The Tc5 Family of Transposable Elements in *Caenorhabditis elegans*

John J. Collins*,†,[1] and Philip Anderson†

**Department of Biochemistry and Molecular Biology, University of New Hampshire, Durham, New Hampshire 03824, and †Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706*

Manuscript received December 10, 1993
Accepted for publication March 9, 1994

ABSTRACT

We have identified Tc5, a new family of transposable genetic elements in the nematode *Caenorhabditis elegans*. All wild-type varieties of *C. elegans* that we examined contain 4–6 copies of Tc5 per haploid genome, but we did not observe transposition or excision of Tc5 in these strains. Tc5 is active, however, in the *mut-2* mutant strain TR679. Of 60 spontaneous *unc-22* mutations isolated from strain TR679, three were caused by insertion of Tc5. All three Tc5-induced mutations are unstable; revertants result from precise or nearly precise excision of Tc5. Individual Tc5 elements are similar to each other in size and structure. The 3.2-kb element is bounded by inverted terminal repeats of nearly 500 bp. Eight of the ten terminal nucleotides of Tc5 are identical to the corresponding nucleotides of Tc4. Further, both elements recognize the same target site for insertion (CTNAG) and both cause duplication of the central TNA trinucleotide upon insertion. Other than these similarities to Tc4, Tc5 is unrelated to the three other transposon families (Tc1, Tc3 and Tc4) that transpose and excise at high frequency in *mut-2* mutant strains. Mechanisms are discussed by which four apparently unrelated transposon families are all affected by the same *mut-2* mutation.

TRANSPOSABLE genetic elements, or transposons, are widespread in nature. The cardinal feature of these elements is their ability to transpose to new chromosomal sites and to excise from sites they occupy. These events can alter gene expression and gene product function in a variety of ways (FINNEGAN 1989a). Consequently, transposable element activity is a significant source of spontaneous mutation in diverse organisms, including bacteria, fungi, plants and animals (BERG and HOWE 1989).

Eukaryotic transposons can be divided into two major groups based on their mode of transposition. Retrotransposons, so named because of their structural and functional similarity to vertebrate retroviruses, transpose via an RNA intermediate (BOEKE *et al.* 1985). Many elements of this type, including Ty elements in yeast (BOEKE 1989) and *copia*, *gypsy* and *412* elements in Drosophila (BINGHAM and ZACHAR 1989) are bounded by long direct terminal repeats (LTRs) of several hundred base pairs. Others, such as mammalian LINE elements (HUTCHISON *et al.* 1989) and Drosophila I elements (FINNEGAN 1989b) lack such terminal repeats. Transposable elements of the other major group transpose via a DNA intermediate and typically have inverted terminal repeats of several to several hundred base pairs. Some members of this group, such as FB elements of Drosophila (BINGHAM and ZACHAR 1989), TU elements of sea urchins (HOFFMAN-LIEBERMAN *et al.* 1989), and the Tc4 element family of *Caenorhabditis elegans* (YUAN *et al.* 1991), consist almost exclusively of two long inverted repeats. These are commonly referred to as "foldback" elements. However, most elements in this group have an internal region between the inverted repeats that contains one or more long open reading frames (ORF). Examples include: *hobo* (BLACKMAN and GELBART 1989), *mariner* (HARTL 1989) and *P* elements (ENGELS 1989) in Drosophila species; *Ac-Ds*, *Spm* and *Mu* elements in maize (FEDEROFF 1989); Tam1, Tam2 and Tam3 in *Antirrhinum majus* (COEN *et al.* 1989); and Tc1 (MOERMAN and WATERSTON 1989), Tc2 (RUVOLO *et al.* 1992) and Tc3 (COLLINS *et al.* 1989) in *C. elegans*.

In several cases, internal ORFs have been shown to encode transposases, proteins required in *trans* for transposition and excision (BAKER *et al.* 1986; MASSON *et al.*, 1987; KARESS and RUBIN 1984; VOS *et al.* 1993; VAN LUENEN *et al.* 1993). The activity of an entire transposable element family often responds to expression of its transposase. For example, the strain-specific and tissue-specific behavior of Drosophila *P* elements results from transcriptional (MISRA and RIO 1990) and post-transcriptional (LASKI *et al.* 1986) control of transposase expression. This knowledge provides a framework for understanding the control of transposition at one level, the transposon family. However, available evidence indicates that the typical eukaryotic genome harbors many families of transposable elements, each family consisting of several dispersed copies. We know very little about what mechanisms, if any, act at a higher level to regulate multiple transposon families over the entire genome.

Four active transposon families have been identified in *C. elegans*: Tc1 (EIDE and ANDERSON 1985a; MOERMAN *et al.* 1986), Tc2 (LEVITT and EMMONS 1989), Tc3

[1] To whom correspondence should be sent at the University of New Hampshire.

(COLLINS *et al.* 1989), and Tc4 (YUAN *et al.* 1991). Each of these elements is found in multiple copies in the genomes of all isolates of *C. elegans* examined. However, their genetic activity is restricted to certain strains. For example, Tc1 is responsible for most spontaneous mutations isolated in *C. elegans* variety "Bergerac" (EIDE and ANDERSON 1985a) but is inactive in the germ line in variety "Bristol" (EIDE and ANDERSON 1985b; MOERMAN and WATERSTON 1984). Tc1 is highly active in somatic cells in all genetic backgrounds examined. In Bergerac, excision of Tc1 is about 1000-fold more frequent in the soma than in the germ line (EMMONS and YESNER 1984; EIDE and ANDERSON 1988).

We have described mutants in which the activity of Tc1 in the germ line is greatly elevated in comparison to Bergerac (COLLINS *et al.* 1987). The mutator phenotype exhibited by these strains results primarily from increased transposition of Tc1. For this reason, the mutator strain TR679 and its derivatives have been widely used for "transposon tagging" efforts in *C. elegans*. We enlisted TR679 for a different purpose, to search for new transposons. We reasoned that the *C. elegans* genome might harbor transposons that had gone undetected because they do not actively transpose in wild-type genetic backgrounds. If such elements were more active in mutator mutants, like Tc1, we should be able to identify them using genetic methods. To test this, we used the *unc-22* gene as a "transposon trap," analyzing a large collection of *unc-22* mutations isolated from TR679 for ones caused by insertions of transposons other than Tc1. As previously reported (COLLINS *et al.* 1989), this approach led to the discovery of the Tc3 family of elements. We also detected insertions of Tc4, a family of elements first identified in a study of non-Tc1 insertion mutations in *unc-86* and *ced-4* isolated from TR679 (YUAN *et al.* 1991). In this report we describe genetic and molecular properties of another *C. elegans* transposon family uncovered using this strategy, the Tc5 family.

MATERIALS AND METHODS

Nematode strains and genetic methods: Wild-type strains Bristol N2, Bergerac (EM1002), DH424 and TR403 have been described (BRENNER 1974; EMMONS *et al.* 1983; LIAO *et al.* 1983; COLLINS *et al.* 1989). Mutator strain TR679 [genotype *mut-2(r459)*] is an *unc-54*$^+$ revertant of strain TR674 [genotype *unc-54(r323::Tc1); mut-2(r459)*] (COLLINS *et al.* 1987). TR674 is an EMS-induced mutant isolated on the basis of the high frequency of Tc1 excision in this strain. TR679 exhibits a mutator phenotype because *mut-2(r459)* also activates transposition of Tc1 and other *C. elegans* transposons including Tc3 (COLLINS *et al.* 1989), Tc4 (YUAN *et al.* 1991) and Tc5 (this study). TW186 [genotype *mut-2(r459)*] was generated by crossing TR679 with Bristol males and identifying an F_2 clone exhibiting the Him (*h*igh *i*ncidence of *m*ales) phenotype conferred by *mut-2(r459)*. Methods for nematode culture and genetics were described by BRENNER (1974). Our strategy for isolating spontaneous *unc-22* mutants was described in detail previously (COLLINS *et al.* 1989). Briefly, *unc-22* mutants were identified in populations of TR679 by screening for nicotine-induced twitching (MOERMAN and BAILLIE 1979). Only a single *unc-22* mutant was retained from each population.

