

Eliminated sequences with different copy numbers clustered in the micronuclear genome of *Tetrahymena thermophila*

Theodore C. White*, M. Raafat El-Gewely**, and Sally Lyman Allen

Division of Biological Sciences, The University of Michigan, Ann Arbor, MI 48109, USA

Summary. As the ciliated protozoan Tetrahymena thermophila develops a new macronucleus (MAC) from products of its micronucleus (MIC), several repetitive sequences are eliminated from the MAC genome. Four MIC DNA clones containing repetitive sequences that are eliminated from the MAC were obtained. One clone contains a representative from each of three families of eliminated sequences. One, present in 200–300 copies in the MIC, is almost completely eliminated from the MAC. A second, present in approximately 50 copies in the MIC, is scattered throughout the genome, although up to half of the family members examined could be localized to chromosome 2. Approximately one tenth of the members of this less repetitive family persist in the MAC while the rest are eliminated. The third type of eliminated sequence has three to four members, all of which are eliminated from the MAC. Three of the members are located on three of the five MIC chromosomes, and one could not be mapped. This sequence is clustered with the other two families of sequences in at least three of the four sites. All three types of eliminated sequences are found in similar arrangements in the MIC of several different inbred strains of T. thermophila.

Introduction

Until recently, an organism's genome was thought to be unaltered during development. However, recent molecular analysis has shown that several DNA rearrangements occur during normal development. Rearrangements have been found in such diverse systems as immunoglobin switching in mammals (Tonegawa 1983), antigenic variation in Trypanosomes (Borst and Cross 1982) and mating type interconversion in yeast (Nasmyth 1982). Each of these rearrangements may involve some sort of DNA elimination. More extensive DNA elimination has been shown cytologically to be a normal part of development in several systems including ciliated protozoa (Prescott and Murti 1974; Yao and Gorovsky 1974), Ascarus (Wilson 1928), crustaceans

(Beerman 1977) and insects (M.J.D. White 1973). Thus, DNA elimination may play an important role in development, either by inactivating a previous function or by activating a new function necessary for further development, as is the case in immunoglobin switching. Further studies of eliminated sequences may elucidate other functions of these sequences in development.

DNA elimination has been shown to occur during macronuclear development in the ciliate, *Tetrahymena thermophila* (Yao and Gorovsky 1974). This protozoan contains two structurally and functionally distinct nuclei in a common cytoplasm. The micronucleus (MIC) is a diploid germinal nucleus containing five pairs of chromosomes. The MIC undergoes meiosis and mitosis yet is transcriptionally inactive. The macronucleus (MAC) is a somatic nucleus with approximately 23 times as much DNA as the MIC. It divides amitotically and is transcriptionally active. The MAC develops from products of the MIC during the sexual process of conjugation and the new MAC is propagated vegetatively until it is destroyed during the next conjugation (reviewed in Allen and Gibson 1973).

MAC development from products of the MIC involves many processes. Overall, the total DNA content of the nucleus increases at least 23 times. At the same time, several molecular events occur: (1) the DNA is fragmented into subchromosomal pieces approximately 600 kb long (Yao and Yao 1981); (2) the 18S and 26S rDNA locus is excised, becomes palindromic and is amplified (reviewed in Yao et al. 1979); (3) the DNA contains a high proportion of adenine bases, 0.8% of which become methylated during MAC development (Gorovsky et al. 1973), while (4) the chromatin composition is altered (reviewed in Gorovsky 1980); and (5) many repetitive sequences are eliminated during MAC development (Yao and Gorovsky 1974).

Eliminated sequences were first studied in Tetrahymena using hybridization kinetics. Yao and Gorovsky (1974) showed that 10%–20% of the MIC genome, mostly repetitive sequences, is eliminated or underreplicated in the MAC. The rest of the genome persists in the MAC. Iwamura et al. (1979) reached similar conclusions.

Eliminated sequences have recently been studied using recombinant DNA techniques. Yao (1981) has studied an eliminated sequence at one end of the rDNA locus that contains a repetitive element as well as a unique element. The sequence is associated with the excision of the rDNA locus and this excision leaves free DNA ends in the MAC. No rejoining of the flanking ends occurs.

^{*} Present address: Netherlands Cancer Institute, Plesmanlaan 21, 1066 CX Amsterdam, The Netherlands

^{**} Present address: Department of Biological Chemistry, Medical School, The University of Michigan, Ann Arbor, MI 48109, USA

Yao (1982) has also studied a cloned segment of MIC DNA which contains four families of repetitive sequences, all of which are eliminated during MAC formation. This clone also contains the (C₄A₂)_n repeat which is added to the ends of MAC fragments. These families are not tandemly arrayed but are clustered throughout the genome. Since no part of the clone persists in the MAC, the rearrangement associated with the elimination of these repetitive families could not be studied. Brunk et al. (1982) have also studied sequences associated with (C₄A₂)_n. They have analyzed DNA clones which contain (C₄A₂)_n, a repetitive family (X-H), and unique sequences. Both the repetitive sequences and the unique sequences are eliminated from the MAC. Brunk et al. (1982) also identified clones which contain eliminated sequences as well as sequences destined for the MAC. The DNA clones which were used to study the terminal and interstitial deletions also contain MIC sequences which are eliminated from the MAC during the deletion process.

Yao and his associates (1984) recently studied a cloned segment of MIC DNA containing repetitive sequences which are eliminated by deletion. Three regions of the clone were eliminated. In two cases, the two flanking regions were rejoined as the sequence was eliminated in the MAC. Similar deletions have recently been shown to occur on both sides of the tubulin gene (Callahan et al. 1984).

In this study, we have analyzed the arrangement of a cloned segment of MIC DNA that contains three different families of eliminated sequences. Each of the three families has a different level of repetition in the MIC, indicating that the MIC genome contains eliminated sequences with a wide spectrum of repetition, and each of the families is eliminated to different degrees, indicating that the members of an eliminated family are not always eliminated. The three families are also clustered together in several places in the MIC genome. This clustering of three distinct types of eliminated sequences occurs in similar arrangements in several different inbred strains of Tetrahymena.

Materials and methods

Cell lines. The cell lines used in this study are listed in Table 1. The genesis of and relationship between inbred strains, A, B, C, D, C2, C3, and D/1 (I₁) is elaborated in some detail by Allen et al. (1984). The first four strains were derived from crosses of wild isolates or of a wild isolate to strain B, and then inbred by selfing for some 16-25 generations. Strains C2 and C3 were generated by genomic exclusion of an AC hybrid crossed to C*. Genomic exclusion, an abnormal form of conjugation, results in the production of homozygous germinal nuclei by the diploidization of haploid nuclei following meiosis (Allen 1967). Strain D/1 is congenic with strain D; its construction has been discussed previously (Allen and Lee 1971). The "young population" was derived from a cross of $D \times D/1$. The young population refers to cells obtained from a population of exconjugant cells which had been purified from nonmating cells by magnetic columns (Bruns et al. 1980). The nullisomic strains were constructed by Bruns et al. (1983).

