Proc. Natl. Acad. Sci. USA
Vol. 78, No. 5, pp. 3015–3019, May 1981
Cell Biology

All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3′ terminus

(inverted terminal repeats/chromosome processing/gene excision/DNA replication)

LAWRENCE A. KLOBUTCHER, MARSHAL T. SWANTON, PIERLUIGI DONINI*, AND DAVID M. PRESCOTT

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309

Contributed by David Marshall Prescott, January 26, 1981

ABSTRACT **In hypotrichous ciliates, all of the macronuclear DNA is in the form of low molecular weight molecules with an average size of ≈2200 base pairs. Total macronuclear DNA from four hypotrichs has been shown to have inverted terminal repeats by direct sequence analysis. In *Oxytricha nova*, *Oxytricha* sp., and *Stylonychia pustulata*, this terminal sequence may be written as**

$$5'\text{-}C_4A_4C_4A_4C_4\ldots$$
$$3'\text{-}G_4T_4G_4T_4G_4T_4G_4T_4G_4\ldots$$

In *Euplotes aediculatus*, the sequence is similar but differs in the lengths of the duplex region (28 base pairs) and of the putative 3′ extension (14 base pairs). Also in *Euplotes*, a second common sequence of 5 base pairs ($\begin{smallmatrix}T\text{-}T\text{-}G\text{-}A\text{-}A\\A\text{-}A\text{-}C\text{-}T\text{-}T\end{smallmatrix}$) occurs internal to the terminal repeat and a 17-base-pair heterogeneous region:

$$5'\text{-}C_4A_4C_4A_4C_4A_4C_4(X)_{17}T\text{-}T\text{-}G\text{-}A\text{-}A\ldots$$
$$3'\text{-}G_2T_4G_4T_4G_4T_4G_4T_4G_4T_4G_4(X)_{17}A\text{-}A\text{-}C\text{-}T\text{-}T\ldots$$

The length of the terminal repeat sequence for *O. nova* was confirmed in cloned macronuclear DNA molecules.

Ciliated protozoa of the order *Hypotrichida* have two types of nuclei—a small transcriptionally inactive (germinal) micronucleus and a large DNA-rich macronucleus, which contains the entire functional vegetative genome. Whereas the DNA of the micronucleus is large (≫150 × 10^6 daltons) and is contained in chromosomes, the transcriptionally active macronuclear genome is achromosomal; the DNA exists as a heterogeneous distribution of molecules ranging in size from 400 to ≈20,000 base pairs (bp) with an average of 2200 bp (1, 2).

The macronucleus is formed *de novo* from a diploid micronucleus after sexual conjugation. In brief, formation of the macronucleus involves the polytenization of micronuclear chromosomes, followed by fragmentation of polytene chromosomes at interbands and degradation of ≈95% of the DNA. The degradation phase eliminates all detectable repetitive DNA and most of the unique sequence DNA. What remains are small DNA molecules, which are extensively replicated to form the mature macronuclear genome (for review, see ref. 3).

This subset of the micronuclear DNA sequences that is retained in the macronucleus contains all of the genetic information (except for mitochondrial) for vegetative growth.

The macronucleus contains ≈24,000 different DNA sequences, present at an average of 1000 copies per macronucleus. Each of these gene-sized molecules is a separate functional genetic unit. Denaturation of macronuclear DNA followed by a short period of reassociation leads to formation of single-stranded circles in more than 80% of the molecules as observed by electron microscopy. The single-stranded DNA is held in a circular configuration by a short duplex region. Therefore, all or nearly all of the gene-sized molecules of the macronucleus have an inverted terminal repeat (4). The sequence of this terminal repeat was shown to be the same in many of the molecules (5), although the sequence reported for the terminal repeat was not correct. We report here the correct sequence and show that the sequence forming the inverted terminal repeat is almost the same in four different hypotrich species. To establish the structure of the termini of native DNA molecules, we determined both the 3′- and 5′-end sequences. We found that all molecules in the four species terminate with a protruding 3′ single-stranded region. The terminal repeat sequence was also confirmed in two cloned macronuclear DNA molecules of *Oxytricha nova*.†

MATERIALS AND METHODS

Culture of Hypotrichous Ciliates and DNA Isolation. *O. nova* and *Stylonychia pustulata* were cultured as described by Swanton *et al.* (6). *Euplotes aediculatus* was grown under non-sterile culture conditions with *Chlorogonium* as the principal food.

