MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 435–443
0270-7306/87/010435-09$02.00/0

Vol. 7, No. 1

Nucleotide Sequence Structure and Consistency of a Developmentally Regulated DNA Deletion in *Tetrahymena thermophila*

CHARLES F. AUSTERBERRY† AND MENG-CHAO YAO†*
Department of Biology, Washington University, St. Louis, Missouri 63130

Received 29 August 1986/Accepted 23 October 1986

DNA deletion by site-specific chromosome breakage and rejoining occurs extensively during macronuclear development in the ciliate *Tetrahymena thermophila*. We have sequenced both the micronuclear (germ line) and rearranged macronuclear (somatic) forms of one region from which 1.1 kilobases of micronuclear DNA are reproducibly deleted during macronuclear development. The deletion junctions lie within a pair of 6-base-pair direct repeats. The termini of the deleted sequence are not inverted repeats. The precision of deletion at the nucleotide level was also characterized by hybridization with a synthetic oligonucleotide matching the determined macronuclear (rejoined) junction sequence. This deletion occurs in a remarkably sequence-specific manner. However, a very minor degree of variability in the macronuclear junction sequences was detected and was shown to be inherent in the mechanism of deletion itself. These results suggest that DNA deletion during macronuclear development in *T. thermophila* may constitute a novel type of DNA recombination and that it can create sequence heterogeneity on the order of a few base pairs at rejoining junctions.

Developmentally regulated DNA rearrangement via chromosome breakage and rejoining is known to occur with important phenotypic consequences in several diverse systems, from the cyanobacterium *Anabaena* sp. (21) to the vertebrate immune system (reviewed in references 26 and 45). Among the genomes most extensively rearranged during development are the macronuclear genomes of several ciliates; within the same nucleus, chromosome breakage is followed by telomere addition in some cases and chromosome rejoining in others (reviewed in references 6, 8, and 32). Some of these rearrangement events result in the elimination of significant amounts of micronuclear (germ line) DNA sequence from the macronuclear (somatic) genome. Studies of DNA renaturation kinetics have suggested that 10 to 20% of the micronuclear genome of the holotrich *Tetrahymena thermophila* (52) and over 95% of the micronuclear genomes of the hypotrichs *Stylonychia mytilus* (4) and *Oxytricha* sp. (34) are eliminated from the macronuclear genomes of these ciliates.

Interstitial deletion (i.e., deletion-rejoining) is known to occur in *T. thermophila* (11, 51, 54) during a specific stage of macronuclear development (5, 9). Such rearrangements appear to be the primary mechanism of sequence elimination (28, 51). Molecular cloning and Southern hybridization experiments suggest that deletion occurs on average once every 30 to 40 kilobases (kb) and that some 5,000 such sites exist in the genome (51). Characterized deletion events appear to be highly regulated, occurring in over 90% of the cells in a conjugating population within a 2-h period (5). The recognition sequences and the mechanism(s) responsible for such extensive, developmentally regulated genome rearrangement, as well as the functions of these deletions, are unknown.

In this paper, we present the sequences involved in a previously characterized 1.1-kb deletion in *T. thermophila* (5, 51). The structure of the junction sequences bears surprisingly little similarity to those of known DNA rearrangement sites in other organisms, which raises the interesting possibility that DNA deletion in *T. thermophila* involves a novel mechanism of recombination. Using hybridization to an oligonucleotide probe, we have also found the deletion process in this region to be remarkably consistent at the nucleotide level, although not without a potentially significant degree of variability. The nucleotide sequence of the deleted and flanking DNA suggests that this rearrangement does not occur within or extremely near a coding region.

MATERIALS AND METHODS

DNA clones. As previously described (51), we isolated recombinant bacteriophages containing part (cTt404) or all (e.g., cTt455) of a unique *Eco*RI fragment from the micronuclear genome of *T. thermophila* B1868-IV and cloned the corresponding macronuclear sequences from the same strain in four independently isolated recombinant phages (cTt1104, cTt1107, cTt1110, and cTt1111). Each of the four macronuclear clones contains a faithful copy of the same unique region of the macronuclear genome (51). cTt1107 was selected for previous studies (5, 51) because it grows better than the others, probably because it retains the vector stuffer fragment (C. Austerberry, unpublished observations). cTt1110 was chosen for sequence determination because the absence of the vector stuffer fragment facilitates gel purification of certain restriction fragments. Restriction fragments of cTt404 and cTt1110 DNAs were prepared and subcloned in plasmid pUC13 as described previously (5).

Sequencing. Most sequencing was done by the base-specific chemical-cleavage method of Maxam and Gilbert (36). CsCl gradient-purified plasmid DNAs were digested with appropriate restriction enzymes and labeled by having the 3′ ends filled in with the large fragment of *Escherichia coli* DNA polymerase. Subsequent procedures for DNA sequencing were as described by Maxam and Gilbert (37). Ambiguities were resolved by sequencing the regions in question by the enzymatic chain termination method (42) as

* Corresponding author.
† Present address: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104.

modified for denatured plasmid templates (25). Sequencing gels 1-ft wide and 3-ft long (ca. 30 by 91 cm) were used for the chemical cleavage method, and gels 20 by 38 cm were used for the enzymatic method. Sequences were analyzed on a VAX computer (Digital Equipment Corp.) by using software written by Wayne Barnes, Biomedical Research Computing Facility, Washington University, and on a Macintosh computer (Apple Computer) by using DNA Inspector II software (Textco).

Inbred *T. thermophila* strains. Strains B1868-III, B1868-IV, B1868-V, CU399, CU401, CU409, CU427, and CU428 were generously provided by P. Bruns, Cornell University, Ithaca, N.Y. Strains A*III, A1873-I, C1564-VI, and F1668-II were generously provided by D. Nanney, University of Illinois, Urbana. Strain C3-368-V was generously provided by E. Blackburn, University of California, Berkeley. Cells were maintained as stocks and grown for mating or DNA isolation as described previously (23).

Caryonidal analysis. Mating was carried out essentially as described by Martindale et al. (35). Exconjugants and caryonides were isolated into individual drops. Caryonidal clones were used to inoculate tube stocks in 1% proteose peptone and were tested for resistance to cycloheximide and 6-methylpurine at 25 and 15 μg/ml, respectively.

Round I of genomic exclusion (2) was carried out by isolating pairs from matings of CU427 or CU428 and A*III. Round I exconjugants were subsequently isolated and established as clonal cultures. Cells derived from the same round I pair were remated in round II. Pairs, exconjugants, and caryonides of the round II matings were subsequently isolated and tested for drug resistance (cycloheximide and 6-methylpurine for CU427 and Cu428, respectively).