Growth of cultures. Cultures were maintained in 1% proteose peptone and were subcultured bimonthly at 23° C before 1979 or monthly at 16° C after 1979 unless frozen. Growth of cultures for nuclear isolation has been described

Table 1. Tetrahymena thermophila strains used in this study

Type of strain	Designation of lineage (names in parentheses are used in text)				
Inbred strains:	A-17686 (A) B-2079X2 (BX2) B-2079X6 (BX6) B-18687N (B7N) B-18687B (B7B)	B-18684 (B4) B-18684b (Bb) C3-3685 (C3) D-25772a (D) ^a D/1(I ₁)-5771 (D/1) ^s			
Nullisomics:	CU 358 CU 359 CU 362 CU 371	CU 372 CU 373 CU 383			
Young population:	DI21481 (DI)				

^a Cultures of these two strains were obtained from frozen samples stored at the American Type Culture Collection (ATCC 30845 and ATCC 30847, respectively)

(Allen et al. 1983). For most strains cultures were grown at 30° C; however, some strains were temperature sensitive and had to be grown at 25° C. Some cultures, especially the nullisomic strains, were contaminated with an unknown fungus. Two fungicides were used to help reduce fungal contamination which can impede nuclear isolation. The two fungicides were fungizone (0.25 mcg/ml) in Antibiotic – Antimycotic (Gibco Laboratories, New York, USA) and mycostatin (5 units/ml, Squibb Nystatin).

Nuclear isolation and DNA preparation. MIC and MAC were purified by the use of Percoll gradients as previously described (Allen et al. 1983). MIC were purified from MAC by successive Percoll gradients until the contamination levels were below 1 MAC in 2000 MIC as estimated by light microscopy. MAC were purified by repeated Percoll gradients until the contamination levels were below 1 MIC in 200 MAC. After nuclear isolation, DNA was purified by CsCl gradient centrifugation as already outlined (Allen et al. 1983).

Clone bank construction and colony hybridization. An incomplete MIC DNA clone bank was constructed using a partial BamHI digest of MIC DNA from a cross of D×D/1. This partially digested DNA was ligated into the BamHI site of pBR322 and the ligated plasmids were transformed into E. coli (RH202; Adams et al. 1979). Ampicillin-resistant, tetracycline-sensitive cells were isolated and stored at -80° C. The clone bank was screened by colony hybridization using nick-translated MIC or MAC DNA as a probe (Grunstein and Hogness 1975).

Plasmids used, their preparation and nick translation. Plasmids used in this study included pT140, pT240, pT1472 and pT1008 which were identified as described above. In addition, seven subclones of pT1008 were constructed: pT1008-A1, pT1008-A2, pT1008-A3, pT1008-A4, pT1008-A5, pT1008-B, and pT1008-C. In the text, these subclones will be referred to as A1, A2, A3, A4, A5, B, and C. One other plasmid used in this study was pDP5, a clone of the SS rDNA gene, which was constructed by D. Peterson (University of Rochester; Pederson et al. 1984). Plasmid DNA was isolated by a modification of the SDS/high salt cleared lysate procedure (Gunsalus et al. 1979), followed by CsCl/

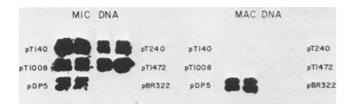


Fig. 1. Reconstruction of colony hybridization of the four DNA clones. *E. coli* containing the plasmids pT140, pT240, pT1008, pT1472 and pBR322 were grown in duplicate on nitrocellulose filters and lysed by a modification of the colony hybridization procedure (Grunstein and Hogness 1975). The filters were probed with nick translated MIC or MAC DNA

ethidium bromide gradient centrifugation in a vertical rotor. DNA was labeled in vitro with α ³²P dATP by a modified nick translation procedure routinely giving about 10⁸ cpm/µg (Maniatis et al. 1975; Rigby et al. 1977). Unincorporated radioactive nucleotides were removed from the samples by chromatography through Sephadex G-50.

Restriction enzyme digestion and gel electrophoresis. Digestions using the equivalent of ten-fold excess of enzyme were carried out by the method of Maniatis et al. (1982) using restriction enzymes purchased from Bethesda Research Laboratories, Gaithersburg, MD, USA, New England Biolabs, Beverley, MA, USA, or Amersham, Arlington Heights, IL, USA. Submarine gel electrophoresis was carried out with 12-inch long gels and a Tris-borate buffer system containing ethidium bromide (Helling et al. 1974; McDonell et al. 1977; Peacock and Dingham 1968) using appropriate size markers, either ϕX 174 DNA digested with HindII or λ DNA digested with AvaI and BgIII. λ DNA digested with AvaI and Bg/II generates fragments of the following sizes (in kb): 14.9, 8.8, 5.5, 4.3, 3.7, 2.44, 2.23, 1.93, 1.64, 1.44, 1.07, 0.54, and 0.48. Dashes in each figure (usually 10.0, 7.5, 5.0, 2.5, 1.0, and 0.5) were estimates of size and are derived from standard curves using the above size markers.

Genomic blotting and hybridization. DNA was transferred from the gels to nitrocellulose filters by the method of Southern (1975) with modifications. Prehybridization, hybridization and washing were all done at 62° C. These conditions were chosen due to the 25% GC content of Tetrahymena DNA (Allen and Gibson 1972). The filters were presoaked in $3 \times SSC$ for one half hour (SSC=0.15 M NaCl, 0.015 M sodium citrate) and then prehybridized for 5 h in $10 \times$ Denhardt's solution (Denhardt 1966) with $3 \times$ SSC. 0.1% SDS and 25 µg/ml of denatured calf thymus DNA. Then the prehybridization mixture was replaced with fresh denatured prehybridization and probe mixture $(1-2\times10^6)$ cpm/ml), and hybridization was continued for at least 36 h. Filters the size of a full sized gel received 4×10^7 cpm. The filters were washed six times in 2×SSC, 1% SDS for 20 min, then once in $0.2 \times SSC$, 1% SDS for 30 min, and finally in 2×SSC for 10 min. The filters were dried for 2 h and exposed to Kodak XAR-5 film at -80° C with an intensifying screen. The length of exposure depended on the amount of radioactivity on the filter.

