SURVEY AND SUMMARY

The evolutionary scrambling and developmental unscrambling of germline genes in hypotrichous ciliates

David M. Prescott*

Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309-0347, USA

Received July 17, 1998; Revised and Accepted November 20, 1998

ABSTRACT

Genes in the germline (micronuclear) genome of hypotrichous ciliates are interrupted by multiple, short, non-coding, AT-rich sequences called internal eliminated segments, or IESs. During conversion of a micronucleus to a somatic nucleus (macronucleus) after cell mating, all IESs are excised from the germline genes and the gene segments, called macronuclear-destined segments, or MDSs, are spliced. Excision of the ∼150 000 IESs from a haploid germline genome in Oxytricha nova requires ∼150 000 recombinant events. In three of 10 genes the MDSs are scrambled. During macronuclear development the MDSs are unscrambled, possibly by folding of the DNA to allow MDSs to ligate in the correct order. The nine MDSs in the actin I gene of O.nova are scrambled in the random order, 3–4–6–5–7–9–2–1–8, and MDS 2 is inverted. The 14 MDSs in the αTP gene of O.nova and Stylonychia mytilus are scrambled in the non-random order, 1–3–5–7–9–11–2–4–6–8–10–12–13–14. The 45 MDSs in the DNA pol α gene are non-randomly scrambled into an odd/even series, with an inversion of one-third of the gene. Additional IESs have been inserted into these three genes during evolution of Oxytricha trifallax, slightly modifying scrambling patterns. The non-random scrambled patterns in the αTP and DNA pol α genes are explained by multiple, simultaneous IES insertions. The randomly scrambled pattern in the actin I gene may arise from an initially non-randomly scrambled pattern by recombination among multiple IESs. Alternatively, IESs inserted sporadically (individually) in a non-scrambled configuration might subsequently recombine, converting a non-scrambled gene into a randomly scrambled one. IESs shift along a DNA molecule, most likely as a result of mutations at MDS/IES junctions. Shifting of IESs has the effect of ‘transferring’ nucleotides from one MDS to another, but does not change the overall sequence of nucleotides in the combined MDSs. In addition to shifting in position, IESs accumulate mutations at a high rate and increase and decrease in length within a species and during speciation. The phenomena of IESs and of MDS scrambling represent remarkable flexibility of the hypotrich genome, possibly reflecting a process of MDS shuffling that facilitates the evolution of genes.

INTRODUCTION

The DNA in hypotrichous ciliates has undergone extraordinary modifications in organization and behavior during evolution. The story of these DNA modifications is far from complete and what they mean for the evolution and genetic operation of ciliates is still not clear. However, at the very least, observations thus far attest to a remarkable flexibility in the organization and processing of DNA in the evolution and function of hypotrich genomes.

IESs AND MDSs

The germline genes in the micronucleus of hypotrichous ciliates are interrupted by multiple, non-coding sequences called internal eliminated segments, or IESs (1). In the example shown in Figure 1a, the germline gene that encodes β telomere-binding protein (βTP) in Oxytricha trifallax is split into seven parts by six IESs, one in the 5′-non-translated leader region and five in the open reading frame (ORF). The 14 MDSs in the DNA pol α gene are non-randomly scrambled into an odd/even series, with an inversion of one-third of the gene. Additional IESs have been inserted into these three genes during evolution of Oxytricha trifallax, slightly modifying scrambling patterns. The non-random scrambled patterns in the αTP and DNA pol α genes are explained by multiple, simultaneous IES insertions. The randomly scrambled pattern in the actin I gene may arise from an initially non-randomly scrambled pattern by recombination among multiple IESs. Alternatively, IESs inserted sporadically (individually) in a non-scrambled configuration might subsequently recombine, converting a non-scrambled gene into a randomly scrambled one. IESs shift along a DNA molecule, most likely as a result of mutations at MDS/IES junctions. Shifting of IESs has the effect of ‘transferring’ nucleotides from one MDS to another, but does not change the overall sequence of nucleotides in the combined MDSs. In addition to shifting in position, IESs accumulate mutations at a high rate and increase and decrease in length within a species and during speciation. The phenomena of IESs and of MDS scrambling represent remarkable flexibility of the hypotrich genome, possibly reflecting a process of MDS shuffling that facilitates the evolution of genes.
are activated. Although removal of IESs, excision of genes from chromosomes and activation of genes are closely related temporally, there is no evidence to suggest that these events are related mechanistically. Some progress has been made in understanding the molecular mechanism of IES excision (2,3) and gene excision/telomere addition (4–6), but nothing is known about the molecular mechanism that maintains the silence of the micronuclear genome.