Three mutants obtained from this screen, *unc-22(r644)*, *unc-22(r741)* and *unc-22(r753)*, are the subject of this report. The original isolate of each of these mutants proved difficult to maintain, presumably due to the accumulation of deleterious mutations caused by the high levels of transposition in this mutator background. To overcome this problem, each mutant was crossed once with the Bristol N2 strain to obtain a derivative with increased fertility. Mutant animals were mated with Bristol males, wild-type F_1 hermaphrodites were picked, and F_2 twitchers that segregated from these animals were cultured individually. For each mutant, a single clone that reproduced well was retained. In each case, the backcrossed derivative does not exhibit the Him phenotype characteristic of mutator strain TR679 (COLLINS *et al.* 1987), indicating it no longer contains *mut-2(r459)*.

Molecular procedures: Our methods for preparation of nematode DNA and Southern blot hybridization were as described previously (EIDE and ANDERSON 1985b), except that hybridization probes were radiolabeled by the random primer method (FEINBERG and VOGELSTEIN 1983). To identify insertion-containing alleles, blots were probed with a combination of λ-DM17 and λ-DM18. Collectively, these clones cover over 25 kb of the *unc-22* gene region and detect most, if not all, transposon insertions that confer a strong *unc-22* mutant phenotype (MOERMAN *et al.* 1986). We cloned the Tc5 element responsible for *unc-22(r644)* as described previously (EIDE and ANDERSON 1985b; COLLINS *et al.* 1989), using λ-EMBL3 as the bacteriophage cloning vector (FRISCHAUF *et al.* 1983) and plasmid TR#26 (see below) as the radiolabeled hybridization probe. We retained one recombinant phage, TR#53, containing the entire Tc5 element plus flanking *unc-22* sequences on both sides.

Restriction fragments were subcloned as described (COLLINS *et al.* 1989), then transformed into *E. coli* host strain DH5α (BRL). Tc5 was subcloned from TR#53 into plasmid pIBI76 (International Biotechnologies, Inc.) as two *Sac*I restriction fragments: the 2.7-kb *Sac*I insert in plasmid clone TR#31 includes the "right" half of Tc5 plus flanking *unc-22* DNA; the 1.8-kb *Sac*I insert in plasmid TR#33 contains the "left" half of Tc5 plus flanking *unc-22* sequences (refer to Figure 2A). Nucleotide sequence analysis of the appropriate regions of TR#31, TR#33 and TR#53 (S. ANDREWS and J. J. COLLINS, unpublished data) confirmed that TR#31 and TR#33 together contain the entire Tc5 element inserted in *unc-22(r644)*. Two restriction fragments from the corresponding region of the wild-type *unc-22* gene were subcloned from λ-DM18 into pIBI76: plasmid TR#26 contains a 1.2-kb *Sac*I restriction fragment spanning this Tc5 insertion site, plasmid TW#30 contains a 3.7-kb *Sac*I restriction fragment adjacent to the insertion site. Restriction fragments from regions of the wild-type *unc-22* gene corresponding to the other two Tc5 insertions were subcloned as well: a 2.9-kb *Bgl*II restriction fragment that flanks the *unc-22(r741::Tc5)* insertion site was subcloned from λ-DM17, yielding plasmid TR#18; a 3.7-kb *Eco*RI restriction fragment that spans the *unc-22(r753::Tc5)* insertion site was subcloned from λ-DM20 (MOERMAN *et al.* 1986), yielding plasmid TW#31. Tc5-Msc2.9, a 2.9-kb *Msc*I restriction fragment containing all of Tc5 except the terminal 115 bp at each end, was obtained by digesting phage TR#53 with *Msc*I and purifying the appropriate fragment from agarose gels using the USBioclean kit (U.S. Biochemical Corp.). Plasmids TR#30 and TR#10, clones of complete Tc1 and Tc3 elements, respectively, have been described (COLLINS *et al.* 1989). The Tc2 clone pCe1022 (LEVITT and EMMONS 1989) and

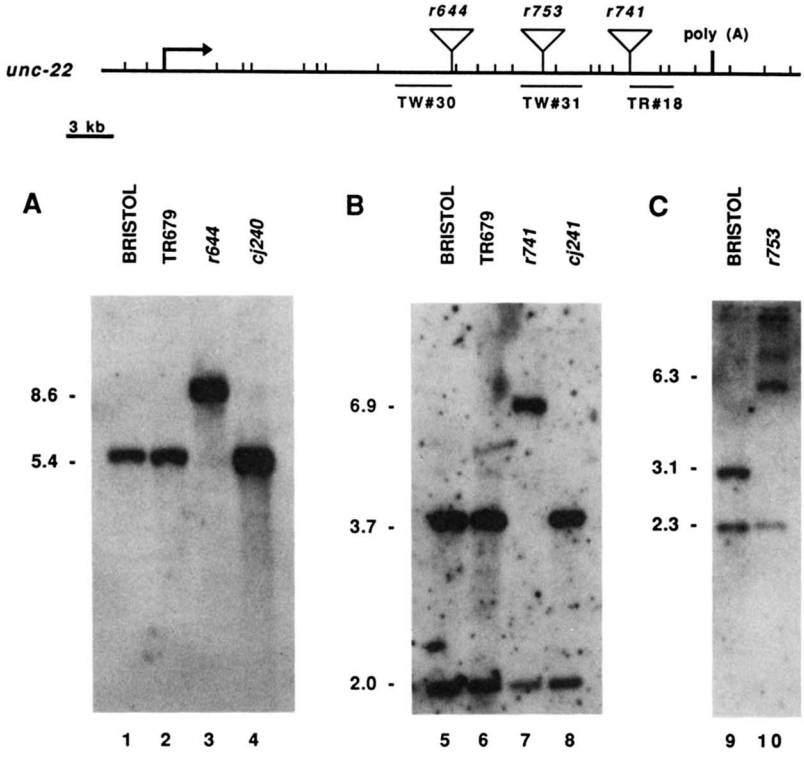


FIGURE 1.—Genomic Southern blot analysis of *unc-22::Tc5* mutants and revertants. The location within *unc-22* of *r644::Tc5*, *r741::Tc5* and *r753::Tc5*, and the corresponding hybridization probes, are indicated at the top. Thin vertical lines represent *Hin*dIII restriction sites. The arrow indicates the start-site and direction of transcription of *unc-22*; the thick vertical line marks the cleavage/polyadenylation site. Wild-type strain "Bristol" and mutator strain TR679 are *unc-22*$^+$. *r644*, *r741* and *r753* are spontaneous *unc-22* mutants isolated from TR679; *cj240* and *cj241* are spontaneous wild-type revertants of *r644* and *r741*, respectively. DNA from each strain was digested with *Hin*dIII and separated on 0.8% agarose gels. Radiolabeled probes were TW#30 (panel A), TR#18 (panel B) and TW#31 (panel C). Sizes, in kilobases, are indicated to the left of each blot.

the Tc6 clone EM#103 (DREYFUS and EMMONS 1991) were provided by S. EMMONS (Albert Einstein College of Medicine, Bronx, New York). Tc4-*n1351*, a plasmid clone of Tc4 (YUAN *et al.* 1991) was provided by H. R. HORVITZ (Massachusetts Institute of Technology, Cambridge, Massachusetts).

A restriction map of Tc5 was determined by comparing the sizes of restriction fragments generated by digesting *unc-22::Tc5*-containing clones (TR#31, TR#33 and TR#53) and the corresponding wild-type *unc-22* clones (TR#26 and TW#30) with various restriction enzymes and double-enzyme combinations. We sequenced the insertional junctions of Tc5 by the SANGER *et al.* (1977) method, as modified for double-stranded plasmid templates (ZAGURSKY *et al.* 1985), using Sequenase (U.S. Biochemical Corp.). The "left" junction (as depicted in Figure 2A) was sequenced in TR#33, using the SP6 Promoter Primer (New England Biolabs). We identified the Tc5 insertion site and terminus by comparing this sequence with the sequence of this region of the wild-type *unc-22* gene (BENIAN *et al.* 1989). To sequence the "right" junction, we designed an oligonucleotide, designated JC1 (a 20-mer obtained from Operon Technologies), complementary to a region of *unc-22* approximately 70 bp right of the insertion site. We used JC1 as a primer to sequence the right junction in TR#31.