Slot blots. Serial dilutions of MIC or MAC DNA were made in TE (10 mM Tris, 1 mM EDTA, pH 7.0). One tenth volume of 3 M NaOH was added to each dilution and the

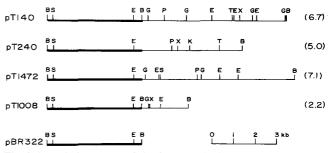


Fig. 2. Restriction maps of the four DNA clones. Restriction sites are designated as follows: B, BamHI; E, EcoRI; G, BglII; K, KpnI; P, HpaI; S, SalI; T, BstEII; X, XbaI. The thicker bar in each clone represents pBR322 sequences. The size of each insert in kb is indicated to the right. pT240 and pT1472 do not contain two BamHI sites. The second site was lost during plasmid construction

mixture was placed at 65° C for 1 h. One volume of 3 M ammonium acetate was then added to the mixtures (Wahl 1983). The mixtures were applied to the slots of the Schleicher and Schuell Minifold II system, which was assembled according to the company's instructions and connected to a vacuum pump. The nitrocellulose filter was removed from the system and baked at 80° C for 2 h. Prehybridization, hybridization and washing of the filters was carried out as described for genomic blots. The filters were then cut into individual slots and counted in a Beckman LS 7000 scintillation counter.

Results

Clone Isolation and Characterization

Colony hybridization was performed on a MIC DNA clone bank to isolate repetitive eliminated sequences. Total MIC or MAC DNA was used as a probe. Four colonies were selected for their intense reaction when probed with MIC DNA and for their minimal reaction when probed with MAC DNA. In the reconstruction experiments shown in Fig. 1, pDP5, a DNA clone containing the Tetrahymena 5S rDNA gene (Allen et al. 1984; Pederson et al. 1984), was used as a control since the 5S rDNA genes exist in approximately 150–200 copies in both MIC and MAC (Allen et al. 1984; Pederson et al. 1984).

The four DNA clones were isolated and mapped with restriction enzymes (see Fig. 2). The clones differ completely in their restriction maps and do not cross hybridize. Clone pT1008 was selected for further analysis because of its smaller size (2.2 kb) and frequency of restriction sites which allowed it to be subcloned, and because genomic blots showed that it was present in few copies in several places in the MAC genome.

pT1008 was subcloned into segments A, B, and C as shown in Fig. 3 (respective approximate sizes in bp: 1310, 560 and 310). Genomic blots indicated that all three subclones were repetitive in the MIC and that subclone A hybridized to several sequences in the MAC. Five separate regions of subclone A (A1 to A5) were further cloned as shown in Fig. 3 (respective approximate sizes in bp: 380, 410, 330, 110, and 80). These seven subclones were used in all subsequent analysis. They do not cross hybridize under the stringency used in the hybridizations shown in Figs. 4–8.

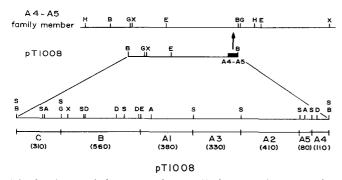


Fig. 3. The restriction map of pT1008's insert and one member of the A4-A5 family. The pT1008 map was constructed by direct analysis of the plasmid. The A4-A5 family member was mapped by measurements of genomic blots of MIC DNA digested with various enzymes which were probed with A4 or A5 (data not shown). The *arrow* indicates the region of the A4-A5 family member which hybridizes to A4 and A5. Restriction sites are designated as follows: A, AluI; B, BamHI; D, DdeI; E, EcoRI; G, Bg/II; H, HindIII; S, Sau3A; X, XbaI. The approximate sizes in base pairs of the subclones of pT1008 are shown in parentheses below the last line

Table 2. Quantitation of the copy number of sequences in MIC and MAC DNA which hybridize to the seven subclones of pT1008 using slot blots^a

DNA used as probe	Relative	e values ^b	Corrected values ^c		
	MIC	MAC	MIC	MAC	
pDP5 ^d	60	77	150	193	
Subclone					
C	128	0.75	320	1.9	
В	126	1.71	315	4.3	
A1	83	0.97	208	2.4	
A2	14	0.97	36	2.4	
A3	34	1.07	85	2.7	
A4	2.1	ND	5.3	ND	
A5	1.1	ND	2.8	ND	

ND=not detectable above background

a 2, 4, and 6 μg of MIC and MAC DNA from cell line DI, along with pT1008 diluted to single copy levels for 3 μg of genomic DNA were applied as slot blots to nitrocellulose filters. The filters were probed with the seven subclones of pT1008 and pDP5. The absolute number of cpm for each slot was determined by scintillation counting. The cpm for each slot was graphed on a linear plot and the value at 3 μg was derived for each set of MIC and MAC slots from the graph

b Relative values were obtained by comparing the estimated value for 3 µg in each set of MIC and MAC slot blots to the value for diluted plasmid which represents single copy sequences in 3 µg of genomic DNA. These relative values estimate the copy number of the sequence in the genome. These relative values correct for any variation in specific activity between probes or for differences in length between the different probes

^c The relative values of the slot blots underestimate the copy number of the 5S rDNA genes by a factor of 2.5. The copy number of the 5S genes was previously determined by Allen et al. (1984) and by Pederson et al. (1984) using other methods. Therefore, the relative values were corrected by a factor of 2.5

^d Average of two sets of MIC and MAC slot blots

Copy number of the seven subclones in MIC and MAC DNA

Slot blots (Gasser et al. 1982) like dot blots (Kafatos et al. 1979) can be used to quantitate the number of copies of a particular sequence in a DNA sample. In this case, slot blots can be used to quantitate the number of copies of each sequence in the MIC and in the MAC. However, because of nuclear cross contamination, and due to nonlinearity in the assay, these values are rough estimates of the true copy numbers of these sequences in the two nuclei. All seven subclones were used as probes against slot blots of MIC or MAC DNA. pDP5 containing the 5S rDNA gene, which is present in 150 copies in MIC DNA (Allen et al. 1984; Pederson et al. 1984) was used as a control probe. Plasmid diluted to single copy concentration was used as a method of standardization. The results are shown in Table 2.

From the data with slot blots of MIC DNA, adjacent subclones C, B and A1 appear to be present in 200–300 copies in the MIC. Adjacent subclones A3 and A2 are present in approximately 50 copies in the MIC while adjacent subclones A5 and A4 are present in three to five copies in the MIC.