Macronuclei were isolated and DNA was purified from *O. nova*, *S. pustulata*, and *E. aediculatus* according to published procedures (6). In some cases, whole cell DNA was used for the sequence analysis of *O. nova*. This is possible because >95% of whole cell DNA in this organism is macronuclear. In addition, because the macronuclear DNA is small (≈1.7 × 10^6 daltons) and the micronuclear DNA is very large (≫1.50 × 10^6 daltons), an even greater percentage of the DNA termini available for labeling and sequence determination are macronuclear in origin. Therefore, whole cell DNA is essentially equivalent to macronuclear DNA in the kinds of studies described in this paper. Whole cell DNA was prepared by suspending cells in 4 vol of 50 mM EDTA/500 mM NaCl. The cells were lysed by the addition of $NaDodSO_4$ to 1% and incubation at 65°C for 15 min. Protein in the preparation was digested by the addition of protease K to 500 μg/ml and incubation at 45°C for 2.5 hr. The preparation was then extracted twice with water-saturated phenol containing 0.1% hydroxyquinoline and once with ether, and the aqueous phase was treated with 2 vol of cold 95%

Abbreviation: bp, base pair(s).

* On leave of absence from the International Institute of Genetics and Biophysics of the Consiglio Nazionale delle Ricerche, Naples, Italy.

† The organism used in work reported from this laboratory before 1978 was isolated from a Boulder pond and was referred to as *Oxytricha* sp. This organism died out and was replaced in work reported in 1978 and since then by an organism isolated from water obtained from North Carolina and designated *O. nova*. The original *O.* sp. and the new *O. nova* are similar in most respects but sufficiently different to suggest that they are different species.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

ethanol to precipitate the DNA. The DNA was suspended in 10 mM Tris·HCl, pH 7.9/0.1 mM EDTA (TE buffer) and digested with RNase A at 100 μg/ml and RNase T1 at 100 units/ml for 1 hr at 37°C. The DNA was isolated by repeating the extraction and precipitation procedure; it was then resuspended in TE buffer, solid CsCl was added to a refractive index of 1.400, and the mixture was centrifuged to equilibrium at 35,000 rpm for 24 hr in a Spinco 75 Ti fixed-angle rotor at 24°C. The DNA band was removed by side puncture of the tube with a syringe, dialyzed twice against TE buffer, and isolated by ethanol precipitation prior to sequence analysis.

Macronuclear DNA was also analyzed from a second strain of *Oxytricha*, previously used in this laboratory, but no longer extant as a laboratory organism. A sample of this DNA was previously isolated as described by Lauth *et al.* (7) and kindly provided by R. Boswell.

Radioactive End Labeling and Sequencing of DNAs. 5′-end labeling of the DNAs with $^{32}PO_4^{3-}$ was done according to the procedure of Maxam and Gilbert (8), as described for protruding 5′ ends. In some instances, labeling was accomplished by using the New England Nuclear 5′-end-labeling system (NEK-006). After labeling, contaminating labeled low molecular weight RNA was removed from the preparation by alkaline hydrolysis. For this, the DNA was made 0.3 M in NaOH and incubated at 37°C for 4–5 hr. The DNA was then neutralized with HCl and separated from mononucleotides by gel filtration on a Sephadex G-50 column equilibrated with 10 mM Tris·HCl, pH 7.5/10 mM NaCl/0.1 mM EDTA/0.1% sodium azide. The DNA was then ethanol precipitated, suspended in distilled water, and used for sequence analysis.

DNA was labeled at the 3′ ends with [α-^{32}P]cordycepin 5′-triphosphate by using the New England Nuclear 3′-end-labeling system (NEK-009).