DNA isolation. Micronuclei and macronuclei were isolated as described previously (23). DNA was purified from isolated nuclei or whole cells of inbred lines as described previously (51). DNA from caryonides and their subclones was purified from whole cells by a rapid method. We have used this method for cultures from 1.5 to 60 ml in volume. Cells were collected by low-speed (400 to 1,100 × *g*) centrifugation and suspended in 0.05 to 0.15 culture volumes of NDS (1% [wt/vol] sodium decyl sulfate, 0.5 M disodium EDTA, 10 mM Tris hydrochloride [pH 9.5]) containing 100 μg of proteinase K per ml. The cells were lysed by incubation at 60 to 65°C for from 2 h to overnight. To the lysate was added an equal volume of 12% polyethylene glycol (average molecular weight, 8,000)–1.2 M NaCl. After being mixed, the samples were kept on ice for at least 1 h. The samples were centrifuged at 12,000 to 15,000 × *g* for 5 min or more (depending on sample volume). The pellets were washed twice with 70% ethanol, dried briefly in vacuo, and suspended in a volume of TE (10 mM Tris hydrochloride [pH 8.0], 1 mM disodium EDTA) approximately equal to the lysate volume. RNA was removed from some preparations by digestion with 100 μg of ribonuclease A per ml in 0.3 M sodium acetate–20 mM disodium EDTA–10 mM Tris hydrochloride (pH 8.0) at 37°C for 1 h. All samples were extracted with equal volumes of phenol equilibrated with sample buffer and then extracted with ether or chloroform (two or more times) to remove phenol. DNA was precipitated with 2 volumes of ethanol, rinsed with 70% ethanol, dried in vacuo, and suspended in TE.

Genomic Southern blots. Nuclear or whole-cell DNAs were digested to completion with *Hin*dIII (New England BioLabs) in buffer recommended by the supplier and electrophoresed in 1% agarose gels as described previously (50). DNA fragments were transferred by capillary blotting to nitrocellulose or Nytran nylon membrane filters (0.45 μm pore size) (both from Schleicher & Schuell) by the method of Southern (43) with slight modifications.

Hybridization probes. The 24-mer 5′-GATTTATA AACAACAATTTGAATG-3′ was made on an Applied Biosystems DNA synthesizer by the Washington University Protein Chemistry Facility and was received in lyophilized form. The oligonucleotide was desalted by chromatography in a Sephadex (Pharmacia, Inc.) G-25 column and stored frozen in TE. The oligonucleotide was labeled with ^{32}P by using T4 polynucleotide kinase (New England BioLabs) as described previously (37). The oligonucleotide was partially purified from other reactants by ethanol precipitation. Typically, ca. 33% of the input radioactivity was recovered in the product.

A pUC13 plasmid subclone containing the *Bgl*II-*Eco*RI fragment in region R of cTt1110 and a gel-purified *Xba*I-*Hin*dIII fragment of this subclone (see Fig. 1) were also used as probes. The *Xba*I-*Hin*dIII fragment was isolated from the plasmid subclone on a low-temperature agrarose gel as previously described (51). Both probes were labeled with ^{32}P by nick translation (40).

Hybridizations with J1110R. Filters were preincubated at the hybridization temperature (see figure legends) in oligonucleotide hybridization solution, which contained 6× SSPE (1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, 1 mM disodium EDTA [pH adjusted to 7.4 with NaOH]), 0.5% (wt/vol) sodium dodecyl sulfate, 0.1 M Tris hydrochloride (pH 8.0), and 2.5× Denhardt solution (15). Kinased oligonucleotide (J1110R) (10^7 dpm/ml) was added and incubated with filters for at least 24 h. Filters were washed in 2× SSPE with 0.1% (wt/vol) sodium dodecyl sulfate at room temperature and subsequently in various concentrations of SSPE at room temperature or 37°C for at least 30 min to obtain the desired stringency.

Rehybridizations with the larger region R probes and subsequent washings were carried out as previously described (51), with SSPE substituting for SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). These conditions (60 to 65°C) removed any remaining J1110R probe.

Filters were exposed to Kodak XAR-5 film at −80°C with intensifying screens.

RESULTS

Sequence comparison of micronuclear and macronuclear DNAs at a rearrangement site. We previously reported the molecular cloning and characterization of a 9.5-kb *Eco*RI fragment from the *T. thermophila* micronuclear genome containing at least three distinct micronucleus-limited sequences (51). The corresponding macronuclear *Eco*RI fragment was also cloned, and restriction maps of the two clones were compared to further localize the micronucleus-limited sequences. Both *Eco*RI fragments were divided into regions specified as L (left), M (middle), and R (right), each micronuclear region containing a micronucleus-limited sequence (5). All three micronucleus-limited sequences were shown to be deleted and degraded and the flanking sequences to be rejoined in the developing macronuclei of conjugating cells (5). We have now determined the sequence of a 1.8-kb portion of the micronuclear region R DNA known to contain a 1.1-kb micronucleus-limited sequence. The sequence of the corresponding macronuclear region R DNA was also determined. The sequencing strategy is shown in Fig. 1.

The results (Fig. 2) reveal that the micronuclear sequence

contains 1.1 kb of DNA that is absent from the macronuclear sequence, as predicted by the restriction map. No other differences were found between the micronuclear and macronuclear sequences. The complete identity between the clones from the two genomes for the sequences immediately flanking the micronucleus-limited sequence, 327 bp flanking the 5′ end and 413 bp flanking the 3′ end, reveals that the deletion event was not accompanied by other rearrangements in this region. The micronucleus-limited sequence is 1,085 bp in length. Both the micronucleus-limited and flanking DNA sequences are rather consistently A+T rich (83% A+T overall).

The junctions of the micronucleus-limited sequence lie within a 6-base-pair (bp) direct repeat (5′-TAAACA-3′; Fig. 2). The circled A nucleotides in both repeats (nucleotides 328 and 1413) define the actual boundaries of the micronucleus-limited sequence. One of these two A nucleotides is present in the macronucleus; the sequence does not reveal which one was deleted and which was retained. Nonetheless, it is clear that neither repeat was entirely deleted.

Arrows in Fig. 2 identify the three additional 5′-TAAACA-3′ sequences in the 1,835 nucleotides examined; the complementary strand contains no 5′-TAAACA-3′ sequences (data not shown). Restriction mapping results suggest that none of these additional sites are used as alternative deletion junctions (C. Austerberry, unpublished observations).

Figure 3 shows the micronuclear and macronuclear junctions and highlights some of the local sequence structure. The terminal 6-bp direct repeats in which the strand breakages occur are identified by a shaded background. We also noted that the first few base pairs immediately adjacent to the outside of the direct repeats (arrows) are similar in sequence on a given strand but opposite in 5′-3′ orientation. We shall refer to these structures as reverse repeats to distinguish them from inverted repeats (in which the same 5′-3′ sequence occurs on opposite strands and may potentially form intrastrand secondary structures).

