A transposon-related palindromic repetitive sequence from *C.elegans*

David H.Dreyfus and Scott W.Emmons
Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

Received November 26, 1990; Revised and Accepted March 8, 1991 EMBL accession no. X55356

ABSTRACT

A family of transposon-like sequences in the *C. elegans* genome is described. This family, termed the Tc6 family, consists mostly of conserved, 1.6 kb elements. Four Tc6 elements or partial elements have been cloned and the DNA sequences of three were determined. One appears to be a complete element of 1603 nucleotides, consisting of a palindrome of 765 nucleotides, with a central, non-palindromic region of 73 nucleotides. Another has an identical structure except for an internal deletion. A third is a partial element terminating at a probable internal restriction site used for cloning. A fourth clone contained portions of the Tc6 sequence juxtaposed to non-Tc6 sequences. All *C. elegans* strains examined contain 20–30 Tc6 elements. The ends of Tc6 elements are conserved and have sequence similarity to the ends of *C. elegans* transposons Tc1 and Tc3. The ends of Tc6 elements also have sequence similarity to the heptamer portion of the immunoglobulin and T-cell receptor recombination signal sequence, raising the possibility of wide phylogenetic conservation of the recombination mechanism. Tc6 elements also share sequence motifs with plant-pathogenic viroid RNA's, possibly indicative of a Tc6 RNA replicative phase.

INTRODUCTION

We have characterized repetitive sequences of the *C. elegans* genome with the aim of identifying the transposable elements of this organism. Many aspects of the biology of *C. elegans*, including organization and function of the genome, are being investigated in a number of laboratories (1), and provide a useful background for this research. Studies on the genome include the construction of a complete physical map, correlated with the genetic map (2, 3). A physical map will allow a thorough investigation of the structure and relationships of the various classes of unique, repetitive, and transposable genomic sequences.

C. elegans transposable elements have been identified both as strain polymorphisms and as the agents of spontaneous mutations. Results of studies on these sequences have been reviewed by Moerman & Waterston (4). The most prominent and well-studied *C. elegans* transposon is Tc1, a 1.6 kb mobile element with 54 nucleotide inverted terminal repeats and an open reading frame encoding a putative product of 273 amino acids (5,6,7). Tc1 is the prototype of a family of similar 1.6 kb elements, including the TCb1 and TCb2 elements of the nematode *C. briggsae* (8), the HB elements of *Drosophila melanogaster* (9,10,11), and the UHU elements of Hawaiian *Drosophila* species (12). Two additional elements in the *C. elegans* genome, Tc2 (13), and Tc3 (14), have been proposed to share with Tc1 components of a common transposition regulatory pathway, and the Tc3 element shares terminal nucleotides with Tc1.

We report here the characterization of a putative new transposon family identified as the basis of a strain polymorphism. Elements of this new family also share terminal nucleotides with Tc1. The structure of the new elements is that of an inverted repeat, which contains no significant open reading frames, and which is, apart from the shared terminal sequences, unrelated in sequence to Tc1. The structure of these palindromic elements is also unlike previously described palindromic transposons of *Drosophila* (15) and sea urchin (16), but has suggestive similarity to pathogenic RNA molecules.

MATERIALS AND METHODS

Nematode Strains and Transposable Elements

C. elegans strain Bristol (N2) was obtained from D. Hirsh; Bergerac (BO) from the Caenorhabditis Genetics Center; Ga1, Ga9, and Pa1 from J. Rand, C. Johnson and R. Russell; TR679 from J. Collins (17); and RW7406 from I. Mori (18). Nematodes were cultured on petri plates by standard methods (19).

We obtained *C. elegans* transposons Tc3 and Tc5 from the laboratory of P. Anderson, and Tc4 from the laboratory of R. Horvitz. Tc1 and Tc2 are from this laboratory.

Isolation of Tc6 Elements

Cloned DNA fragments used in this work are shown in Figure 1.

Tc6.1 was isolated from a library of genomic DNA of the *C. elegans* strain Bergerac BO. The library was constructed by double cleavage with restriction enzymes *Bam*HI and *Sal*I and ligation to the plasmid vector pBR322 (20), which was digested with the same two enzymes. Recombination deficient *E. coli* strains HB101 and DH5a were used to propagate the library (see 21 for genotypes). The library was probed with plasmid EM # 99, containing a restriction fragment from the *C. elegans* Bristol strain subcloned from plasmid EM # 12 (pCe14 of 22). The probe was radiolabelled by nick-translation. The screen yielded plasmid EM # 58.

Tc6.2 and Tc6.3 were isolated from a library of size-fractionated (2kb to 10kb) *Eco*RI fragments of *C. elegans* strain Bristol (N2) DNA in plasmid pUC18 (23). The probe was the *Bam*HI/*Xho*I fragment of plasmid EM # 58, which was purified from vector sequences and labelled by the random priming technique of Feinberg & Vogelstein (24). Tc6.4 was isolated from a library constructed by Felsenstein & Emmons (25), consisting of 15−20 kb *Bam*HI partial digestion products of *C. elegans* Bristol DNA in the λ1059 vector. The probe was the insert from EM # 58, excised from the plasmid and labelled by nick-translation. Library screening techniques were as described in Maniatis *et al.* (26). Filters were hybridized to probe using the 50% formamide buffer described in Emmons *et al.* (22), at 37°C (corresponding to approximately 20°C below the melting point (Tm) of A/T rich *C. elegans* DNA).