DNA sequence reactions and DNA sequencing gels were essentially as described by Maxam and Gilbert (8), with minor modifications. The base-specific modification reactions for guanine, cytosine plus thymine, and cytosine were done at 23°C for 5 min. The guanine plus adenine reactions were either at 37°C for 30 min or at 23°C for 60–70 min.

The terminal nucleotide of 5′-end-labeled DNA was independently determined by digesting a portion of the labeled material to completion by incubation at 37°C for 3 hr in a solution containing snake venom phosphodiesterase at 0.4 mg/ml, DNase I at 0.2 mg/ml, 40 mM Tris·HCl (pH 8.5), 2 mM $CaCl_2$, 20 mM $MgCl_2$, and each of the four 5′-dNMPs at 0.2 mg/ml (1). The mononucleotides were then separated by chromatography on polyethyleneimine-cellulose thin-layer plates (9), and the labeled nucleotides were detected by autoradiography. For quantification, the labeled spots were cut out of the chromatographic plate and radioactivity was determined by scintillation counting.

Construction and Sequence Analysis of Macronuclear Recombinant Plasmids. Containment conditions for recombinant DNA experiments were as prescribed by the National Institutes of Health *Guidelines for Recombinant DNA Research*. The method used to construct recombinant plasmids containing macronuclear DNA of *O. nova* (pOH series) has been described (6). Briefly, the plasmid vehicle pBR322 was linearized by digestion with the restriction enzyme *Pst* I. The plasmid was then tailed with a small number of deoxyguanosine residues, and the macronuclear DNA was tailed with a similar number of deoxycytidine residues. These DNAs were mixed and allowed to reassociate and then used to transform *Escherichia coli* HB101, selecting for transformants on the basis of the plasmid-encoded tetracycline-resistance gene. This procedure reconstructs *Pst* I sites at the ends of the plasmid vehicle.

Two recombinant plasmids with macronuclear DNA segments—pOH6 and pOr2A—were chosen here for terminal sequence analysis. pOr2A (6) contains the genes coding for the 19S and 25S rRNAs. These plasmid DNAs were isolated and purified by the procedure of Clewell (10).

To analyze the terminal sequences of the macronuclear DNA inserts, the recombinant plasmids were first digested with *Pst* I, which, in these cases, produces an intact macronuclear DNA insert. This material was then 3′ end labeled and digested with a second restriction enzyme that cut at least once within the macronuclear DNA insert. The restriction fragments were separated on a 20-cm horizontal 1% agarose gel, using the buffer system of Loening (11), run at 100 V for 4 hr. The labeled end fragments of the macronuclear DNA insert were identified by autoradiography and purified from the gel by a combination of hydroxyapatite and DE52 chromatography. Alternatively, the restriction fragments were separated on and purified from a 1% low melting point agarose (Bethesda Research Laboratories, Rockville, MD) gel as described by Kuhn *et al.* (12).

RESULTS

Nucleotide Sequence at the 5′ Ends of Total Macronuclear DNA Molecules. The 5′ ends of macronuclear DNA molecules from *O. nova* and *E. aediculatus* were labeled with $^{32}PO_4^{3-}$, and a portion was digested to completion with snake venom phosphodiesterase and DNase I. The terminal nucleotide was then identified by thin-layer chromatography. For *O. nova* and *E. aediculatus*, at least 70% of the label was found in dCMP.

The 5′-end-labeled DNAs of *O. nova* and *E. aediculatus* were analyzed for terminal nucleotide sequence. The results of typical sequence-determination experiments carried out with *O. nova* are shown in Fig. 1. As expected, both macronuclear DNAs gave banding patterns corresponding to nucleotide sequences that are identical at both ends of all or almost all the molecules, followed by an uninterpretable ladder where there is no common sequence. The length and organization of the terminal sequences were closely related but not identical in the two species. The 5′ end of *O. nova* DNA terminated with the 20-nucleotide sequence 5′-$C_4A_4C_4A_4C_4$, followed internally by a heterogeneous sequence. The 5′-terminal sequence found by us in *O. nova* is identical to the terminal sequence reported by Oka *et al.* (13) for *S. pustulata*.