IESs are randomly distributed in micronuclear genes, occurring in the 5′ leader, the ORF, and/or the 3′ trailer. IESs are generally short; >95% are <100 bp (7). The shortest IESs detected so far in hypotrichs are 5 bp in length (8), while the longest is 595 bp (7). IESs are AT-rich (75–100% AT), unique sequences with no consistent distinguishing structural characteristics, e.g. palindromes or consensus sequences. Pairs of 2 to 7 bp repeat sequences that are part of the MDSs (Fig. 1c) are present at the junctions of an IES with its adjacent MDSs. These repeats probably help direct MDS splicing. IESs are excised during macronuclear development in Euplotes crassus by staggered cuts at the two IES/MDS junctions such that the IES and one copy of the repeat pair are removed (2). The other copy of the repeat remains at the splice site between the two adjacent MDSs.

SCRAMBLED MICRONUCLEAR GENES

The structures of 10 different micronuclear genes and their macronuclear counterparts have been completely characterized in various species of Oxytricha and Stylonychia (1,8–9,17). In three of these genes, encoding actin I, α telomere-binding protein (αTP) and DNA polymerase α (DNA pol α), multiple MDSs have been rearranged into a scrambled order during evolution.

The actin I gene

The actin I gene in Oxytricha nova (Fig. 2f) consists of eight IESs and nine MDSs in the order 3′–4′–5′–6′–7′–9′–2′–1′–8′, with MDS 2 inverted (15). MDSs 3 and 4 are not scrambled relative to each other, but the other seven MDSs are scrambled in largely random order. How did these MDSs become scrambled? One possibility is the relatively simple model in Figure 2. The first step consists of insertion of eight IESs simultaneously or separately in time into the actin I gene (Fig. 2b). This creates nine MDSs in the non-scrambled order 1′–2′–3′–4′–5′–6′–7′–8′–9′. Subsequently, four hypothetical intragenic recombinations among the IESs shown in Figure 2 create the contemporary scrambled state of the actin I germline gene. These postulated scrambling events are not necessarily intragenic as pictured in the model in Figure 2; both alleles in the micronucleus may have participated.

Scrambled genes continue to undergo evolutionary change. This is affirmed by comparing the structures of particular genes among different species. The structural arrangements of MDSs and IESs in the complete actin I gene in two species, O.nova (15) and O.trifallax (11), and in part of the actin I gene in Oxytricha sp.? (Aspen) (7), are similar, but there are several major differences (Fig. 3). Because of rapid accumulation of mutations, corresponding IESs from species to species share no sequence identity and differ considerably in length. In addition, positions of IESs within the actin I gene are shifted by one to many base pairs from species to species. Shifting of IES positions is explained by a mutational model (11) in which nucleotides at MDS/IES junctions are ‘shifted’ one or two at a time between an MDS and an IES. This phenomenon is illustrated in Figure 4a with the junctions between a hypothetical IES and its flanking MDSs. A mutation at the right junction shortens the IES by one base and a single mutation at the left junction lengthens the IES by two bases and shortens the MDS by two bases. As the junctions shift the sequence of the pair of repeats changes, but the coding sequence generated by splicing the MDSs during macronuclear development remains the same. This mutational model of IES shifting is applied in Figure 4b to the junction of MDS 4 with IES 2 and the junction of IES 3 with MDS 5 in the scrambled actin I gene of O.nova, O.trifallax and Oxytricha sp.? (Aspen) (Fig. 4c). The shifting of IES/MDS junctions has major structural consequences for the micronuclear actin I gene. The sequences of the

Figure 1. (a) Diagram of the micronuclear gene encoding βTP in O.trifallax. Six IESs (lines) separate seven gene segments (clear boxes labeled MDS 1–7). The black box is an intron. (b) Diagram of the macronuclear DNA molecule encoding βTP. The vertical lines in the clear box indicate where MDSs have been spliced after IES removal. The striped boxes at the ends are telomeres, with 3′ 16 base overhangs. (c) IES 3 (lower case letters) and its junctions with MDSs 3 and 4 (upper case letters) in the micronuclear βTP gene of O.trifallax. The 4 bp repeat sequence (ACTTC) is underlined. From Prescott and DuBois (7).