RESULTS

Isolation of Tc5 insertions in *unc-22*: We discovered Tc5 using *unc-22* to "trap" transposons that are active in the mutator strain TR679 (COLLINS *et al.* 1987). The details of our rationale and methods have been described (COLLINS *et al.* 1989). Briefly, we isolated a large collection of *unc-22* mutants in the TR679 background and analyzed the gene structure of each with genomic Southern blots to identify ones containing novel inserts. Several features of *unc-22* make it a good gene for this approach: (i) *unc-22* is non-essential; *unc-22* mutations confer a distinctive "twitcher" phenotype (WATERSTON *et al.* 1980). Using a simple screen (MOERMAN and BAILLIE 1979), twitcher mutants are easy to identify in large populations of wild type animals. (ii) *unc-22* spans over 40 kb (BENIAN *et al.* 1993) making it a good target for transposon insertion. It is a preferred target for Tc1 insertion (MOERMAN *et al.* 1986); it might be for other transposons as well. (iii) Molecular clones of *unc-22* are available (MOERMAN *et al.* 1986) for analyzing the gene structure of mutant alleles.

We isolated sixty spontaneous *unc-22* mutants in a manner that insured they represent independent mutational events (see MATERIALS AND METHODS). We used genomic Southern blots to identify insertion-containing alleles and to determine the size and restriction map of each insertion. We reported previously that, by these criteria, 43 of 60 tested *unc-22* mutants contain insertions of Tc1, while three contain insertions of Tc3 (COLLINS *et al.* 1989). Subsequent studies have shown that two mutants contain insertions of Tc4 (C. PARHAM and J. J. COLLINS; unpublished results).

Three of the remaining dozen mutants, *unc-22(r644)*, *unc-22(r741)* and *unc-22(r753)*, contain insertions of approximately 3.2 kb within *unc-22*. These insertions are copies of a family of transposons that we call Tc5. Genomic Southern blots of *r644*, *r741* and *r753* are presented in Figure 1, lanes 3, 7 and 10, respectively. Each of the probes used for these blots (TW#30, lanes 1–4; TR#18, lanes 5–8, TW#31, lanes 9 and 10) hybridizes to only a small region of *unc-22*

surrounding the corresponding Tc5 insertion site, as illustrated at the top of Figure 1. The restriction enzyme used (*Hin*dIII) does not cut within Tc5. We performed similar Southern blots with a variety of restriction enzymes. This enabled us to determine the location of each Tc5 insertion in *unc-22* with greater precision and to compare the structures of the three inserted elements. We draw three conclusions from these experiments: (1) the Tc5 insertion sites in *r644*, *r741* and *r753* are widely separated in *unc-22* (the position within *unc-22* of each Tc5 insertion is shown at the top of Figure 1), (2) the restriction maps of the three elements derived from this analysis were indistinguishable from each other and (3) based on size and restriction map, these elements are distinct from other known *C. elegans* transposons: Tc1, Tc2, Tc3, Tc4 and the transposon-like sequence Tc6.

Germ-line excision of Tc5 leads to reversion of Tc5-induced mutations: *unc-22(r644::Tc5)*, *unc-22(r741::Tc5)* and *unc-22(r753::Tc5)* are unstable in the TR679 background, reverting spontaneously to *unc-22*$^+$ at high frequency. We detected wild-type revertants in the original isolate of each *unc-22::Tc5* mutant when populations contained approximately 1000 animals. We analyzed the gene structures of two revertant alleles: *unc-22(cj240)* and *unc-22(cj241)* are wild-type revertants of *unc-22(r644::Tc5)* and *unc-22(r741::Tc5)*, respectively. Figure 1 shows that both *cj240* (lane 4) and *cj241* (lane 8) resulted from precise or nearly precise excision of Tc5 from *unc-22*. These events depend on the mutator activity present in TR679. When we crossed *r644::Tc5*, *r741::Tc5* and *r753::Tc5* into a genetic background lacking this mutator (see MATERIALS AND METHODS) reversion was no longer observed among over 500,000 animals screened for each mutant.

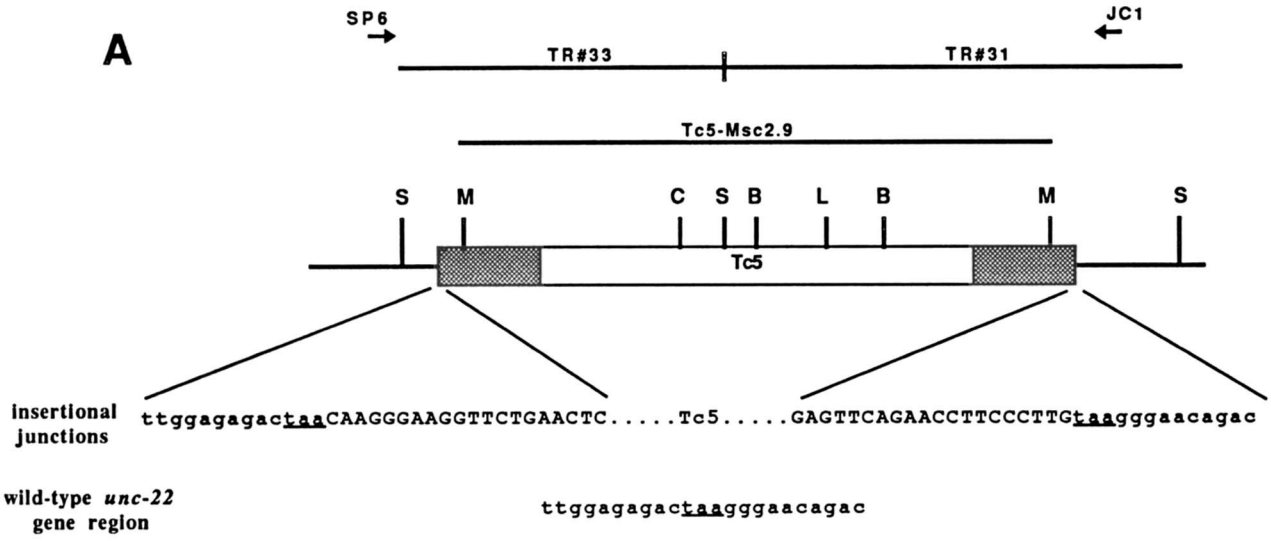
Somatic excision of Tc5: Figure 1 demonstrates that Tc5 does not exhibit the high frequency of somatic excision characteristic of Tc1. Tc1 elements excise from sites of insertion at very high frequency in somatic cells. Somatic excisions are so frequent that the "empty sites" can be detected as restriction fragments of wild-type mobility on genomic Southern blots (EMMONS and YESNER 1984). For a typical Tc1 insertion site, these empty sites account for several percent of the total DNA molecules (EMMONS *et al.* 1986). Consistent with this molecular evidence, genetic methods indicate that excision of Tc1 is about 1000-fold more frequent (per cell division) in the soma than the germ line (EIDE and ANDERSON 1988).

Restriction fragments of wild-type mobility are not evident on genomic Southern blots of *unc-22::Tc5* mutants (Figure 1; lanes 3, 7 and 10). Even upon prolonged exposure of autoradiograms, we do not detect somatic excision for any of the Tc5 insertions. Apparently, somatic excision of Tc5 does not occur at the same high frequency observed for Tc1. However, we cannot rule out the possibility that somatic excision of Tc5 occurs by a mechanism that generates empty sites of heterogeneous size and, therefore, below the limits of detection.