The sequence at the 5′ end of *E. aediculatus* was more complex and consisted of an initial stretch of 28 nucleotides with the sequence 5′-$C_4A_4C_4A_4C_4A_4C_4$, followed by a heterogeneous sequence of 17 nucleotides, followed, in turn, by a sequence of 5 nucleotides with the sequence T-T-G-A-A. Internal to this short common sequence, the sequence again became heterogeneous. In short, all or nearly all macronuclear DNA molecules in *E. aediculatus* terminate at both 5′ ends with the sequence

$$5'\text{-}C_4A_4C_4A_4C_4A_4C_4(X)_{17}\text{T-T-G-A-A}\ldots$$

Lawn (5) has reported a terminal sequence for an *Oxytricha* species, using a different sequence-determination method, that bears little resemblance to the sequence we have found in *O. nova*. We therefore reexamined the 5′-terminal sequence of this *Oxytricha* species, using part of the same DNA preparation used by Lawn. We confirmed Lawn's finding that the 5′-terminal nucleotide is dCMP, but in contrast to the sequence described by Lawn, we found the same terminal sequence as in *O. nova*—i.e., 5′-$C_4A_4C_4A_4C_4$.

Nucleotide Sequence at the 3′ Ends of Total Macronuclear DNA. Macronuclear DNA prepared from the four organisms discussed above were labeled at their 3′ ends, and their sequences were determined. *O. nova*, *O. sp.*, and *S. pustulata*

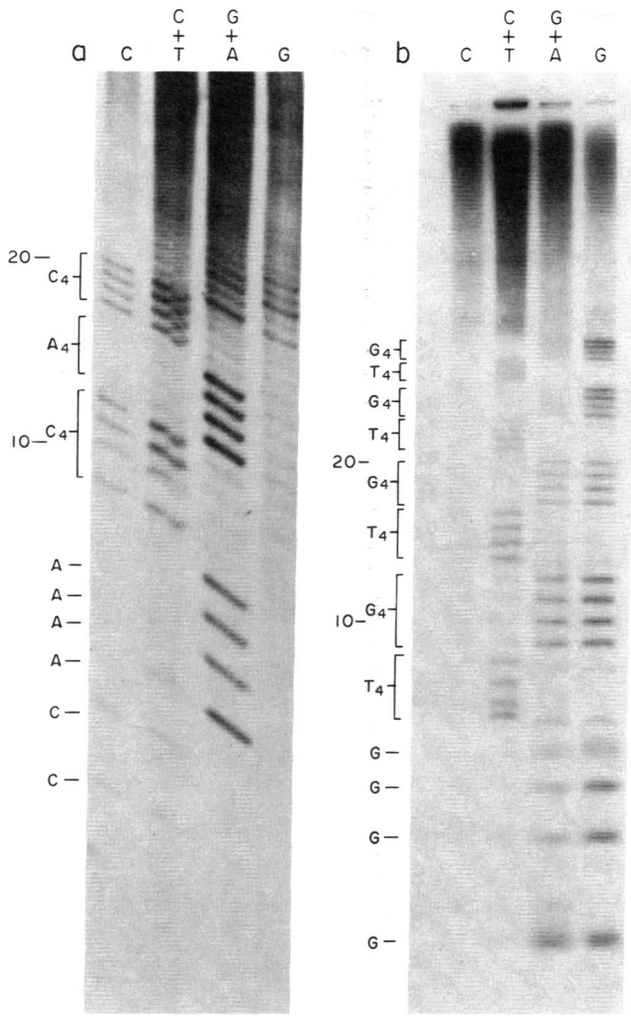


FIG. 1. Terminal sequences of *O. nova* macronuclear DNA molecules. (*a*) Twenty percent sequencing gel of 5′-labeled macronuclear DNA. The two terminal nucleotides (cytidines) are not visible on this gel. (*b*) Twenty percent sequencing gel of 3′-labeled macronuclear DNA.

