Characterization of Chromosome Fragmentation in Two Protozoans and Identification of a Candidate Fragmentation Sequence in *Euplotes crassus*

Scott Everet Baird  
*Wright State University - Main Campus*, scott.baird@wright.edu

Lawrence A. Klobutcher

Follow this and additional works at: [https://corescholar.libraries.wright.edu/biology](https://corescholar.libraries.wright.edu/biology)

Part of the Biology Commons, Genetics Commons, Medical Sciences Commons, and the Systems Biology Commons

**Repository Citation**
[https://corescholar.libraries.wright.edu/biology/162](https://corescholar.libraries.wright.edu/biology/162)

This Article is brought to you for free and open access by the Biological Sciences at CORE Scholar. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.
Characterization of chromosome fragmentation in two protozoans and identification of a candidate fragmentation sequence in Euplotes crassus.

S E Baird and L A Klobutcher

Genes Dev. 1989 3: 585-597
Access the most recent version at doi:10.1101/gad.3.5.585

References
This article cites 39 articles, 20 of which can be accessed free at: http://genesdev.cshlp.org/content/3/5/585.full.html#ref-list-1

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.
Characterization of chromosome fragmentation in two protozoans and identification of a candidate fragmentation sequence in *Euplotes crassus*

Scott E. Baird and Lawrence A. Klobutcher

Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06032 USA

Following the sexual cycle, hypotrichous ciliated protozoans fragment a set of their micronuclear chromosomes to generate the thousands of short, linear DNA molecules present in the transcriptionally active macronucleus. We have used a hybrid selection procedure to examine macronuclear DNA molecules for subtelomeric length heterogeneity to determine whether chromosome fragmentation occurs at unique or multiple sites. The results suggest that multiple, but closely spaced, chromosome fragmentation sites are used by *Oxytricha nova*. In contrast, *Euplotes crassus* uses unique chromosome fragmentation sites in a reproducible manner to generate the ends of macronuclear DNA molecules. Additional studies compared DNA sequences in the vicinity of chromosome fragmentation sites in an attempt to define cis-acting sequences that direct the fragmentation process. A conserved sequence was found near chromosome fragmentation sites in *E. crassus*. The location of the conserved sequence suggests that chromosome fragmentation involves staggered cuts of the micronuclear DNA molecules.

[Key Words: Chromosome fragmentation; DNA rearrangement; macronuclear development; *Oxytricha nova; Euplotes crassus*]

Received January 17, 1989; revised version accepted March 14, 1989.

The transcriptionally active macronuclear genome of unicellular hypotrichous ciliated protozoa, such as *Oxytricha nova* and *Euplotes crassus*, consists of a highly amplified set of small, linear DNA molecules. For example, in *O. nova* there are ~20,000 different linear DNA molecules, with an average size of 2 kbp, each present in 1000 or more copies (Swanton et al. 1980). Each of these linear DNA molecules appears to contain a single coding unit (e.g., Helftenbein 1985), so that they are often called macronuclear genes. The macronuclear DNA molecules represent a portion of the genome of the second nucleus present in each cell, the micronucleus.

The micronuclear genome has a chromosomal arrangement but is transcriptionally inert during vegetative growth of the organism. During each sexual phase of the life cycle, the macronucleus is destroyed and a new one is generated from a mitotic copy of the micronucleus via a complex series of genome rearrangement events, including chromosome fragmentation and DNA elimination (for reviews, see Klobutcher and Prescott 1986; Steinbruck 1986). At the cytological level, the development of the new macronucleus begins with endoreplication of the micronuclear chromosomes, resulting in the formation of polytene chromosomes. The polytene chromosomes then are fragmented, and the resulting pieces are encased in vesicles. Once the vesicles form, the majority of the DNA in the developing macronucleus is degraded. Finally, the vesicle structures break down, and the remaining DNA undergoes several rounds of replication, resulting in the mature vegetative macronucleus. Comparisons of cloned segments of micronuclear and macronuclear DNA indicate that precursors of macronuclear DNA molecules reside at internal regions of the micronuclear chromosomes and generally are clustered in the genome (Klobutcher et al. 1986; Jahn et al. 1988a, b). Moreover, the micronuclear precursors lack the telomeric sequences characteristic of macronuclear DNA molecules (repeats of the octanucleotide 5'-C4A4-3'), indicating that telomeres are added to macronuclear DNA molecules following their excision from the micronuclear genome (Oka and Honjo 1983; Dawson and Herrick 1984; Klobutcher et al. 1984).

 Similar DNA rearrangement processes occur during macronuclear development in other ciliated protozoa,
such as *Tetrahymena* and *Paramecium* but on a more limited level (for review, see Blackburn and Karrer 1986). For example, only a few hundred chromosome fragmentation events occur during macronuclear development in *Tetrahymena*, resulting in macronuclear DNA molecules with an average size of ~600 kbp.

Chromosome fragmentation is one of the major events of macronuclear development in the ciliates and, as such, has been the subject of a number of studies. To a first approximation, chromosome fragmentation is a reproducible process. That is, at the level of resolution of agarose gel electrophoresis and Southern hybridization, all copies of most genes are contained on similarly sized macronuclear DNA molecules and the size does not vary through independent episodes of macronuclear development (Steinbruck 1983; Cartinhour and Herrick 1984; Roth and Prescott 1985). However, two types of observations indicate that variability in the use of chromosome fragmentation sites can occur. The first is alternative processing of the micronuclear DNA. Families of macronuclear DNA molecules have been noted in *Oxytricha fallax* (Cartinhour and Herrick 1984), *O. nova* (Klobutcher et al. 1988), and *Paramecium* (Forney and Blackburn 1988), whose members share a common core sequence, but have different lengths of DNA (differing by >1 kbp) adjacent to the telomeric repeat sequences. In cases where the micronuclear organization of such macronuclear families has been examined, the results indicate that alternative use of chromosome fragmentation sites along the length of the micronuclear chromosome is responsible for generating such families (Herrick et al. 1987; Klobutcher et al. 1988). Second, evidence for an extended chromosome ‘fragmentation domain,’ as opposed to a unique fragmentation site, has been obtained based on terminal sequence analyses of multiple independent clones of particular macronuclear DNA molecules. In *O. fallax*, macronuclear DNA molecules have been found to differ by up to 33 bp in their subtelomeric regions (Herrick et al. 1987), suggesting that multiple chromosome fragmentation sites exist within a defined region of the micronuclear chromosome. In *Paramecium*, the fragmentation domains appear to be much larger, as subtelomeric length heterogeneity of up to 800 bp has been observed (Baroin et al. 1987; Forney and Blackburn 1988).

