High fidelity developmental excision of Tec1 transposons and internal eliminated sequences in *Euplotes crassus*

S.Lorraine Tausta⁺, Leah R.Turner, Linda K.Buckley and Lawrence A.Klobutcher*

Department of Biochemistry, University of Connecticut Health Center, Farmington, CT 06030, USA

Received April 10, 1991; Revised and Accepted May 18, 1991

ABSTRACT

Following the sexual phase of its life cycle, the hypotrichous ciliate *Euplotes crassus* transforms a copy of its chromosomal micronucleus into a transcriptionally active macronucleus containing short, linear, gene-sized DNA molecules. Tens of thousands of DNA breakage and joining, or splicing, events occur during macronuclear development. The DNA removed by such events includes transposon-like elements, referred to as Tec1 elements, as well as segments of unique sequence DNA, termed internal eliminated sequences (IESs). Both types of elements are bounded by short direct repeats. In the current study, a polymerase chain reaction (PCR) and DNA sequencing strategy has been used to examine the fidelity of excision of two Tec1 elements and three IESs. In all cases, the vast majority of excision events were found to be precise, with one copy of the terminal direct repeats retained at the empty site in the macronuclear DNA molecule. These results, in combination with previous studies that have characterized the excised DNA elements, indicate that the two products of excision (the free element and the macronuclear DNA molecule) share DNA sequences. This suggests that excision events are initiated by staggered cuts in the chromosomal DNA.

INTRODUCTION

A number of ciliated protozoans have been found to undergo a dramatic genome rearrangement process as part of their normal life cycles (reviewed in 1,2). These unicellular organisms are able to alter their DNA structure, yet maintain genetic continuity, because each cell possesses two distinct types of nuclei. Micronuclei contain DNA in the form of chromosomes and divide by mitosis, but are by and large transcriptionally inactive. The micronuclear genome is unrearranged and serves as the germ line. The macronucleus, on the other hand, is transcriptionally active and contains subchromosomal DNA molecules that represent a subset of the micronuclear sequences.

Following the sexual phase of the life cycle (conjugation), the macronucleus is destroyed and a new macronucleus is generated from a copy of the micronucleus through a series of DNA rearrangement events. The hypotrichous ciliates, such as *Euplotes crassus*, display an extreme form of genome rearrangement (1). Macronuclear development begins in these organisms with a number of rounds of DNA replication resulting in the formation of polytene chromosomes. The polytene chromosomes are then fragmented in concert with the formation of vesicular structures, and up to 95% of the micronuclear genome complexity is eliminated. The vesicles then break down, and the remaining DNA molecules undergo additional replication resulting in the mature macronuclear genome. As a result of the chromosome fragmentation and DNA amplification events, the mature macronucleus contains 10,000–20,000 short linear DNA molecules (avg. size of ~2 kbp), each present at 1,000 or more copies. Each linear macronuclear DNA molecule contains a single coding region, as well as the information required for its expression and replication.

Comparisons of cloned macronuclear DNA molecules and the regions of the micronuclear genome from which they are derived indicate that DNA breakage and joining, or splicing, events also occur during macronuclear development in the hypotrichs. Two general classes of DNA sequences eliminated by DNA breakage and joining have been identified. The first are internal eliminated sequences (IESs), which are short segments of DNA ranging in size from 14 to ~500 bp (3–6). IESs are generally dissimilar in primary sequence, but share two organizational features. Each IES is bounded by a 2–6 bp direct repeat and often has short imperfect inverted repeats near its termini. Following IES excision, one copy of the terminal direct repeat appears to be retained in the macronuclear DNA molecule. More recently, large transposon-like elements have been found to undergo developmental excision in the hypotrichs *Oxytricha fallax* (7–9) and *E. crassus* (10–12). In *E. crassus*, approximately 10,000 copies of a 5.5 kbp element referred to as Tec1 are present in the micronuclear genome and undergo excision during macronuclear development (10–12). Tec1 elements have 698 bp inverted repeats, with 3–4 bp direct repeats at their termini (11,12). Finally, the non-hypotrichous ciliated protozoa

* To whom correspondence should be addressed

⁺ Present address: Department of Biochemistry and Genetics, The Connecticut Agricultural Experiment Station, New Haven, CT 06504, USA

Tetrahymena and *Paramecium* are also known to undergo similar DNA breakage and joining events during macronuclear development, suggesting that the process is widespread among the ciliates (reviewed in 2).

Based on the terminal sequence organization of IESs, a link between IES removal and a developmentally controlled transposon excision process has been proposed (4,7,10,13). The basic idea is that IESs are degenerate copies of transposable elements whose excision is controlled by other intact elements in the micronuclear genome. The micronucleus is a perfect haven for transposable elements because it is transcriptionally inactive, which allows element insertions anywhere in the genome without immediate phenotypic consequences. The discovery of both IESs and transposon-like elements that undergo developmental excision in species such as *E. crassus* has lent support to this notion. Moreover, recent work has uncovered additional similarities between *E. crassus* IESs and Tec1 elements. Both IESs and Tec1 elements have been found to exist as circular forms following their developmental excision during the polytene chromosome stage of macronuclear development (12,13). The junctions of the free circular forms of IESs are unusual in that they appear to contain both of the terminal direct repeats as well as some of the same nucleotides that are retained in the mature macronuclear DNA molecule (13). It is likely that the circular Tec1 elements possess the same type of junction (12). Finally, the direct repeats defining the ends of all *E. crassus* IESs and Tec1 elements sequenced to date can be viewed as the dinucleotide 5'TA3' (11,12; the direct repeats of some elements can be viewed as being longer [see 11], but, to simplify the discussion, we will consider the shared 5'TA3' as the direct repeat).