From the data with slot blots of MAC DNA, subclones A5 and A4 do not hybridize to MAC DNA at detectable levels. Subclones A2 and A3 are present in the MAC in two or three copies, and subclones C, B and A1 are present in the MAC in two to four copies, much lower than their levels in the MIC.

Genomic blots of MIC and MAC DNA

MIC and MAC DNA were digested with restriction enzymes, and the fragments were separated by electrophoresis through an agarose gel and blotted to nitrocellulose filters. The filters were cut into strips and probed with the seven subclones. Figure 4 shows the results of genomic blots of MIC DNA when the DNAs were digested with HindIII. Similar results were obtained when the DNAs were digested with EcoRI (data not shown). Estimates of copy number for the different sequences are similar in magnitude to those observed in the slot blot analysis. Subclones C, B and A1 are present on hundreds of fragments in the MIC DNA, as determined by the intensity and number of bands seen in blots exposed for 1 day, or longer (a 5-hour exposure is shown in Fig. 4). A2 and A3 are present on tens of fragments (a 1-day exposure is shown in Fig. 4), and A4 and A5 are present on only a few fragments (a 4-day exposure is shown in Fig. 4; see also Fig. 8 where four distinct fragments can be seen clearly for the A4 probe. Thus, relative to each other, these families differ significantly in copy number. The high copy number seen for subclones C, B, and A1 is not due to incomplete digestion of the DNA, since (1) the experiment was repeated several times; (2) a 10-fold excess of enzyme was used to digest the DNA; and (3) a positive control was included to test for active enzyme.

In the MAC genomic blots in Fig. 5, subclone C does not hybridize to any fragments. This does not agree with the slot blot analysis which suggests that C is present in two copies in the MAC. However, subclone C has not hybridized to genomic blots of MAC DNA digested with *EcoRI*, *HindIII*, or *MboI* (data not shown). Subclones B and A1 both hybridize to one major fragment in the MAC, a number lower than the copy number predicted from the slot blots. Nevertheless, the data from slot blots and genom-

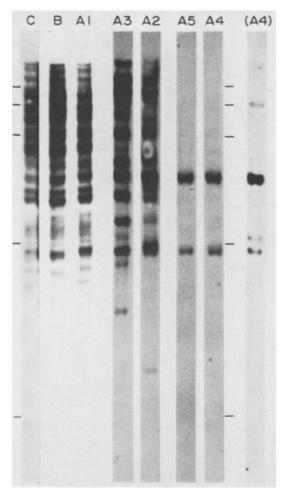


Fig. 4. Genomic blots of MIC DNA probed with the seven subclones. Each lane contains $5\,\mu g$ of MIC DNA from DI digested with HindIII. The fragments were separated by electrophoresis through 0.7% agarose gels and blotted to nitrocellulose. The nitrocellulose was cut into strips and each strip was probed with a different subclone. Dashes on the sides of the autoradiograms represent 10.0, 7.5, 5.0, 2.5 and 1.0 kb. These size estimates were derived from standard curves using λ DNA digested with AvaI and Bg/II, and ϕX 174 DNA digested with HindII. The exposure time for autoradiographs of filters probed with C, B, and A1 was 5 h; for A2 and A3, the exposure time was 1 day and for A4 and A5, the exposure time was 4 days. Other experiments with longer exposure times did show a third and fourth band (7.8 and 2.5 kb) for A4, as indicated by (A4), the genomic blot on the right, and a third band (7.8 kb) for A5 (data not shown). The four bands revealed by A4 are seen even more clearly in Fig. 8

ic blots are in agreement in suggesting that subclones C, B and A1 are drastically reduced in the MAC compared to the hundreds of copies seen in the MIC which hybridize to these fragments. Subclone A2 hybridizes to only a few fragments in the MAC compared to the 15 to 30 fragments seen in the MIC. Subclone A3 hybridizes to one fragment in the MAC which is the same size as a weakly hybridizing fragment found in the A2 pattern. The A2 and A3 copy numbers are similar though not identical to those observed in the slot blot analysis of these subclones. Neither A4 nor A5 detectably hybridized to any fragments of MAC DNA as predicted from the slot blots.

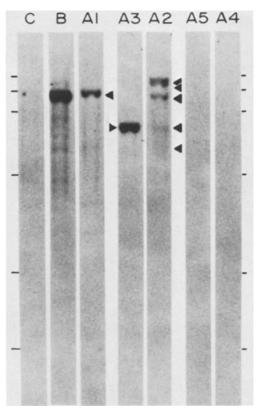


Fig. 5. Genomic blots of MAC DNA probed with the seven subclones. Each lane contains 5 µg of MAC DNA from DI cut with HindIII. The DNA fragments were separated by electrophoresis through 0.7% agarose gels and blotted to nitrocellulose. The nitrocellulose was cut into strips and each strip was probed with a different subclone. All plasmids hybridized to similar extents. The dashes on the sides represent 10.0, 7.5, 5.0, 2.5, 1.0 and 0.5 kb. These size estimates were derived from standard curves using λ DNA digested AvaI and BgIII, and ϕX 174 DNA digested with HindII. The background in subclone B's lane is due to MIC contamination. Arrowheads indicate bands observed in each strip

The major repeating unit of pT1008

MIC DNA was digested with several different enzyme combinations. Genomic blots of the DNA were probed with pT1008 to determine which parts of pT1008 define a repeating unit of the MIC. The results are shown in Fig. 6. The restriction enzymes used are enzymes that cut within pT1008. If the region of DNA between two restriction sites is part of a repeating unit, then an intense band of the size predicted from the restriction map should be present in the genomic blot. For example, if all of subclone C is part of a repeating unit, then a BamHJI/BglII double digestion should (and does) give an intense band since BamHI and BgIII sites are the endpoints of subclone C. The genomic blots in Fig. 6 show that the restriction sites BglII, XbaI and EcoRI are part of a repeating unit. The BamHI site at the C end of pT1008 is also part of the repeating unit. However, the BamHI site at the A4 end of pT1008 (labeled "B2" in Fig. 6) does not appear to be part of the repeating unit, since double digestions with BamHI and another enzyme do not give intense bands of the sizes predicted by the restriction map. These results suggest that subclones C, B, and probably A1 are part of a repeating unit. In