It should be noted that models other than variable chromosome fragmentation can explain the observations discussed above. For instance, chromosome fragmentation may occur at single defined sites along the chromosome, but variable amounts of exonucleolytic digestion of the resulting ends prior to telomere addition could generate the observed terminal length heterogeneity of macronuclear DNA molecules. However, to simplify the discussion we will consider the entire series of events leading to the generation of a DNA end capable of receiving a telomere as ‘chromosome fragmentation.’

In this study we examine macronuclear DNA molecules of *O. nova* and *E. crassa* for terminal length heterogeneity to determine whether chromosome fragmentation occurs at unique sites in these organisms. As an alternative to sequencing many independent clones of macronuclear DNA termini, a hybrid selection procedure was developed to isolate and size the termini of the many copies of a particular macronuclear DNA molecule in a cell line. Multiple, closely spaced, chromosome fragmentation sites appear to be used in the processing of the two ends of a macronuclear gene precursor in *O. nova*. In contrast, six macronuclear DNA termini in *E. crassa* were found to be invariant in size, implying that chromosome fragmentation occurs at unique sites in this organism. Additional studies compared the sequences of regions of the micronuclear chromosomes subject to fragmentation in an attempt to define cis-acting DNA sequences that direct the chromosome fragmentation process in these two organisms. A conserved sequence element was identified near chromosome fragmentation sites in *E. crassa*, suggesting a model of the chromosome fragmentation event.

**Results and discussion**

**Evidence for the use of multiple chromosome fragmentation sites by *O. nova***

An initial indication that *O. nova* uses multiple chromosome fragmentation sites was obtained during the analysis of a 2.85-kb macronuclear DNA molecule referred to as C3. As has been observed for a number of macronuclear DNA molecules in this species, two versions of C3, defined by restriction site differences (Fig. 1A), exist in the macronucleus. In a previous study, micronuclear and macronuclear recombinant clones of the two versions of C3 were isolated, and the chromosome fragmentation sites used to generate one of the macronuclear C3 DNA molecules defined by DNA sequence analysis (Klobutcher et al. 1986; Fig. 1A,B). To determine whether the second version of the C3 gene used the same chromosome fragmentation sites, selected regions of macronuclear clone LMAC3-G and micronuclear clone LMIC2-5 were sequenced (Fig. 1). The results indicate that the macronuclear C3 gene version represented by clone LMAC3-G is 8 bp longer at its left end and 3 bp shorter at its right end, compared with the version of C3 represented by clone pMAC3 [Fig. 1B].

The above results indicate that the micronuclear chromosomes containing the two different C3 gene precursors were fragmented at different sites during macronuclear development. One explanation of these results is that multiple chromosome fragmentation sites are used in this organism. The use of multiple fragmentation sites is formally possible within a clonal cell line because polytenization of the micronuclear chromosomes is an early event during macronuclear development, providing many DNA strands that can undergo fragmentation. Alternatively, chromosome fragmentation may be precise for each version, the difference in the observed positions arising from the small numbers of base changes that define the two versions (Fig. 1B).

To distinguish between these alternatives, we developed a variation on the hybrid selection procedure (Parnes et al. 1981) that allows the isolation and sizing of
terminal restriction fragments from the many copies of a particular macronuclear gene that exist in the amplified macronucleus (see Materials and methods). In brief, the procedure entails the immobilization of a recombinant clone containing the macronuclear gene to be analyzed on a small nitrocellulose filter. Total macronuclear DNA is treated so as to generate a small terminal restriction fragment radioactively labeled at one end for the macronuclear DNA molecule under analysis. The labeled total macronuclear DNA is then hybridized to the cloned DNA on the filter, and DNA fragments that have formed specific hybrids with the immobilized cloned DNA are eluted and analyzed on acrylamide sequencing gels using sequenced DNA as size markers. This procedure allows us to detect differences in terminal restriction fragment length of as little as 1 bp. In most experiments, the hybrid selection procedure was capable of detecting a subclass of terminal fragments if they composed at least 10% of the population.

To examine terminal length heterogeneity of the two versions of C3 by hybrid selection, HindIII-digested total macronuclear DNA was labeled by the Escherichia coli DNA polymerase I fill-in reaction (Maniatis et al. 1982) and hybridized to a nitrocellulose filter containing pMAC3 DNA. This strategy takes advantage of variable HindIII sites in the two versions of C3 (Fig. 1), so as to select distinctly sized terminal fragments from each end of each version. With the exception of the large left terminal HindIII fragment derived from the LMAC3-G-like versions of C3, the terminal fragments are sufficiently small so as to be precisely sized on DNA sequencing gels.

Hybrid-selected DNA fragments derived from the right end of LMAC3-G-like versions of C3 were first sized by examining DNA fragments in the 148-bp size range (Fig. 2A). A 148-bp fragment was observed, which was expected based on the sequence of clone LMAC3-G (Fig. 1A), along with fragments of 145, 143, and 140 bp. The additional three fragments represent successively shorter terminal fragments derived from other LMAC3-G-like versions of C3. The intense band of ~230 bp (Fig. 2A) was found to be a doublet upon shorter autoradiographic exposure (data not shown) and corresponds to the two strands of the internal HindIII fragment of LMAC3-G-like versions of C3 (Fig. 1A). The fact that only two bands are observed in this region of the gel indicated that nonspecific exonuclease degradation is not occurring during the hybrid selection procedure and is not the explanation for the observed terminal fragment length heterogeneity.

