gene clusters in a number of nullisomic strains (Dr. Sally Allen, personal communication) indicates that clusters are found on four of the five Tetrahymena chromosomes.

Sequence heterogeneity among 5S genes is greater between than within 5S gene clusters

Every 5S gene in all the cloned clusters except for λ515 contains a Sal I site. None of the genes in λ515 contains a Sal I site. This, together with the pattern in the right 2 lanes of Figure 4 indicating that about 10 clusters are not cut by Sal I and about 20 are cut completely by Sal I, suggests that the sequence variants containing a modified Sal I site (residues 110 through 115 in the gene) are grouped in specific clusters rather than being distributed randomly among all clusters. This in turn

suggests either that dispersal of single, slightly divergent gene copies preceeded the production of tandem repeats during the evolution of 5S genes in Tetrahymena or that rectification occurs more rapidly among the 5S genes within a single cluster than between clusters. Since Luehrson et al. (28) found that some Tetrahymena 5S transcripts do not contain C at position 115 as required for Sal I cleavage, at least some of these variant clusters must be transcriptionally active.

Sequences surrounding 5S ribosomal gene tandem repeats

To see whether different clusters share common bordering sequences long enough to cross hybridize at moderate criteria, DNA from 2 λ clones (λ517 and λ529) was subcloned (Figure 5) and hybridized to restriction digests of macro- and micronuclear DNA (Figure 6). Two of the subclones -- one containing DNA from an upstream region (pDP9) and one containing DNA from a downstream region (pDP10) -- hybridize only to the parental band suggesting that they contain only single copy DNA. A third clone (pDP7) containing sequences lying upstream of the cluster in λ517 hybridizes to the parent genomic fragment strongly, and weakly to about 10 other fragments in both macro- and micronuclei. The weak hybridization could result from poor homology, or because the region containing the repeated DNA is much smaller than the whole cloned insert (about 2.1 KB). When the nitrocellulose filter was rewashed at progressively higher criteria all the fragments eluted at about the same temperature (data not shown), supporting the latter explanation.

Several lines of evidence argue that the repeat element does not play a role in regulating 5S gene expression in Tetrahymena. If this sequence were somehow involved in the differential regulation of macro- and micronuclear 5S genes by an epigenetic mechanism, it might be expected to be associated with most, if not all of the 5S clusters which are presumably active in macronuclei. Yet, the repeat element does not hybridize to the other 4 λ clones (data not shown) and there are only about 10 copies of the repeat element whereas there are roughly 30 widely spaced 5S clusters. If this sequence were involved in rearangement-mediated gene activation, it might be expected to show a consistent difference between macro- and micronuclei (e.g., all its copies to either be involved in DNA rearrangements or to be lost during macronuclear development). Yet, many of the fragments to which the repeat element hybridizes are identical in size in macro- and micronuclei.

DISCUSSION

Sequence of the Tetrahymena 5S genes

The 5S gene sequence of Tetrahymena thermophila is 34% divergent with respect to the somatic 5S gene of Xenopus borealis (30). Curiously, the divergence from the Xenopus sequence is slightly greater (38%) in the region 50 to 83 base pairs from the start of transcription where deletion experiments locate the Xenopus gene's promoter (18,19). However, only one out of the 9 residues identified as important for promoter binding by guanosine methylation experiments (31) differs between Tetrahymena and Xenopus. This, together with the observation by R. Hallberg (personal communication) that Xenopus oocyte extracts will transcribe the 5S genes of Tetrahymena, argues strongly that Tetrahymena 5S genes contain an internal promoter.

The Tetrahymena 5S gene transcript made by the Xenopus extract is 120 nucleotides in length, (R. Hallberg, personal communcation) suggesting that the mature RNA termini (28) correspond to the termini of the primary in vivo transcript. A cluster of at least 4 T residues flanked by GC rich residues is required to efficiently terminate transcription of the Xenopus 5S gene (33). The Tetrahymena 5S RNA terminates after the first T in the sequence 5' . . . CCTTTTTATTTTTTTGTC . . . 3' suggesting that transcription termination of 5S genes operates similarly in the two organisms. The fact that this T-rich tract is palindromic to sequences further downstream in the spacer (Figure 1) is intriguing but at present cannot be related to known features of 5S gene transcription.