Southern Blotting

DNA for Southern hybridization was transferred to Nytran (Schleicher and Schuell, Keene, NH) and hybridized in the 50% formamide buffer described in Emmons *et al.* (22). After labelling by nick translation or the random priming technique, probes were hybridized to filters at 37°C (Tm−20°C) for initial characterization of the Tc6 family shown in Figure 2, and elsewhere except as noted otherwise. For lower stringency hybridizations, blots were hybridized to probe at 31°C (Tm−26°C) and 26°C (Tm−31°C) to detect divergent Tc6 elements.

Oligonucleotide SWE2, consisting of the sequence 5′-TGGT-CCTGATGGCTACCGTG-3′ was labeled with T4 polynucleotide kinase as described in Maniatis *et al.* (26) and hybridized to filters as described in Geliebter *et al.* (27). A hybridization temperature of 52°C was used to correspond to Td-12°C [Td = 2(A+T) + 4(G+C), (28)].

For rehybridization, probe was removed from filters by soaking in distilled water at 60°C for 2 minutes; complete removal was confirmed by exposure of the stripped blot to film.

DNA Sequencing and Sequence Analysis

All DNA sequences were determined by the method of Henikoff (29) from multiple overlapping clones of both strands using an Exonuclease III digestion protocol (Erase-a-base, Promega Biologicals). DNA and protein sequences were analyzed by means of The Staden Sequence Analysis Package developed at the Biomathematics Computation Laboratory of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and implemented on the Albert Einstein College of Medicine UNIX system. The algorithms used by this Sequence Analysis Package for analyzing open reading frames are described in reference 30. The sequence of Tc6.1 (Figure 4) has been assigned the following EMBL Data Library accession number: X55356 C. elegans repeat element DNA.

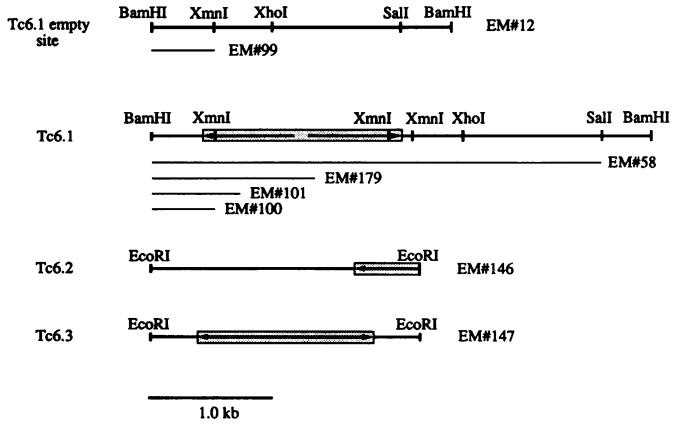

Figure 1. Clones and subclones of Tc6.1 and its corresponding empty site, and clones of Tc6.2 and Tc6.3. The hatched box designates the Tc6 element, the arrows indicate its palindromic structure.

RESULTS

Isolation of the Tc6.1 element and identification of the Tc6 repetitive family

A 1603-base repetitive DNA sequence, denoted Tc6.1, was identified as the cause of a restriction fragment length polymorphism between *C. elegans* strain Bergerac (BO) and other *C. elegans* strains. This strain polymorphism was originally detected in earlier experiments in which randomly-cloned *Bam*HI restriction fragments of the Bristol (N2) strain were used as probes in Southern hybridization experiments with genomic DNA from various nematode strains (22). Plasmid EM # 12 (pCe14 of 22), containing an approximately 2.4 kb Bristol *Bam*HI fragment (Figure 1), hybridized to an approximately 4 kb fragment in Bergerac (BO) DNA. Figure 2A shows this polymorphism as revealed in an *Eco*RI digest. Tc6.1 was defined as DNA present in the Bergerac fragment (denoted by bxP1::Tc6), not present in the corresponding Bristol fragment (denoted bxP1). The slower migrating Tc6.1-containing allele was also not present in DNA of all additional wild isolates of *C. elegans* tested (Figure 2A).