In a similar manner, it was possible to assess terminal length heterogeneity of the two ends of pMAC3-like versions of C3 macronuclear DNA molecules. For the left ends, fragments of 325, 326, 329, 333, and 334 bp were observed (Fig. 2B). A fragment of 325 bp is predicted from the sequence of clone pMAC3. For the right end,
Figure 2. Autoradiographs of sequencing gels containing hybrid-selected DNA representative of the termini of O. nova C3 macronuclear DNA molecules. C3 terminal HindIII fragments from total macronuclear DNA digested with HindIII and 3' end-labeled, were hybrid selected using macronuclear clone pMAC3. (A) Hybrid-selected DNA fragments (lane HS) derived from the right termini of LMAC3-G-like copies of C3 are seen at 140, 143, 145, and 148 bp. The band at ~230 bp represents an internal HindIII fragment. Sequence standards [lanes G + A and C + T] are base-specific, chemical-cleavage reactions of a 2.1-kbp HindIII-Smal fragment of clone LMAC3-G, which represents the region of LMAC3-G beginning at the right-most HindIII site of the insert (Fig. 1A) and extending through the right terminus into vector DNA, labeled at its 3' end. (B) Hybrid-selected DNA fragments derived from the right and left termini of pMAC3-like versions of C3. Fragments seen in the 325–334-bp size range are derived from left ends of pMAC3-like copies of C3, whereas the fragments at 375 and 378 bp represent two distinct right termini. Size standards [lanes G + A and C + T] consist of a 0.5-kbp HindIII-PvuII fragment of pMAC3, representing the left terminus of the insert, 3'-end-labeled at the HindIII site and sequenced by the chemical cleavage method.

The expected fragment of 378 bp was observed along with a 375 bp fragment (Fig 2B).

These results indicate that terminal length heterogeneity exists for the three version-specific ends of macronuclear C3 genes examined. We suggest that this is the result of variability in the use of chromosome fragmentation sites. An alternative explanation might be that the telomeric sequences can vary in length. We do not consider this a likely explanation of the results for the following two reasons: (1) Sequencing of O. nova total macronuclear DNA is consistent with a telomeric structure of invariant length [Klobutcher et al. 1981]; (2) 19 cloned O. nova macronuclear termini have been sequenced [e.g., Klobutcher et al. 1981, 1986, 1988, and unpubl., Boswell et al. 1982, Greslin et al. 1988, Harper and Jahn 1989]. All but three, which may be cloning artifacts, possessed the correct telomere length, implying that telomeres of abnormal length are infrequent. A second alternative explanation of our results, which we cannot rigorously exclude, is that telomeric heterogeneity resulted from deletion, insertion, or recombination events during vegetative growth of the organism.

Assuming that our hybrid selection results exclusively reflect use of variable chromosome fragmentation sites, the positions of the sites that would be used to generate the ends of the two C3 versions are indicated in Figure 1B. It is evident that some of the fragmentation sites are version specific, whereas others are used for both versions. The results obtained for the C3 gene are similar to those obtained for O. fallax and Paramecium, where all termini examined to date indicate the use of multiple fragmentation sites [Baroin et al. 1987; Herrick et al. 1987; Forney and Blackburn 1988]. The primary difference of the C3 gene results is that the fragmentation sites are closely clustered, separated by a maximum of 9 bp. Fragmentation domains are somewhat larger in O. fallax, where DNA termini differing by 33 bp have been observed, and much larger in Paramecium, where fragmentation appears to occur over intervals hundreds of base pairs in length. Our results also extend these studies in that they indicate that fragmentation occurs only at a subset of the positions present within the fragmentation domain.

E. crassus uses unique chromosome fragmentation sites in a reproducible manner

More extensive studies on the fidelity of the chromosome fragmentation process were carried out on the hypotrichous ciliate E. crassus. Partially inbred strains of this organism have been generated that greatly reduce the version problem and simplify the analysis. In addition, it is possible to obtain viable offspring from this organism, so that the question of reproducibility could be addressed.

Five macronuclear DNA molecules, for which we have isolated and characterized macronuclear and micronuclear clones [S.E. Baird et al., in prep.], were examined using the hybrid selection procedure. The five macronuclear DNA molecules are referred to as V1, V2, V3, D7, and D8 and have sizes of 3.85, 0.94, 1.80, 1.65, and 1.55 kbp, respectively [Fig. 3]. The precursors of D7 and D8 are adjacent in the micronuclear genome, separated by a 36-bp intergenic spacer. Similarly, the precursors of V1, V2, and V3 are clustered in an independent region of
Chromosome fragmentation in two protozoans

Figure 3. Restriction maps of recombinant clones of *E. crassus* macronuclear DNA molecules and their micronuclear precursors. (A) Restriction maps of the inserts of clones LEMACD7 and LEMACD-8, which contain copies of the D7 and D8 macronuclear DNA molecules, respectively. Also shown is the region of clone LEMICD that contains the micronuclear precursors of these two macronuclear DNA molecules. In the micronuclear clone map, regions giving rise to macronuclear DNA molecules are shown as boxes, and spacers separating adjacent precursors as single lines. Internal eliminated sequences, which are removed from precursors during macronuclear development, are indicated by black ovals in cases where their precise position is unknown, and by black boxes in cases where they have been localized by sequence analysis. (*) The position of the variable HaeIII restriction site that defines the two forms of D8 macronuclear DNA molecules. Predicted sizes of terminal restriction fragments isolated in hybrid selection experiments are given below the maps in base pairs. (A) AluI, (B) BglII, (E) EcoRI, (H) HindIII, (Ha) HaeIII, (Hh) HhaI, (P) PstI, (Pv) PvuII, (Ss) SstI, (X) XbaI, and (Xm) XmnI. Not all AluI, HaeIII, HhaI, PvuII, PstI, and XmnI sites are shown. Note that the EcoRI sites in parentheses at the ends of the macronuclear clone inserts result from the addition of linkers during the cloning process and are not present in the native macronuclear DNA molecules. (B) Restriction maps of the inserts of clones pMACV1-4, pMACV2, and pMACV3, which contain copies of the V1, V2, and V3 macronuclear DNA molecules, respectively. Also shown is the restriction map of the region of micronuclear clone LEMICV that contains the precursors of the three macronuclear DNA molecules. Other aspects of the figure are as described above.

the micronuclear genome. The precursors of V2 and V3 are separated by a 400-bp intergenic spacer, whereas the right end of V1 overlaps the left end of V2 by 6 bp. Genomic restriction mapping experiments indicate that only single versions of V1, V2, V3, and D7 exist in the macronucleus and micronucleus of *E. crassus* strain G1. Two forms or versions of D8 macronuclear DNA molecules exist in strain G1, which differ in the presence and absence of a HaeIII restriction endonuclease recognition site [Fig. 3A], but genetic studies have demonstrated that the two forms are allelic in this instance [Baird and Klo butcher 1988].