In this study we have examined the regions of macronuclear DNA molecules from which IESs or Tec1 elements have been removed ('empty sites') to determine if the excision events are precise and reproducible. This is of interest for two reasons. First, at least some IESs reside within coding regions (3,4,6,8), so that

one would expect that excision events must be precise to form functional genes. However, DNA breakage and joining events occur during the polytene chromosome stage of macronuclear development (12,13). Because multiple precursors of each macronuclear DNA molecule are present in each cell, all excision events need not be precise to form at least some functional macronuclear DNA molecules. Second, as noted above, the free, circular, excised IESs and Tec1 elements appear to share some nucleotides with the other product of excision: the mature macronuclear DNA molecule. These shared nucleotides may be the result of the excision mechanism. Alternatively, it is possible that the macronuclear DNA molecule and free circular elements do not share nucleotides. The boundaries of IESs and Tec1 elements have in most cases been operationally defined by sequencing and comparing single micronuclear and macronuclear clones. If the excision process is inexact, the cloned macronuclear DNA molecules previously sequenced could represent minor forms, with most macronuclear DNA molecules having different excision boundaries such that they do not share nucleotides with the majority of the circular, excised elements.

We have used polymerase chain reaction (PCR;14) strategies to amplify the empty sites resulting from the removal of two Tec1 elements and three IESs, and performed DNA sequencing on the PCR products to determine if excision events are variable. The results indicate that the vast majority of IES and Tec1 excision events occur with the same boundaries and leave behind one of the terminal direct repeats of the eliminated element.

MATERIALS AND METHODS

E. crassus Cell Lines

Cells were grown using the alga *Dunaliella salina* as the food source as described by Roth, et al. (15), except that vitamin B12 was omitted from the artificial sea water. The following combinations of *E. crassus* strains were used in matings (16) to generate cells undergoing macronuclear development: CC51

Figure 1. Maps of the regions of the micronuclear genome analyzed. A.) Restriction map of micronuclear clone LEMICD. Areas that give rise to macronuclear DNA molecules (D7 and D8) are shown as open boxes, regions external to macronuclear precursors are shown as cross-hatched boxes, IESs are shown as black boxes when their sequences are known and as black ovals when they are not. Tec1 elements are shown as stippled boxes. The sizes of sequenced IESs is indicated in base pairs. The approximate positions in the clones complementary to oligonucleotides used as PCR primers are indicated by arrowheads, which point in the direction the oligonucleotides prime DNA synthesis. Note that the Tec1 element is not drawn to scale, and the left end of the Tec1 element does not actually reside in clones LEMICD. B.) Restriction map of micronuclear clone LEMICV. The regions of the clone that give rise to the V1, V2, and V3 macronuclear DNA molecules are indicated. Other aspects are as described above. B = Bgl II, E = Eco RI, H = Hind III, P = Pst I, Pv = Pvu II, Ss = Sst I, X = Xba I, and X = Xmn I. Pst I, Pvu II, and Xmn I sites have not been mapped in LEMICD.

(mating type III; mt^3/mt^1) $\times$ CC55 (mating type I; mt^1/mt^1), CC102 (mating type III; mt^3/mt^1) $\times$ ST2 (mating type I; mt^1/mt^1), and 8 $\times$ 24. Strains 8 and 24 were kindly provided by Dr. C. Jahn (Univ. of Illinois, Chicago). The ST series of clonal cell lines were derived from matings of strains CC51 and CC55 as described by Baird and Klobutcher (17).

Recombinant Clones

The micronuclear clones used in these studies, LEMICD and LEMICV, are described in Baird et al., (11; Figure 1). Both LEMICD and LEMICV contain the precursors of multiple macronuclear DNA molecules, and recombinant clones of these macronuclear DNA molecules (pMAC or LMAC clones) have been isolated and characterized (11).

The regions of LEMICD and LEMICV, and the macronuclear DNA molecules derived from them, that have been sequenced are described in Baird et al. (11). These sequences are contained in the GenBank/EMBL Sequence Databases under accession #'s M28498–M28506. All sequences were based on recombinant clones derived from *E. crassus* strain G1.

Preparation of Cellular DNAs

Total cellular DNA (whole cell DNA) was isolated from vegetative or developing cells as previously described (18). Whole cell DNA was isolated from the series of ST cell lines approximately 20 vegetative cell generations after they had completed macronuclear development.