The termini of the region R micronucleus-limited DNA are not inverted repeats of one another, nor are they related in any other apparent way (Fig. 3). Compact inverted repeats with 1-bp separations lie outside each junction in the nearby flanking retained DNA and are shown in Fig. 3 as stem structures arbitrarily displayed with the maximum number of Watson-Crick base pairs. Several additional short direct, reverse, and inverted repeat structures exist throughout the micronucleus-limited and flanking sequence (data not shown).

The direct repeat sequence $(CCCCAA)_n$ found at the ends of macronuclear DNA molecules (7, 53), and permutations of this sequence, occur in multiple copies internally in some micronucleus-limited DNA (10, 12, 49, 50, 54). Only a single hexanucleotide sequence resembling the telomeric repeat unit is contained in the region we analyzed, and it lies roughly in the center of the micronucleus-limited DNA (underlined in Fig. 2; nucleotides 1066 to 1071). On a random basis, one such hexanucleotide would be expected in every 30 kb of *Tetrahymena* DNA. Whether it has any biological significance in this particular instance is not clear.

Deletion in region R is highly sequence specific. Previous restriction mapping studies (5, 51; unpublished observations) showed that each occurrence of deletion in region R removed the same length of DNA (1.1 kb) to within the resolution of these experiments, approximately 20 to 50 bp. Knowledge of the junction sequence formed in macronuclear region R by one deletion event allowed us to begin characterizing the precision of this process at the nucleotide level

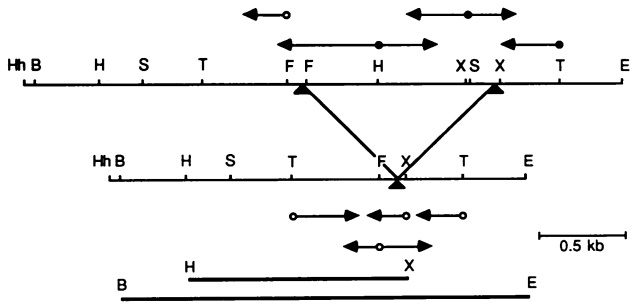

FIG. 1. Nucleotide sequencing strategy. Cloned micronuclear (cTt404; top) and macronuclear (cTt1110; bottom) DNAs from region R of strain B1868-IV were sequenced by the strategy represented by arrows over the micronuclear and under the macronuclear restriction maps. Open circles indicate that some (<10) bases most proximal to the restriction site were not determined in the particular experiment represented by the arrow; arrows with solid circles indicate that all bases, up to and including the restriction site, were determined. The triangles and connecting diagonal lines show the locations of deletion junctions. The heavy bars beneath the macronuclear map show the DNAs labeled by nick translation for use as probes of region R. Abbreviations: Hh, *Hha*I; B, *Bgl*II; H, *Hin*dIII; S, *Sau*3A; T, *Taq*I, F, *Hin*fI; X, *Xba*I; E, *Eco*RI.

as well. Of particular importance is knowledge of whether any variation in the junction sequences resulting from DNA deletion in region R is due to (i) differences among genomes prior to rearrangement or (ii) epigenetic variability inherent in the rearrangement process itself. The techniques of caryonidal analysis and genomic exclusion in *T. thermophila* provide an excellent system for distinguishing genetic from epigenetic variation (reviewed in reference 3). The macronuclei (and micronuclei) of caryonides are derived from two identical zygotic nuclei, one formed in each conjugating cell through a reciprocal exchange process. Genomic exclusion (2) allows the construction of completely homozygous zygotic nuclei in both conjugants. Because the genomes undergoing rearrangement in the second round of genomic exclusion matings are homozygous, any diversity at a particular locus in the resulting macronuclear genomes, whether within one caryonide or between caryonides, must reflect variability in the rearrangement process itself.

Hybridization experiments with oligodeoxyribonucleotide probes have been used to distinguish between sequences differing by as little as 1 bp in genomic DNAs bound to filters (reviewed in reference 29). We therefore probed blots of *Hin*dIII-digested DNAs from various strains of *T. thermophila* with a 24-mer (J1110R; Fig. 3) matching the sequence surrounding the junction in region R of phage clone cTt1110. The junction is at the center of the oligonucleotide. Any hybrids between the oligonucleotide probe and the filter-bound DNAs that contained mismatches, especially in the center of the oligonucleotide, would have significantly lower thermal stability than perfectly matched hybrids would (47). The hybridization conditions used in all experiments were of low stringency, allowing imperfect hybrids to form. Stringency was increased in steps through a series of washes at progressively lower ionic strengths, with film exposures obtained after each wash. The same filters were finally rehybridized with a larger probe from cTt1110 (Fig. 1) comprising a significant portion of macronuclear region R; the larger probe should detect all junction-containing fragments regardless of the precise sequence at the junction. We probed Southern blots of genomic DNAs from caryonides of normal matings (CU427 × CU428), genomic exclusion