The termini of Tc5 are inverted repeats similar to those of Tc4: To investigate in more detail the structure and genomic organization of Tc5 and its relationship to other *C. elegans* transposons, we cloned the Tc5 element inserted in *unc-22(r644::Tc5)*. We made a genomic library of this strain in a bacteriophage lambda vector and identified clones that hybridized with TR#26, a clone of the wild-type *unc-22* gene region spanning this Tc5 insertion site. The Tc5-containing region was subcloned from one such phage into a plasmid vector. Two restriction fragments, containing opposite halves of the inserted element plus flanking *unc-22* sequences, were cloned separately. Together, the resulting clones, TR#31 and TR#33, contain the entire Tc5 element, as illustrated in Figure 2A and described in MATERIALS AND METHODS.

To investigate the nature of the Tc5 termini and target site, we sequenced the insertional junctions of *unc-22(r644::Tc5)* contained in TR#31 and TR#33, as described in MATERIALS AND METHODS and illustrated in Figure 2A. The insertional junction sequences shown in Figure 2A reveal several interesting features. The termini of Tc5 are perfect inverted repeats of nearly 500 bp (only the terminal 20 nucleotides at each end of the element are shown in Figure 2A). The extreme termini of Tc5 are remarkably similar to the ends of the inverted repeats of Tc4. The termini of these two elements are aligned in Figure 2B; 8 of 10 terminal nucleotides are identical in Tc4 and Tc5. Otherwise, no significant similarities are apparent between the termini of Tc5 and the ends of other *C. elegans* transposons (see DISCUSSION and Figure 5). Tc5 inserted into the target site sequence "CTAAG" in *unc-22(r644)*. Whereas a single copy of the central TAA trinucleotide is present in the wild-type gene, TAA trinucleotides flank the inserted element. Thus, Tc5 causes a three base pair target site duplication upon insertion.

The Tc5 insertion of *unc-116(e2281::Tc5)* occurs at the target sequence "CTCAG" (PATEL *et al.* 1993), and TCA trinucleotides flank the inserted element. Comparison of these insertion site sequences with all sequenced Tc4 insertion sites, including one each in *unc-86* and *ced-4* (YUAN *et al.* 1991) and two independent insertions at the same site in *unc-33* (LI and SHAW 1993), suggests that both Tc4 and Tc5 recognize the interrupted palindrome CTNAG (where N = G, A, T or C) for insertion, and cause duplication of the central TNA trinucleotide upon insertion. [Note that due to the palindromic nature of this target site sequence, it is also possible that Tc5 begins with an A and ends with a T, and only the N of the target site is duplicated upon insertion. This caveat was noted previously for Tc4 by LI and SHAW (1993).] No further similarities are apparent between the Tc5 insertion site and the sites of insertion of Tc4.



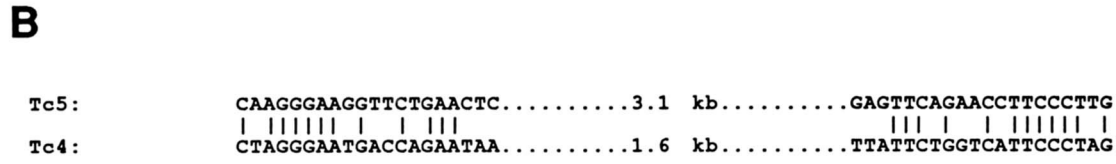


FIGURE 2.—Structure of Tc5 and its insertional junctions. (A) A restriction map of the Tc5 element contained in phage TR#53 is shown for the restriction enzymes *Sac*I, S; *Msc*I, M; *Bam*HI, B; *Sal*I, L; and *Cla*I, C. Restriction enzymes tested that do not cut within Tc5 are: *Eco*RI, *Hin*dIII, *Kpn*I and *Sma*I. The shaded regions at each end of Tc5 represent its terminal inverted repeats. In flanking *unc-22* DNA (solid bar), only the *Sac*I restriction sites used to subclone Tc5 are shown. The restriction fragment "Tc5-Msc2.9" and the restriction fragments contained in plasmid subclones TR#31 and TR#33 are shown above the map. The position and orientation of primers used to sequence the insertional junctions are indicated by the arrows above these clones. The sequences of the insertional junctions of *unc-22(r644::Tc5)* are shown below the map. Tc5 DNA is upper case, *unc-22* DNA is lower case. The target site duplication is underlined. The sequence of the wild-type *unc-22* target site region is shown at the bottom. (B) The sequences of the termini of Tc5 and Tc4 are aligned for comparison. All sequences in panels A and B are shown 5′–3′.

Tc5 is unrelated to other *C. elegans* transposons: To investigate the relatedness of Tc5 to other *C. elegans* transposons, we tested Tc5 for hybridization to molecular clones of Tc1 (TR#30), Tc2 (pCE1022), Tc3 (TR#10), Tc4 (Tc4-*n1351*) and the repetitive, transposon-like sequence Tc6 (EM#103)–see MATERIALS AND METHODS for a full description of each clone. We prepared a Southern blot containing in adjacent lanes restriction fragments from 100 ng of TR#30, pCE1022, TR#10, Tc4-*n1351*, EM#103 and the Tc5-containing plasmids TR#31 and TR#33. This blot was hybridized with radiolabeled Tc5-Msc2.9, a 2.9-kb *Msc*I restriction fragment containing the entire Tc5 element except for the terminal 115 bp at each end (Tc5-Msc2.9 is illustrated in Figure 2A). No hybridization was detected between Tc5 and any of the other elements; strong hybridization was observed to TR#31 and TR#33. Based on these results, we conclude that Tc5 shares no significant nucleotide sequence similarity with other *C. elegans* transposon families.

Tc5 is a dispersed repetitive element: We investigated the genomic organization of Tc5 elements in different wild-type *C. elegans* varieties. Each of these strains is a distinct, geographically separate isolate; hence, each represents a unique wild-type genetic background. We digested genomic DNA from several wild-type strains with *Eco*RI, a restriction enzyme that does not cut within Tc5. Figure 3 presents a Southern blot of wild-type strains Bristol N2, Bergerac, DH424 and TR403 probed with radiolabeled Tc5-Msc2.9. Tc5 is present in 4–6 copies in these strains. The pattern of hybridizing bands is very similar for each strain, suggesting that some Tc5 elements reside at common genomic sites in the different strains examined. This similarity in copy number and position indicates that Tc5 elements have not moved at significant levels since these strains diverged.

For each of the strains in Figure 3A, the Tc5-hybridizing bands are of comparable intensity, indicating that different Tc5 elements in the genomes of these wild-type strains have retained a high degree of

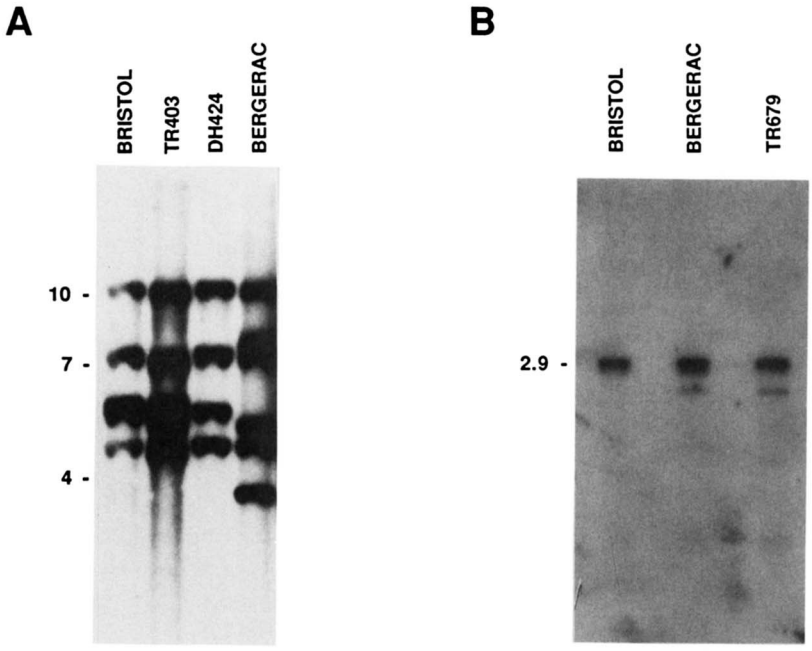


FIGURE 3.—Genomic organization of Tc5 elements. Genomic DNAs were digested with the indicated enzyme and separated on 0.8% agarose gels. Blots were probed with radiolabeled restriction fragment Tc5-Msc2.9 (see MATERIALS AND METHODS). Sizes, in kilobases, are shown to the left of each blot. (A) Southern blot analysis of wild-type strains Bristol, TR403, DH424 and Bergerac. Genomic DNA from each strain was digested with *Eco*RI. (B) Southern blot analysis of wild-type strains Bristol and Bergerac and mutator strain TR679. Genomic DNA from each strain was digested with *Msc*I.