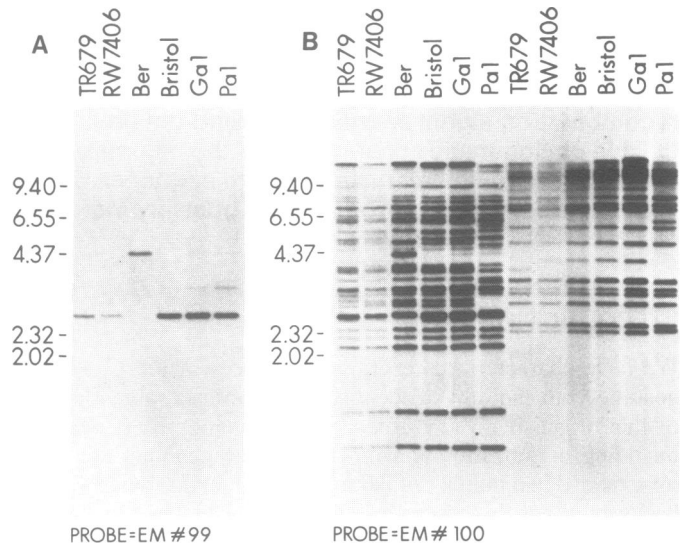

Figure 2. The Tc6 repetitive family. A: Genomic Southern hybridization with empty site clone EM # 99 as probe, showing the bxP1::Tc6 polymorphism. Genomic DNA of various strains was cleaved with *Eco*RI. This filter is the left half of the filter shown in B, which was first hybridized to a Tc6 probe, then stripped and rehybridized to EM # 99. B: Genomic Southern hybridization probed with clone EM # 100, showing the Tc6 repetitive family. Clone EM # 100 contains the 101 terminal nucleotides of Tc6; an identical pattern was obtained with a full-sequence Tc6 probe. Left hand 6 lanes: DNA cleaved with *Eco*RI; right hand 6 lanes: DNA cleaved with *Eco*RV. Strains Bergerac (Ber), Bristol, Gal, and Pal are independent wild isolates of *C. elegans*. Strains TR679 and RW7406 are Bristol/Bergerac hybrid strains.

The Bergerac filled site containing Tc6.1 sequences was isolated from a clone bank of Bergerac DNA fragments using a subclone of EM # 12 (EM # 99) as the probe, as described in Materials and Methods. Tc6.1 sequences hybridized to approximately 24 distinct genomic fragments at Tm-20°C in all *C. elegans* strains tested (Figure 2B), and did not hybridize to DNA of other nematode transposons Tc1, Tc2, Tc3, Tc4 and Tc5 (data not shown). At lower hybridization temperatures (Tm-26°C) approximately twice as many hybridizing bands were detected in *C. elegans* DNA (Figure 3), indicating that additional related sequences were present in the genome. A repetitive family of similar copy number was also detected in nematode species *C. briggsae* (Figure 3).

The pattern of Tc6-hybridizing bands was relatively constant in the strains tested, the wild isolate Pa1 being the most divergent. Among the *Eco*RV digests, the pattern of bands from the strain TR679 is missing one band and has one additional band when compared to the patterns of Bristol and Bergerac (Figure 2B, right). As TR679 is a hybrid derived from a cross between these two strains (17), these polymorphisms might be due to transposition or excision events involving Tc6 that occurred during or subsequent to strain construction, although other rearrangements, possibly involving other transposons, could also account for the differences.

Nucleotide sequence of Tc6.1

The complete nucleotide sequence of Tc6.1 is presented in Figure 4. The limits of the element were defined by comparison of the sequence in Bergerac (BO) to that at the corresponding genomic position in Bristol (not shown). Tc6.1 is a near-perfect palindrome with a non-palindromic central region. The palindromic arms of Tc6.1 are 765 nucleotides in length, and contain a single mismatch. The central non-palindromic region is of 73 nucleotides. The palindromic arms contain a sequence of 23 nucleotides that is directly repeated with 1 mismatch (nucleotides 7−29 and 192−214).

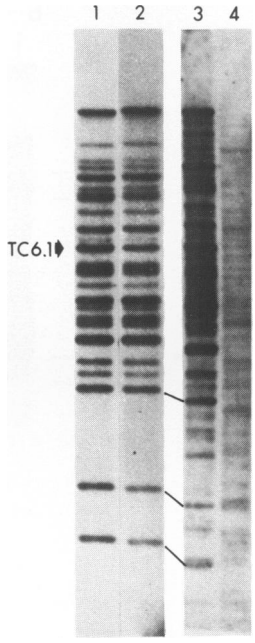

Figure 3. Degree of conservation of Tc6 family. Lane 1: *Eco*RI-digested Bergerac DNA hybridized under standard conditions (Tm-20°C) to a probe containing the 101 terminal nucleotides of Tc6.1 (EM # 100). Lane 2: Same filter as shown in Lane 1, after stripping and rehybridization under the same conditions to a probe containing all Tc6.1 sequences (EM # 179). No additional bands are seen when internal as well as terminal nucleotides are included in the probe, demonstrating the conserved structure of Tc6 elements. Lanes 3 and 4: *Eco*RI-digested DNA of *C. elegans* Bergerac (Lane 3) and *C. briggsae* (Lane 4), hybridized at Tm−26°C to a probe containing all Tc6.1 sequences (EM # 179). Extra bands indicate the presence of additional variant elements in both species.