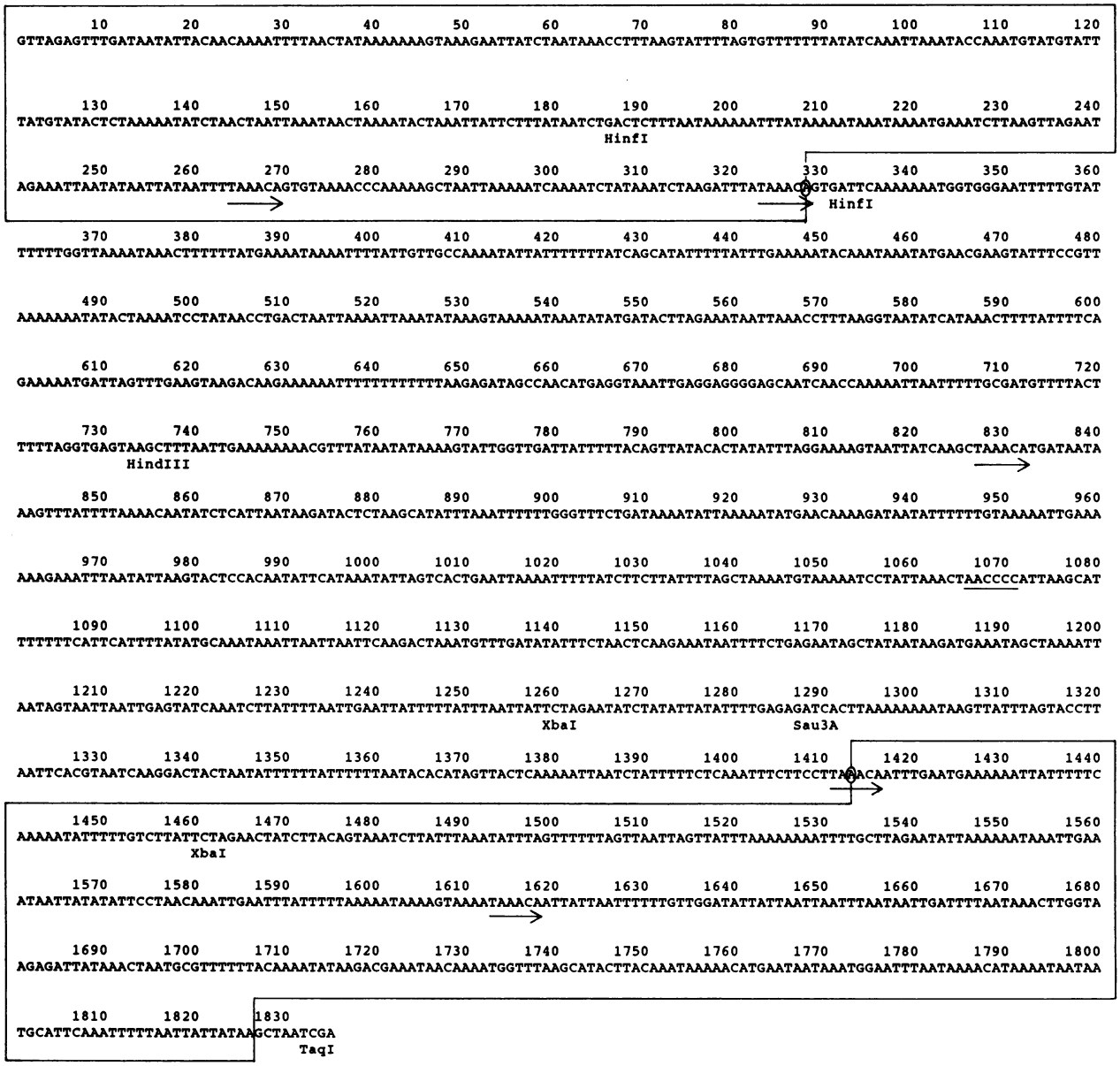

FIG. 2. Nucleotide sequences of micronuclear and macronuclear region R DNAs. The entire sequence determined in micronuclear clone cTt404 is shown. The two blocks of sequence enclosed by lines exist as one uninterrupted sequence in macronuclear clone cTt1110; in micronuclear clone cTt404, these sequence blocks are separated by 1,085 bp of micronucleus-limited sequence as shown. The deletion breakpoints are adjacent to the circled A nucleotides (see Results). The last 9 nucleotides shown are presumed to be identical in the macronucleus as well (e.g., the *Taq*I site is retained) but were only determined for cTt404. Arrows indicate the TAAACA repeats containing the junctions and the three other TAAACA sequences found. The only telomerelike sequence found, a nonrepeated hexanucleotide (AACCCC), is underlined (see Results).

matings (Cu427 or Cu428 × A*III), and several inbred strains by this method.

Many bands appeared after low-stringency washes of J1110R hybridizations to all samples (data not shown). Washes at 37°C in dilutions of SSPE at 0.25× (45 mM NaCl) or 0.15× (27 mM NaCl) removed all probe with the exception of one band, which remained in most lanes. Results obtained from selected strains are shown in Fig. 4. The remaining band always corresponded in size (2.8 kb) to the macronuclear *Hin*dIII fragment created by deletion of micronucleus-limited DNA from region R (51; unpublished observations); this was confirmed when the same filters were later rehybridized with the larger probe and the same 2.8-kb band appeared in every lane (Fig. 4, upper panel). J1110R did not hybridize under stringent conditions to the micronuclear DNA cloned in cTt404 or to genomic micronuclear DNA from any of several strains tested (data not shown). We conclude that the J1110R probe hybridized under stringent conditions specifically to the region R junction created by deletion during macronuclear development.

We expect that the nonvariant macronuclear junction sequences (those hybridizing with J1110R under stringent conditions) differ from the cloned and sequenced junction in cTt1110 by no more than 1 or 2 nucleotides, if at all, while

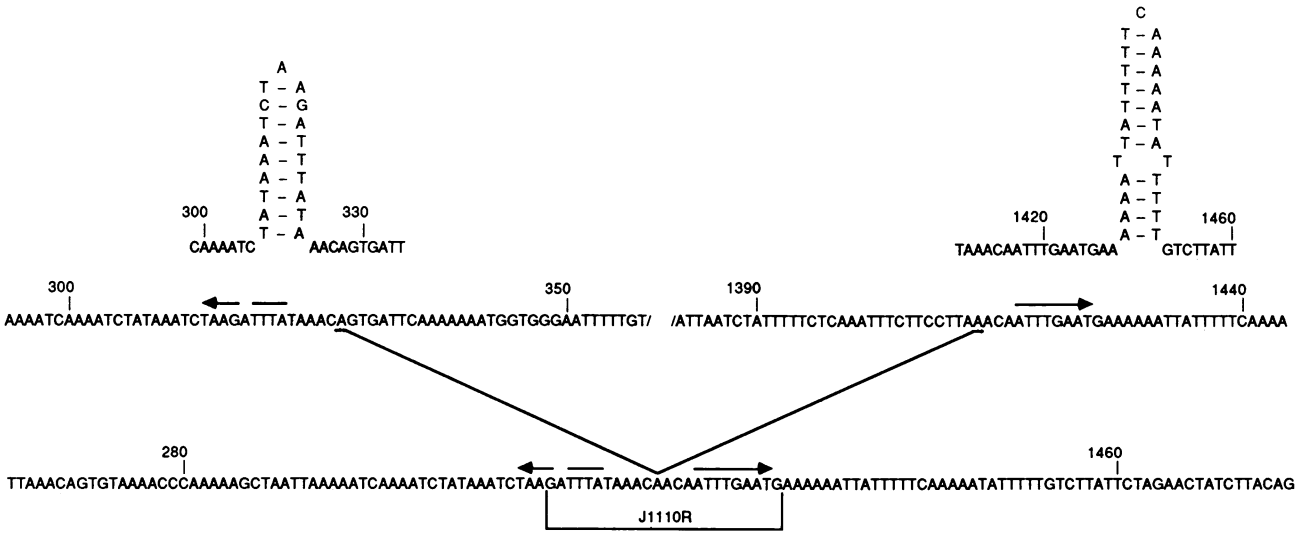

FIG. 3. Sequence structure of deletion junctions. The macronuclear sequence in cTt1110 (lower) may be derived from the micronuclear sequence in cTt404 (upper) by simple deletion of the micronucleus-limited sequence. The deleted sequence includes one of the two underlined A bases in the TAAACA direct repeats (shaded background). Inverted repeats near the junctions are arbitrarily shown as stems containing the maximum number of Watson-Crick base pairs. Reverse repeats immediately flanking the direct repeats are indicated by arrows. The sequence of J1110R, a synthetic oligonucleotide probe used to characterize the consistency of the deletion process (see text), is indicated on the macronuclear sequence.