Figure 2. A model describing the origin of scrambling of the actin I gene in O.nova. (a) The micronuclear gene prior to invasion by IESs. (b) Insertion of eight IESs, creating nine non-scrambled MDSs. (c) Inversion and translocation of MDS 2 to the 3′-end of the gene. (d) Translocation of MDSs 2 and 1 to the 3′-end of the gene. (e) Translocation of MDS 8 to the 3′-end of the gene. (f) Translocation of MDS 6 to a position between MDSs 4 and 5, producing the contemporary arrangement of MDSs in the actin I gene. MDSs are clear blocks and IESs are lines between MDSs.
The actin I gene in *O.nova*, *O.trifallax* and *Oxytricha* sp.? (Aspen). The three genes are scrambled in similar, largely random patterns. MDS 2 is inverted (→) in all three species. The *O.trifallax* gene contains an additional IES and MDS. MDSs are clear blocks. IESs are lines between MDSs. TAS, telomere addition site. From Prescott and DuBois. The scrambled pattern of the gene with a closed or open loop of A T-rich, non-coding DNA (most of the germline genome is made up of such sequences) is illustrated in Figure 4, beginning with an odd/even pattern. In *O.nova*, 5 mutations show how a single mutation subtracts two bases from MDS 4 and adds them to the left end of IES 2 in *O.trifallax*. At the right end of IES 3 in *O.nova* two mutations shorten MDS 5 by two bases and shorten the IES by two bases in *O.trifallax*. Additional hypothetical mutations show how the sequences at junctions in *Oxytricha* could progress to the junction sequences in *Oxytricha* sp.? (Aspen). The sequences of the repeat pairs progressively change and the lengths of the MDSs change as the junctions shift. An overall effect is that coding sequence is ‘transferred’ from MDS 4 to MDS 5, a process that does not change the coding sequence. Removal of the IES during macronuclear development yields the same sequence in all three species. All of the mutational changes could have occurred in the opposite direction and in a different order among the organisms. The *O.nova* data are from Greslin et al. (25), the *Oxytricha* data are from DuBois and Prescott (11) and the *Oxytricha* sp.? (Aspen) data are from Prescott and DuBois. The scrambled αTP gene

Unlike the actin I gene, the scrambled pattern of the gene encoding αTP is non-random. The micronuclear αTP gene consists of 14 MDSs in *O.nova* (13) and in *Styloynchia mytilus* (8) in the non-random order 1–3–5–7–9–11–2–4–6–8–10–12–13–14 (Fig. 5a). The gene in *Oxytricha* contains 17 MDSs in the order 1–3–5–7–10–12–2–4–6–8–9–11–13–14–15–16–17. (An intron interrupts MDS 2 between codons 35 and 36 in all three species.) From the similarity of the scrambled patterns, it is evident that the αTP gene became scrambled in an ancestor of the three species in the pattern currently shared by *O.nova* and *S.mytilus*. The pattern in *Oxytricha* is somewhat different and very likely resulted from insertions of IESs 10, 14 and 15 after divergence of *Oxytricha* from *O.nova* and *S.mytilus*. IES 10 in *Oxytricha* has been inserted into the original MDS 8, dividing it into original MDS 8 and 9. As a result, MDS and IES numberings change in *Oxytricha* for all MDSs and IESs that numerically follow MDS 8 and IES 10. IESs 14 and 15 were likely inserted into the original MDS 13, creating MDSs 14–16 in *Oxytricha*. Unscrambling and ligation of MDSs produces three very similar macronuclear genes (Fig. 5b).