nucleotide sequence similarity. To investigate the degree of length heterogeneity within the Tc5 family, we digested genomic DNA with the restriction enzyme *Msc*I, and probed Southern blots with radiolabeled Tc5-Msc2.9. *Msc*I cuts twice in the cloned copy of Tc5, once in each inverted terminal repeat (see the restriction map of Tc5; Figure 2A). If all genomic copies of Tc5 are similar in structure, this probe should detect a single 2.9-kb *Msc*I restriction fragment on genomic Southern blots. Figure 3B presents the results for wild-type strains Bristol and Bergerac, and mutator strain TR679. For each strain, Tc5-Msc2.9 hybridized almost exclusively to a single 2.9-kb *Msc*I fragment. Weak hybridization to a second *Msc*I fragment of approximately 2.5 kb is evident for strains Bergerac and TR679. This may reflect heterogeneity in structure or *Msc*I restriction sites among Tc5 elements in these strains. The relative intensity of hybridization to the two fragments suggests that such heterogeneity is likely restricted to a single member of the Tc5 family. The presence of this second Tc5-hybridizing fragment in both the Bergerac and TR679 strains is consistent with the ancestry of TR679: it was isolated from a strain with a Bristol/Bergerac hybrid genetic background. Collectively, we conclude from these results that most or all Tc5 elements in the genomes of these strains are similar in size and sequence to the cloned element.

Tc5 is activated by the *mut-2* mutator: Several lines of evidence indicate that Tc5 elements transpose and excise in the germ line of TR679, but not in wild-type strains. (i) To date, no Tc5-induced mutations have been isolated in wild-type genetic backgrounds, while several have been recovered in TR679. Over 250 spontaneous mutations isolated in wild-type strains Bristol, Bergerac, DH424 and TR403 have been analyzed at the molecular level. These include mutations in *unc-22*, *unc-54* and *lin-12* (EIDE and ANDERSON 1985a; EIDE and ANDERSON 1985b; GREENWALD 1985; MOERMAN *et al.* 1986; COLLINS *et al.* 1987, D. EIDE and P. ANDERSON, unpublished data; A. M. LOZEAU and J. J. COLLINS, unpublished data). None of these mutations resulted from insertion of Tc5. By comparison, at least five spontaneous, Tc5-induced mutations have been isolated in TR679, including three in *unc-22* (Figure 1A), and one each in *mec-7* (SAVAGE and CHALFIE 1989) and *unc-116* (PATEL *et al.* 1993). (ii) All three *unc-22::Tc5* mutations are unstable (due to excision of Tc5; Figure 1A) in the TR679 background, but not when crossed into a genetic background lacking the *mut-2* mutator activity present in this strain. (iii) The copy number and genomic position of Tc5 elements is very similar in different wild-type strains (Figure 3A), indicating that very little transposition and excision of Tc5 has occurred in these strains since they diverged.

To investigate in more detail the level of Tc5 activity in the *mut-2* mutator background, we examined the genomic distribution of Tc5 elements in different sublines derived from TR679. Each subline was established from a single TR679 animal and propagated independently for several generations before isolating DNA for analysis. DNAs were digested with *Eco*RI and Southern blots were probed with Tc5-Msc2.9 to assess Tc5 copy number and position. Differences among these sublines reflect heterogeneity within the TR679 population from which they were derived, indicating that Tc5 elements are probably highly active in this strain. Figure 4 presents a genomic Southern blot of TR679, TW186 (a Bristol-backcrossed derivative of TR679; see MATERIALS AND METHODS), and five independent sublines derived from TR679. Wild-type strains Bristol and Bergerac are shown

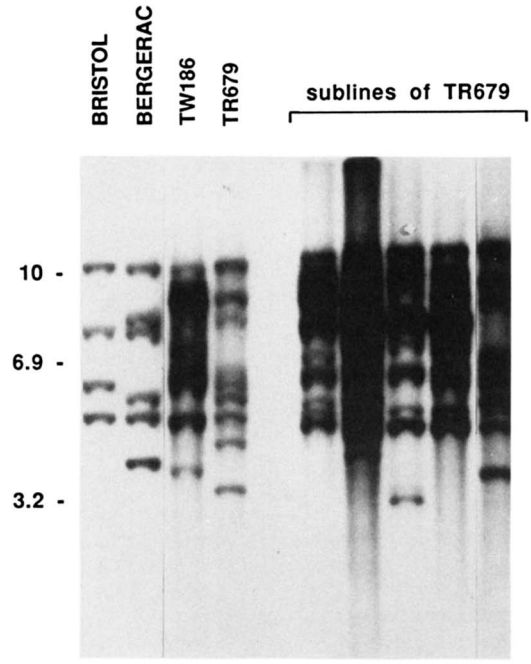


FIGURE 4.—Activation of Tc5 by *mut-2(r459)*. Genomic Southern blot of wild-type strains Bristol and Bergerac, *mut-2* mutants TW186 (see MATERIALS AND METHODS) and TR679, and five independent sublines derived from TR679. DNA from each strain was digested with *Eco*RI and separated on a 0.8% agarose gel. The hybridization probe was Tc5-Msc2.9. Sizes, in kilobases, are shown to the left of the blot.

for comparison. TR679 and its derivatives are heterogeneous with respect to Tc5 copy number and position. We estimate that TR679 and TW186 each contain eight or nine copies of Tc5, while the copy number in different sublines varies from 8 to 13. Further, each strain exhibits a different pattern of Tc5-hybridizing bands, reflecting a unique genomic complement of Tc5 elements. These results indicate that transposition and excision of Tc5 are frequent in TR679. We believe the observed differences are the consequence of Tc5 transposition and excision because these strains contain *additional* copies of Tc5 compared to wild-type strains. The pattern of Tc5-hybridizing bands for strain TR445, the non-mutator grandparent of TR679, is indistinguishable from that of Bristol (data not shown). This *increase* in Tc5 copy number cannot be explained by chromosomal rearrangements or insertion of other *C. elegans* transposons in or near copies of Tc5.

One additional observation suggests that transposition of Tc5 is very frequent in the *mut-2* mutator background. For each subline shown in Figure 4, the different Tc5-hybridizing bands are of variable intensity, in contrast to the uniform intensity of hybridization to different genomic copies of Tc5 observed in wild-type strains. We believe this variability reflects heterogeneity in Tc5 copy number and position for different individuals within a subline, resulting from Tc5 transposition and excision during propagation of that subline.

DISCUSSION

We describe here the discovery and characterization of Tc5, a family of transposable elements in *C. elegans*. Several properties of Tc5 establish that it is a transposable element: (1) Tc5 transposes and excises. Three spontaneous *unc-22* mutations contain Tc5 inserted within the gene. These Tc5-induced mutations are unstable; reversion is due to excision of Tc5. (2) The termini of Tc5 are perfect inverted repeats. Notably, the ends of these inverted repeats are similar to the ends of the inverted repeats of another *C. elegans* transposon, Tc4. (3) Tc5 appears to recognize the same target sequence for insertion (CTNAG) as Tc4, and both elements cause duplication of the central TNA trinucleotide upon insertion. (4) Tc5 is a dispersed, repetitive sequence in the *C. elegans* genome. (5) The copy number of Tc5 is greater in the mutator strain TR679 than in the wild-type strains from which TR679 was derived.