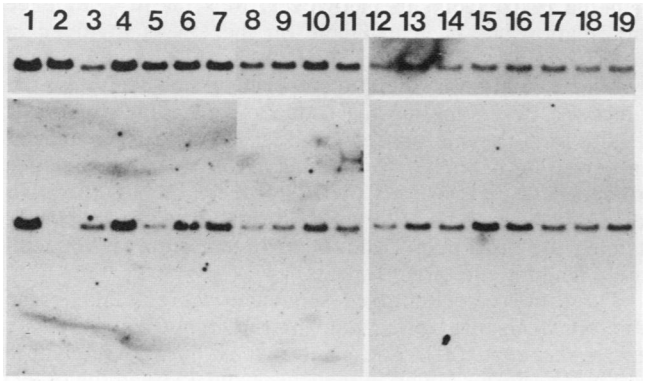

FIG. 4. Consistency of DNA deletion in region R. Macronuclear or whole-cell (>90% macronuclear) (1 to 5 μg) DNAs from inbred strains (lanes 1 to 3), caryonides from CU427 × CU428 matings (lanes 4 to 11), and caryonides from CU427 × A*III genomic exclusion matings (lanes 12 to 19) were digested with *Hin*dIII, fractionated in an agarose gel, blotted onto a nitrocellulose filter, and hybridized sequentially with two probes. The intial hybridizations (lower panel) were with ca. 10^7 dpm of kinased oligonucleotide J1110R per ml at 42°C; the filters were washed at 37°C in 0.25× SSPE (lanes 1 to 11) or 0.15× SSPE (lanes 12 to 19) and exposed for 39 h (lanes 1 to 11) or 94 h (lanes 12 to 19). The same filters were subsequently hybridized with ca. 10^6 dpm of nick-translated plasmid per ml containing a *Bgl*II-*Eco*RI fragment of cTt1110 comprising almost all of macronuclear region R (Fig. 1) under standard conditions (see Materials and Methods) and exposed for 1 to 2 days. The same 2.8-kb band previously appearing in most lanes after hybridization with J1110R (lower panel) appeared in all lanes (upper panel). Lanes: 1, B1868-IV macronuclear DNA; 2, C3-368-V macronuclear DNA; 3, CU428 whole-cell DNA; 4 to 7 and 8 to 11; whole-cell DNAs from CU427 × CU428 caryonides 1A to 1D and 2A to 2D, respectively; 12 to 15 and 16 to 19, whole-cell DNAs from CU427 × A*III round II caryonides 1A to 1D and 4A to 4D, respectively. Within each caryonidal set, caryonides A and B are sister caryonides, as are caryonides C and D.

the variant junctions could be highly mismatched. Washes of filters in 0.5× SSPE at 37°C were already of sufficient stringency to remove all probe from the variant junctions, as described below and shown in Fig. 5. Higher-stringency washes in 0.15× SSPE at 37°C, however, failed to remove J1110R from any of the nonvariant junctions (see, e.g., Fig. 4). Washes at only slightly lower ionic strength (0.10× SSPE) at 37°C removed all J1110R from the DNAs of all strains, including B1868-IV, as well as from the perfectly matched DNA in cTt1110 (data not shown). Thus, the sequences at the nonvariant junctions are indistinguishable from that in cTt1110 and are most probably identical to it. The same approach used with another oligonucleotide (20-mer) probe of similar base composition was capable of distinguishing between two filter-bound DNAs, one containing a perfect match to the oligonucleotide probe and the other known to contain a 1-bp mismatch (M.-C. Yao and C.-H. Yao, submitted for publication). We did not, however, observe lower thermal stability in hybrids between J1110R and any of the nonvariant junctions than in perfectly matched hybrids.

*Hin*dIII-digested macronuclear DNAs from 12 inbred strains were analyzed by oligonucleotide hybridization. Seven of these strains hybridized with J1110R under stringent conditions (B1868-IV, F1668-II, CU399, CU401, CU409, CU427, and CU428), and five did not (A1873-I, B1868-III, B1868-V, C1564-VI, and C3-368-V). Results for strains B1868-IV, C3-368-V, and CU428 are shown in Fig. 4 (lanes 1 to 3). We also analyzed DNAs from 40 caryonides (10 sets of four each) from the following matings: A1873-I × B1868-III (one set), CU427 × CU428 (three sets), CU427 × A*III (round II of genomic exclusion, four sets), and CU428 × A*III (round II of genomic exclusion; two sets). All 40 caryonidal DNA samples contained 2.8-kb *Hin*dIII fragments that hybridized with J1110R under stringent conditions; four sets are shown in Fig. 4 (lanes 4 to 19). All caryonides therefore contained sequences closely related or identical to the 24-bp sequence surrounding the macronu-

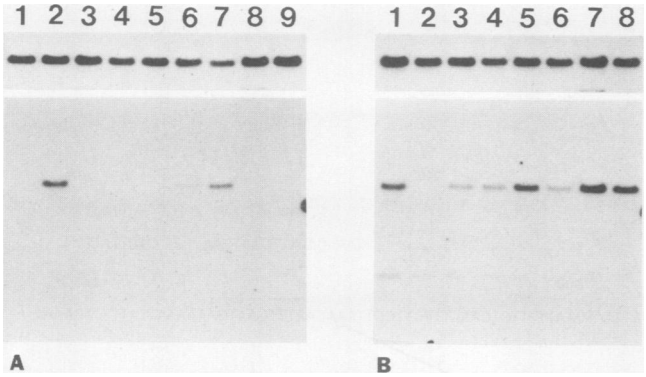

FIG. 5. Assortment of variant macronuclear region R junctions. Whole-cell DNAs (5 μg) from postassortment subclones of caryonide 1B from CU427 × CU428 matings (A) or of caryonide 1A from CU427 × A*III genomic exclusion matings (B) were digested and electrophoresed as in the legend to Fig. 4, blotted onto nylon filters, and hybridized sequentially with two probes as in the legend to Fig. 4. The initial hybridizations with 10^7 dpm of J1110R per ml (lower panel) were at 37°C; the filters were washed at 37°C in 0.5× SSPE and exposed for 70 h. A smaller fragment apparently unrelated to region R (most visible in lane 1 of panel B) is detectable in all lanes washed under these conditions; washes of higher stringency, such as 0.25× SSPE at 37°C (see, e.g., Fig. 4, and data not shown) removed this band without reducing the intensity of the 2.8-kb band containing region R DNA. The same filters were subsequently hybridized with 2.7×10^5 dpm of a nick-translated *Hin*dIII-*Xba*I fragment of macronuclear region R (Fig. 1) per ml, as described in Materials and Methods, and exposed for 45 h.

clear region R junction cloned from strain B1868-IV. Such a high frequency of occurrence suggests that the junctions identified in this study are the major ones involved in deletion events in region R and that the deletion mechanism is highly consistent at the nucleotide level.