Unlike the actin I gene, none of the MDSs in αTP are inverted and most of the MDSs in the αTP gene are distributed in a strikingly non-random, odd/even pattern. In *O.nova* and *S.mytilus* the exceptions are the non-scrambled MDSs 12–14 and in *Oxytricha* the exceptions are the non-scrambled (relative to one another) MDSs 8 and 9 and non-scrambled MDSs 12–15. Like the actin I gene, IESs that correspond from species to species are completely different in sequence, differ considerably in length and have shifted positions along the DNA. As a consequence, the lengths of MDSs, and therefore the distribution of coding sequence among the MDSs, are changed, but these MDS changes do not alter the coding sequence. This is illustrated in Figure 4 with MDSs 4 and 5 in the actin I gene in three species of *Oxytricha*. Although MDSs 4 and 5 differ in length from species to species, removal of the IES and one copy of the repeat during macronuclear development yields the same sequence in all three species. A model of the origin of non-random scrambling

The odd/even pattern of MDSs in the αTP gene implies that scrambling most likely took place in a single event in an ancestor common to the three species. A model of this event is shown in Figure 6, beginning with an αTP gene that lacks IESs. Alignment of the gene with a closed or open loop of AT-rich, non-coding DNA (most of the germline genome is made up of such sequence), followed by multiple, simultaneous recombinations between the αTP gene and the AT-rich loop (Fig. 6c), creates an odd/even series of scrambled MDSs 1–12. In the model the pairs of repeats are created essentially as described previously (7) for a non-scrambled gene with IESs. First, MDSs are formed by staggered cuts in the αTP gene in Figure 6a, as illustrated for
The DNA pol α gene

The gene encoding DNA pol α represents a third pattern of scrambling, one that is more complex than those in the actin I and αTP genes. The DNA pol α gene consists of 45 MDSs and 44 IESs in O.nova (13) and 51 MDSs and 50 IESs in O.trifallax (Fig. 7; 18). Almost all of the MDSs are arranged in a non-random, odd/even pattern that is disrupted by inversion of approximately one-third of the gene with respect to the other two-thirds in both species. Like the actin I and αTP genes, a few MDSs (e.g. MDSs 26 and 27, MDSs 44 and 45 and MDSs 1–3 in the DNA pol α gene of O.nova) are not scrambled relative to one another. Essentially the same model for the origin of non-random scrambling of MDSs in the αTP gene in Figure 6 also accounts for the non-random, odd/even pattern of MDSs in the DNA pol α gene. The final step in scrambling in O.nova is recombination between the IESs flanking MDSs 1 and 4 to create the large inversion. The second inversion point is in DNA flanking one or the other end of the gene. As in the αTP and actin I genes, the non-scrambled MDSs probably originated by insertion of individual IESs in unrelated events (sporadic insertions) after gene scrambling, e.g. splitting of an original MDS 26 in the DNA pol α gene in O.nova (Fig. 7) into non-scrambled MDSs 26 and 27 by insertion of a new IES.

The six additional MDSs in the DNA pol α gene of O.trifallax are most parsimoniously explained by insertion of the six IESs after evolutionary divergence between O.nova and O.trifallax. Four new MDSs were created in O.trifallax, after the initial, non-random scrambling of the DNA pol α gene, by a mechanism that preserved the odd/even pattern of MDSs in that region of the gene. In a model of this mechanism (Fig. 8; 12), the original MDS 6, which is still present as MDS 6 in the DNA pol α gene of O.nova, was split into MDSs 6a–6c in O.trifallax by the insertion of two new IESs. Because of the inversion in the gene between MDS 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8). Recombination between the two new IESs and the IES between original MDSs 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8). Recombination between the two new IESs and the IES between original MDSs 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8). Recombination between the two new IESs and the IES between original MDSs 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8). Recombination between the two new IESs and the IES between original MDSs 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8). Recombination between the two new IESs and the IES between original MDSs 1 and 4, the molecule can be folded into a hairpin such that the MDSs in each branch of the hairpin face in the same direction with respect to transcription (arrows in Fig. 8).