Prior to carrying out hybrid selection analyses, it was necessary to characterize the telomeric structure of *E. crassus* macronuclear DNA molecules. Based on sequencing the termini of cloned macronuclear DNA molecules, which were blunt-ended during the cloning process, the telomeres contain a 28-bp double-stranded region of C4A4 repeats. The existence of a 14-bp 3'-single-stranded tail was deduced from sequencing total macronuclear DNA labeled at 3' ends [data not shown], indicating the following telomeric structure:

\[
5' - C_4T_4C_4A_4T_4G_4T_4G_4A_4C_4N_7\text{ttgaa}.. \\
3' - G_2T_4G_4T_4G_4T_4G_4T_4G_4T_4G_4N_7\text{sacctt}..
\]

In the sequencing of total DNA, the sequence 5'-TTGAA-3' was also readable above the general background at a position 17 bp from the telomeric repeats [discussed further below]. The deduced telomeric structure is identical to that obtained previously for the related hypotrich *Euplotes aediculatus* [Klobutcher et al. 1981].

The left end of the V3 gene was first examined for length heterogeneity by digesting *E. crassus* G1 whole-cell DNA with HindIII and labeling the resulting fragments at their 3' ends using the *E. coli* DNA polymerase I fill-in procedure. Labeled V3 gene fragments then were selected, using the macronuclear clone pMACV3. The left terminal HindIII fragment of V3 was predicted to be 269 bp in length based on the sequence of clone pMACV3, and a single fragment of this size was observed in the hybrid selection analysis [Fig. 3A]. The analysis was then repeated on *E. crassus* strain NeC10, and three cell lines generated from a mating of
Figure 4. Autoradiographs from hybrid selection experiments on the *E. crassus* V3, D8, and D7 macronuclear DNA molecules. (A) Hybrid selection of left terminus–*HindIII* fragments of V3 from cell lines G1, NcC10, CC51, CC54, and CC57. In each case, terminal fragments were hybrid-selected from *HindIII*-digested total cellular DNA labeled at its 3' ends using clone pMACV3. Sizes of hybrid-selected fragments are indicated in base pairs. Size standards (lanes G + A and C + T) consist of a 3'-end-labeled 0.28-kbp *HindIII*–*BamHI* (vector site) fragment of clone pMACV3, which is essentially equivalent to the hybrid-selected fragments, sequenced by the chemical-cleavage method. (B) Hybrid selection of the left terminus–*BglII* fragments of D8 macronuclear DNA molecules. Terminal fragments were hybrid-selected from total cellular DNA, 3' end labeled at *BglII* sites, using a plasmid subclone containing the 1.15-kbp *EcoRI* fragment of LEMACD8-8. Size markers (lanes A,T) consist of the 0.20-kbp *BglII*–left terminus fragment of LMACD8-8 cloned in M13mpl1 and sequenced by the dideoxy chain-termination procedure. (C) Hybrid selection of the right terminus–*EcoRI* fragments of D7 macronuclear DNA molecules. Terminal fragments were hybrid selected from total cellular DNA, 3' end labeled at *EcoRI* sites, using a plasmid subclone containing the 0.25-kbp *EcoRI* fragment of LEMACD7. Size standards (lanes A,T) consist of an M13mpl1 clone containing the 0.25-kbp *EcoRI* fragment of LEMACD7 sequenced by the dideoxy chain-termination procedure.

G1 × NcC10, (lines CC51, CC54, and CC57). In each case, a single fragment of 269 bp was observed (Fig. 4A), indicating that a single fragmentation site is used to generate the left end of the macronuclear V3 gene and that this site is reproducibly used in subsequent rounds of macronuclear development.

Similar analyses were performed on the left end of D8 and the right end of D7. For D8, the left terminal *BglII* fragment, which is predicted to be 195 bp in length, was hybrid-selected from the same five cell lines used above. In each case, a single fragment of 195 bp was detected (Fig. 4B). For D7, the right terminal *EcoRI* fragment with a predicted size of 240 bp was hybrid-selected from G1, NcC10, CC54, and CC57 total cellular DNA. Hybrid selection resulted in a single fragment with a size of 239 bp for each of the four cell lines (Fig. 4C), which is 1 bp smaller than expected. This is likely to be the result of slight gel mobility differences between the hybrid-selected fragments and the size markers, which can occur as the result of differences in base composition (Landick et al. 1984). The markers used in this case were an M13 clone of the 240-bp right terminal *EcoRI* fragment of D8 sequenced by the dideoxy chain-termination method. Although the majority of the sequence standards should have the identical base composition of the hybrid-selected DNA, some M13 vector sequences are also contained on the standards that could affect mobility. Overall, however, the results indicate that single fragmentation sites are used to generate both the right end of D7 and the left end of D8.