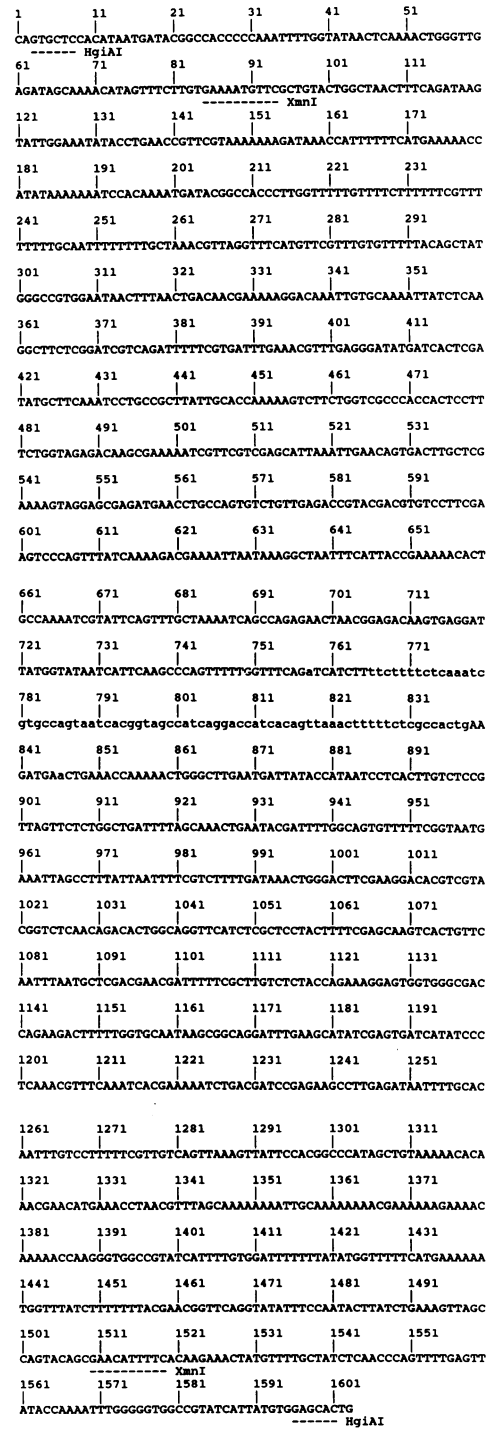

Figure 4. Nucleotide sequence of Tc6.1. Upper case letters indicate bases within the Tc6.1 palindromic arms. Lower case letters indicate bases within the 73-base central non-palindromic region. Symmetrical *Xmn*I and *HgiA*I restriction enzyme recognition sequences cited in the text are indicated.

The Tc6.1 sequenced was analyzed by means of the Staden sequence analysis program (30). In contrast to large palindromic transposon sequences from other species (15,16), Tc6 sequences do not have a highly repetitive internal sequence organization, and are not similar to viral enhancer sequences (16). The coding capacity of Tc6.1 is limited to small open reading frames encoding potential polypeptides of less than 70 amino acids. No potential open reading frame product of greater than 40 amino acids was found to have significant similarity to sequences in Genbank, functionally conserved DNA binding protein sequences, or to reverse transcriptase proteins. The open reading frames present were judged unlikely to encode protein products on the basis of four independent codon frequency algorithms available within the Staden sequence analysis package.

Many Tc6 elements have a structure similar to that of Tc6.1

Tc6.1 was used to isolate two Tc6 elements from the genome of the *C. elegans* Bristol strain, as described in Materials and Methods (Figure 1). The DNA sequence of both elements was determined. The clone of Tc6.2 contained a segment with 93% sequence identity to a palindromic arm of Tc6.1. The Tc6.1-like segment was truncated at nucleotide 604 at an *Eco*RI site used for cloning, suggesting that the genomic Tc6.2 element contains an *Eco*RI site at this position due to a G to T transversion at nucleotide 602. The clone of Tc6.3 contained a 1.4 kb element with 97% sequence identity to Tc6.1. This element was identical in structure to Tc6.1 except for the replacement of nucleotides 657 to 849 by the nucleotides AAA, giving a net deletion of 190 nucleotides. This deletion included the non-palindromic central region of Tc6.1.

Tc6.1, Tc6.2, and Tc6.3 were cloned in plasmid vectors, and were stable in recA$^-$ bacterial hosts. The insert in each clone corresponded in size to a genomic fragment present in a genomic Southern hybridization (data not shown). We were unable to isolate 1.6 kb palindromic Tc6 elements in extensive screens of two clone banks in phage λ vectors, probably because a palindromic structure of the size of Tc6 inhibits growth of the phage (31). Screens of a λ1059 clone bank resulted in the isolation of a Tc6 element of a different structure, denoted Tc6.4, which is described further below.