Deletion in region R is not invariant at the nucleotide level. The identification of inbred strains whose macronuclei contain no region R junctions closely related to J1110R (e.g., C3-368-V; Fig. 4, lane 2) shows that variation in these macronuclear junctions exists. All of the caryonides from genomic exclusion matings contained nonvariant junctions. Comparisons between the lanes containing caryonidal DNAs in the lower panel of Fig. 4, however, show that the intensities of the bands are not the same in every case. These quantitative differences are not due to differences in the total amount of region R DNA present in the lanes, as shown by rehybridization of the samples with the larger probe (Fig. 4, top panel). The simplest explanation is that there is intracaryonidal variability, i.e., individual cells may contain heterogeneous populations of junction sequences, some matching the J1110R sequence better than others. Deletion in region R occurs at a stage of macronuclear development in which the genome copy number is four to eight C (5); thus at least four region R deletion events occur in each macronuclear anlagen. Caryonides showing reduced hybridization to J1110R could contain proportionately fewer region R junction sequences identical or nearly identical to J1110R.

If indeed the deletion in region R can create different junctions in a given macronucleus, the various junctions would be expected to assort during vegetative growth (reviewed in reference 8), if such junctions are viable. We tested this hypothesis with two caryonides selected on the basis of reduced hybridization of their DNAs with J1110R: caryonide 1B of the CU427 × CU428 matings (Fig. 4, lane 5) and caryonide 1A of the CU427 × A*III genomic exclusion matings (Fig. 4, lane 12). Cultures of these two caryonidal cell clones were grown vegetatively for approximately 40 generations, and several postassortment subclones were established. Whole-cell DNAs from the subclones were analyzed by hybridization (Fig. 5) with J1110R (lower panel) and a larger probe (upper panel). We found, as predicted by our hypothesis, that the original caryonides generated vegetative descendants with various proportions of junctions capable of hybridizing with J1110R under stringent conditions. DNAs from some postassortment cell clones hybridized as well as or better than the original caryonidal DNAs (Fig. 5A, lane 2 and Fig. 5B, lanes 1, 5, 7, and 8), and others hybridized less well or not at all. Hybridization with the larger probe (Figure 5, upper panel) showed that the total number of macronuclear DNA molecules containing region R was relatively constant in the postassortment subclones and hence was presumably under normal copy number control.

DISCUSSION

Restriction mapping studies of DNA in micronuclei, mature macronuclei, and developing macronuclear anlagen of *T. thermophila* identified DNA deletion events that are site specific, developmental stage specific, and highly efficient (5, 51). These observations, combined with data suggesting that thousands such deletions occur throughout the genome (51), raised questions as to the function of these events and the mechanisms responsible for such extensive, highly regulated genome rearrangements. However, thus far, no sequence information has been available on the DNAs involved in deletion in *T. thermophila*. We have now characterized the region R deletion at the nucleotide level and identified the junctions by comparing the sequences of cloned region R DNAs from the micronuclear and macronuclear genomes of inbred strain B1868-IV. Surprisingly little repeated-sequence structure is apparent, upon comparison of the two micronuclear junctions, other than small direct repeats containing the breakpoints and small reverse repeats immediately flanking the direct repeats.

Three internal eliminated sequences were found in the micronuclear form of a gene in the hypotrichous ciliate *Oxytricha nova* (31). These internal eliminated sequences are much shorter (49 bp in two cases and 32 bp in the other) than the region R micronucleus-limited sequences in *T. thermophila*. The termini of each of these internal eliminated sequences are bounded by short direct repeats of 3 or 4 bp, with one copy deleted and the other retained in the macronucleus. No reverse repeats flank the outsides of the direct repeats. The termini of these internal eliminated sequences, however, are inverted repeats. Terminal inverted repeats are also characteristic of transposable elements (reviewed in references 16, 20, and 30). Such repeats have been shown to be required for efficient transposition-independent excision of some transposons in *E. coli* (17, 19). Terminal inverted repeats, however, are not recognizable in the region R micronucleus-limited sequence.

Short (5- to 8-bp) direct repeats alone are frequent endpoints of small spontaneous deletions in *E. coli* (18), but longer (>10-bp) repeats are much more frequently involved in deletions of over 700 bp (1) and in regulated, efficient deletions such as phage λ excision (33) and *nifD* gene rearrangement in *Anabaena* spp. (21). Mechanisms relying heavily on sequence homology between deletion junctions seem unlikely in the case of the region R deletion in *T. thermophila*. The inverted repeats with 1-bp separations

found outside both micronuclear junctions are possible protein-DNA recognition sites for site-specific recombination. The extremely A+T-rich (83% A+T) sequence throughout the deleted and flanking DNA contains several similar palindromes, however, which suggests that caution be exercised in interpreting the significance of such structures.

The oligonucleotide hybridization studies provide important information on the deletion process in region R. Most significantly, they show that the junctions identified on the basis of sequence characterization of DNAs cloned from strain B1868-IV are not rare or artifactual, but rather are representative of junctions found in the macronuclear DNA of strain B1868-IV and many other strains as well, including all caryonidal strains studied here. Because deletion in region R occurs early in macronuclear development before much endoreplication of the genome has taken place (5), a limited number of homologous chromosomes (between four and eight) go through this rearrangement in each caryonide. In every one of the 40 caryonides tested, at least one of these four to eight chromosomes acquired, after deletion, a rejoining junction identical or nearly identical to the junction sequence cloned from strain B1868-IV. The caryonidal DNA exhibiting the least intense hybridization to J1110R (CU427 × CU428 1B) may have had a minority of these junctions, but the other caryonidal DNAs hybridized with severalfold-greater intensity. Most caryonidal DNAs hybridized with J1110R at a standard, maximal intensity equal to that of the seven hybridizing inbred strains. The site specificity of deletion in *T. thermophila*, at least in region R, is highly precise at the single-nucleotide level.

The uniformity and sequence specificity of deletion from region R in *T. thermophila* are greater than those observed in most known examples of DNA rearrangement in eucaryotes. Deletion in region R could not involve the extent of junctional site diversity and junctional insertion diversity characteristic of immunoglobulin gene rearrangement (reviewed in reference 45). Transposable-element excision is also quite variable. Precise excision of P elements in *Drosophila melanogaster* does occur (39) but is probably less frequent than imprecise excision. Genetic and standard Southern hybridization experiments measuring the frequency of imprecise excision of P elements produced minimum estimates of 35% in one study (46) and 75% in another (14). Events indistinguishable from precise transposon excision at the resolution of such experiments are often shown to be imprecise when sequence analysis is performed. Such is the case in the plant transposable elements studied to date (reviewed in reference 41). The oligonucleotide method we used here could also be used as a rapid and accurate way to determine the precision of large numbers of transposable element excision or other DNA rearrangements.