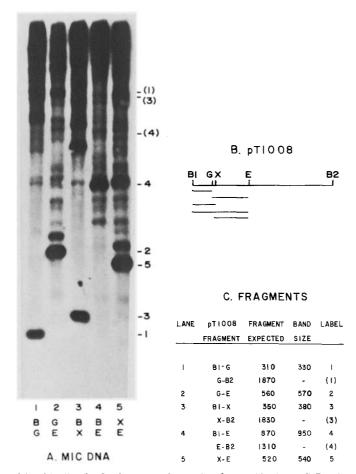


Fig. 6A-C. The Major repeating unit of pT1008. A MIC DNA from DI was digested as described, and the fragments were separated by electrophoresis through a 1.8% agarose gel and blotted to nitrocellulose. The filter was probed with pT1008. Each of the lanes (1–5) contains 5 µg of MIC DNA digested with two different restriction enzymes. The two enzymes are listed below the lane. (B, BamHI; E, EcoRI; G, BglII, X, XbaI). The numbers on the right are explained in the column designated "label" in the table shown in part C. B A restriction map of pT1008 shows the restriction sites of importance. Lines below the restriction map are DNA fragments which were detected as major bands. C A list of pT1008 fragments and their sizes. When a major band was detected in the genomic blots, that band was labeled with its lane number. The estimated size of the band is listed next to the corresponding pT1008 fragment. These size estimates were derived from standard curves using ϕX 174 DNA digested with *HindII* in one lane and λ DNA digested with AvaI and Bg/III in another lane. Numbers in parentheses correspond to the expected pT1008 fragments shown to the right of part A, which were not detectable as major bands in the genomic blot

a similar type of analysis, genomic blots of MIC DNA digested with Sau3A do not show a major band of the size of A2 or A3, indicating that neither A2 nor A3 are part of the major repeating unit of pT1008 (data not shown).

Another way of defining the C-B-A1 family is to use each subclone as a probe against genomic blots of MIC DNA. The DNA is digested with enzymes such as *Hin*dIII which do not cut within pT1008 or with enzymes such as *Eco*RI which do cut within the clone. The results are shown in Fig. 7. Subclones C, B, and A1 give nearly identical patterns when the MIC DNA is digested with *Hin*dIII. Sub-

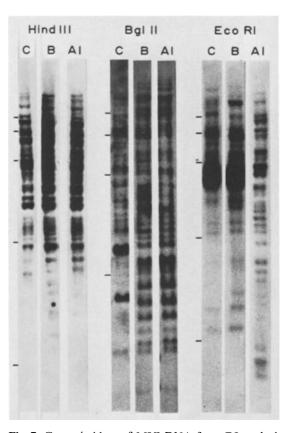


Fig. 7. Genomic blots of MIC DNA from DI probed with three of the subclones. Each lane contains 5 μ g of MIC DNA digested with the corresponding enzyme. All the lanes for one enzyme were run in the same gel and blotted together. The nitrocellulose filter was then cut into strips and probed with the indicated subclones. Dashes on the left of each set of filters represent 10, 7.5, 5, 2.5, and 1.0 kb. These size estimates were derived using λ DNA digested with AvaI and BgIII, and ϕX 174 DNA digested with HindII

clones C and B give nearly identical patterns which are different from the A1 pattern when MIC DNA is digested with EcoRI, consistent with the fact that there is an EcoRI site between A1 and B. Similarly, when MIC DNA digested with BgIII (Fig. 7) or XbaI (data not shown) was used in genomic blots and probed with the three subclones, subclone C's pattern was different from subclone B's pattern, consistent with the fact that there is a BgIII site between C and B and there is an XbaI site 50 base pairs from the BgIII site (see Fig. 3). Subclone A1's pattern was similar to subclone B's pattern in the BgIII (Fig. 7) and XbaI (data not shown) digests, as expected.

The A2-A3 and A4-A5 families

Subclones A2 and A3 can be used in a similar way to show that together they represent a different, less repetitive family. When MIC DNA digested with *HindIII* or *EcoRI* is used in genomic blots and the blots are probed with A2 and A3, the A2 and A3 patterns are similar though not identical (see Fig. 4). When MIC DNA digested with *Sau3A* is used in genomic blots, A2 and A3 give very different patterns, as expected since A2 and A3 are separated by a *Sau3A* site (data not shown).

Subclones A4 and A5 also give similar patterns when they are used as probes against genomic blots of MIC DNA

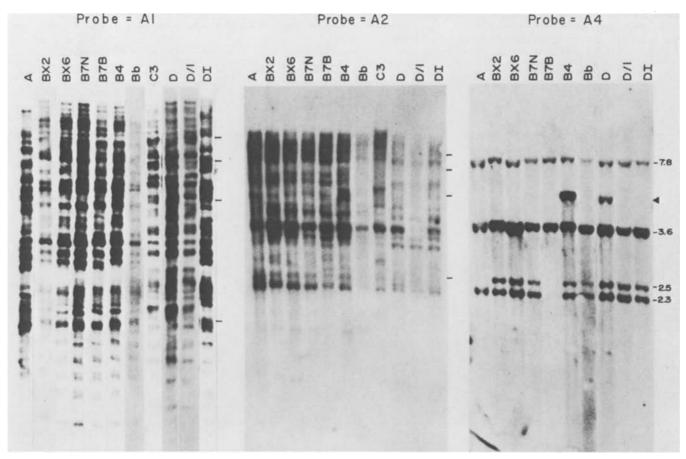


Fig. 8. Genomic blots of the MIC DNAs from various inbred strains. Each lane contains 5 μ g of MIC DNA digested with HindIII. The fragments were separated by electrophoresis through 0.7% agarose gels and blotted to nitrocellulose. Each nitrocellulose filter was probed with a different subclone (A1, A2 or A4). Dashes at the right of the blots of A1 and A2 are markers for 10.0, 7.5, 5.0 and 2.5 kb. The bands seen when the genomic blots are probed with A4 are labeled with their sizes in kb. These size estimates were derived from standard curves using λ DNA digested with AvaI and BgIII, and ϕ X 174 DNA digested with HindII. C3's MIC DNA was also probed with A4 and contains the four A4 bands (data not shown). The extra band (arrowhead) seen in B4 and D (and very faintly in B7N) probed with A4 are fragments of DNA which are also detected when genomic blots are probed with pBR322. Therefore, these MIC DNA preparations appear to be contaminated with plasmid DNA. MIC and MAC preparations with plasmid contamination

digested with *EcoRI* or *HindIII* (see Fig. 4). A4 hybridizes to four bands (7.8, 3.6, 2.5, and 2.3 kb), which are more clearly seen in Fig. 8. A5 hybridizes to three of these bands (7.8, 3.6, and 2.3 kb).