Additional analyses were performed on the right end of V1 and the left end of V2. For each of these ends, an alternative labeling procedure was required in the hybrid selection procedure, as there were no convenient restriction sites that could be labeled by the *E. coli* DNA polymerase I fill-in procedure. To examine the right end of V1, total cellular DNA was first labeled at its 3' ends by adding an α-32P-labeled cordycepin residue, using the enzyme terminal deoxynucleotidyl transferase (Tu and Cohen 1980). The labeled DNA was then digested with *AluI*, generating an expected right terminal fragment of 237 bp (including the cordycepin residue), and hybridized to a filter containing immobilized pMACV1-4 DNA. For cell lines NcC10, CC51, CC54, and CC57, a
single band of 237 bp was observed as predicted [Fig. 5A; the second band observed on the gel is discussed below]. For cell line G1, however, a band was seen at 236 bp, as well as 237 bp. We suspected that the shorter hybrid-selected fragments might have resulted from exonucleolytic 'nibbling' of the total cellular DNA prior to its use in the hybrid selection procedure. In contrast to the previous experiments, the labeling procedure employed here results in the detection of the DNA strand that contains the 5'-G4T4-3' telomeric repeat sequence [referred to as the 3' strand]. The 3' strand is extended 14 bp relative to the 5'-C4A4-3' strand and is potentially more accessible to attack by single-stranded exonucleases. To determine whether nibbling was the cause of the above result, an alternative labeling procedure that removes the 3' single-stranded tails of macronuclear DNA molecules was employed [Herrick and Wesley 1978]. Strain G1 total cellular DNA was first treated with T4 DNA polymerase in the presence of dATP, dTTP, and dCTP. Under these conditions, the 3' → 5' exonuclease activity removes the single-stranded tail, as well as the first four G residues of the double-stranded region of the telomere. The molecules were then labeled and blunt-ended by the addition of [α-32P]dGTP to the reaction. Following AluI digestion and hybrid selection, a single fragment of 222 bp was observed as expected, based on the labeling method [Fig. 5C]. Thus, it appears that only a single fragmentation site is used to generate the right end of V1 in cell line G1, as well as the other cell lines examined.

In the analyses of V1 right ends, a second band of ~330 bp also was observed [Fig. 5A]. Although the left

---

**Figure 5.** Autoradiographs of hybrid selection experiments of terminal fragments from V1 and V2 macronuclear DNA molecules of *E. crassus* cell lines G1, NcC10, CC51, CC54, and CC57. (A) Hybrid selection of the right terminus-AluI fragments of V1. Total cellular DNA was 3'-end-labeled with 32P-labeled cordycepin and digested with AluI, and hybrid selection was carried out using clone pMACV1-4. In addition to right terminal fragments of 236 and 237 bp, the band at ~ 330 bp represents left terminal fragments of V1. Size standards [lanes A, T, C] consist of an M13mp11 clone containing the 1.2-kbp EcoRI-PstI fragment of pMACV1-4 sequenced by the dideoxy chain-termination procedure. (B) For hybrid selection of the left terminus HhaI-fragments of V2, total cellular DNA was 3'-end-labeled with 32P-labeled cordycepin, digested with HhaI, and hybridized to a filter containing pMACV2 DNA. Size standards [lanes A, T] consist of an M13mp11 clone of the 800-bp EcoRI-BglII fragment of pMACV2 sequenced by the dideoxy chain-termination procedure. (C) Hybrid selection of the V1 termini and the V2 left terminus from cell line G1 DNA blunt-ended and 3'-end-labeled using T4 DNA polymerase. Expected sizes of terminal fragments are 15 bp smaller than in A and B due to removal of the 14 bp of single-stranded telomere and the lack of a cordycepin residue. Size standards [lanes C + A and C + T] consist of an M13mp10 clone of a 0.28-kbp HaeIII–EcoRI [left end] fragment of pMACV2 sequenced by the dideoxy chain-termination procedure. The size markers used in A and B were calibrated using the M13mp10 clone of the 0.28-kbp HaeIII–EcoRI fragment of pMACV2 [C]. The latter clone produces standards that consist of the G4T4-containing strand of V2, and are thus more closely related to the DNA strands isolated during hybrid selection.
end of V1 has not been sequenced, a terminal AluI fragment of ~350 bp is expected on the basis of restriction mapping of clone pMACV1-4. The second band represents the left end of V1, and the results indicate that it is also generated by fragmentation at a unique site.

To examine the left end of V2, total cellular DNAs were 3'-end-labeled using cordycepin and digested with HhaI to yield an expected fragment of 194 bp. Only fragments of 194 bp were observed in hybrid-selected DNA from cell lines NeC10, CC54, and CC51 [Fig. 5B]. Cell lines G1 and CC57 displayed the 194-bp fragment, as well as 193-bp fragments. The smaller fragments again appear to be the result of exonucleolytic nibbling during DNA isolation, as only a single band of the expected size was observed when G1 DNA was labeled by the T4 polymerase/blunt-ending procedure [Fig. 5C].

Overall, six ends of E. crassus macronuclear DNA molecules were examined in the above analyses and all were found to be generated by fragmentation at unique sites that are used reproducibly during independent episodes of macronuclear development. The results also indicate that the length of the telomeric repeats is invariant in this organism. The hybrid selection experiments on the left end of the D8 macronuclear DNA molecule were especially significant. As mentioned above, two allelic forms of D8 have been identified that differ in the presence or absence of a HaeIII restriction endonuclease recognition site. Cell lines G1, NeC10, and CC54 are all heterozygous for the two forms of D8, whereas CC51 and CC57 are homozygous for the HaeIII minus and plus forms of D8, respectively [Baird and Klobutcher 1988]. The isolation of single fragments of the expected size in the hybrid selection experiments indicates that both alleles use the same fragmentation site in a reproducible manner.

As we have no reason to expect that the macronuclear DNA molecules examined in the above experiments are in any sense unusual relative to other macronuclear DNA molecules, the results indicate that chromosome fragmentation at unique sites is a general property of this organism. E. crassus is the first ciliated protozoan to be involved in the resolution of overlapping genes in O. nova. The precursors of the C3 macronuclear DNA molecule appeared to overlap the precursor of a 5.0-kbp macronuclear DNA molecule [C4] by 5 bp [Fig. 1A; Klobutcher et al. 1988]. The conclusion in this case was not definitive, as the macronuclear C4 clone analyzed was of a different version than the macronuclear C3 and micronuclear clones. More recently we have found that the macronuclear precursors of V1 and V2 overlap by 6 bp in E. crassus [Fig 3; S.E. Baird et al., in prep.]. In this instance, multiple versions are not a problem, as single forms of V1 and V2 exist in E. crassus strain G1.