all yielded the same sequence, 3′-$G_4T_4G_4T_4G_4T_4G_4T_4G_4$, followed by heterogeneous sequences (see Fig. 1). Thus, the common 3′-terminal sequence is complementary to the corresponding 5′ sequence but is 16 nucleotides longer. The sequence obtained for *E. aediculatus* reflected the corresponding 5′ sequence for that organism in that it consisted of two interpretable segments separated by 17 nucleotides of heterogeneous sequences (Fig. 2). The 3′-G_4T_4 repeat was also longer than the 5′-C_4A_4 repeat, as in the other three organisms. Thus, the sequence obtained for this organism was 3′G_2T_4-$G_4T_4G_4T_4G_4T_4G_4T_4G_4(X)_{17}$A-A-C-T-T ....

Nucleotide Sequence at 3′ Ends of Cloned Macronuclear DNA Molecules. Two recombinant DNA clones—pOr2A and pOH6—containing *O. nova* macronuclear DNA segments, were also partially sequenced. Each clone contains a native macronuclear DNA segment inserted into the *Pst* I site of the plasmid pBR322 by the guanosine-cytidine tailing method. The inserts can be excised intact from the plasmid with the restriction enzyme *Pst* I. Both of these clones appear to contain a full-size macronuclear DNA molecule. This was determined by hybridizing labeled recombinant plasmid to Southern blots (14) of native macronuclear DNA. In each case, the plasmid hybridized to a size class of native macronuclear DNA that was identical (±50 bp) to the size of the cloned macronuclear DNA insert (ref. 6; R. Boswell, personal communication).

Each of the recombinant plasmids was digested with *Pst* I, and the 3′ termini were labeled. This material was then digested with a second restriction enzyme that cut at least once within the macronuclear DNA insert, and each end of the macronuclear insert was purified and its sequence was determined. The sequencing strategy for each clone is shown in Fig. 3.

Three of the four ends of the macronuclear DNA inserts showed an essentially identical terminal sequence (Figs. 3 and 4). The 3′ DNA sequence for these ends begins with the reconstructed *Pst* I site (A-C-G-T-C); this is followed by the stretch of cytidine residues introduced during the cloning procedure and then by 36 residues of the G_4T_4 repeat—i.e., a sequence identical to that found for total macronuclear DNA of *O. nova*. Beyond this point, the sequence of each end becomes unique and specific for each molecule.

A fourth end, on the 870-bp fragment obtained by *Pst* I/*Hin*dIII digestion of pOr2A, was similarly arranged, differing only in having a G-G-G-G-G sequence (position 39–43) within the G_4T_4 repeat sequence (see Fig. 3). This could indicate some heterogeneity in the ends of macronuclear DNA molecules. Alternatively, this result could be an artifact introduced in the cloning procedure.