We discovered Tc5 in our analysis of 60 spontaneous *unc-22* mutants isolated in the TR679 mutator strain. Because of the high frequency of Tc1 transposition in this strain (COLLINS *et al.* 1987), it is widely used for "transposon tagging" efforts in *C. elegans*. For this reason, information on the spectrum of mutational events recovered in this strain is of general interest. Our collection of 60 *unc-22* mutants includes: 43 caused by Tc1 insertion; three caused by Tc3 insertion (COLLINS *et al.* 1989); two caused by Tc4 insertion (C. PARHAM and J. J. COLLINS, unpublished results); and three caused by Tc5 insertion (Figure 1). The remaining nine *unc-22* mutants exhibited no restriction fragments of altered mobility on genomic Southern blots (data not shown). However, it remains possible that this group includes additional transposon-induced alleles that escaped our detection. The probe we used to screen Southern blots for insertions (see MATERIALS AND METHODS) covers approximately 25 kb of the *unc-22* gene; the entire gene spans over 40 kb of genomic DNA (BENIAN *et al.* 1993). In any case, it is clear that our methods allowed us to detect most, if not all, of the insertions that occurred (51 of a possible 60).

Do these results reflect general features of the TR679 mutator background or are they unique to *unc-22*? Results from other studies indicate that the preponderance of Tc1 insertions observed at *unc-22* is not a general property of all *C. elegans* genes. For example, Tc5 insertions into *mec-7* (C. SAVAGE and M. CHALFIE, personal communication) and *unc-116* (PATEL *et al.* 1993) were identified among populations of TR679 that failed to yield any Tc1-induced alleles of these genes. Similarly, mutant hunts with TR679 identified Tc4-induced alleles of *ced-4*, *unc-86* and *unc-33* (YUAN *et al.* 1991; LI *et al.* 1992); Tc3-induced alleles of *unc-86* (FINNEY *et al.* 1988) and a Tc2-induced allele of *gld-2* (A. JONES and T. SCHEDL, personal communication). These same screens identified few, if any, Tc1 insertions into the same genes.

Collectively, these results demonstrate that the frequencies of transposon-induced mutants vary greatly from element-to-element and from gene-to-gene. In Drosophila, *hobo* and *P* elements exhibit very different patterns of insertion into chromosome *3* (SMITH *et al.* 1993). For any given gene, however, the apparent frequency of transposon-induced mutants may severely underestimate the true frequency of transposon insertion. Tc1 and Tc3 are efficiently spliced from many, if not most, mutant pre-mRNAs (RUSHFORTH *et al.* 1993; BENIAN *et al.* 1993; A. RUSHFORTH and P. ANDERSON, unpublished observations; M. MILLS, J. GLASNER and J. J. COLLINS, unpublished observations). Hence, not all insertions of these elements cause mutant phenotypes. Tc4 can also be spliced from mutant pre-mRNAs (LI *et al.* 1992; LI and SHAW 1993), suggesting that it may often insert into genes without phenotypic consequence as well. These results indicate that the apparent gene-specificity of insertion might be related to the ways in which transposons are spliced from transcripts, rather than the specificity of insertion *per se*.

We define Tc5 as a distinct *C. elegans* transposon family based on its failure to hybridize with other known transposons in this organism and its unique size, copy number and genome organization. The Tc5 family is apparently homogeneous; most or all genomic copies of Tc5 are similar in size and structure (Figure 3B). Two other *C. elegans* transposon families, Tc1 and Tc3, also exhibit striking structural homogeneity (EMMONS *et al.* 1983; LIAO *et al.* 1983; COLLINS *et al.* 1989). Even the Tc4 family, originally described as heterogeneous in structure (YUAN *et al.* 1991), may actually consist of two structurally homogeneous subfamilies, Tc4 and Tc4v (LI and SHAW 1993). Each of these *C. elegans* transposons is bounded by inverted repeats. In other organisms, a common feature of transposons with inverted terminal repeats is the heterogeneous structure of individual family members. For example, the *Ac-Ds* transposon family in maize consists of full-length *Ac* elements and "defective" *Ds* elements, containing internal deletions and other rearrangements (FEDEROFF 1989). Similarly, many *P* elements in the Drosophila genome contain internal deletions (O'HARE and RUBIN 1983). ENGELS *et al.* (1989) propose that these internally deleted *P* elements result from incomplete repair of double-strand breaks generated by *P* element excision. The remarkable structural homogeneity of *C. elegans* transposon families might reflect their lack of significant transposition and excision in wild-type strains, resulting in few, if any, gap repair-induced deletion events. Alternatively, the gene conversion process responsible for double-strand gap repair might be less prone to interruption in *C. elegans* than other organisms.

Among transposable elements with terminal inverted repeats, a large number of elements from a wide range of species share common sequences at their termini.

Element	Species	Inverted repeat termini
Tgm	*Glycine max*	CACTATTAGAAA
Tam1	*Antirrhinum majus*	CACTACAACAAA
Tam2	*Antirrhinum majus*	CACTACAACAAA
Spm	*Zea mays*	CACTACAAGAAA
Tst1	*Solanum tuberosum*	CAGGGGCGTAT
dTph1	*Petunia hybrida*	CAGGGGCGGAGC
Bg-rbg	*Zea mays*	CAGGG
Ac	*Zea mays*	CAGGGATGAAA
P	*Drosophila melanogaster*	CA-TGATGAAAT
hobo	*Drosophila melanogaster*	CA--GA-GAACT
Tc5	*Caenorhabditis elegans*	CAAGG--GAAGG
Tc4	*Caenorhabditis elegans*	CTAGG--GAAGG
HB1	*Drosophila melanogaster*	CAGTGC-GAAGG
Tc1	*Caenorhabditis elegans*	CAGTGCTGGCCA
Tc3	*Caenorhabditis elegans*	CAGTG-TGGGAA
Tc1b	*Caenorhabditis briggsae*	CAGTACTGGCCA
Tc6	*Caenorhabditis elegans*	CAGTGCTCCACA
Uhu	*Drosophila heteroneura*	CAGTGTCTTACA

FIGURE 5.—A comparison of the inverted repeat termini of 18 different transposable elements. All sequences are 5′-3′, representing the "left" end of each element. The termini of *Ac* and *Tst1* are imperfect inverted repeats; mismatches are underlined. Data for these elements are from the following references: Tc1 (ROSENSWEIG *et al.* 1983); Tc3 (COLLINS *et al.* 1989); Tc4 (YUAN *et al.* 1991); Tc5 (Figure 2A); Tc6 (DREYFUS and EMMONS 1991); Tcb1 (HARRIS *et al.* 1988); *Uhu* (BREZINSKY *et al.* 1990); HB1 (BRIERLEY and POTTER 1985); *hobo* (MCGINNIS *et al.* 1983); *P* (O'HARE and RUBIN 1983); *Ac* (POHLMAN *et al.* 1984); *Spm* (PERIERA *et al.* 1986); *Bg-rbg* (HARTINGS *et al.* 1991); Tam1 (BONAS *et al.* 1984); Tam2 (UPADHYAYA *et al.* 1985); *Tgm* (VODKIN *et al.* 1983); *dTph1* (GERATS *et al.* 1990); and *Tst1* (KOSTER-TOPFER *et al.* 1990).

Figure 5 presents a comparison of the inverted repeat termini of 18 transposons, including: Tc1, Tc3, Tc4, Tc5 and Tc6, a palindromic transposon-like sequence in *C. elegans* (DREYFUS and EMMONS 1991); the Tcb1 element from the related nematode *Caenorhabditis briggsae* (HARRIS *et al.* 1988), and 12 different transposable elements from fruit flies, maize, snapdragons, soybeans, petunia and potato. With the exception of Tc4, each of these elements begins with the dinucleotide CA (CT for Tc4). In many cases, more extensive similarities are evident, including some recognized previously (DORING and STARLINGER 1986; COLLINS *et al.* 1989). These similarities may reflect underlying similarities in the mechanisms of transposition and excision, resulting from convergent evolution or common ancestry of these elements. Their presence in such diverse species suggests that these elements might have spread by horizontal transmission, as recently demonstrated for Drosophila *P* elements (KIDWELL 1992), and suggested for the Tc1/mariner "superfamily" of transposons (HENIKOFF 1992; ROBERTSON 1993).