Many Tc6 elements have a 1.6 kb palindromic structure similar to that of Tc6.1, suggesting that Tc6.1 represents a prototypical Tc6 element. This was shown by Southern blotting analysis of *C. elegans* genomic DNA digested with restriction endonucleases *HgiA*I and *Xmn*I, enzymes which cut near the termini of Tc6.1, yielding fragments of 1.6 kb and 1.4 kb respectively (Figure 4). When genomic DNA was digested with *HgiA*I, Tc6 hybridization was largely limited to a single 1.6 kb band (Figure 5); with *Xmn*I a prominent 1.4 kb band was present, together with additional fainter bands due to hybridization of terminal sequences (data not shown). Therefore, many genomic Tc6 elements have conserved terminal sequences, containing sites for these restriction enzymes, separated by a conserved number of nucleotides. A small number of other bands in Figure 5 indicated the presence of variant elements, such as the element Tc6.3.

Evidence for conservation of overall structure was also obtained by comparing hybridization on a genomic Southern of a probe containing the terminal 101 nucleotides of Tc6.1 (EM # 100) to that of a probe containing the entire Tc6.1 sequence (EM # 179). Both probes gave the same pattern of bands with similar intensities (Figure 3), showing that terminal sequences and internal

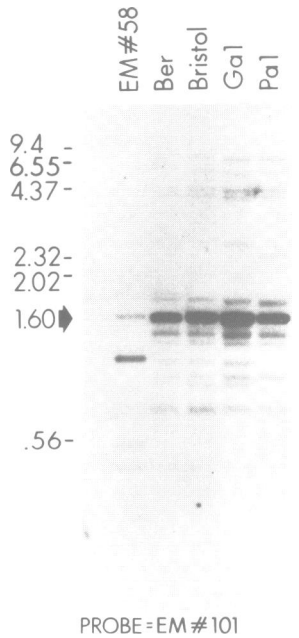

Figure 5. Conserved structure of genomic Tc6 elements. Southern hybridization of genomic DNA of several strains, cleaved with restriction endonuclease *HgiA*I (see Figure 4), and hybridized to EM # 101. A small amount of *HgiA*I-digested DNA of plasmid EM # 58 (denoted pCe1006 in the Figure) is included to show the size of the *HgiA*I fragment of Tc6.1 (arrow). The lower 4 bands in the pCe1006 lane contain genomic sequences flanking Tc6.1, and vector sequences.

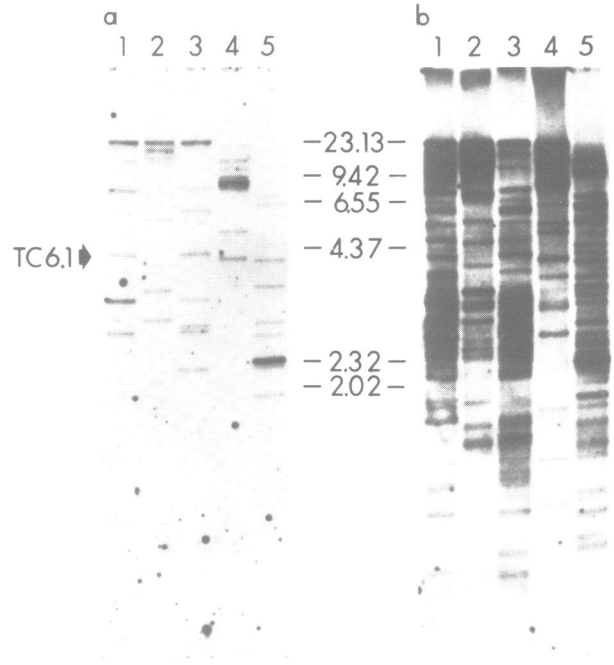

Figure 6. Presence of the central, non-palindromic sequence in Tc6 elements. Bergerac DNA was digested with 1: *Eco*RI, 2: *Eco*RV, 3: *Eco*RI/*Eco*RV, 4: *Bam*HI, 5: *Bam*HI/*Cla*I, and probed with a: an oligonucleotide specific to the central, non-palindromic region (SWE2, Materials and Methods), and b: a probe containing all Tc6 sequences (EM # 179). Comparison of a and b demonstrates that no bands hybridize to SWE2 which do not also hybridize strongly to the full-sequence probe, indicating that the central non-palindromic region is always associated with Tc6 arms. Exact correspondence of the bands in A to bands in B could be confirmed by laying one film on top of the other.

sequences are associated in the genome. Further evidence for the conserved palindromic structure of Tc6 elements was obtained by digesting genomic DNA with nuclease S1 after brief reannealing. A single predominant species of .75 kb was resistant to the enzyme (data not shown).

We wished to know whether, in addition to conserved length and overall sequence, Tc6 elements also had conserved central non-palindromic regions. To investigate whether the 73 base internal non-palindromic sequence of Tc6.1 was invariably present together with Tc6 palindromic arms, an oligonucleotide was synthesized to hybridize to nucleotides 792 to 811 of the central non-palindromic region of Tc6.1 (SWE2, see Materials and Methods). When used as a probe in a genomic Southern hybridization, this oligonucleotide hybridized to a subset of the fragments that hybridized to a probe containing all Tc6 sequences (Figure 6). Therefore, the 20 nucleotide SWE2 sequence does not occur in other sequence contexts, consistent with the 73 nucleotide non-palindromic sequence always being present together with Tc6 palindromic arms. The number of bands present when the oligonucleotide is used as the probe (approximately 10) gives a lower limit for the number of Tc6 elements in the genome. Additional elements may be present with central regions diverged such that they no longer hybridize to the short probe sequence. Such elements could account for the greater number of bands present when the probe containing all Tc6 sequences is used. Additional bands would also result from Tc6 elements such as Tc6.2 that contain restriction sites for the enzymes used in the analysis.