A prerequisite for the model in Figure 8 is folding of the DNA pol α gene into a hairpin, as shown in Figure 10. Such folding might be prescribed by alignment of the homologous sequences in the pairs of repeats at the ends of numerically consecutive MDSs. This places the newly created MDSs in position for recombination that transfers a new MDS (MDS 6b in Fig. 8) to the opposite branch of the hairpin.

Comparison of DNA pol α gene structure between O.nova and O.trifallax demonstrates that this gene has undergone the same type of evolutionary changes that occurred in the actin I and αTP genes. Corresponding IESs are completely different in sequence, differ substantially in length and have shifted along the DNA molecule. IES shifting has changed the lengths of corresponding MDSs, in some cases by >50 bp. Last, six more IESs have been inserted into the O.trifallax gene. Thus, there is a consistent pattern of evolutionary change in all three scrambled genes.
Figure 6. A hypothesis to account for the origin of the non-randomly scrambled structure of the micronuclear αTP genes in O.nova, S.mytilus and O.trifallax. (a) A loop of AT-rich DNA is aligned with an αTP gene that contains no IESs. (b) Staggered cuts are made in the αTP gene, followed by recombination of sections of AT-rich DNA with the single-stranded overhangs resulting from staggered cuts. Fill-in of single-strand gaps to create repeat pairs (R1, R2, etc.) (c) Insertion of multiple IESs at staggered cuts in the αTP gene (as shown in b) creates 11 IESs separating MDSs 1–12. MDSs 13 and 14 are created by separate insertions. (d) The regular odd/even pattern of MDSs in O.nova and S.mytilus evolve into the O.trifallax pattern by insertion of three additional IESs, dividing MDS 8 into MDSs 8 and 9 and dividing MDS 13 into MDSs 14–16. MDS 14 in O.nova/S.mytilus becomes MDS 17 in O.trifallax. From Prescott et al. (8).

Oxytricha trifallax contains one, three and six more IESs in the actin I, αTP and DNA pol α genes, respectively, than does O.nova. Similarly, the non-scrambled genes encoding βTP and histone H4 (ORF only) contain six and one IESs, respectively, in O.trifallax compared with three and none in O.nova. The greater number of IESs in five out of the five genes characterized so far implies that O.trifallax (82 IESs) has a greater propensity for insertion of IESs than does O.nova (68 IESs).

Scrambling of the DNA pol α gene presents another complexity. MDSs 29–31–33–35–37–39–41–43, encoding 230 bp of the 4479 bp in the ORF in O.nova, became separated from the main body of the gene after scrambling had taken place. These missing MDSs are not present in the first several kilobases flanking either end of the gene and their chromosomal location remains unknown. This same portion of the gene, consisting of 193 bp in MDSs 34–36–38–40–42–44, is also missing from the main body of the gene in O.trifallax. Translocation of this segment of the gene has occurred in an ancestor common to O.nova and O.trifallax after the scrambling event that created the odd/even pattern. A large inversion could have displaced the missing MDSs more than several kilobases into the DNA flanking either end of the gene. Most of these missing MDSs in O.trifallax have been identified by PCR as a group in the genome, but their chromosomal location relative to the bulk of the gene is still unknown. The difference in the amount of translocated coding sequence (230 versus 193 bp) can be accounted for by IES shifting subsequent to scrambling. IES shifting has had the effect of “transferring” coding sequence from the translocated MDSs to MDSs in the main body of the gene or vice versa, as illustrated by the “transfer” of coding sequence between MDS 4 and 5 in the actin I gene in Figure 4.

ORIGIN OF RANDOM SCRAMBLING

IESs may have been originally inserted into the actin I gene in sporadic fashion, without producing scrambling. This must have happened in the non-scrambled βTP, C2 and R1 genes of O.nova, the βTP gene of O.trifallax and S.mytilus and the MA83-A gene of Oxytricha fallax. Random scrambling of the actin I gene may have subsequently been produced by intragenic IES recombination as already described in Figure 2. Alternatively, the actin I gene
Figure 7. Diagrams of the scrambled DNA pol α genes in *O. trifallax* and *O. nova*. The genes continue on two lines. MDSs are clear boxes and IESs are lines between MDSs. MDSs 32 and 51 of the *O. trifallax* gene are drawn as open-ended boxes because the sequences at the 5’- and 3’-ends have not been determined. MDSs 29, 31, 33, 35, 37, 39, 41 and 43 in the *O. nova* gene and MDSs 34, 36, 40, 42, 44, 46 and 48 are not present with the main body of the gene. An inversion is present in the same location within the gene in both species. The second inversion point is not known but must be present in the DNA that flanks one or the other end of the micronuclear gene. A suggested location in the DNA flanking MDS 43 in *O. nova* is shown in Figure 10. From Hoffman and Prescott (12).