Tc5 exhibits intriguing similarities with Tc4. First, the terminal ten nucleotides of these two elements are nearly identical (Figure 2B). For many transposable elements, sequences in the terminal inverted repeats are known to be required in *cis* for transposition and excision (for example, KARESS and RUBIN 1984). Such sequences serve as binding sites for element-encoded transposases (MIZUUCHI 1992) and, for Drosophila *P* elements, a host-encoded protein of unknown function

(RIO and RUBIN 1988; KAUFMAN *et al.* 1989). Second, Tc5 and Tc4 recognize the same target site sequence (CTNAG), and both elements cause duplication of the central TNA trinucleotide upon insertion (Figure 2B; YUAN *et al.* 1991). Many transposable elements cause duplication of short stretches of target site sequence upon insertion. Often, the length of duplicated sequence is characteristic for a particular transposable element, reflecting the mechanism of transposition of that element. These similarities suggest that Tc5 and Tc4 might transpose by a common mechanism, in response to the same transposase. We previously suggested that Tc1 and Tc3 elements might share a common transposase, based on similarities between their termini and insertion sites (COLLINS *et al.* 1989).

It seems unlikely that a single transposase is responsible for transposition of Tc1, Tc3, Tc4 and Tc5. No extended similarities common to the termini of all four elements are apparent in Figure 5. Further, the target sites recognized by Tc1 and Tc3 do not resemble the CTNAG target sequence for insertion of Tc4 and Tc5, and insertion of Tc1 and Tc3 cause duplication of the dinucleotide TA (EIDE and ANDERSON 1988; MORI *et al.* 1988; COLLINS *et al.* 1989), compared to the trinucleotide TNA for Tc4 and Tc5. Additional, more direct, evidence indicates that each of these elements might encode its own transposase. Tc1 and Tc3 encode proteins that promote transposition of the corresponding element when expressed from a heterologous promoter in transgenic animals (VOS *et al.* 1993; VAN LUENEN *et al.* 1993). In these experiments, each transposase was specific for the element that encodes it by two criteria: activation of somatic transposition and binding to terminal inverted repeat sequences. Inspection of their deduced amino acid sequences reveals that the proteins encoded by Tc1 and Tc3 are quite similar (28% identity overall–ROSENZWEIG *et al.* 1983; D. SCHNEIDER, J. J. COLLINS and P. ANDERSON, unpublished results). Tc4 has the structure of a foldback element and contains no significant protein coding capacity. However, a variant subfamily of Tc4 elements, Tc4v, contain an ORF capable of encoding a protein of 537 amino acids (LI and SHAW 1993). Our nucleotide sequence analysis of Tc5 reveals an ORF that could encode a 532 amino acid protein that is similar to the putative Tc4v-encoded protein (33% identity overall–P. OLSEN, S. ANDREWS and J. J. COLLINS, submitted for publication). These predicted proteins do not resemble the transposase proteins encoded by Tc1 and Tc3. In this light, the similarities in termini and insertion sites discussed above probably reflect similarities between the element-specific transposases encoded by Tc1 and Tc3, and Tc4 and Tc5, rather than shared transposases.

While each of these elements is activated in the mutator strain TR679 (COLLINS *et al.* 1987, 1989; YUAN *et al.* 1991; Figure 4) their activities are not regulated in an identical manner. Specifically, Tc1 is genetically active in the wild-type variety Bergerac (EIDE and ANDERSON 1985a; MOERMAN *et al.* 1986), but genetic activity of Tc3, Tc4 and Tc5 have been detected only in the *mut-2* mutant background. In addition, Tc1 excises at very high frequency in somatic cells (EMMONS and YESNER 1984), but no such events have been detected for Tc3 (COLLINS *et al.* 1989), Tc4 (C. PARHAM and J. J. COLLINS, unpublished results) or Tc5 (Figure 1). This difference is not surprising if Tc3, Tc4 and Tc5 require *mut-2(r459)* for activity, since the effects of this mutation are limited to the germ line (COLLINS *et al.* 1987).

How does a single mutation, *mut-2(r459)*, activate at least four different families of transposable elements? Evidence discussed above indicates that each of these elements probably encodes its own transposase. The *mut-2(r459)* mutation might affect a host-encoded protein that interacts with or regulates the expression of each of these transposases. Alternatively, this mutation might affect a general process that influences the activity of many transposable element families. For example, chromatin structure might be altered in a way that causes a general increase in transposition frequencies. Consistent with a chromatin-level effect, *mut-2(r459)* also causes increased incidence of *X* chromosome nondisjunction (COLLINS *et al.* 1987).

In conclusion, we have identified a new family of transposable elements in *C. elegans*, the Tc5 family. Transposition of Tc5 is activated in *mut-2* mutant strains, adding another transposon to the repertoire available for transposon tagging efforts using these strains. At least one gene, *unc-116*, has been cloned already by tagging with Tc5 (PATEL *et al.* 1993). Further investigation of the mechanisms that regulate Tc5 elements and, in particular, the molecular basis of their activation by *mut-2(r459)* may provide insight into the global control of transposition in eukaryotic genomes.

We thank D. MOERMAN, R. WATERSTON, S. EMMONS, A. LEVITT, D. DREYFUS, J.-Y. YUAN and R. HORVITZ for providing molecular clones; G. BENIAN for sharing *unc-22* sequence data before its publication; C. SAVAGE, M. CHALFIE, N. PATEL, J. MANCILLAS and T. SCHEDL for sharing unpublished information on transposon-induced alleles isolated from TR679; and J. SHAW for discussions regarding the similarity between putative Tc4v-encoded and Tc5-encoded proteins. We acknowledge BETSY FORBES, ALICE RUSHFORTH, KAREN RASMUSSEN, KEN NORMAN and CHRISTI PARHAM for their contributions to this study; JEREMY GLASNER for help with the figures; and RICK COTE and members of the COLLINS laboratory for critically reading the manuscript. This work was supported by U.S. Public Health Service grants GM47570 (J.J.C.) and GM41807 (P.A.). This is scientific contribution 1851 from the New Hampshire Agricultural Experiment Station.

LITERATURE CITED

BAKER, B., J. SCHELL, H. LORZ and N. V. FEDEROFF, 1986 Transposition of the controlling element *Activator* on tobacco. Proc. Natl. Acad. Sci. USA **83:** 4844–4848.

BERG, D. E., and M. M. HOWE, 1989 *Mobile DNA*. American Society for Microbiology, Washington, D.C.

BENIAN, G. M., J. KIFF, N. NECKELMAN, D. G. MOERMAN and R. H. WATERSTON, 1989 Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. Nature **342:** 45–50.

BENIAN, G. M., S. W. L'HERNAULT and M. E. MORRIS, 1993 Additional sequence complexity in the muscle gene, *unc-22*, and its encoded protein, twitchin, of *Caenorhabditis elegans*. Genetics **134:** 1097–1104.

BINGHAM, P. M., and Z. ZACHAR, 1989 Retrotransposons and the FB transposon from *Drosophila melanogaster*, pp. 485–502 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

BLACKMAN, R. K., and W. M. GELBART, 1989 The transposable element *hobo* of *Drosophila melanogaster*, pp. 523–529 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

BOEKE, J. D, 1989 Transposable elements in *Saccharomyces cerevisiae*, pp. 335–374 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

BOEKE, J. D., D. J. GARFINKEL, C. A. STYLES and G. R. FINK, 1985 Ty elements transpose through an RNA intermediate. Cell **40:** 491–500.

BONAS, U., H. SOMMER and H. SAEDLER, 1984 The 17 kb *Tam1* element of *Antirrhinum majus* induces a 3 bp duplication upon integration into the chalcone synthase gene. EMBO J. **3:** 1015–1019.

BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94.

BRIERLEY, H. L., and S. S. POTTER, 1985 Distinct characteristics of loop sequences of two *Drosophila* foldback transposable elements. Nucleic Acids Res. **13:** 485–500.

BREZINSKY, L., G. V. L. WANG, T. HUMPHREYS and J. HUNT, 1990 The transposable element *Uhu* from Hawaiian *Drosophila*–a member of the widely dispersed class of Tc1-like transposons. Nucleic Acids Res. **18:** 2053–2059.