Although most genomic Tc6 elements resemble Tc6.1, one clone containing an insert of 15 kb, was isolated that contained Tc6-related sequences in a different sequence context. This clone, containing an element denoted Tc6.4, was isolated by screening a clone bank of genomic DNA of *C. elegans* Bristol (N2) in a phage λ vector with a Tc6-specific probe, as described in Materials and Methods. The Tc6.4 clone was shown to contain Tc6-hybridizing restriction fragments in multiple copies. Multi-copy Tc6-hybridizing restriction fragments of similar size were shown to be present in the *C. elegans* genome, indicating an absence of rearrangement on cloning. Examples of such multi-copy fragments can be seen in Figure 6a, lane 4 (dark band of 8.0 kb) and lane 5 (dark band of 2.3 kb). Examination by electron microscopy of the Tc6.4 clone revealed a double-stranded foldback region of 5.0 kb, indicating the presence of a 10 kb palindromic sequence (not shown). A limited further characterization of this clone was pursued. The restriction map was consistent with the presence of a large palindrome, but not with a model in which this palindrome was made up of subrepeats of 1.6 kb palindromic elements. By DNA sequence analysis it was found that the palindrome contained at least 3 copies of segments of Tc6 joined to non-Tc6 sequences (32).

```
Tc6.1  ccatgtaCAGTGCTCCACATAATGATACGGCCACCCCCAAATTTTGGTATAA          CTCAAAACTGGGTTGA
Tc6.2  atttata..........G.................................ACATAAAAAA.................
Tc6.3  actaata.....................................................  ................
Tc1          ta.......GGC..A..A....TCCA.TTTTGGTTT...GT..G...          ..TTTTT..
Tc3          ta..... .GGGA.AGT.CTATA..AC....
```

Figure 7. Comparison of terminal sequences of Tc6, Tc1 and Tc2. Upper case letters denote transposon sequences, lower case letters denote flanking sequences, underlined bases denote the conserved flanking TA dinucleotides. Dots represent bases identical to corresponding bases of Tc6.1; spaces are introduced to allow for maximum alignment.

Tc6 termini are similar to Tc1 termini

Comparison of the termini of Tc6.1, Tc6.2, and Tc6.3, revealed that Tc6 sequences have conserved termini with similarity to termini of the *C. elegans* transposable elements Tc1 and Tc3 (Figure 7). For the three similar Tc6 elements, Tc6.1, Tc6.2, and Tc6.3, sequence identity among the elements, the end of the inverted repeat, and sequences found at the polymorphic 'filled' site bxP1:Tc6 but lacking at the 'empty' site bxP1, defined the same nucleotide as the end of the Tc6 element. The Tc6.1 element is inserted at a TA dinucleotide, and all three Tc6 elements are flanked by TA dinucleotides, either because they duplicate a target TA upon insertion, or because the elements themselves terminate in TA.

The ends of Tc6 elements are identical for 7 nucleotides with the ends of Tc1, and for 5 nucleotides with the ends of Tc3. Near the ends there are additional significant sequence similarities. Tc1 and Tc3 also insert at and appear to duplicate a TA dinucleotide, this being a property of the class of Tc1-like transposons generally (12).

DISCUSSION

A DNA strain polymorphism allowed us to identify a family of repetitive elements in the *C. elegans* genome that resemble transposable elements. The copy number of Tc6 elements was 20 to 30 as judged by Southern hybridization at Tm-20°C, although additional Tc6-related sequences were evident under less stringent hybridization conditions. All *C. elegans* strains appeared to have similar numbers of Tc6 elements, and many Tc6 elements had a conserved 1.6 kb length. The ends of Tc6 elements were conserved and shared sequence similarity to two other characterized *C. elegans* transposons, Tc1 and Tc3. Like other transposons with terminal sequences related to Tc1 (12), Tc6 elements were inserted at and appeared to duplicate a TA dinucleotide.

Because of their conserved structure with ends consisting of Tc1-like sequences, and their association with strain polymorphisms, we propose that Tc6 elements are transposable. However, we did not obtain direct evidence that Tc6 elements are presently active in transposition, and no spontaneous mutation involving a Tc6 sequence has been reported. Evidence for transposition might be obtained by analyzing the restriction fragment polymorphisms present in strain TR679 (Figure 2).

From the presence of shared terminal nucleotides, we infer that Tc1 and Tc6 transposition might occur by a similar transposition pathway involving common trans-acting factors.