Variant rejoined junction sequences, i.e., those not hybridizing with J1110R, are created by deletion in region R in at least some developing macronuclei. Assortment of subclones containing only variant macronuclear junctions from caryonides containing both variant and nonvariant macronuclear junctions suggests that the inbred strains containing only variant macronuclear junctions also could have arisen by assortment. This does not exclude the possibility of germ line sequence variation at the micronuclear region R junctions among those inbred strains. Our data show, however, that both variant and nonvariant macronuclear junctions can be formed in the same macronucleus from a completely homozygous germ line through the process of deletion. The variation was not detected in restriction mapping experiments, and therefore the sequence heterogeneity must not be very extensive. Complete characterization of the variant junctions awaits seqencing of these junctions. Variation on a larger scale, detectable by restriction mapping, is characteristic of some other similar rearrangement sites in *T. thermophila* (5, 28, 48). One example is a deletion within 3 kb of the region R deletion, in region M of the same *Eco*RI fragment (5). These two classes of variability, on the order of a few versus several hundred base pairs, may be distinct in mechanism and function. If the region R deletion is representative of other deletions occurring during macronuclear development in *T. thermophila*, the potential phenotypic effects of this small-scale variation could be enormous if such deletions occur in coding regions or other functional DNA sensitive to minor sequence changes.

We do not believe that protein is encoded by any of the micronuclear or macronuclear DNA we have sequenced in region R. The data available cannot unambiguously prove this point, however; only one stop codon is known in *T. thermophila* (13, 24, 27); therefore, open reading frames are, in theory, frequent. When compared with the base composition of known coding regions in *T. thermophila* (>40% G+C; M. Gorovsky, personal communication), the base composition of the sequenced region R DNA strongly suggests that it is noncoding. The function of the micronucleus-limited DNA in *T. thermophila* is unknown; if it has a micronuclear function, this function may be related to activities of the micronuclear genome not shared by the macronuclear genome (reviewed in reference 22), e.g., mitosis and meiosis. Another intriguing possibility is that the deletion of micronucleus-limited DNA is directly required for the major function of the macronuclear genome, i.e., gene expression. The micronuclear genome appears not to be expressed during vegetative growth (38; reviewed in reference 22). DNA deletion during macronuclear development may release genes from some form of constitutive repression. We should note that an approximately 1.3-kb, poly(A)$^+$ RNA appears to be transcribed from DNA lying within 1 kb to the left (5′) end of the sequence shown in Fig. 2 (M. Altschuler and M.-C. Yao, unpublished observations). The recently developed ability to transform *T. thermophila* by microinjection of DNA (44) may allow questions of function to be investigated experimentally. It would be informative to observe the effect of DNA not normally retained in the macronucleus on the expression of nearby genes. The mechanism of deletion may also be investigated; if micronuclear DNA introduced into young macronuclear anlagen is shown to be faithfully rearranged along with the endogenous DNA, then the fate of DNA appropriately mutagenized in vitro might reveal what sequence structures are important in the deletion process.

ACKNOWLEDGMENTS

We thank M. Altschuler and L.-J. Wu for providing the caryonidal samples, E. Slattery for purifying the oligonucleotide, R. Snyder for assistance with the hybridizations, and M. Altschuler, D. Berg, E. Slattery, and M. Tondravi for advice on the manuscript. We also thank M. Gorovsky, University of Rochester, Rochester, N.Y., for communicating unpublished information.

This work was supported by a grant from the National Institute of General Medical Sciences (GM26210). M.-C.Y. is the recipient of a National Institutes of Health Research Career Development Award (HD00547). C.F.A. was supported by a National Institutes of Health Predoctoral Training Grant (GM07067) and a National Science Foundation Graduate Fellowship.