The sequenced clone was derived from Bam HI digested DNA and thus contains 30 base pairs of one gene, and 90 base pairs of an adjacent 5S gene. Our finding that heterogeneity at the Sal I sequence occurs between rather than within 5S gene clusters suggests that our composite 5S gene sequence is identical to the sequence of each gene in the cluster from which the 5S clone was derived. If this is the case, our sequence supports the secondary structure model proposed for the 5S RNA (28). This model shows residues 2 and 117, and 4 and 115 paired in "helix I" of the RNA, which predicts that the sequence heterogeneity seen at these 4 residues would be limited in any one gene to combinations resulting in perfect base pairing. The sequence here contains C and G at residues 2 and 117, and G and C at residues 4 and 115. Paired substitutions that maintain base pairing in other regions of the 5S RNA of Neurospora have been reported and discussed (34). It should also be noted that while our composite 5S gene sequence has a repeat length of 286 bp,

restriction endonuclease analyses indicate that there is length heterogeneity
in 5S repeat units which range in size from about 250 bp to about 290 bp
(unpublished observations).

Number of 5S genes in Tetrahymena

Data shown here indicate that Tetrahymena macro- and micronuclei both
contain about 150 5S genes per haploid genome. This agrees reasonably with
earlier saturation hybridization estimates of up to 350 genes per macronuclear
haploid genome and a slightly lower level for micronuclei (16; unpublished
observations). The somewhat different estimates may reflect differences in
the methods used to derive them. The G1 value of the macronucleus is
approximately 45C, and each haploid macronuclear genome is probably fragmented
into about 200 pieces (see 12, 13, 35). At nuclear division macronuclear
chromosomes segregate randomly (reviewed in 36) which could result in one
daughter cell receiving more than its 45C share of certain genes. One
daughter cell might therefore achieve a growth advantage over the other,
leading eventually to a population of cells which have moderately amplified
certain sequences. The gene counting experiments described here rely on an
internal standard for calibration, and amplification would not have been
detected unless it was selective for a few 5S gene clusters. Thus the
possibility of preferential amplifcation of certain sequences in macronuclei
during vegetative growth, together with slight uncertainties about the degree
of loss of sequence complexity during macronuclear development (27) can
explain why the observed saturation hybridization values for macronuclei are
somewhat higher than those obtained from micronuclei (16; see Fig. 3).

Organization of the 5S genes in Tetrahymena

The Tetrahymena genome contains several single 5S genes in addition to
those clustered in tandem repeats. Except for Dictyostelium (37, 38) and
yeast (39, 40) where the 5S genes are linked to the large ribosomal genes, and
Neurospora where mainly dispersed single copies of the 5S genes are found
(34), a tandem repeat type of organization is common for these genes in
eukaryotes (reviewed in 18, 41; see also 42). The single genes could be
nontranscribed pseudogenes. However, they possess characteristic restriction
sites and the one isolated in the clone pDP6 supports transcription of a full
length 5S RNA in vitro (R. Hallberg, personal communication), suggesting that
they are functional.

Cloning results and nullisomic mapping suggest that the 5S gene clusters
are widely separated in the genome. Therefore, the mechanism which activates
5S genes during macronuclear development cannot be a single large scale change

affecting a single, large 5S gene locus.

Sequences surrounding 5S gene clusters

The sequences which surround clusters of 5S genes in Tetrahymena
macro- and micronuclei are largely unrelated and unique. The one exception
found consists of a short moderately repeated sequence within the clone pDP7
whose homologues appear not to be associated with other 5S genes. Thus, there
are no extensive 5S regulatory sequences adjacent to the 5S gene clusters.
Similar results have been found for dispersed, multigene families in other
organisms (34, 43-45). Our hybridization experiments would not have detected
short consensus sequences of the kind specific to histone genes (45) or those
preceeding functionally related RNA polymerase II transcribed genes (46).
They would have detected sequences like the "Bam Island" sequences which
preceed the large ribosomal genes in Xenopus and may be important for promoter
activity (47). Finally, while the downstream sequence analyzed abuts the 5S
genes, the subclones containing sequences lying upstream of the gene clusters
extend only to within about 200 base pairs of the 5S genes. Common sequences
closer to the 5' side of gene clusters would not have been detected.