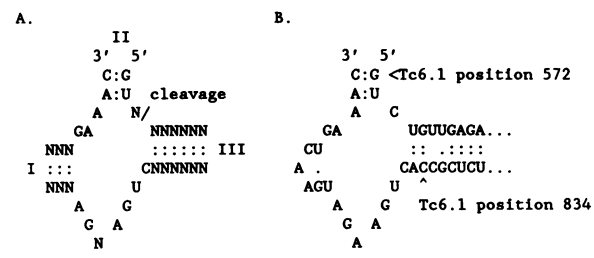

Figure 8. A. Consensus hammerhead of Uhlenbeck (40). B. Hypothetical hammerhead formed in Tc6 RNA. : indicates AU or GC base pairing; . indicates a GU base pair.

However, Tc6 activity is not identical to Tc1, and we conclude that additional internal transposon sequences must also be important. For example, Tc1 copy number, but not Tc6 copy number, has been amplified ten-fold in certain nematode strains, presumably because of the presence in these strains of *mutator* loci, which activate transposition of Tc1 in the germline (5,18). Tc6 elements are present in similar numbers in all strains, regardless of the presence or absence of Tc1 mutators. In addition, Tc6 elements do not undergo high frequency somatic excision, a reaction characteristic of Tc1 elements (33). DNA fragments from empty insertion sites resulting from somatic excision of Tc1 are easily detectable in a Southern hybridization with a probe specific for flanking sequences; we have been unable to consistently detect such fragments for Tc6 (data not shown).

Tc6 is unlike transposable elements of the foldback class in other organisms (reviewed in 34) in not having an internal structure made up of repeats of a short sequence. It contains no open reading frames with the properties of expressed genes. In attempting to account for the presence and conserved palindromic structure of Tc6 elements, we have considered the possibility that Tc6 may transpose by a mechanism involving RNA. Noting the similarity of the Tc6 palindromic structure to that of hepatitis delta virus (35) and to plant viroids and satellite RNA's (virusoids) (36), we compared the sequence of Tc6.1 to these elements. The Tc6.1 sequences 329-GAAAAAGGA, and 1360-AAAAAAA-CGAAAAAAGAAA in the plus-strand (the plus-strand is given in Figure 4), and 774-AGAAAAGAAAAAGA in the minus-strand, resemble the polypurine sequence of potato spindle tubor viroid (PSTV) AGAAAAGAAAAAAGAAGG; polypurine sequences are conserved in known viroids and have been proposed to be related to a reverse transcriptase initiation site (37). The probability of the 11 nucleotide match beginning at nucleotide 774 occuring in a random sequence with the base composition of Tc6 is .003. The Tc6.1 plus-strand sequence 1157-GC-AATAAG is present in the conserved pathogenicity-determining sequence UGGCAAUAAGGA of PSTV (38); the probability of this occurance is 0.03.

One characteristic of viroids is that they contain catalytic sequences that result in RNA cleavage. Such catalytic sequences can typically be folded into a so-called 'hammerhead' configuration (39). We found that a hammerhead configuration that contained all of the conserved nucleotides described by Uhlenbeck (40) could be obtained by folding the Tc6.1 sequence (Figure 8). Part of the putative Tc6 hammerhead, the plus-strand sequence 834-CACTGAAGATG, is identical to a sequence located at the same hammerhead position within the plus-strand hammerhead of the satellite RNA of lucerne transient streak virus (39). Taken altogether, these sequence similarities are consistent with a possible connection between Tc6 and RNA genetic elements, suggesting that Tc6 might have an RNA replicative phase, followed by reverse transcription and integration of DNA copies via the Tc1 transposase. Structural similarities between transposons and viroids have previously been pointed out by Kiefer *et al.* (36). We examined the *in vitro* stability of RNA transcripts of Tc6.1 but obtained no evidence for self cleavage (32). Significant Tc6 transcription was not detectable in Northern hybridization analysis of RNA from nematodes.

We further found that the ends of Tc6 elements, as well as Tc1 and Tc3 elements, share sequence similarity with mammalian immune system recombination signal sequences. For example, the flanking and terminal nucleotides of Tc6.1 (taCAGTGCTCC-A...) are identical at 11/12 positions to the nucleotides at one end of the DNA segment separating the V and J regions of the mouse immunoglobulin κ light chain gene $V_\kappa 21C$ (|CACAG-TGCTCCA, | designates the point of recombination) (41). The conserved heptamer consensus sequence CACAGTG found at all immune system recombination junctions (41) is identical at 6/7 positions to the ends of Tc6, Tc1, and Tc3 elements, when the conserved flanking TA dinucleotide is included (taCAGTG...) (Figure 7). This similarity may be evidence for a common, widely-distributed underlying cellular recombination mechanism.

ACKNOWLEDGEMENTS

We thank Drs. S. Baird, K.-L. Chow, and D. Fitch for their comments on the manuscript. This work was supported by National Institutes of Health Grant GM38174 to SE. DHD is supported by National Institutes of Health grant T32 GM07288. SE is the recipient of the Harry Winston Established Fellowship for Research of the American Heart Association. Some nematode strains used in this study were provided by the Caenorhaditis Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR).