Developing macronuclei were isolated from a mating of strains CC102 $\times$ ST2 and DNA was purified based on procedures

originally described by Roth, et al. (15) with minor modifications (19). For the PCR analyses using DNA from 50 hr developing macronuclei, the DNA was further purified to remove contaminating DNA from the old macronuclei. This was achieved by running the DNA from 50 hr developing macronuclei on a 1% low melting point agarose (LMP agarose; Life Technologies, Inc.) gel and cutting out the band of high molecular weight developing macronuclear DNA. The DNA in the gel slice was digested in situ with Bgl II, and purified from the gel as described (20). A Southern hybridization analysis of the resulting DNA preparation indicated that less than 2% of the empty sites in this preparation were the result of contamination with mature macronuclear DNA molecules.

For the PCR analysis of the 144 bp V3 IES empty sites in 50 hr DNA from developing macronuclei, a preparation of 50 hr whole cell DNA derived from a mating of strains 8 and 24 was digested with Hind III and run on a 1% LMP agarose gel. DNA in the 3–5 kbp size range was excised, purified (20), and used for PCR. The D7 Tec1 empty site was expected to reside on a 4.1 kbp Hind III fragment in DNA from 50 hr developing macronuclei (11,19). Southern hybridization indicated that there was less than 5% contamination with mature macronuclear DNA molecules in this DNA preparation.

Polymerase Chain Reaction (PCR)

The following 20 base oligonucleotides were synthesized using a Cyclone dual column DNA synthesizer (Biosearch, Inc., San Rafael, CA) for use as PCR primers:

Figure 2. Analyses of the empty sites generated by the excision of the D7 and V1 Tec1 elements. A.) Sequences at the ends of the Tec1 elements in the micronuclear precursors of the D7 and V1 macronuclear DNA molecules. Bases that will form the macronuclear DNA molecule are shown in upper case, while the ends of the Tec1 elements are shown in lower case. The 5'TA3' direct repeats at the ends of the Tec1 elements are highlighted, with the left direct repeats arbitrarily indicated as the ones retained in the macronuclear DNA molecules. The sequence of the left end of the D7 Tec1 element has not been determined. Both sequences are in reverse orientation relative to the published sequences (11) to aid the comparison to the empty site analyses (below). Note that one of the direct repeats of the V1 Tec1 element was indicated incorrectly in a previous publication (11) and has been corrected in the figure. B.) Sequences of the PCR products derived from the D7 and V1 Tec1 element empty sites in DNA from 50 hr developing macronuclei and macronuclear DNA from clonal vegetative cell lines (ST cell lines). The position of the 5'TA3' direct repeat that is retained at the empty site is indicated. For the D7 Tec1 element empty sites, PCR was carried out on 50 hr whole cell DNA isolated from a mating of cell lines CC55 and CC51 using primers DTMH and SD7/D8, while PCR of the ST cell lines utilized oligonucleotides DTMH and DTME. Dideoxy sequencing was performed using DTMH as the primer. For the V1 Tec1 element empty sites, PCR was performed using oligonucleotides TMA and TMC. Oligonucleotide TMA was 5'-end labeled prior to PCR, and the products were sequenced by the chemical method. The 50 hr DNA used for this analysis was isolated from developing macronuclei of a mating of cell lines CC102 and ST2.

DTMH: 5' ATTTTAGCATTACCTAACGG 3' DTME: 5' TACCACAAGTGAAGAATATG 3'
 VOMG: 5' ATTTTCATACCATAACCAATG 3' VICG: 5' GAACAAATAATGGGTAAAAG 3'
 VIMC: 5' GAAACAGTTTGTGTGAGAAC 3' VIMT: 5' CTAAGCCAAGTTCCAATAT 3'
 TMA: 5' AGCCTCAGAGATCTTAGAGA 3' TMC: 5' CTGCAGACAGTATTGAAGAC 3'
 V2MG: 5' TTACTTAATAAACAGGTTTCG 3' V2MT: 5' TTACAACACCTAATACACAT 3'
 V3MA: 5' CAAAATGGTATGTATAGATA 3' V3MC5' AAAATGGACAGGACTAAGAC 3'
 SD7/D8: 5' AGCCATAATGGAATACCTAC 3'

The regions of clones LEMICD and LEMICV complementary to the PCR primers are indicated in Figure 1.

PCR reactions (14) were set up using a GeneAmp polymerase chain reaction kit according to the manufacturer's instructions (Perkin Elmer Cetus, Norwalk, Ct.) either with the *Taq* polymerase included in the kit or with Promega *Taq* polymerase and Promega reaction buffer (Promega, Madison, WI). Twenty-five to thirty cycles of PCR were performed in a Model 50 Tempcycler (Coy Laboratory Products, Inc., Ann Arbor, Michigan). Approximately 1 µg of DNA was used in each PCR reaction. Annealing temperatures for the various combinations of oligonucleotide primers were 45°C, with the exception of V3MA and V3MC (42°C).

Sequencing of PCR Products

Sequencing of the PCR products derived from IES and Tec1 empty sites was performed using either the Maxam and Gilbert chemical method (21) or the dideoxy chain-termination procedure. For the chemical sequencing method (21), one of the oligonucleotide primers was 5'-end-labeled with ³²P using T4 polynucleotide kinase (22) prior to carrying out PCR. Following PCR, the labeled products were run on a LMP agarose gel, purified (20), and sequenced essentially as described (21).