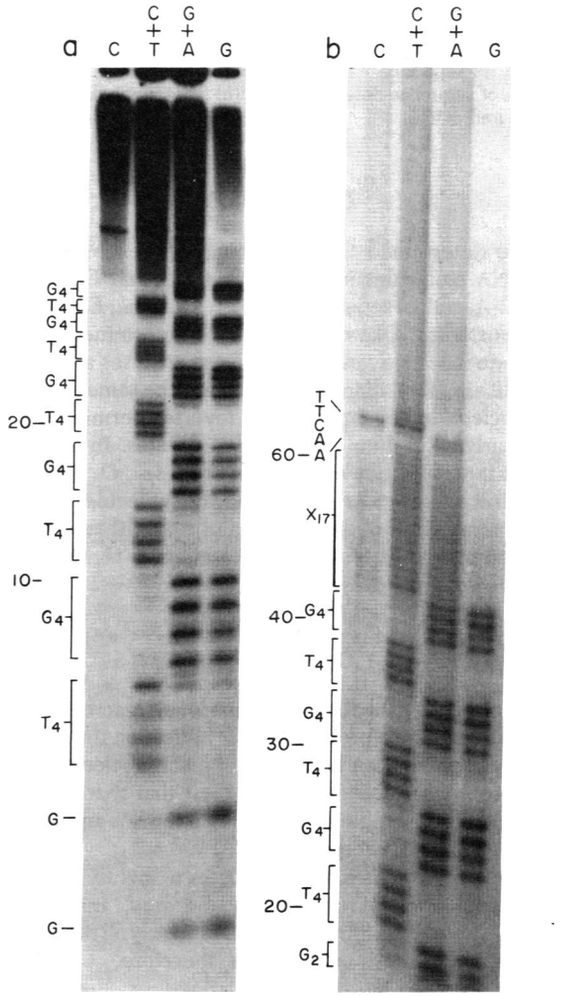
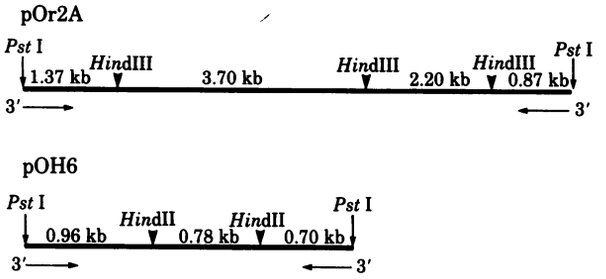


FIG. 2. 3′-Terminal sequence of *E. aediculatus* macronuclear DNA. (*a*) Twenty percent sequencing gel, showing first 42 nucleotides from 3′ end. (*b*) Eight percent sequencing gel, showing nucleotides 17–64.

a



b

pOr2A, 1370BP PstI-HindIII FRAGMENT

3' ACGTCCCCCCCCCCCCCCCCCCCCGGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGATCTTACTTA...5'

pOr2A, 870BP PstI-HindIII FRAGMENT

3' ACGTCCCCCCCCCCCCCCCCCCCGGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGTCGTTACTTCA...5'

pOH6, 960BP PstI-HindII FRAGMENT

3' ACGTCCCCCCCCCCCCCCCCCCCCCCGGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGCTCAGAATTT...5'

pOH6, 700BP PstI-HindII FRAGMENT

3' ACGTCCCCCCCCCCCCCCCCCCCCCCCGGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGCTCTAAGTCT...5'

FIG. 3. 3′ sequences of *O. nova*-cloned macronuclear DNA molecules. (*a*) Sequencing strategy of pOr2A and pOH6. Molecules were 3′-end labeled at their terminal *Pst* I restriction sites. Each molecule was then digested with the enzyme indicated and then each end was isolated and its sequence was determined. Horizontal arrows indicate directions of sequence determination. (*b*) 3′ sequences of the termini of pOr2A and pOH6.

DISCUSSION

We have determined the 5′-terminal sequence of the gene-sized DNA molecules in *O. sp.*, *O. nova*, and *E. aediculatus* and the 3′-terminal sequence in *O. sp.*, *O. nova*, *E. aediculatus*, and *S. pustulata*. Oka *et al.* (13) have determined the 5′-terminal sequence in *S. pustulata*. These analyses are consistent in establishing that all four species share the same or nearly the same nucleotide sequence in the inverted terminal repeat in all or nearly all of their gene-sized molecules. In addition, the simplest interpretation of the data for *O. sp.*, *O. nova*, *S. pustulata*, and *E. aediculatus* is that all the molecules contain a 3′ single-stranded tail of guanosines and ribosylthymines.

The arrangement of ends of molecules in *O. sp.*, *O. nova*, and *S. pustulata* is thus

$$5'\text{-}C_4A_4C_4A_4C_4\ldots G_4T_4G_4T_4G_4T_4G_4T_4G_4\text{-}3'$$
$$3'\text{-}G_4T_4G_4T_4G_4T_4G_4T_4G_4\ldots C_4A_4C_4A_4C_4\text{-}5'$$

The sequence data obtained from cloned macronuclear DNA segments of *O. nova* provide further indication that the 5′-C_4A_4 repeats are present on both ends of macronuclear DNA molecules. Moreover, these results suggest that there is little heterogeneity in the length of the terminal repeats among different macronuclear DNA sequences.