REFERENCES

1. Emmons, S.W. (1988) In W. Wood (ed.) The genome, in *The Nematode* Caenorhabditis elegans. Cold Spring Harbor Laboratory, pp47–79.
2. Coulson, A., Sulston, J., Brenner, S., and Karn, J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7821–7825.
3. Coulson, A., Waterston, R., Kiff, J., Sulston, J., and Kohara, Y. (1988) Genome linking with yeast artificial chromosomes. Nature 335, 184–186.
4. Moerman, D.G., and Waterston, R.H. (1989) In D.E. Berg and M.M. Howe (eds.) Mobile elements in *Caenorhabditis elegans* and other nematodes, *Mobile DNA*. Amer. Soc. Microbiol., pp537–556.
5. Emmons, S.W., Yesner, L., Ruan, K.S., and Katzenberg, D. (1983) *Cell* **32**, 55–65.
6. Liao, L.W., Rosenzweig, B., and Hirsh, D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3585–3589.
7. Rosenzweig, B., Liao, L., and Hirsh, D. (1983) *Nucleic Acids Res.* 11, 4201–4209.
8. Harris, L.J., Prasad, S., and Rose, A.M. (1990) *J. Mol. Evol.* **30**, 359–369.
9. Brierly, H.L., and Potter, S.S. (1985) *Nucleic Acids Res.* **13**, 485–500.
10. Harris, L.J., Baillie, D.L., and Rose, A.M. (1988) *Nucleic Acids Res.* 16, 5991–5998.
11. Henikoff, S., and Plasterk, R.H.A. (1988) *Nucleic Acids Res.* **16**, 6234.
12. Brezinsky, L., Wang, G.V.L., Humphreys, T., and Hunt, J. (1990) *Nucleic Acids Res.* **18**, 2053–2059.
13. Levitt, A., and Emmons, S.W. (1989) *Proc. Natl. Acad. Sci. USA* 86, 3232–3236.
14. Collins, J., Forbes, E., and Anderson, P. (1989) *Genetics* **121**, 47–55.
15. Potter, S., Truett, M.A., Phillips, M., and Maher, A. (1980) *Cell* **20**, 639–647.
16. Liebermann, D., Liebermann, B.H., Weinthal, J., Childs, G., Maxson, R., Mauron, A., Cohen, S.N., and Kedes, L. (1983) *Nature* **306**, 342–347.
17. Collins, J., Saari, B., and Anderson, P. (1987) *Nature* **328**, 726–728.
18. Mori, I., Moerman, D.G., and Waterston, R.H. (1988) *Genetics* **120**, 397–407.
19. Brenner, S. (1974) *Genetics* **77**, 71–94.
20. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., and Boyer, H.W. (1977) *Gene* 2, 95–113.
21. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (eds) (1987) *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience.
22. Emmons, S.W., Klass, M.R., and Hirsh, D. (1979) *Proc. Natl. Acad. Sci. USA* 76, 1333–1337.
23. Norrander, J., Kempe, T., and Messing, J. (1983) *Gene* **26**, 101–106.
24. Feinberg, A.P., and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
25. Felsenstein, K.M., and Emmons, S.W. (1987) *J. Molec. Evolu.* **25**, 230–240.
26. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning*. Cold Spring Harbor Laboratory

27. Geliebter, J., Zeff, R.A., Schulze, D.H., Pease, L.R., Weiss, E.H., Mellor, A.L., Flavell, R.A., and Nathenson, S.G. (1986) *Mol. Cell. Biol.* **6**, 645–652.
28. Suggs, S.V., Hirose, T., Miyake, T., Kawashima, E.H., Johnson, M.J., Itakura, K.,and Wallace, R.B. (1981) *ICN-UCLA Symp. Mol. Cell. Biol.* 231, 683.
29. Henikoff, S. (1984) *Gene* **28**, 351–359.
30. Staden, R. (1984) NAR 12, 551–567.
31. Leach, D.R.F., and Stahl, F.W. (1983) *Nature* **305**, 448–451.
32. Dreyfus, D. (1990) Characterization of Tc6 repetitive DNA sequences in the genome of the nematode *Caenorhabditis elegans*. PhD Thesis, Yeshiva University.
33. Emmons, S. W., and Yesner, L. (1984) *Cell* **36**, 599–605.
34. Finnegan D.J. (1985) *Int-Rev-Cytol.* **93**, 281–326.
35. Wang, K.S., Choo, Q.L., Weiner, A.J., Ou, J.H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L., and Houghton, M. (1986) *Nature* **323**, 508–513.
36. Riesner, D., and Gross, H.J. (1985) *Ann. Rev. Biochem.* **54**, 531–564.
37. Kiefer, M.C., Owens, R.A., and Diener, T.O. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6234–6238.
38. Dinter-Gottlieb, G. (1986) *Proc. Natl. Acad. Sci. USA* 83, 6250–6254.
39. Forster, A.C., and Symons, R.H. (1987) *Cell* **49**, 211–220.
40. Uhlenbeck, O.C. (1987) *Nature* **328**, 596–600.
41. Akira, S., Okazaki, K., and Sakano, H. (1987) *Science* **238**, 1134–1138.