When dideoxy sequencing was employed, DNA was purified from the PCR reactions and sequenced as described by Kuskawa et al. (23), except that the sequencing reagents were derived from a Sequenase 2.0 kit (United States Biochemicals, Cleveland, OH). The sequencing primer consisted of one of the oligonucleotides used for PCR. In some cases, the sequencing reactions were subjected to terminal deoxynucleotidyltransferase extension as described by Fawcett and Bartlett (24) in order to reduce artifact banding.

RESULTS

Strategy for Examining the Empty Sites Generated by Tec1 and IES Excision Events

We examined the fidelity of IES and Tec1 element excision in two previously cloned segments of micronuclear DNA, LEMICD and LEMICV (11; Fig. 1). These regions of the micronuclear genome give rise to all or part of five separate macronuclear DNA molecules (referred to as D7, D8, V1, V2, and V3), and contain numerous IESs. Degenerate Tec1 elements are also present in the precursors of the D7 and V1 macronuclear DNA molecules (11; Fig. 1).

The strategy for examining the fidelity of element excision involved PCR amplification (14) of empty sites in cellular DNAs followed by sequencing of the empty site regions. If IES and Tec1 element excision were precise, we expected to observe a clear, unambiguous sequencing ladder through the empty site region. Alternatively, if excision events were variable, the sequence would be clear until the empty site region is encountered. At this point, the sequencing ladder would break down owing to the presence of multiple out of phase sequences.

Empty sites were examined in two types of *E. crassus* cellular DNA preparations. First, PCR was used to amplify and analyze empty sites in DNA from 50 hr developing macronuclei (50 hr after mating was initiated). At this time, most, if not all, of the IESs and TECs have been removed, but the fragmentation of the chromosomes has not yet occurred (12,19). Because these DNA preparations were derived from a population of cells undergoing macronuclear development, the PCR/sequencing procedure allowed us to examine the fidelity of excision events in many different cells simultaneously. Second, PCR/sequencing analyses were performed on whole cell DNAs (essentially equivalent to macronuclear DNA) derived from a series of clonal cell lines (the ST cell lines) that had recently completed macronuclear development and had passed through fewer than 20 vegetative divisions. In these cases, we were analyzing all the excision events that occurred in a single cell during the previous round of macronuclear development. The DNAs from the clonal cell lines were analyzed to uncover any potential minor variants that may have been obscured in the analyses of DNA from the entire population of developing cells. At least six clonal cell lines were examined in each case to better increase the chance of identifying minor variants.

Analysis of Tec1 Element Empty Sites

The empty sites formed by the excision of Tec1 elements from the micronuclear precursors of the D7 and V1 macronuclear DNA molecules (Fig. 1) were first examined. In the case of the D7 Tec1 empty sites, PCR was carried out on 50 hr whole cell DNA using oligonucleotide primers DTMH and SD7/D8 (Fig. 1). The two primers bracket the empty site resulting in its amplification. However, because the SD7/D8 primer is complementary to the eliminated spacer DNA separating the micronuclear precursors of the D7 and D8 macronuclear DNA molecules, only the empty sites in chromosomal DNA (primarily from the developing macronucleus) will be amplified, and not those in the old mature macronuclear DNA molecules that are also present in the 50 hr whole cell DNA preparation. Sequencing of the PCR products produced from 50 hr whole cell DNA produced a single clear sequence (Fig. 2) that was identical to that previously obtained from a cloned D7 macronuclear DNA molecule (11). That is, one of the terminal 5'TA3' direct repeats was retained at the empty site. PCR and sequencing was then carried out on six of the clonal ST cell lines (ST1, 2, 3, 4, 5 and 11) using the DTMH and DTME primers (Fig. 1). All displayed a single, clear empty site sequence that retained one of the Tec1 direct repeats (Fig. 2 and data not shown).

Similar analyses were performed on the Tec1 empty sites in the V1 macronuclear DNA molecule. In this case, PCR was carried out using oligonucleotides TMA and TMC (Fig. 1). The empty sites in 50 hr DNA isolated from purified developing macronuclei and in whole cell DNA from six of the ST cell lines (ST1, 2, 3, 4, 5, and 6) all produced a clear sequence that again retained one of the terminal 5'TA3' direct repeats (Fig. 2 and data not shown). No heterogeneity in the sequences was evident. Thus, these results provide no evidence for the use of multiple boundaries during the Tec1 excision process.

Analysis of IES Empty Sites

We next examined the empty sites generated by the excision of the 374 bp, 31 bp, and 144 bp IESs from the micronuclear precursors of the V1, V2, and V3 macronuclear DNA molecules, respectively (Fig. 1). For the 144 bp V3 IES empty sites, PCR

was carried out using the V3MA and V3MC oligonucleotide primers (Fig. 1). As was the case for the *Tec1* empty sites, both the DNAs from 50 hr developing macronuclei and six ST cell lines (ST1, 2, 3, 4, 8, and 9) produced a single clear sequence that retained one of the 5'TA3' direct repeats of the 144 bp IES (Fig. 3 and data not shown). The sequences obtained conform exactly to the previously sequenced V3 macronuclear clone (11).