Clustering of the three families

These three families were identified from adjacent fragments of pT1008. To prove that this clustering of families on pT1008 is not an artifact of cloning, A4 and A5 were used as probes against genomic blots of MIC DNA digested with various enzyme combinations (data not shown). One member of the A4-A5 family hybridizes more intensely than the others. The regions surrounding A4-A5 on this member were mapped with restriction enzymes, and the map is shown on the top line in Fig. 3. The restriction map closely resembles the pT1008 restriction map, indicating that pT1008 was not a cloning artifact and that the three families are clustered together in at least one location in the MIC genome. The slight differences in map distances between restriction sites are probably due to the difference between sizing of fragments seen in genomic blots and sizing plasmid fragments.

Clustering of these three families may occur in several places in the genome. In genomic blots of MIC DNA digested with *HindIII* or *EcoRI* (Fig. 4), several fragments which hybridize to the A2 and A3 probes are the same sizes as some of the fragments which hybridize to the A1 probe, suggesting that members of both families may be present together. Likewise, three of the four fragments which hybridize to the A4–A5 probe are similar in size to some of the fragments which hybridize to the A2–A3 probe, suggesting that members of these two families may be present together.

Chromosomal localizations of A2, A4, and A5 family members

Nullisomic strains of Tetrahymena lack both copies of a chromosome or part of a chromosome in their MIC (Bruns et al. 1983). These strains have been used to map genetic loci as well as cloned DNA segments to particular places in the MIC genome (Bruns 1982). In this analysis, MIC DNA was prepared from each of the nullisomic strains listed in Table 1. Genomic blots were made from the MIC DNAs after digestion with *HindIII*, and the blots were

Table 3. Presence or Absence of A2, A4 or A5 sequences in nullisomic strains

Size in kb	Strain	Strain (and chromosome(s) missing)						Chromosomal		
		(-3) (-4) (-5)	(-2) (-3) (-5)	(-3)	(-1R) (-2R)	(-1R) (-2R) (-3)	(-1R) (-2L)	(-4)		location
	DI	CU 358	CU 359	CU 362	CU 371	CU 372	CU ^d 373	CU 383	BX2	
A4 and A5 Sequences ^a										
7.8	c	+	+	c	_	_		+	c	1R
3.6	+	_	c	+	+		+	+	+	e
2.5	+	_	_	subset.	+	_	+	+	+	3
2.3	+	+	_	+	_	_	-	+	+	2R
A2 sequences										
4.9	+	+	+	+	+	c	+	+	+	1L
4.4	+	_	_	- 	+	+	+	+	+	5
4.1	+	+	+	+	+	+	+	+	+	1L
3.6^{b}	+	_	+	+	+	+	+	+	+	e
3.4	+	+	_	+	+	+	c	+	+	2L
3.3	+	+	_	+	+	+	c	+	+	2L
2.9	+	+	+	+	_	+	_	+	+	1R
2.8	+	+		+	+	+	+	+	+	2L
2.4	+	_	+	+	+	+	+	+	+	e
2.3 ^b	+	+	_	+		_	_	+	+	2R
1.7	_	+	_	+	_	_	_	+	+	2R
1.2	+	+	_	+	-	_	_	+	+	2R

^a The 7.8, 3.6 and 2.3 fragments, produced by digestion of MIC DNA with *HindIII*, are common to A4 and A5

probed with A2, A4 or A5. The results are presented in Table 3.

The four A4 bands can be localized to specific regions of the chromosome. One band (2.3 kb) maps to chromosome 2, probably to the right arm, despite the ambiguity of the CU 373 data. [The chromosomal characterization of CU 373 was not completely certain according to P.J. Bruns, personal communication.] The second band (2.5 kb) maps to chromosome 3. The third band (3.6 kb) behaves oddly since it is missing from CU 358 only. This band represents the segment of MIC DNA cloned into pT1008. The fourth band (7.8 kb) is probably on the right arm of chromosome 1, although it is hard to detect in some of the genomic blots. Subclone A5's bands are the 2.3 kb, 3.6 kb and 7.8 kb bands of A4. Subclone A2 has over 20 bands in the MIC DNA. The twelve bands that were mapped appear to be distributed on chromosomes 1, 2 and 5, with six of these twelve bands on chromosome 2 (Table 4). The A2 bands not mapped were higher in molecular weight and were not separated well enough for definite chromosomal assignments.

Genomic blots and of the inbred strains

Several different strains of Tetrahymena have been inbred for many years. They are variously related as described previously (Allen et al. 1984). The three families of sequences were tested against genomic blots of MIC DNA from four of the inbred strains (A, B, C3, and D) including

Table 4. Distribution of certain A2 sequences on the five MIC chromosomes

Chromosome or chromosome arm	Sizes of fragments on that chromosome or chromosome arm	Total number		
1L	4.9, 4.1	2		
1R	2.9	1		
2L	3.4, 3.3, 2.8	3		
2R	2.3, 1.7, 1.2	3		
3		0		
4		0		
5	4.4	1		
a	3.6, 2.4	2		
	Tot	tal = 12		

^a Bands are missing in CU 358 only

six different cell clones from inbred strain B (Fig. 8). All MIC DNAs were digested with *Hin*dIII. All three families were present in the MIC of the inbred strains in approximately the same patterns. The C-B-A1 family (tested by hybridization with subclone A1) and the A2-A3 family (tested by hybridization with subclone A2) were similar in all inbred strains, although there were a few differences between strains. The A4-A5 family (tested with A4) was

b The A4/A5 sequences (3.6 and 2.3) are contained in the A2 sequences

^c Band not visible, obscured or faint

The characterization of CU 373 is not completely certain

e Band missing in CU 358 only

also present in all inbred strains. Of the four major bands of the A4 pattern, the 3.6 kb and 7.8 kb bands were present in all the inbred strains. Inbred strain A lacked the 2.5 kb band and B-18687B lacked both the 2.5 and the 2.3 kb bands.

Discussion

The analysis of seven subclones of a cloned segment of MIC DNA has identified three types of eliminated sequence. The seven subclones, designated A1 to A5, B and C, were separated into three distinct repetitive families based on their degree of repetition in MIC DNA as determined in slot blots and the similarity of their hybridization patterns when used as probes against genomic blots of MIC DNA digested with various restriction enzymes. The three families of sequences were found despite the lack of cross homology between any of the subclones.

The first type of eliminated sequence is represented by subclones C, B, and A1. It is present in 200–300 copies in the MIC genome and almost totally eliminated from the MAC. Repetitive MIC sequences with the same degree of reiteration which are completely eliminated from the MAC have been described previously (Brunk et al. 1982; Yao 1981; Yao 1982; Yao et al. 1984). This family is different because one member of the family is not eliminated and is retained in the MAC. Hybridization to this MAC sequence is not due to partial homology or to unique sequences interspersed within the subclones (unpublished observations). This sequence is an actual family member which persists in some but not all of the developing macronuclei (White and Allen 1985).