Figure 8. A model for the introduction of additional non-randomly scrambled MDSs into the DNA pol α gene in *O. trifallax*. An inversion is present between MDSs 1 and 4, so the molecule is drawn as a hairpin to put MDSs in the same right–left direction with respect to transcription. Arrows indicate the 5’→3’ direction of the coding strand. (A) Two new IESs are inserted into original MDS 6 (carets). (B) MDS 6 has been divided into MDSs 6a–6c. (C) Two recombinations between the two new IESs and the IES between MDSs 5 and 7 inserts MDS 6b between MDSs 5 and 7. (D) MDS 6a becomes MDS 6, MDS 6c becomes MDS 8, and MDS 6b becomes MDS 7, maintaining the odd/even pattern of MDSs of the *O. trifallax* gene. From Hoffman and Prescott (12).
Figure 9. Possible origin of the largely random, scrambled arrangement of MDSs in the actin I gene of *O. nova*. (a) The gene was originally scrambled in an odd/even pattern by the mechanism shown in Figure 6. (b) Recombination 1 between IES 3 and flanking DNA displaces MDSs 7 and 2 to the right of MDS 8. (c) Recombination 2 inverts MDS 2 and inserts it between MDSs 8 and 7. (d) Recombination 3 inserts MDS 1 between MDSs 2 and 7. (e) Insertion of a new IES in MDS 3 splits MDS 3 into MDS 3 and a new MDS 4. Old MDSs 5, 4, 6, 8 and 7 are now renumbered as 6, 5, 7, 9 and 8 to yield the current structure of the actin I gene in *O. nova*.

may have been originally scrambled in a non-random pattern by the concerted mechanism illustrated for the αTP gene in Figure 6. Subsequently, the gene evolved to a random pattern as described in Figure 9. By a series of three transpositions, inversion of MDS 2 and insertion of an additional IES, an original odd/even pattern of 1–3–5–7–2–4–6–8 could have been converted to the current pattern of 3–4–6–5–7–9–2–1–8 in the actin I gene. Analysis of the micronuclear actin I gene in other hypotrich species might provide a test of this model of how the scrambled state of this gene in *O. nova* evolved.

CONVERSION OF A MICRONUCLEUS INTO A MACRONUCLEUS

Extrapolation from the number of IESs in seven micronuclear genes in *O. nova* gives an estimate of ~150 000 IESs/haploid genome. Removal of these IESs and assembly of the macronuclear genome during development from micronuclear DNA requires four types of processes (19). (i) Removal of the IESs via ~150 000 recombination events. Since IES excision occurs during the polytene chromosome stage of macronuclear development, the number of recombinational excisions of IESs is much greater than 150 000, potentially as great as ~10^7. (ii) In at least some hypotrichs, one or more large families of transposon-like elements are also excised and destroyed during the polytene chromosome stage (20–22). (iii) In association with IES excision, scrambled MDSs are unscrambled and ligated. (iv) Spacer DNA between genes in micronuclear chromosomes is selectively destroyed, releasing each gene as an individual, short DNA molecule to which telomeric sequences are added.