The empty sites of the 374 bp V1 IES were examined using the primers V1MC and V1MT for the PCR reactions. Sequencing of the PCR-amplified DNAs derived from either purified developing macronuclear at 50 hr or from any of the ST cell lines (ST1, 2, 3, 4, 5, and 8) again revealed a clear sequence and the retention of only one of the direct repeats (Fig. 3 and data not shown). However, in some of the sequences, such as that derived from 50 hr DNA, there was difficulty in determining the second base following the 5'TA3' direct repeat. Bands in both the A and T lanes were observed at this position in the sequencing ladder. The clonal ST cell lines displayed either only the expected T at this position (ST2 and 3), only a A (ST5), or both an A and T (ST1, 4, and 8; Fig. 3 and data not shown). Additional positions in the sequencing ladders behaved in a similar manner (Fig. 3 and data not shown).

There are two potential explanations for the observation of an A and/or T residue following the direct repeat. First, the results could be explained if the boundaries of the excision event were sometimes shifted two base pairs to the right. Based on a previously sequenced micronuclear clone (Fig. 3A; 11), the nucleotide immediately following both the direct repeats is a T, while the second nucleotide after the left direct repeat is expected

to be an A ('within' the IES) and the second nucleotide following the right direct repeat is expected to be a T. The second possible explanation for the A/T base discrepancy was the presence of an allelic difference at the second base following the right direct repeat of the V1 IES. The cell lines in the laboratory are only partially inbred and, therefore, are not homozygous at all loci (e.g., see 17).

To test the latter possibility, we determined the sequences at the right ends of the 374 bp IES and flanking DNA in the micronuclei of *E. crassus* cell lines CC51 and CC55, which are the parents of the ST cell lines. This was done by using oligonucleotides V1CG and VOMG (Fig. 1) to PCR amplify the right ends of 374 bp IESs in vegetative whole cell DNAs isolated from cell lines CC51 and CC55, followed by direct sequencing of the PCR products using VOMG as the primer for dideoxy sequencing. Each parental cell line displayed both an A and a T at the second base following the right direct repeat of the V1 IES, as well as two bands at the other positions previously observed in the empty site analysis (data not shown). This is consistent with the two cell lines being heterozygous at the V1 locus. The behavior of the two putative alleles is consistent with Mendelian segregation, as the observed sequences indicate the ST offspring are either homozygous for one of the two alleles (displaying only an A or a T) or heterozygous (displaying both an A and a T). Moreover, the additional sequence polymorphisms behave in the homozygotes as one would expect for two alleles. That is, these cell lines either exclusively display a sequence identical to the previously sequenced macronuclear clone (cell lines ST 2 and 3) or entirely the alternate sequence (ST5). Thus,

Figure 3. Analyses of the empty sites generated by the excision of the 144 bp V3 and 374 bp V1 IESs. A.) Sequences at the ends of the 144 bp and 374 bp IES present in the micronuclear precursors of macronuclear V3 and V1, respectively. Bases that will form the macronuclear DNA molecules are shown in upper case, while the ends of the IESs are shown in lower case. The 5'TA3' direct repeats at the ends of the IESs are highlighted, with the left direct repeats arbitrarily indicated as the ones retained in the macronuclear DNA molecules. Bases that differ in the second allele of macronuclear V1 are indicated below the sequence. The sequence of the 144 bp V3 IES is shown in reverse orientation relative to the published sequence (11) to aid the comparison to the empty site analyses (below). B.) Sequences of the PCR products derived from the 144 bp V3 IES and 374 bp V1 IES empty sites in DNA from 50 hr developing macronuclei and macronuclear DNA from clonal vegetative cell lines (ST cell lines). The position of the 5'TA3' direct repeat that is retained at the empty site is indicated. For the 374 bp V3 IES empty site, a sequence polymorphism 2 bp following the empty site is also indicated. A second sequence polymorphism is indicated by arrowheads. For the 144 bp V3 IES empty sites, PCR was carried out using oligonucleotides V3MC and V3MA as primers, and oligonucleotide V3MA was used as the primer for dideoxy sequencing. The 50 hr DNA in this instance was gel-purified 3–5 kbp *Hind* III fragments from 50 hr whole cell DNA from a mating of cell lines 8 and 24 (see Materials and Methods). For the 374 bp V1 IES analyses, PCR was carried out using oligonucleotides V1MC and V1MT as the primers, and oligonucleotide V1MC was used as the primer for dideoxy sequencing. The 50 hr DNA was from developing macronuclei of a mating of cell lines CC102 and ST2.

the results are most easily interpreted as indicating that two alleles of V1 exist and that the boundaries of the IES are for the most part constant, with one direct repeat retained following excision.