Figure 6 presents three possible models for the resolution of macronuclear DNA molecules whose precursors overlap in the micronucleus. In the first model, the fragmentation event involves the generation of a staggered double-stranded cut of the DNA, much like some restriction endonucleases, generating ends with single-stranded tails. A DNA synthesis step to fill-in the ends prior to telomere addition would result in macronuclear DNA molecules that share the region defined by the staggered cut. The second and third models rely on the fact that chromosome fragmentation occurs after the formation of polytene chromosomes in the developing macronucleus, so that many DNA strands are available. In the second model, chromosome fragmentation occurs at either of two sites that define the region of overlap. Only those ends that receive the region of overlap can serve as substrates for telomere addition, giving rise to DNA molecules that reach the mature macronucleus. The molecules that lack the overlap are not substrates for telomere addition and are degraded at later stages of macronuclear development. The third model employs multiple chromosome fragmentation sites, with all of the resulting ends serving as substrates for telomere addition and reaching the mature macronucleus. In this last case, the appearance of overlapping precursors would result from analyzing macronuclear clones that were derived from different DNA strands during development; i.e., the two products derived from the fragmentation of a single strand of DNA would not share terminal sequences.

The hybrid selection experiments on the V1 and V2 genes allow us to rule out at least one of these models for E. crassus. Both V1 and V2 display no terminal heterogeneity, so that model three cannot be the mechanism by which these two macronuclear DNA molecules are generated. Both models I and II, however, give rise to macronuclear DNA molecules with discreet ends, so both remain possibilities for the resolution of overlapping precursors in this organism. The hybrid selection data are less informative for O. nova, where multiple fragmentation sites are observed. On the surface, model III is the most consistent with the results, as it relies on multiple fragmentation sites. However, if one incorporates the use of multiple fragmentation sites into both the first and second models, aspects of each might also be involved in the resolution of overlapping genes in O. nova. For instance, a modified model I with multiple staggered cuts that are all filled in prior to telomere addition is equally consistent with the data.
Identification of conserved DNA sequences in the vicinity of chromosome fragmentation sites

The overall specificity of the chromosome fragmentation process implies that cis-acting DNA sequences exist that define the sites of chromosome fragmentation during macronuclear development. In an attempt to define such a putative fragmentation signal, we have compared previously determined sequences of regions of micronuclear DNA subject to fragmentation in O. nova and E. crassus. The analysis included cases where 100 bp of sequence flanking both sides of the fragmentation site have been determined. This 200-bp interval is likely to contain any cis-acting fragmentation sequence for the following two reasons: (1) The intergenic spacers separating adjacent macronuclear gene precursors in the micronucleus are generally small and in only one instance exceed 100 bp [Klobutcher et al. 1986; 1988; S.E. Baird et al., in prep.]; (2) coding sequences appear to occupy all but the terminal 100–200 bp of most macronuclear DNA molecules e.g., Kaine and Spear 1982; Helftenbein 1985; see also Klobutcher and Prescott 1986).

Eight fragmentation sites have been sequenced in O. nova [Klobutcher et al. 1984, 1986, 1988; Ribas-Aparicio et al. 1987]. We have been unable to detect any significant sequence features shared by all eight of these fragmentation sites. The O. nova fragmentation sequences also lacked significant similarity to the 15-bp chromosome breakage sequence (Cbs) that Yao et al. [1987] have found to be associated with chromosome fragmentation sites in the ciliate Tetrahymena. Overall, the situation for O. nova is similar to that reported for Paramecium [Baroin et al. 1987; Forney and Blackburn 1988] and O. fallax [Herrick et al. 1987], where no significant sequence features in the vicinity of fragmentation sites have been observed.

The sequences of six E. crassus chromosome fragmentation sites obtained from micronuclear clones LEMICD and LEMICV [S.E. Baird et al., in prep.] are shown in Figure 7A. To simplify Figure 7, only the 50 bp preceding the chromosome fragmentation sites and the first 50 bp of sequence that form the subtelomeric regions of the respective macronuclear DNA molecules are shown. Also shown are the subtelomeric sequences of the two ends of the 1.15-kbp E5 macronuclear DNA molecule.

A potential candidate for a chromosome fragmentation signal was the conserved sequence 5'-TTGAA-3'. As noted above, this short sequence was readable above background at a position 17 bp from the end of the telomeric repeats when total, end-labeled macronuclear DNA was sequenced. An identical result was obtained previously for E. aediculatus [Klobutcher et al. 1981], and it was suggested that all ends of macronuclear DNA molecules possess this sequence 17 bp from the telomere addition site (Fig. 7A). For the three cases where the sequences of both ends of macronuclear DNA molecules were determined [V2, D8, and E5], the conserved sequence is found at only one end of each macronuclear DNA molecule.

The failure to find this conserved sequence at both ends of macronuclear DNA molecules suggests that it is not important for telomere structure per se. Instead, we believe that this pentanucleotide forms part of a larger consensus sequence that exists either within the macronuclear gene precursor or upstream of the fragmentation site in inverted orientation. The existence of such a larger consensus initially was indicated by significant similarities in the V1 left, D8 right, and V3 left fragmentation sequences just upstream of the chromosome fragmentation site (Fig. 7A). The core of these similar sequences consists of the sequence 5'-TTCAC-3', located 11 bp upstream of the chromosome fragmentation site [5'-ATCAGA-3' for V3 left]. As this is the complement of the conserved sequence 17 bp from the telomere addition site of macronuclear DNA molecules, we were prompted to reorient the fragmentation sequences based on the maximum homology to 5'-TTCAC-3' 11 bp prior to the fragmentation site, or 5'-TTGAA-3' 17 bp past the fragmentation site. When this was done [Fig. 7B], it was possible to derive a consensus sequence of 5'-ATCAGA\%YAT-3' just upstream of the chromosome fragmentation sites (Fig. 7B).