The second type of eliminated sequence is represented by subclones A2 and A3 and is present in approximately 50 copies in the MIC. Sequences with this level of repetition have not previously been reported. More interesting is the fact that several members of this family (one to five) persist in the MAC.

The third type of eliminated sequence is present in three to four copies in the MIC. Again, sequences with this level of repetition have not been previously described, although unique sequences which are eliminated from the MIC have been identified (Brunk et al. 1982). This third family of sequences is totally eliminated from the MAC.

Analysis of these three types of sequences suggests that eliminated sequences can exist at any level of repetition from unique to several hundred copies. This analysis also suggests that all members of a repetitive family may not be eliminated but that one, or a few, may be retained in the MAC. The observation of partial elimination could represent normal variation in the elimination process or it could represent infrequent mistakes in elimination during MAC development (White and Allen 1985).

These three types of eliminated sequences do not appear to be scattered randomly throughout the genome. When genomic blots of MIC DNA are digested with *Hin*dIII or *Eco*RI and probed with the subclones, the following correlations can be drawn. At least three of the four fragments which hybridize to the A4–A5 probe are of the same sizes as some of the fragments which hybridize to the A2–A3 probe. Likewise, several of the fragments which hybridize to the A2–A3 probe are the same sizes as some of the fragments which hybridize to the C-B-A1 probe. When a member of one family is of the same fragment size as a

member of another family, it may be purely by chance or it may be due to the members of the two families being clustered on the same fragment. The fact that not just some but most of the fragments of the A2–A3 family match the size of fragments in the C-B-A1 family, and the fact that the families are known to be clustered in clone pT1008, strongly argue for clustering in the genome. In the case of such clustering, some clusters such as that in pT1008 would consist of complete elements C-B-A1-A3-A2-A5-A4, whereas others would consist of incomplete elements such as C-B-A1-A3-A2 and C-B-A1, for example. Some clusters may appear to be incomplete due to the presence of extra restriction sites within those clusters.

Clustering has been observed before by Yao (1982) for two different repetitive families both of which were present in approximately the same number of copies. The clustering was determined by plaque hybridizations to λ recombinant clones containing approximately 15 kb of MIC DNA. Clustering was detected as hybridization of both repetitive families to the same plaque. The clustering observed here with the three families suggests that the three families are tightly clustered to the same restriction fragments. Moreover, this clustering involves repetitive sequences with different degrees of repetition.

Examining the positions of the three repetitive families on pT1008 allows some conclusions about the sequences which persist in the MAC. While the entire length of the C-B-A1 family is unknown, one end of the C-B-A1 family member lies within or just beyond subclone A1 in pT1008. The other end of the C-B-A1 family member must lie at the C end but outside of clone pT1008 since the Bam HI site at the end of pT1008 is included in the repeating unit (see Fig. 6). Subclone C does not hybridize to MAC sequences while subclones A1 and B hybridize to HindIII fragments of the same size in the MAC. Assuming that the C, B and A1 sequences are on the same HindIII fragment in the MIC, this implies that the sequences of the C-B-A1 family that persist in the MAC may be associated with an end of the repetitive unit, and possibly with the elimination process itself. The other two families cannot be analyzed in this manner since both A2 and A3 are associated with ends of the repetitive family and none of the sequences within the A4-A5 unit are retained in the MAC.

Families of MIC-specific sequences which appear to be completely eliminated have been found to map on several chromosomes (Karrer 1983; Yao 1982). Mapping to several chromosomes is also observed for the families we have studied that are not completely eliminated. Thus, the A2-A3 family is present on at least three of the five chromosomes, with preference for chromosome 2. Subclones containing the C-B-A1 family were not used to localize specific bands to specific chromosomes; however, members of the C-B-A1 family are probably located on more than one chromosome since several of the A1 sequences appear to be contiguous to the A2 sequences. Three members of the A4-A5 family map to three different MIC chromosomes. The remaining member which matches pT1008 is missing in CU 358 only. This odd behavior was also observed for three 5S rDNA gene clusters (Allen et al. 1984) and for another A2 family member (see Table 3). This is probably a very specific region of a chromosome which has been deleted from CU 358, yet is present in all the other nullisomics tested.

The analysis of the inbred strains indicates that these sequences are organized similarly in each strain, showing them to be quite stable. Minor strain variations are not unexpected, given the distant relationship of some of the strains and the number of times that they have been inbred (Allen et al. 1984). The A4–A5 family does show strain differences: the 2.5 kb fragment is absent from the strain A clone and the 2.5 and 2.3 kb fragments are absent from B-18687B. Since these two clones have been in culture for 16 years, it is also possible that the absence of these bands reflects deletions in the MIC chromosomes as a result of ageing occurring during asexual reproduction. This is supporting by the analysis of the 5S rDNA clusters which also show instability in their MIC sequences after prolonged culture (Allen et al. 1984).

In summary, three families of repetitive sequences have been identified that have different levels of repetition in the MIC and that are eliminated to different degrees in the MAC. These three families appear to be clustered together in several regions of the MIC genome. These clusters are also present in similar arrangements in several different strains of Tetrahymena. The clustering of eliminated sequences may be a necessary part of the elimination process, ensuring the inactivation of a MIC function or the activation of a MAC function.

Acknowledgements. This work was supported by research grants, GM-27703 from the National Institute of General Medical Sciences, U.S. Public Health Service, and PCM-8300502 from the National Sciences Foundation. T.C.W. was supported in part by fellowships from the Rackham Graduate School and from the University of Michigan Cancer Institute. We thank Almuth H. Tschunko for her comments on the manuscript. We also thank Peter J. Bruns, David Nanney and Les Jenkins for supplying several of the strains used in this study, Elizabeth Lau for help with the initial clone bank construction and colony hybridizations, Ning C. McLaren for the colony hybridization reconstructions, and Marc Orbach for developing and Mark Swancutt for perfecting the procedure for preparative magnetic columns used to select exconjugant populations.