Finally, the empty sites generated by the removal of the 31 bp IES in the precursor of macronuclear V2 were examined. The V2 IES empty site was PCR amplified with the primers V2MG and V2MT (Fig. 1). In all seven ST cell lines tested (ST1, 2, 3, 4, 5, 6, and 11), a clear sequence was obtained that contained one of the 5'TA3' direct repeats at the empty site (Fig. 4 and data not shown). As was the case for the V1 empty site analyses, some of the cell lines had two bands at some positions in the sequencing ladders or had an unexpected band at these same positions (Fig. 4). This again indicates the presence of two alleles. Indeed, the pattern of V2 allele segregation agrees precisely with that observed for V1, which is to be expected as the two loci are adjacent in the micronucleus (Fig. 1).

An unexpected result was found upon sequencing the PCR products derived from the V2 IES 'empty sites' in the 50 hr DNA

Figure 4. Analyses of the empty sites generated by the excision of the 31 bp V2 IES. A.) The sequence of the 31 bp V2 IES and flanking macronuclear DNA are shown. Bases that will form the V2 macronuclear DNA molecule are shown in upper case, while the IES is shown in lower case. The 5'TA3' direct repeats at the ends of the IES are highlighted, with the left direct repeat arbitrarily indicated as being the one retained in the macronuclear DNA molecule. Bases that differ in the second allele of macronuclear V2 are indicated below the sequence (note that allelic differences also appear to exist within the 31 bp IES, but are not shown). B.) Sequences of the PCR products derived from the 31 bp V2 IES empty sites in DNA from 50 hr developing macronuclei and macronuclear DNA from clonal vegetative cell lines (ST cell lines). The position of the 5'TA3' direct repeat that is retained at the empty site in the ST cell lines is indicated. For 50 hr DNA, the position of the retained IES is indicated. Arrowheads denote sequence polymorphisms. The 50 hr DNA was isolated from developing macronuclei of a mating of cell lines CC102 and ST2. PCR of all DNAs was carried out using the V2MG and V2MT oligonucleotides as primers. For the 50 hr DNA, the oligonucleotide V2MG was 5'-end labeled prior to PCR and the products were sequenced by the chemical method. The PCR products of the ST cell lines were sequenced by the dideoxy method using the V2MG oligonucleotide as a primer.

from developing macronuclei. At 50 hours, the sequence of the entire V2 IES was still observed in the PCR products (Fig. 4). This indicates that the 31 bp IES had not yet been excised in contrast to the other IES and *Tec1* elements examined. Some sequence polymorphisms also appear to be present within the 31 bp IES, but all could not be unambiguously resolved due to the use of the chemical sequencing method in this analysis.

Sensitivity of the PCR/Sequencing Technique

The sequencing gels from the above analyses revealed only a single readable sequence suggesting that excision is precise. Analyses were performed to determine the sensitivity of the technique in regard to detecting different types of heterogeneity in the excision process. The first possibility considered was that a major sequence is present, but a second sequence arising from an excision event with different boundaries was also present as a minor form. This minor form would give rise to essentially the same sequence, except that the addition or deletion of bases at the empty site would result in a shift of its sequence on the gel. To examine this possibility, selected sequencing gels were quantitated by analyzing the radioactivity that occurred in the G lane using a Betascope 603 blot analyzer (Betagen, Corp., Waltham, MA). Significant amounts of radioactivity were observed only at the expected G positions, and there was no evidence of a second out of phase sequence (Fig. 5). We have also carried out reconstruction experiments that involved mixing two different base-specific sequencing reactions of cloned DNA. The results indicated that a second minor sequence would have been detected visually on autoradiographs if it represented as little as 5% of the total (data not shown).

The second situation considered was that there are a large number of different empty sites in macronuclear DNA molecules aside from the major type, but each is present as a very small percentage of the total population. This situation would likely result in a general increase in the background on the sequencing gel owing to the many similar but shifted sequences. To obtain

Figure 5. Quantitation of a sequencing gel. The graph displays the amount of radioactivity vs. distance in the G lane (top to bottom) of the sequencing gel generated during the analysis of the D7 *Tec1* empty sites in DNA from cell line ST1 (see Fig. 2). The scan was performed using a Betascope 603 blot analyzer (Betagen, Corp., Waltham, MA). Black dots denote the positions at which an autoradiograph of the gel contained a visible band in the G lane. The arrowhead indicates the position of the 5'TA3' direct repeat that remains at the empty site.

a maximum estimate for the percentage of such variant forms in the population, selected sequencing gels were again quantitated using the Betagen scanner. In this case we determined the amount of radioactivity in a visible band (e.g. a 'G') and the amount of radioactivity in another lane (e.g., the 'C' lane) at the same position in the sequencing ladder. The latter value was then used to calculate the amount of background radioactivity in all four lanes of the sequencing gel. This analysis indicated that minor sequences could, at most, account for 22% of the molecules in the population. In all likelihood, this is an overestimate for the percentage of minor forms present, as the quality of the sequence information obtained in the empty site analyses was in most instances equivalent to that observed when sequencing cloned segments of DNA.