However, *E. aediculatus* possesses a slightly different terminal sequence arrangement and has, in addition, a 5′-T-T-G-A-A / 3′-A-A-C-T-T sequence internally in all or nearly all molecules. The arrangement of these molecules is, therefore,

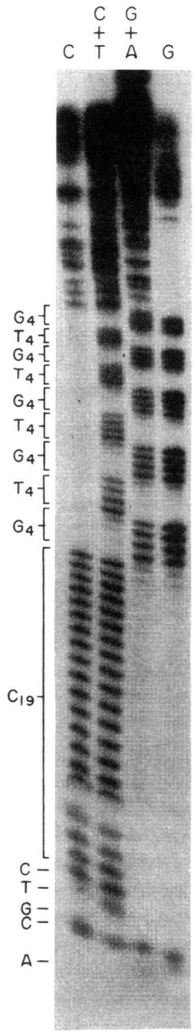


FIG. 4. 3′-Terminal sequence of the 1370-bp terminal *Pst* I/*Hin*dIII fragment of pOr2A. Twenty percent sequencing gel.

Thus, the 3′ tail in *E. aediculatus* is shorter by two guanosines than those in the other three organisms. Internal to the 3′ tail, three of the organisms share a 20-bp sequence of

$$5'\text{-}C_4A_4C_4A_4C_4$$
$$G_4T_4G_4T_4G_4.$$

In *E. aediculatus*, the repeat is continued an additional 8 bp.

A 3′ single-stranded tail is only one of several possible structures that can be derived from the data. For example, the ends of macronuclear DNA molecules could be made flush ended by shifting the base-pairing register and forming an internal G_4T_4 loop. Such a structure might be favored *in vivo* by specific protein interactions. However, for purpose of discussion, we will consider the terminal structure to be a 3′ tail.

The two primary questions are (*i*) What function does the terminal sequence serve and (*ii*) what is the significance of the 3′ single-stranded tail? A third question, with which we will not deal further, concerns the invariant sequence of 5′-T-T-G-A-A located 37 bp inward from the 5′ ends in the molecules of *E. aediculatus*.

$$5'\text{-}C_4A_4C_4A_4C_4A_4C_4(X)_{17}\text{T-T-G-A-A}\ldots\text{T-T-C-A-A}(X)_{17}G_4T_4G_4T_4G_4T_4G_4T_4G_2\text{-}3'$$
$$3'\text{-}G_2T_4G_4T_4G_4T_4G_4T_4G_4T_4G_4(X)_{17}\text{A-A-C-T-T}\ldots\text{A-A-G-T-T}(X)_{17}C_4A_4C_4A_4C_4A_4C_4\text{-}5'.$$

In considering the first question, we are naturally struck not only by the fact that all or nearly all of the different gene-sized molecules in a given species (e.g., 24,000 different molecules in *O. nova*) have the same terminal sequence but also by the observation that four different organisms share the same duplex terminal sequence with the exception that the sequence is 8 bp longer in *E. aediculatus*. This extraordinary evolutionary conservation of the sequence, both among the molecules of one species and among the macronuclear genomes of different species, suggests that stringent maintenance of this particular sequence is essential to its function. Such absence of sequence drift among the molecules of a single species, as well as among different species, could imply, for example, a function involving recognition of the sequence by a specific protein(s).

Of the various possibilities, a role of the terminal sequence in transcription seems the most unlikely. Transcription is a polar process, and the sequence is present at both ends of the molecules. The inverted arrangement of the terminal sequence is reminiscent of sequence arrangements in transposable genes (15). It is therefore conceivable that the inverted repeats allow the transposition of genes in the chromosomes in the micronucleus and serve no purpose in the macronucleus.