This consensus sequence is moderately predictive of chromosome fragmentation sites. The 5080 bp of E. crassus DNA that we have sequenced [S.E. Baird et al., in prep. and unpubl.] were searched with the consensus sequence, asking for sites that contained perfect matches of the six highly conserved residues [indicated above in boldface type] and an overall 85% identity to the extended consensus. Five of the seven fragmentation sites were picked up in the search, along with three other sites that are not subject to chromosome fragmentation. Thus, although the derived consensus sequence has some predictive value in identifying fragmentation sites, it is not a perfect indicator and suggests that other
The fragmentation site that showed the least similarity to the derived consensus was the one near the right end of V2 (Fig. 7B). This fragmentation site is also unusual in possessing a high degree of G/C strand asymmetry. In the 100 bp of the V2 right fragmentation sequence shown in Figure 7A, there are 44 G residues and 59 A residues. The fragmentation site that showed the least similarity to the derived consensus was the one near the right end of V2 (Fig. 7B). This fragmentation site is also unusual in possessing a high degree of G/C strand asymmetry. In the 100 bp of the V2 right fragmentation sequence shown in Figure 7A, there are 44 G residues and 59 A residues.
only a single C residue. Moreover, the G/C asymmetry actually extends to the entire 400-bp intergenic spacer that separates the right end of V2 from the left end of V3 in the macronucleus. Beginning at the left end of V2 and containing toward V3, the first 167 bp of the intergenic spacer contain 82 C residues and only 8 G residues, whereas the last 233 bp contain 93 G residues and 7 C residues [S.E. Baird et al., in prep.]. A similar strand asymmetry has been observed in the terminal inverted repeats of a transposon-like element in the micronuclear genome of *O. fallax* [Herrick et al. 1985], and two of the *O. nova* fragmentation sequences display G/C strand asymmetry either upstream or downstream of the chromosome fragmentation site [see C2 left and C3 right/C4 left in Klobutcher et al. 1984 and 1988, respectively]. The significance of these unusual regions is unknown, but they may be involved in defining a second class of fragmentation sites.

The position of the consensus sequence suggests a model of the chromosome fragmentation event (Fig. 7C). When the consensus is located in sequences destined to form part of the macronuclear DNA molecule [e.g., E5 left], telomeres are added at a position 17 bp from the highly conserved 5'-A/TCTAAA-3' of the consensus sequence. When the consensus sequence resides outside the macronuclear gene precursor [e.g., D8 right], the subtelomeric region of the macronuclear DNA molecule begins 11 bp from the highly conserved pentanucleotide. This dual arrangement of the consensus sequence could be accommodated if the initial fragmentation event results in a staggered cut in the DNA generating a 6-bp 5' overhang, which is subsequently filled-in prior to telomere addition [Fig. 7C]. Whether the DNA molecules to the left and/or right of the fragmentation site ultimately appear in the mature macronucleus could be controlled by other factors. For instance, some of the DNA ends may not be recognized as substrates for telomere addition and, hence degraded later in development, or some of the resulting DNA molecules may lack origins of DNA replication. The staggered-cut model is particularly attractive because it is consistent with the resolution of V1 and V2, which overlap by 6 bp in the micronucleus. In this case, the consensus sequence resides in the V2 precursor [Fig. 7B] but would be sufficient to generate the ends of two macronuclear DNA molecules. It will be of interest to determine whether other overlapping macronuclear gene precursors in *E. crassus* overlap by 6 bp as predicted by this model.

Although the above model postulates a 5' overhang, the data are consistent with the fragmentation event generating molecules with a 6-bp 3' overhang. In this case, telomeric repeats might be added directly to the 3' overhang, and the 5' gap filled-in later. Such a scenario is attractive because the telomere terminal transferase activities that have been detected in protozoa are capable of adding telomeric repeats to the 3' ends of oligonucleotides that correspond to the natural 3' ends of telomeres [Greider and Blackburn 1985; Zahler and Prescott 1988; Shippen-Lentz and Blackburn 1989]. Although not yet demonstrated, it is possible that this activity also recognizes DNA molecules with a specific terminal structure and is responsible for the de novo telomere addition that occurs during macronuclear development.

The ability to detect the consensus sequence was aided greatly by its defined positions relative to the unique chromosome fragmentation sites used by *E. crassus*. It may be that similar weak consensus sequences exist in organisms such as *Paramaecium* and *Oxytricha*, but have been difficult to detect because of the use of multiple chromosome fragmentation sites in these organisms [Baroin et al. 1987, Herrick et al. 1987; Forney and Blackburn 1988]. In comparison to the *Tetrahymena* Cbs element, the conserved *E. crassus* sequence displays two significant differences. First, the 15-bp Cbs element is strictly conserved at all *Tetrahymena* chromosome fragmentation sites [Yao et al. 1987], whereas the *E. crassus* consensus sequence is comparatively loose and appears to tolerate a number of base mismatches. This may be related to the greater number of fragmentation events in *E. crassus*. Second, the *Tetrahymena* Cbs element always resides within the DNA eliminated during macronuclear development and not within the macronuclear DNA molecule that results from the fragmentation event [Yao et al. 1987].

Whether or not the consensus sequence defined for *E. crassus* actually specifies the position of chromosome fragmentation will clearly require further studies. The possibility of adapting procedures that have been used successfully for DNA transformation of *Tetrahymena* [Tondravi and Yao 1986] to hypotrichous ciliates, in conjunction with in vitro mutagenesis of fragmentation sites, may make it possible to address directly the involvement of our putative chromosome fragmentation signal. Alternatively, the ability to prepare large numbers of *E. crassus* cells synchronously proceeding through macronuclear development [Roth et al. 1985] may allow the development of an in vitro chromosome fragmentation system to test the involvement of the fragmentation consensus sequence.

**Materials and methods**

*E. crassus* cell lines and DNA isolation

The *E. crassus* cell lines used in these studies were derived from cells originally isolated by Dr. P. Luporini [University of Camerino, Italy], from San Terenzio, Italy. Strain G1 is a partially inbred mating type I cell line [mtU/mtl], for a discussion of the mating type system, see Heckmann 1964). Other cell lines used were NcC10, a mating type III cell line [mtU/mtl], and cell lines CC51, CC54, and CC57, which were derived from a mating of G1 and NcC10 as described previously [Baird and Klobutcher 1988]. *E. crassus* cell lines were grown and maintained in supplemented artificial seawater [Instant Ocean, Aquarium Systems; Mentor, Ohio], using the marine algae *Dunaliella salina* as a food source, as described by Roth et al. [1985].