DISCUSSION

IES and Tec1 Excision Events Are Highly Reproducible

We have examined the empty sites in macronuclear DNA molecules, or their precursors, generated by the developmental excision of three IESs and two Tec1 transposon-like elements. In each case, our results indicate that most, if not all, excision events occur in the same manner and that one of the terminal direct repeats of the element is retained in the macronuclear DNA molecule. The excision process is very reproducible, as multiple cell lines generated from independent episodes of macronuclear development all displayed the same empty sites. It is likely that the IESs and Tec1 elements examined are representative of other elements in the micronuclear genome, as the cloned segments of micronuclear DNA in which they reside were essentially randomly selected (11), with no bias towards selecting elements that must undergo precise excision. Indeed, the three IESs examined do not appear to reside in coding regions, so that it is less likely that precise excision is required to preserve gene function in these cases.

Our results are similar to those reported by Hunter et al. (9). They used a PCR-based approach to provide evidence that the excision of both an IES and a TBE1 transposon-like element is highly reproducible in *Oxytricha fallax*. Their study, however, involved the analysis of a single vegetative cell line, and it remained formally possible that macronuclear DNA molecules with variant empty sites had assorted during vegetative growth. Since our analyses employed DNA isolated from cells very near the time of excision events, or cells that had recently completed macronuclear development, it is very unlikely that assortment of variantly spliced forms is occurring. Overall, these studies indicate that developmental DNA excision events are, in general, precise in the hypotrichs. The situation appears to be much different in *Tetrahymena thermophila*. Two developmentally excised DNA segments have been examined in detail in this organism. Both were found to generate different classes of empty sites that differ by a few base pairs (25,26).

Our results also provide a further indication that *E. crassus* is very fastidious in regard to carrying out the DNA rearrangement events that occur during macronuclear development. We had previously shown that one of the other major DNA rearrangement events of macronuclear development, chromosome fragmentation/telomere addition, occurs in a precise manner in this organism (27). All other ciliate species examined have been found to display microheterogeneity at chromosome fragmentation/telomere addition sites (8,27,28).

If the IES and Tec1 element excision process(es) are indeed

transposon based, precise excision is a major advantage in the context of the nuclear dualism displayed by the ciliates. The ability to precisely excise during the course of macronuclear development would allow the elements to insert essentially anywhere in the micronucleus, including coding regions, without deleterious consequences for the cell. This feature would presumably remove one of the constraints on maintaining a high copy number of a transposable element in the micronuclear genome.

Implications for the Mechanism of IES and Tec1 Excision

Previous work has demonstrated that both IESs and Tec1 elements exist as circular forms following their excision in the developing macronucleus (12,13). The majority of free IES circles were found to retain both 5'TA3' direct repeats at their circle junctions, separated by 10 bp that are apparently shared with the mature macronuclear DNA molecule (13; Fig. 6). The free circular Tec1 elements also appear to have the same type of structure (13, C.Jahn, pers. comm.). Our current results indicate that the vast majority of macronuclear DNA molecules retain one of the element's direct repeats at the site of excision. Thus, variability in the excision process can be ruled out as an explanation for the structure of the free circle junctions, and it is now clear that the major products of excision share sequence information.

The shared nucleotides suggest that staggered cuts in the DNA may initiate the excision process, and that these staggered cuts are filled-in in at least one of the excision products prior to the joining of the DNA strands. The precise nature of the staggered cuts is still obscure, as unequal numbers of base pairs from the regions flanking the left and right sides of the IES appear to be retained at the circle junction (13). More recently, we have obtained evidence that indicates two distinct sequences may be present between the direct repeats at the circle junctions: one composed primarily of the sequence flanking the left end of the IES, and the other primarily of right flanking sequence (L.Klobutcher, L.Turner, and J.LaPlante, unpublished results). This suggests that a novel molecular mechanism may be involved in IES and Tec1 excision. We are currently carrying out

Figure 6. Diagrammatic representation of the products of IES and Tec1 excision events. The top line shows an element (black rectangle) in its micronuclear chromosomal context, with its 5'TA3' terminal direct repeats indicated. Following excision, the empty site in the macronuclear precursor retains one of the element direct repeats. A free circular form of the element is also generated, which contains both direct repeats at the circle junction plus 10 bp that are shared with the macronuclear precursor.

additional analyses to more precisely define the nature of the circle junction and its origin, as this should provide further insights into the mechanism of excision.

Excision of the 31 bp V2 IES Occurs After 50 hr of Development

An unexpected result was obtained in the analysis of the empty sites left by the 31 bp V2 IES in DNA from developing macronuclei at 50 hr. Unlike the other IESs and Tec1 elements examined, the V2 IES was still present in the 50 hr DNA (Fig. 4). This result is not an artifact resulting from variation in the rate of macronuclear development because the 50 hr preparation of developing macronuclear DNA used in this analysis was the same one used in the analyses of the V1 Tec1 and 374 bp IES empty sites, and these elements are clearly excised by 50 hr. The latter results are also consistent with previous studies on the timing of IES and Tec1 excision (12,19). Although the results clearly indicate that removal of the V2 IES is delayed relative to other IESs, we do not yet know when it is excised during development. The small size of this IES, coupled with the fact that chromosome fragmentation and telomere addition begin at approximately 55 hrs of development (19), have made it difficult to design strategies to determine precisely when it is removed.