A related but more attractive possibility is that the terminal sequences serve as guides for the excision of genes from chromosomes during macronuclear development. In pursuing this idea it is important to determine whether 5′-C_4A_4 repeat sequences occur in micronuclear DNA, in particular at the termini of those sequences retained in the macronucleus. If this proves to be the case, it will also be important to determine whether there is a continuation of the 5′-C_4A_4 sequence in the adjacent micronuclear DNA, arranged, for example, to produce a palindrome with the macronuclear DNA terminus. The adjacent micronuclear sequences could provide major clues to the function of the macronuclear inverted terminal repeats.

Finally, with regard to the significance of the 5′-C_4A_4 sequence we note that the distantly related ciliate *Tetrahymena* contains extrachromosomal genes for rRNA that contain terminal sequences consisting of repeats of 5′-C_4A_2 (16), suggesting that the formation of extrachromosomal rDNA in this organism might occur by a mechanism related to gene excision in hypotrichs.

The second major question concerns the 3′ single-stranded tails on both ends of all molecules in the four organisms considered. At first consideration, it might be supposed that these tails are generated during excision of the genes from the chromosomes, for example, by an enzyme that makes cuts staggered by 16 bases in *S. pustulata*, *O. sp.*, and *O. nova* and by 14 bases in *E. aediculatus*. This seems an unlikely possibility because the 3′ tails are still present after the excised molecules have replicated many times during vegetative growth. This maintenance of the 3′ tails through the DNA replication process suggests instead that such tails are somehow *connected* with DNA replication. For example, 3′ tails might more appropriately be viewed as 5′ gaps. Such 5′ gaps might be generated by removal of RNA primers from the 5′ ends of newly replicated duplexes. It may be, in fact, that *in vivo* the 5′ gaps are occupied by RNA primers. Such putative RNA segments would have been destroyed by the alkali digestion used during the DNA sequencing procedure. Similarly, the 5′ gap might be occupied by a DNA fragment separated from the main section of the molecule by a nick or small gap. Such a fragment might have been lost prior to sequence determination.

In any event, the hypotrichs offer an especially favorable opportunity for studying the replication of linear DNA molecules, particularly for analysis of the so-called 5′ primer problem.

We gratefully acknowledge the help of Drs. D. Pribnow, C. Napoli, and S. Eisenberg for teaching us the sequencing techniques and of Robert Boswell and Michael Keeley for other useful techniques. We thank Gayle Prescott for the skillful typing of the manuscript. This work was supported by National Institutes of Health Postdoctoral Fellowship 1F32 GM07578 to L.A.K. and National Institute of General Medical Sciences Grant GM19199 to D.M.P. and M.T.S. The Group of Microbiology and General Pathology of the University of Rome, Rome, Italy, generously provided support for P.D.

1. Lawn, R. M., Heumann, J. M., Herrick, G. & Prescott, D. M. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 483–492.
2. Swanton, M. T., Heumann, J. M. & Prescott, D. M. (1980) *Chromosoma* **77**, 217–227.
3. Prescott, D. M., Murti, K. G. & Bostock, C. J. (1973) *Nature (London)* **242**, 597–600.
4. Wesley, R. D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 678–682.
5. Lawn, R. M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4325–4328.
6. Swanton, M. T., Greslin, A. F. & Prescott, D. M. (1980) *Chromosoma* **77**, 203–215.
7. Lauth, M. R., Spear, B. B., Heumann, J. & Prescott, D. M. (1976) *Cell* **7**, 67–74.
8. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
9. Randerath, K. & Randerath, E. (1967) *Methods Enzymol.* **12A**, 323–347.
10. Clewell, D. B. (1972) *J. Bacteriol.* **110**, 667–676.
11. Loening, J. E. (1969) *Biochem. J.* **113**, 131–138.
12. Kuhn, S., Anitz, H. J. & Starlinger, P. (1979) *Mol. Gen. Genet.* **167**, 235–241.
13. Oka, Y., Shiota, S., Nakai, S., Nishida, Y. & Okubo, S. (1980) *Gene* **10**, 301–306.
14. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
15. Calos, M. P. & Miller, J. H. (1980) *Cell* **20**, 579–595.
16. Blackburn, E. H. & Gall, J. G. (1978) *J. Mol. Biol.* **120**, 33–53.