*O. nova* strain H10, a clonal cell line, was grown using the algae *Chlorogonium* as a food source [Swanton et al. 1980].

DNA isolation

Total cellular DNA and macronuclear DNA were isolated from the hypotrichous ciliated protozoa using established methods.
[Lauth et al. 1976; Klobutcher et al. 1981, 1984]. Bacterial plasmid DNAs were prepared using either the SDS–lysis or the alkaline–lysis procedures [Godson and Vapnek 1973; Maniatis et al. 1982]. Recombinant bacteriophage were purified on glycerol step gradients, and DNA was prepared as described by Maniatis et al. [1982].

**Restriction endonuclease digestions and agarose gel electrophoresis**

DNAs were digested with restriction enzymes in accordance with the manufacturers’ directions [New England Biolabs, Inc., Beverly, Massachusetts; Bethesda Research Laboratories, Bethesda, Maryland]. DNAs were sized by electrophoresis through agarose or low-melting-point agarose gels prepared and run in TBE buffer [89 mM Tris base, 89 mM boric acid, 2 mM EDTA (pH 8.0)].

**DNA fragments analysis**

DNA fragments were cloned into bacteriophage M13 mp8, mp9, mp10, or mp11 and sequenced by the dideoxy chain-termination method [Sanger et al. 1977] using kits purchased from New England Biolabs or United States Biochemical Corp. (Cleveland, Ohio) and α-thio-32P-labeled dATP as the labeled nucleotide triphosphate. Alternatively, the chemical cleavage sequencing method was employed using DNA fragments labeled at a 3' end [Maxam and Gilbert 1980].

The terminal sequences of the 1.15-kbp E5 macronuclear DNA molecule were determined from clone EMACS. 3'-end-labeled terminal fragments of the EMAC5 insert were generated and sequenced by the chemical cleavage method.

DNA sequences were analyzed using an IBM PC equipped with IBI/Pustell Sequence Analysis software (International Biotechnologies, New Haven, Connecticut). Alternatively, a Homo sapiens data processor, equipped with a complex optical scanner, was employed.

**DNA labeling**

Macronuclear DNA or total cellular DNA used in hybrid selection was radioactively labeled by several procedures. In most cases, 10 µg of DNA was first digested with a restriction enzyme that generated 3'-ends and then labeled by filling in the resulting gaps using E. coli DNA polymerase I, along with one α-32P-dNTP and other unlabeled dNTPs [Maniatis et al. 1982]. Alternatively, the natural 3' termini of macronuclear DNA molecules were radiolabeled by the addition of an α-32P-labeled cordycepin [New England Nuclear, Boston, Massachusetts] residue, using the enzyme terminal deoxynucleotidyl transferase [Tu and Cohen 1980]. In some cases, the DNA was labeled via minor modifications of a procedure originally described by Herrick and Wesley [1978], which removes the 3' single-stranded tail of the telomeres of native macronuclear DNA molecules and introduces α-32P-labeled nucleotides at the 3' end of the double-stranded region of the telomeres. For this procedure, 10–15 µg of DNA was added to a reaction mixture consisting of 2.5 mM dTTP, 2.5 mM dATP, 16.6 mM NH4SO4, 67 mM Tris-Cl [pH 8.0], 6.7 mM MgCl2, and 10 mM β-mercaptoethanol. Ten units of bacteriophage T4 DNA polymerase were added and the mixture incubated at 37°C for 15 min. Then, 200 µCi of α-32P-labeled dGTP were added to the reaction mixture, and incubation continued for 15 min more. As a final step, the reaction mixture was made 250 mM dGTP and incubated for an additional 5 min to ensure that blunt-ended telomeres were formed. The labeled DNA was ethanol-precipitated prior to use in the hybrid selection protocol.

**Hybrid selection**

Hybrid selection was carried out using 1-cm² nitrocellulose filters containing 5 µg of a recombinant plasmid clone harboring a copy of the target macronuclear DNA molecules. The filters were prepared by first soaking them in 6× SSC [1× SSC = 0.15 M NaCl, 0.015 M Na-citrate, pH 7.0] and then blotting them to remove excess moisture. Plasmid DNA was brought to a volume of 50 µl in distilled water, denatured by heating at 100°C, and immediately applied to the filters. The filters were allowed to air-dry, baked at 80°C for 2 hr under vacuum, and stored at room temperature until use.

Prehybridization and hybridization of the filters were carried out essentially as described previously [Boswell et al. 1982], except that the hybridizations were performed at 65°C in 1.5-mI microcentrifuge tubes using 1 ml of hybridization solution. Following hybridization for 20–24 hr with labeled genomic DNA, the filters were removed from the microcentrifuge tubes and washed twice for 30 min at 65°C in 50 ml of 2× SSC, 0.5% SDS, followed by two additional washes in 0.1× SSC 0.5% SDS. In cases where the hybrid-selected DNA fragments were expected to be <150 bp, the second pair of washes was done in 0.2× SSC, 0.5% SDS. Following the washes, the filter was rinsed in 0.1× SSC at room temperature transferred to a microcentrifuge tube containing 150 µl of TE [10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0], and boiled for 3 min to elute the specifically bound, hybrid-selected DNA fragments. The TE was removed and saved, and the boiling step repeated using 150 µl of fresh TE solution. The TE solutions were then pooled, and hybrid-selected DNAs recovered by ethanol precipitation using 5 µg of E. coli tRNA as carrier. The hybrid-selected DNAs were resuspended in 5–10 µl of loading buffer and analyzed by electrophoresis on 6% or 8% acrylamide–urea DNA sequencing gels [Maxam and Gilbert 1980], using sequenced DNAs as size standards.

**Acknowledgments**

We would like to thank Dr. Ann Cowan and Ms. S.L. Tausta for their critical reading of the manuscript. The work was supported by U.S. Public Health Service grant GM-33277 and a grant from the University of Connecticut Research Foundation.

**References**