Several lines of evidence indicate that rearrangements which occur near
5S gene clusters during macronuclear development are random in position with
respect to the 5S genes and are not related to 5S gene transcription. Our
observation that about 1/4 of the restriction fragments which contain 5S gene
clusters and average about 5 kb in size, are rearranged during macronuclear
development (17; Figures 2 and 3) is reminiscent of the observation (48) that
similar rearrangements occur, on the average, once every 30 kb in randomly
isolated clones from a micronuclear library. Thus the frequency of
rearrangements near 5S genes is about average with respect to the whole
genome. If the rearrangements served to bring in promoter elements or to
nucleate a specific chromatin structure involving the 5S genes, one might
expect them to occur at uniform distances from the 5S gene clusters or perhaps
with a single polarity with respect to the clusters. Neither of these
expectations is true since rearrangements occur both up- and downstream of
clusters (Figure 4), and examples can be found of fragments flanking 5S genes
which are several kilobase pairs in length and are not rearranged (Figures 4
and 5), as well as others much smaller which are rearranged (Figure 4).

Regulation of the 5S genes in Tetrahymena macro- and micronuclei

We have ruled out nearby DNA sequence rearrangements, action at large
homologous sequence blocks flanking 5S gene clusters, and propagation of
signals across one or a few compact 5S gene loci as the basis for the

difference in 5S gene expression in macro- and micronuclei. The problem of maintaining identical 5S genes in different states of activity in Tetrahymena is similar to the one faced by Xenopus, whose somatic tissues maintain active somatic type and inactive oocyte type 5S genes. In Xenopus, there is evidence that differences in chromatin structure may play an important role in that discrimination (49-51). Recent studies in Xenopus (52,53) also indicate that the oocyte type 5S genes from somatic nuclei remain inactive, even when the chromatin containing them has been reduced to less than a kilobase in length by nuclease digestions, suggesting that the 5S genes in Xenopus are regulated by local rather than global changes. The dispersed organization of the 5S gene clusters in Tetrahymena, coupled with the absence of any consistent difference between macro- and micronuclear 5S genes, suggests that similar mechanisms may be operating in Tetrahymena.

REFERENCES
1. Gorovsky, M. A. (1973) J. Protozool. 2, 19-25.
2. Bruns, P. J. and Brussard, T. B. (1974) Genetics 78, 831-41.
3. Bruns, P. J. and Brussard. T. B. (1981) Science 213, 549-51.
4. Kaney, A. R. and Speare, V. J. (1983) Exp. Cell Res. 143, 461-7.
5. Gorovsky, M. A., Yao, M.-C., Keevert, J. B. and Pleger, G. L. (1975) in Methods in Cell Biology, D. M. Prescott, ed., 9, 311-27 Academic Press, New York.
6. Gorovsky, M. A. (1980) Ann. Rev. Genet. 14, 203-39.
7. Bromberg, S., Pratt, K. and Hattman, S. (1982) J. Bacteriol. 150, 993-6.
8. King, B. O. and Yao, M.-C. (1982) Cell 31, 177-82.
9. Pan, W.-C. and Blackburn, E. H. (1981) Cell 23, 459-66.
10. Pan, W.-C., Orias, E., Flacks, M. and Blackburn, E. H. (1982) Cell 28, 595-604.
11. Pratt, K. and Hattman, S. (1981) Molec. and Cell Biol. 1, 600-8.
12. Yao, M.-C. (1981) Cell 24, 765-74.
13. Yao, M. C. and C. H. Yao. (1981) Proc. Natl. Acad. Sci. USA 78, 7436-9.
14. Yao, M.-C., Blackburn, E. and Gall, J. G. (1981) J. Cell Biol. 90, 515-20.
15. Yokoyama, R. and Yao, M.-C. (1982) Chromosoma 85, 11-22.
16. Kimmel, A. R. and Gorovsky, M. A. (1976) Chromosoma 54, 327-37.
17. Kimmel, A. R. and Gorovsky, M. A. (1978) Chromosoma 67, 1-20.
18. Korn, L. J. (1982) Nature 295, 101-5.
19. Korn, . J. and Bogenhagen, D. F. (1982) in The Cell Nucleus, H. Busch and L. Rothblum, eds., Vol. XII, Academic Press, New York.
20. Pederson, D. S. (1983) Thesis, University of Rochester, Rochester, New York.

21. Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981) Analytical Biochem. 112, 295-8.
22. Zasloff, M., Ginder, G. D. and Felsenfeld, G. (1978) Nucleic Acids Res. 5, 1139-52.
23. Maxam, A. M. and Gilbert, W. (1980) in Methods in Enzymology, Lawrence Grossman and Kivie Moldave, eds., 65, 499-560, Academic Press, New York.
24. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-7.
25. Wahl, G. M., Stern, M. and Stark, G. R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-7.
26. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Cold Spring Harbor, New York.
27. Yao, M.-C. and Gorovsky, M. A. (1974) Chromosoma 48, 1-18.
28. Leuhrsen, K. R., Fox, G. E. and Woese, C. R. (1980) Current Microbiology 4, 123-6.
29. Gorovsky, M. A., Hattman, S. and Pleger, G. L. (1973) J. Cell Biol. 56, 697-701.
30. Peterson, R. C., Doreing, J. L. and Brown, D. D. (1980) Cell 20, 131-41.
31. Sakonju, S. and Brown, D. D. (1982) Cell 31, 395-405.
32. Trifonov, E. N. and Sussman, J. L. (1980) Proc. Natl. Acad. Sci. USA 77, 3816-20. Deleted in proof.
33. Bogenhagen, D. F. and Brown, D. D. (1981) Cell 24, 261-70.
34. Selker, E. U., Yanofsky, C., Driftmier, K., Metzenberg, R. L., Alzner-DeWeerd, B. and RajBhandary, U. L. (1981) Cell 24, 819-28.
35. Katzen, A. L., Cann, G. M., and Blackburn, E. H. (1981) Cell 24, 313-20.
36. Nanney, D. L. and Preparata, R. M. (1979) J. Protozool. 26, 2-9.
37. Maizels, N. (1976) Cell 9, 431-8.
38. Cockburn, A. F., Newkirk, M. J. and Firtel, R. A. (1976) Cell 9, 605-13.
39. Bell, G. I., DeGennaro, L. J., Gelfand, D. H., Bishop, R. J., Valenzuela, P. and Rutter, W. J. (1977) J. Biol. Chem. 252, 8118-25.
40. Philippsen, P., Thomas, M., Kramer, R. A. and Davis, R. W. (1978) J. Mol. Biol. 123, 387-404.
41. Long, E. O. and Dawid, I. B. (1980) Ann. Rev. Biochem. 49, 727-64.
42. Hart, R. P. and Folk, W. R. (1982) J. Biol. Chem. 257, 11706-11.
43. Zasloff, M. and T. Santos (1980) Proc. Nat'l. Acad. Sci. USA. 77, 5668-5672.
44. McKeown, M., A. R. Kimmel and R. H. Firtel (1981) in ICN-UCLA Symp. Mol. Cell Biol., D. Brown, ed., XXIII, 107-116, Academic Press, New York.
45. Hentschel, C. C. and M. L. Birnstiel (1981) Cell 24, 301-313.
46. Davidson, E. H., H. T. Jacobs and R. J. Britten (1983) Nature 301:468-471.
47. Moss, T. (1982) Cell 30, 835-42.
48. Yao, M.-C., J. Choi, S. Yokoyama, C. Austerberry and C.-H. Yao. (1983) Submitted for publication.
49. Korn, L. J. and J. B. Gurdon (1981) Nature 289, 461-465.
50. Bogenhagen, D. F., Wormington, W. M. and Brown, D. D. (1982) Cell 28, 413-21.
51. Gottesfeld, J. and Bloomer, L. S. (1982) Cell 28, 781-91.
52. Gurdon, J. B., Dingwall, C., Laskey, R. A. and Korn, L. J. (1982) Nature 299, 652-53.
53. Reynolds, W. F., Bloomer, L. S. and Gottesfeld, J. M. (1983) Nucleic