There are at least two potential explanations for the late removal of the V2 IES. First, there may be a continuum of times for the excision of particular elements. Although Tec1 elements and IESs appear related in many respects, Tec1 elements are excised prior to IESs (12,19). We have suggested that the same enzymatic machinery may be involved in both excision processes, but differences in the affinity of the element recognition sites may give rise to the differences in timing (13). According to this type of model, the 31 bp V2 IES might have a particularly poor recognition signal, so that its excision is delayed relative to the other IESs. Alternatively, there may be more than one type of DNA excision process occurring during macronuclear development. The 31 bp V2 IES is quite distinct from the other *E. crassus* IESs both in regard to its small size and its sequence, which resembles degenerate telomeric C₄A₄ repeats (Fig. 4A). In *Oxytricha nova* two classes of IESs have been defined on the basis of different, short regions of conserved sequence near their termini (4). Two or more classes of excised sequences may also exist in *Tetrahymena thermophila*. Godiska and Yao (29) have identified DNA sequences required for the developmental excision of one eliminated DNA segment in *Tetrahymena*, but similar sequences are not observed at other sites of DNA elimination. If two or more classes of excised DNA sequences exist in *E. crassus*, it will be of interest to determine if completely different sets of enzymatic machinery exist for their excision, or whether the two systems share components.

ACKNOWLEDGEMENTS

We thank Dr. A. Cowan for her comments on the manuscript. This work was supported by Public Health Service grant GM33277 and a University of Connecticut Health Center Research Advisory Committee grant to L.A.K.

REFERENCES

1. Klobutcher, L.A. and Prescott, D.M. (1986) In Gall, J. (ed.), *The Molecular Biology of Ciliated Protozoa*. Academic Press, New York, pp. 111–154.
2. Yao, M.-C. (1989) In Berg, D.E. and Howe, M.M. (eds.), *Mobile DNA*. American Society for Microbiology, Washington, D.C., pp. 715–734.

3. Klobutcher, L.A., Jahn, C.L. and Prescott, D.M. (1984) *Cell*, **36**, 1045–1055.
4. Ribas-Aparicio, R.M., Sparkowski, J.J., Proulx, A.E., Mitchell, J.D. and Klobutcher, L.A. (1987) *Genes & Dev.*, **1**, 323–336.
5. Herrick, G., Cartinhour, S.W., Williams, K.R. and Kotter, K.P. (1987) *J. Protozool.*, **34**, 429–434.
6. Greslin, A.F., Prescott, D.M., Oka, Y., Loukin, S.H. and Chappell, J.C. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6264–6268.
7. Herrick, G., Cartinhour, S., Dawson, D., Ang, D., Sheets, R., Lee, A., and Williams, K. (1985) *Cell*, **43**, 759–768.
8. Herrick, G., Hunter, D., Williams, K. and Kotter, K. (1987) *Genes & Dev.*, **1**, 1047–1058.
9. Hunter, D.J., Williams, K., Cartinhour, S. and Herrick, G. (1989) *Genes & Dev.*, **3**, 2101–2112.
10. Jahn, C.L., Nilles, L.A., and Krikau, M.F. (1988) *J. Protozool.*, **35**, 590–601.
11. Baird, S.E., Fino, G.M., Tausta, S.L. and Klobutcher, L.A. (1989) *Mol. Cell. Biol.*, **9**, 3793–3807.
12. Jahn, C.L., Krikau, M.F. and Shyman, S. (1989) *Cell*, **59**, 1009–1018.
13. Tausta, S.L. and Klobutcher, L.A. (1989) *Cell*, **59**, 1019–1026.
14. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) *Science*, **230**, 1350–1354.
15. Roth, M., Lin, M. and Prescott, D.M. (1985) *J. Cell Biol.*, **101**, 79–84.
16. Heckmann, K. (1964) *Z. Vererbungsl.*, **95**, 114–124.
17. Baird, S.E. and Klobutcher, L.A. (1988) *J. Protozool.*, **35**, 459–465.
18. Klobutcher, L.A., Swanton, M.S., Donini, P. and Prescott, D.M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3015–3019.
19. Tausta, S.L. and Klobutcher, L.A. (1990) *Nucleic Acids Res.*, **18**, 845–853.
20. Kuhn, S., Fritz, H.-J. and Starlinger, P. (1979) *Mol. Gen. Genet.*, **167**, 235–241.
21. Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
22. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor, New York.
23. Kusukawa, N., Uemori, T., Asada, K. and Kato, I. (1990) *Biotechniques*, **9**, 66–72.
24. Fawcett, T. W. and Bartlett, S. G. (1990) *Biotechniques*, **9**, 46–48.
25. Austerberry, C. F. and Yao, M.-C. (1987) *Mol. Cell. Biol.*, **7**, 435–443.
26. Austerberry, C. F., Snyder, R. O. and Yao, M.-C. (1989) *Nucleic Acids Res.*, **17**, 7263–7272.
27. Baird, S. E. and Klobutcher, L. A. (1989) *Genes & Dev.*, **3**, 585–597.
28. Forney, J. D. and Blackburn, E. H. (1988) *Mol. Cell. Biol.*, **8**, 251–258.
29. Godiska, R. and Yao, M. C. (1990) *Cell*, **61**, 1237–1246.