LITERATURE CITED

1. **Albertini, A. M., M. Hofer, M. P. Calos, and J. H. Miller.** 1982. On the formation of spontaneous deletion: the importance of short sequence homologies in the generation of large deletions. Cell **29**:319–328.
2. **Allen, S. L.** 1967. Cytogenetics of genomic exclusion in *Tetrahymena*. Genetics **55**:797–822.
3. **Allen, S. L., and I. Gibson.** 1973. Genetics of *Tetrahymena*, p. 307–373. *In* A. M. Elliott (ed.), Biology of *Tetrahymena*. Donden, Hutchinson, Ross, Inc., Stroudsburg, Pa.
4. **Ammermann, D., G. Steinbruck, L. von Berger, and W. Hennig.** 1974. The development of the macronucleus in the ciliated protozoan *Stylonychia mytilus*. Chromosoma **45**:401–429.
5. **Austerberry, C. F., C. D. Allis, and M.-C. Yao.** 1984. Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. Proc. Natl. Acad. Sci. USA **81**:7383–7387.
6. **Blackburn, E. H.** 1986. Structure and formation of telomeres in holotrichous ciliates. Int. Rev. Cytol. **99**:29–47.
7. **Blackburn, E. H., and J. G. Gall.** 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. J. Mol. Biol. **120**:33–53.
8. **Brunk, C. F.** 1986. Genome reorganization in *Tetrahymena*. Int. Rev. Cytol. **99**:49–83.
9. **Brunk, C. F., and R. K. Conover.** 1985. Elimination of micronuclear specific DNA sequences early in anlagen development. Mol. Cell. Biol. **5**:93–98.
10. **Brunk, C. F., S. G. S. Tsao, C. H. Diamond, P. S. Ohashi, N. N. G. Tsao, and R. E. Pearlman.** 1982. Reorganization of unique and repetitive sequences during nuclear development in *Tetrahymena thermophila*. Can. J. Biochem. **60**:847–853.
11. **Callahan, R. C., G. Shalke, and M. A. Gorovsky.** 1984. Developmental rearrangements associated with a single type of expressed alpha-tubulin gene in *Tetrahymena*. Cell **36**:441–445.
12. **Cherry, J. M., and E. H. Blackburn.** 1985. The internally located telomeric sequences in the germ-line chromosomes of *Tetrahymena* are at the ends of transposon-like elements. Cell **43**:747–758.
13. **Cupples, C. G., and R. E. Pearlman.** 1986. Isolation and characterization of the actin gene from *Tetrahymena thermophila*. Proc. Natl. Acad. Sci. USA **83**:5160–5164.
14. **Daniels, S. B., M. McCarron, C. Love, and A. Chovnick.** 1985. Dysgenesis-induced instability of rosy locus transformation in *Drosophila melanogaster*: analysis of excision events and the selective recovery of control element deletions. Genetics **109**:95–117.
15. **Denhardt, D. T.** 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641–646.
16. **Doring, H.-P., and P. Starlinger.** 1984. Barbar McClintock's controlling elements: now at the DNA level. Cell **39**:253–259.
17. **Egner, C., and D. E. Berg.** 1981. Excision of transposon Tn5 is dependent on the inverted repeats but not on the transposase function of Tn5. Proc. Natl. Acad. Sci. USA **78**:459–463.
18. **Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller.** 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **126**:847–857.
19. **Foster, T. J., V. Lundblad, S. Hanley-Way, S. M. Halling, and N. Kleckner.** 1981. Three Tn10-associated excision events: relationship to transposition and role of direct and inverted repeats. Cell **23**:215–227.
20. **Georgiev, G. P.** 1984. Mobile genetic elements in animal cells and their biological significance. Eur. J. Biochem. **145**:203–220.
21. **Golden, J. W., S. J. Robinson, and R. Haselkorn.** 1985. Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. Nature (London) **314**:419–423.
22. **Gorovsky, M. A.** 1973. Macro- and micronuclei of *Tetrahymena pyriformis*: a model system for studying the structure and function of eukaryotic nuclei. J. Protozool. **20**:19–25.
23. **Gorovsky, M. A., M.-C. Yao, J. B. Keevert, and G. L. Pleger.** 1975. Isolation of micro- and macronuclei of *Tetrahymena pyriformis*. Methods Cell Biol. **9**:311–327.
24. **Hanyu, N., Y. Kuchino, and S. Nishimura.** 1986. Dramatic events in ciliate evolution: alteration of UAA and UAG termination codons due to anticodon mutations in two *Tetrahymena* $tRNAs^{Gln}$. EMBO J. **5**:1307–1311.
25. **Hattori, M., and Y. Sakaki.** 1986. Didexoy sequencing method using denatured plasmid templates. Anal. Biochem. **152**:232–238.
26. **Hood, L., M. Kronenberg, and T. Hunkapiller.** 1985. T cell antigen receptors and the immunoglobulin supergene family. Cell **40**:225–229.
27. **Horowitz, S., and M. A. Gorovsky.** 1985. An unusual genetic code in nuclear genes of *Tetrahymena*. Proc. Natl. Acad. Sci. USA **82**:2452–2455.
28. **Howard, E. A., and E. H. Blackburn.** 1985. Reproducible and variable genomic rearrangements occur in the developing somatic nucleus of the ciliate *Tetrahymena thermophila*. Mol. Cell. Biol. **5**:2039–2050.
29. **Itakura, K., J. J. Rossi, and R. B. Wallace.** 1984. Synthesis and use of synthetic oligonucleotides. Annu. Rev. Biochem. **53**:323–356.
30. **Kleckner, N.** 1981. Transposable elements in prokaryotes. Annu. Rev. Genet. **15**:341–404.
31. **Klobutcher, L. A., C. L. Jahn, and D. M. Prescott.** 1984. Internal sequences are eliminated from genes during macronuclear development in the ciliated protozoa *Oxytricha nova*. Cell **36**:1045–1055.
32. **Kraut, H., H. J. Lipps, and D. M. Prescott.** 1986. The genome of hypotrichous ciliates. Int. Rev. Cytol. **99**:1–28.
33. **Landy, A., and W. Ross.** 1977. Viral integration and excision: structure of the lambda *att* sites. Science **197**:1147–1160.
34. **Lauth, M. R., B. B. Spear, J. Heumann, and D. M. Prescott.** 1976. DNA of ciliated protozoa: DNA sequence diminution during macronuclear development of *Oxytricha*. Cell **7**:67–74.
35. **Martindale, D. W., C. D. Allis, and P. J. Bruns.** 1982. Conjugation in *Tetrahymena thermophila*: a temporal analysis of cytological stages. Exp. Cell Res. **140**:227–236.
36. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
37. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
38. **Mayo, K. A., and E. Orias.** 1985. Lack of expression of micronuclear genes determining two different enzymatic activities in *Tetrahymena thermophila*. Differentiation **28**:217–224.
39. **O'Hare, K., and G. M. Rubin.** 1983. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34**:25–35.
40. **Rigby, P. W., J. M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
41. **Saedler, H., and P. Nevers.** 1985. Transposition in plants: a molecular model. EMBO J. **4**:585–590.
42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
44. **Tondravi, M. M., and M.-C. Yao.** 1986. Transformation of *Tetrahymena thermophila* by microinjection of ribosomal RNA genes. Proc. Natl. Acad. Sci. USA **83**:4369–4373.
45. **Tonegawa, S.** 1983. Somatic generation of antibody diversity. Nature (London) **302**:575–581.
46. **Voelker, R. A., A. L. Greenleaf, H. Gyurkovics, G. B. Wisely, S. Huang, and L. L. Searles.** 1984. Frequent imprecise excision among reversions of a P element-caused lethal mutation in *Drosophila*. Genetics **107**:279–294.
47. **Wallace, F. B., J. Shaffer, R. F. Murphy, J. Bonner, T. Hirose, and K. Itakura.** 1979. Hybridization of synthetic oligodexoyribonucleotides to ϕ χ174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. **6**:3543–3556.
48. **White, T. C., and S. L. Allen.** 1986. Alternative processing of

sequences during macronuclear development in *Tetrahymena thermophila*. J. Protozool. **33**:30–38.

49. **Yao, M.-C.** 1982. Elimination of specific DNA sequences from the somatic nucleus of the ciliate *Tetrahymena*. J. Cell Biol. **92**:783–789.
50. **Yao, M.-C., E. H. Blackburn, and J. G. Gall.** 1981. Tandemly repeated CCCCAA hexanucleotide of *Tetrahymena thermophila* ribosomal DNA is present elsewhere in the genome and may be related to the alteration of the somatic genome. J. Cell Biol. **90**:515–520.
51. **Yao, M.-C., J. Choi, S. Yokoyama, C. F. Austerberry, and C.-H. Yao.** 1984. DNA elimination in *Tetrahymena*: a developmental process involving extensive breakage and rejoining of DNA at defined sites. Cell **36**:433–440.
52. **Yao, M.-C., and M. A. Gorovsky.** 1974. Comparison of the sequences of the macro- and micronuclear DNA of *Tetrahymena pyriformis*. Chromosoma **48**:1–18.
53. **Yao, M.-C., and C.-H. Yao.** 1981. Repeated hexanucleotide CCCCAA is present near free ends of the macronuclear DNA of *Tetrahymena*. Proc. Natl. Acad. Sci. USA **78**:7436–7439.
54. **Yokoyama, R., and M.-C. Yao.** 1984. Internal micronuclear DNA regions which include sequences homologous to macronuclear telomeres are deleted during development in *Tetrahymena*. Nucleic Acids Res. **12**:6103–6116.