Little is known about how these DNA processing events are directed. IES removal requires that the excision mechanism must be able to identify the two ends of the IES. The mechanism must be able to cut the DNA at the MDS/IES junctions and splice the MDSs. The 2–7 bp repeats flanking IESs that separate non-scrambled MDSs (Fig. 1) contain inadequate information for IES recognition. No other discernible, consistent structural features of IESs or MDS/IES junctions in hypotrichs have been observed. In contrast, repeats that flank IESs between scrambled MDSs are longer (average 11 bp) than those between non-scrambled MDSs (average 4 bp). Theoretically, the longer repeats are sufficient to guide folding of the DNA molecule directed by alignment of repeat pairs [as proposed for unscrambling of the actin I, αTP, and DNA pol α (Fig. 10) genes]. In this case, homologous recombination between the two aligned repeats in a pair eliminates the IES plus one copy of the repeat and splices MDSs in the orthodox order. In addition to solving the unscrambling problem, the folding model may also eliminate the need for specific recognition of IESs. IES removal simply becomes a by-product of MDS splicing. However, the stiffness of the DNA duplex must be reduced, perhaps by proteins, because most of the IESs and many of the MDSs are too short to fold without assistance. The folding model may not apply in non-scrambled situations because the repeats are too short to guide folding unambiguously. This is illustrated by the 292 bp IES 6 in the βTP gene of *O. trifallax* (Fig. 1a). IES 6 is flanked by the 3 bp repeat AGT, a trinucleotide occurring five times within the IES, and at 5, 81 and 91 bp downstream of the IES and 44 bp upstream. Additional information of unknown form and source is required to direct the excision mechanism to the correct trinucleotides. Last, the difference in
repeat lengths between scrambled and non-scrambled MDSs suggests that the two configurations of MDSs and IESs might be processed by fundamentally different mechanisms.

An alternative to the folding model is suggested by observations on IES removal in \textit{Paramecium}. In this case the old macronucleus appears to transmit sequence information to the developing macronucleus, perhaps effecting IES excisions by a template mechanism \cite{23}. In hypotrichs, DNA molecules transmitted from the old macronucleus to the developing macronucleus could solve three related requirements in DNA processing: IES excision, unscrambling of MDSs and excision of genes from chromosomes (although no template mechanism for these three processing events is easily visualized).

The presence of IESs and scrambled MDSs in hypotrich genes raises many questions. Did IESs originate as transposons that then degenerated to their current simple structure, as proposed for the old macronucleus to the developing macronucleus could solve three related requirements in DNA processing: IES excision, unscrambling of MDSs and excision of genes from chromosomes (although no template mechanism for these three processing events is easily visualized).

The presence of IESs and scrambled MDSs in hypotrich genes raises many questions. Did IESs originate as transposons that then degenerated to their current simple structure, as proposed for the old macronucleus to the developing macronucleus could solve three related requirements in DNA processing: IES excision, unscrambling of MDSs and excision of genes from chromosomes (although no template mechanism for these three processing events is easily visualized).

**Figure 10.** Folding of the DNA pol α gene of \textit{O. nova} to align pairs of repeat sequences that flank IESs. MDSs are heavy lines and IESs are thin lines. A repeat pair is present wherever MDSs in the two branches of the hairpin overlap at their ends. Recombination between each pair of repeats splices MDSs in the unscrambled order, removes the IESs and removes one copy of the repeat in each pair. From Hoffman and Prescott (12). An inversion point is present between MDSs 1 and 4. The other end of the inversion is shown in DNA flanking MDS 43, but theoretically it could be in flanking DNA at the other end of the gene.

short, flanking repeats. Thus, the IES excision mechanism is vastly different from what is known about transposon excision.

Hypotrichous ciliates are very successful, ubiquitous eukaryotes, thriving in virtually every body of water on Earth. Whether the splitting of germline genes into MDSs by IESs and whether MDS scrambling are innovations that somehow contribute in any way to this success is not obvious. The presence of IESs is innocuous since germline genes are genetically silent and because all IESs are removed when a macronucleus is formed. MDS scrambling, which serves no discernible purpose, is made possible by the presence of IESs. Scrambling is tolerable because its resolution is an integral part of IES removal. At the very least, the phenomena of IESs and of MDS scrambling demonstrate a remarkable flexibility in the organization and behavior of hypotrich DNA sequences. MDS scrambling may, in fact, be only a surface reflection of a more profound process in which MDS shuffling facilitates and accelerates the evolution of genes.

**ACKNOWLEDGEMENTS**

This work was supported by NIGMS Research Grant GM56161 to D.M.P. The author is grateful to Gayle Prescott for preparation of the manuscript and Jason Prescott for preparation of figures.

**REFERENCES**