References

- Adams J, Kinney T, Thompson S, Rubin L, Helling RB (1979) Frequency-dependent selection for plasmid containing cells of Escherichia coli. Genetics 91:627-637
- Allen SL (1967) A rapid means for inducing homozygous diploid lines in *Tetrahymena pyriformis* syngen 1. Science 155:575–577
- Allen SL, Gibson I (1972) Genome amplification and gene expression in the ciliate macronucleus. Biochem Genet 6:293–313
- Allen SL, Gibson I (1973) Genetics of Tetrahymena. In: Elliott AM (ed) Biology of Tetrahymena. Dowden. Hutchinson and Ross, Stroudsburg, pp 307–373
- Allen SL, Lee PHT (1971) The preparation of congenic strains of *Tetrahymena*. J Protozool 18:214–218
- Allen SL, White TC, Langmore JP, Swancutt MA (1983) Highly purified micro- and macronuclei from *Tetrahymena thermophila* isolated by percoll gradients. J Protozool 30:21–30
- Allen SL, Ervin PR, McLaren NC, Brand RE (1984) The 5S ribosomal RNA gene clusters in *Tetrahymena thermophila*: strain differences, chromosomal locations, and loss during micronuclear ageing. Mol Gen Genet 197:244–253
- Beerman S (1977) The diminution of heterochromatic chromosomal segments in Cyclops (Crustacea, Copepoda). Chromosoma 60:297–344
- Borst P, Cross GAM (1982) Molecular basis for trypanosome antigenic variation. Cell 29:291–303
- Brunk CF, Tsao SGS, Diamond CH, Ohasi PS, Tsao NNG, Pearlman RE (1982) Reorganization of unique and repetitive

- sequences during nuclear development in *Tetrahymena thermo-phila*. Can J Biochem 60:847-853
- Bruns PJ (1982) Tetrhymena thermophila. Genetic Maps 2:178–181 Bruns PJ, Møller KM, Leick V (1980) Magnetic purification of mating Tetrahymena. Carlsberg Res Commun 45:29–33
- Bruns PJ, Brussard TB, Merriam EV (1983) Nullisomic Tetrahymena II. A set of nullisomics define the germinal chromosomes. Genetics 104:257–270
- Callahan RC, Shalke G, Gorovsky MA (1984) Developmental rearrangements associated with a single type of expressed α-tubulin gene in Tetrahymena. Cell 36:441–446
- Denhardt D (1966) A membrane-filter technique for the detection of complementary DNA. Biochem Biophys Res Commun 23:641-652
- Gasser SC, Simonsen CC, Schilling JW, Schimke RT (1982) Expression of abbreviated mouse dihydrofolate reductase genes in cultured hamster cells. Proc Natl Acad Sci USA 79:6522–6526
- Gorovsky MA (1980) Genomic organization and reorganization in *Tetrahymena*. Annu Rev Genet 14:203–239
- Gorovsky MA, Hattman S, Pleger GL (1973) [6N]methyl adenine in the nuclear DNA of a eukaryote. *Tetrahymena pyriformis*. J Cell Biol 56:697–701
- Grunstein M, Hogness DS (1975) Colony hybridization: A method for the isolation of cloned DNA's that contain a specific gene. Proc Natl Acad Sci USA 72:3961–3965
- Gunsalus RP, Zurawski G, Yanofsky C (1979) Structural and functional analysis of cloned deoxyribonucleic acid containing the *trpR-thr* regions of the *Escherichia coli* chromosome. J Bacterial 140:106–113
- Helling RB, Goodman HM, Boyer HW (1974) Analysis of endonuclease R · Eco R1 fragments of DNA from lambdoid bacteriophage and other viruses by agarose gel electrophoresis. J Virol 14:1235–1244
- Iwamura Y, Sakai M, Mita T, Muramatsu M (1979) Unequal gene amplification and transcription in the macronucleus of *Tetrahymena pyriformis*. Biochemistry 18:5289–5294
- Kafatos FC, Jones WC, Efstratiadis A (1979) Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucl Acid Res 7:1541–1552
- Karrer KM (1983) Germ line-specific DNA sequences are present on all five micronuclear chromosomes in *Tetrahymena thermo*phila. Mol Cell Biol 3:1909–1919
- Maniatis T, Fritsch EF, Sambrock J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY pp 98–101
- Maniatis T, Jeffrey A, Kleid AG (1975) Nucleotide sequence of the rightward operator of phage λ. Proc Natl Acad Sci USA 72:1184-1188
- McDonell MW, Simon MN, Studier FW (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J Mol Biol 110:119–146
- Nasmyth KA (1982) Molecular Genetics of Yeast Mating Type. Annu Rev Genet 16:439-500
- Peacock AC, Dingham CW (1968) Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose-acrylamide composite gels. Biochemistry 7:668-674
- Pederson DS, Yao MC, Kimmel AR, Gorovsky MA (1984) Sequence organization within and flanking clusters of 5S ribosomal RNA genes in *Tetrahymena*. Nucl Acids Res 12:3003-3021
- Prescott DM, Murti KG (1974) Chromosome structure in ciliated protozoans. Cold Spring Harbor Symp Quant Biol 38:609-618
- Rigby PWJ, Dieckman M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to higher specific activity in vitro by nick translation with DNA polymerase I. J Mol Biol 113:237–251
- Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:593-517
- Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575-581

- Wahl G (1983) Rapid detection of DNA and RNA using Slot Blotting. (Application Update #371) Schleicher and Schuell Inc, Keene
- White MJD (1973) Animal cytology and evolution, 3rd ed. Cambridge University Press, Cambridge, pp 500–546
- White TC, Allen SL (1985) Macronuclear persistence of sequences normally eliminated during development in *Tetrahymena thermophila*. Dev Genet, in press
- Wilson EB (1928) The cell in development and heredity, 3rd ed. Macmillan, New York, NY pp 323-328
- Yao MC (1981) Ribosomal RNA gene amplification in Tetrahymena may be associated with chromosome breakage and DNA elimination. Cell 24:765–774
- Yao MC (1982) Elimination of specific DNA sequences from the somatic nucleus of the ciliate *Tetrahymena*. J Cell Biol 92:783–789
- Yao MC, Gorovsky MA (1974) Comparison of the sequences of

- macro- and micronuclear DNA of Tetrahymena pyriformis. Chromosoma 48:1-18
- Yao MC, Yao CH (1981) Repeated hexanucleotide C-C-C-A-A is present near free ends of macronuclear DNA of *Tetrahymena*. Proc Natl Acad Sci USA 78:7436–7439
- Yao MA, Blackburn E, Gall JG (1979) Amplification of the rRNA genes in *Tetrahymena*. Cold Spring Harbor Symp Quant Biol 43:1293–1296
- Yao MC, Choi J, Yokoyama S, Austerberry CF, Yao CH (1984) DNA elimination in Tetrahymena: A developmental process involving extensive breakage and rejoining of DNA at defined sites. Cell 36:433-440

Communicated by G.R. Smith

Received March 20 / July 2, 1985