COEN, E. S., T. P. ROBBINS, J. ALMEIDA, A. HUDSON and R. CARPENTER, 1989 Consequences and mechanisms of transposition in *Antirrhinum majus*, pp. 413–436 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

COLLINS, J., B. SAARI and P. ANDERSON, 1987 Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. Nature **328:** 726–728.

COLLINS, J., E. FORBES and P. ANDERSON, 1989 The Tc3 family of transposable genetic elements in *Caenorhabditis elegans*. Genetics **121:** 47–55.

DORING, H.-P., and P. STARLINGER, 1986 Molecular genetics of transposable elements in plants. Annu. Rev. Genet. **20:** 175–200.

DREYFUS, D. H., and S. W. EMMONS, 1991 A transposon-related palindromic sequence from *Caenorhabditis elegans*. Nucleic Acids Res. **19:** 1871–1877.

EIDE, D., and P. ANDERSON, 1985a Transposition of Tc1 in the nematode *C. elegans*. Proc. Natl. Acad. Sci. USA **82:** 1756–1760.

EIDE, D., and P. ANDERSON, 1985b The gene structures of spontaneous mutations affecting a *Caenorhabditis elegans* myosin heavy chain gene. Genetics **109:** 67–79.

EIDE, D., and P. ANDERSON, 1988 Insertion and excision of the *Caenorhabditis elegans* transposable element Tc1. Mol. Cell. Biol. **8:** 737–746.

EMMONS, S. W., and L. YESNER, 1984 High frequency excision of transposable element Tc1 in the nematode *Caenorhabditis elegans* is limited to somatic cells. Cell **36:** 599–605.

EMMONS, S. W., L. YESNER, K. S. RUAN and D. KATZENBERG, 1983 Evidence for a transposon in *Caenorhabditis elegans*. Cell **32:** 55–65.

EMMONS, S. W., S. ROBERTS and K. S. RUAN, 1986 Evidence in a nematode for regulation of transposon excision by tissue-specific factors. Mol. Gen. Genet. **202:** 410–415.

ENGELS, W. R., 1989 *P* elements in *Drosophila melanogaster*, pp. 437–484 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

ENGELS, W. R., D. M. JOHNSON-SCHLITZ, W. B. EGGELSTON and J. SVED, 1989 High-frequency *P* element loss in *Drosophila* is homolog dependent. Cell **62:** 515–525.

FEDEROFF, N. V., 1989 Maize transposable elements, pp. 375–411 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

FEINBERG, A., and B. VOGELSTEIN, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FINNEGAN, D. J., 1989a Eukaryotic transposable elements and genome evolution. Trends Genet. **5:** 103–107.

FINNEGAN, D. J., 1989b The I factor and I-R hybrid dysgenesis in *Drosophila melanogaster*, pp. 503–517 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

FINNEY, M., G. RUVKUN and H. R. HORVITZ, 1988 The *C. elegans* cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. Cell **55:** 757–769.

FRISCHAUF, A.-M., H. LEHRACH, A. POSUSTKA and N. MURRAY, 1983 Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:** 827–842.

GERATS, A. G. M., H. HUITS, E. VRIJLANDT, C. MARANA, E. SOUER *et al.*, 1990 Molecular characterization of a nonautonomous transposable element (*dTph1*) of petunia. Plant Cell **2:** 1121–1128.

GREENWALD, I., 1985 *lin-12*, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. Cell **43:** 583–590.

HARRIS, L. J., D. L. BAILLIE and A. M. ROSE, 1988 Sequence identity between an inverted repeat family of transposable elements in *Drosophila* and *Caenorhabditis*. Nucleic Acids Res. **16:** 5991–5999.

HARTINGS, H., C. SPILMONT, N. LAZZARONI, V. ROSSI, F. SALAMINI *et al.*, 1991 Molecular analysis of the *Bg-rbg* transposable element system of *Zea mays* L. Mol. Gen. Genet. **227:** 91–97.

HARTL, D. L., 1989 Transposable element *mariner* in *Drosophila* species, pp. 531–536 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

HENIKOFF, S., 1992 Detection of *Caenorhabditis* transposon homologs in diverse organisms. New Biol. **4:** 382–388.

HOFFMAN-LIEBERMAN, B., D. LIEBERMAN and S. N. COHEN, 1989 TU elements and puppy sequences, pp. 575–592 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

HUTCHISON, C. A., III, S. C. HARDIES, D. D. LOEB, W. R. SHEHEE and M. H. EDGELL, 1989 LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome, pp. 593–617 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

KARESS, R. E., and G. M. RUBIN, 1984 Analysis of *P* transposable element functions in *Drosophila*. Cell **38:** 135–146.

KAUFMAN, P. D., R. F. DOLL and D. C. RIO, 1989 *Drosophila P* element transposase recognizes internal *P* element DNA sequences. Cell **59:** 359–371.

KIDWELL, M. G., 1992 Horizontal transfer of *P* elements and other short inverted repeat transposons. Genetica **86:** 275–286.

KOSTER-TOPFER, M., W. B. FROMMER, M. ROCHA-SOSA and L. WILMITZER, 1990 Presence of a transposon-like element in the promoter region of an inactive patatin gene in *Solanum tuberosum* L. Plant Mol. Biol. **14:** 239–247.

LASKI, F. A., D. C. RIO and G. M. RUBIN, 1986 Tissue specificity of *Drosophila P* element transposition is regulated at the level of mRNA splicing. Cell **44:** 7–19.

LEVITT, A., and S. W. EMMONS, 1989 The Tc2 transposon in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **86:** 3232–3236.

LI, W., and J. E. SHAW, 1993 A variant Tc4 transposable element in the nematode *C. elegans* could encode a novel protein. Nucleic Acids Res. **21:** 59–67.

LI, W., R. K. HERMAN and J. E. SHAW, 1992 Analysis of the *Caenorhabditis elegans* axonal guidance and outgrowth gene *unc-33*. Genetics **132:** 675–689.

LIAO, L. W., B. ROSENZWEIG and D. HIRSH, 1983 Analysis of a transposable element in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **80:** 3585–3589.

MASSON, P., R. SUROSKY, J. KINGSBURY and N. V. FEDEROFF, 1987 Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. Genetics **177:** 117–137.

MCGINNIS, W., A. W. SHERMOEN and S. K. BECKENDORF, 1983 A transposable element inserted just 5′ to a *Drosophila* glue protein gene alters gene expression and chromatin structure. Cell **34:** 75–84.

MISRA, S., and D. C. RIO, 1990 Cytotype control of *Drosophila P* element transposition: the 66 kd protein is a repressor of transposase activity. Cell **62:** 269–284.

MIZUUCHI, K., 1992 Transpositional recombination: mechanistic insights from studies of Mu and other elements. Annu. Rev. Biochem. **61:** 1011–1051.

MOERMAN, D. G., and D. L. BAILLIE, 1979 Genetic organization in *C. elegans*: fine structure analysis of the *unc-22* gene. Genetics **91:** 95–103.

MOERMAN, D. G., and R. H. WATERSTON, 1984 Spontaneous *unc-22* IV mutations in *C. elegans* var. Bergerac. Genetics **108:** 859–877.

MOERMAN, D. G., and R. H. WATERSTON, 1989 Mobile elements in *Caenorhabditis elegans* and other nematodes, pp. 537–556 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

MOERMAN, D. G., G. M. BENIAN and R. H. WATERSTON, 1986 Molecular cloning of the muscle gene *unc-22* in *Caenorhabditis elegans* by Tc1 transposon tagging. Proc. Natl. Acad. Sci. USA **83:** 2579–2583.

MORI, I., G. M. BENIAN, D. G. MOERMAN and R. H. WATERSTON, 1988 The transposon Tc1 of *C. elegans* recognizes specific target sequences for integration. Proc. Natl. Acad. Sci. USA **85:** 861–864.

O'HARE, K., and G. M. RUBIN, 1983 Structure of *P* transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34:** 25–35.

PATEL, N., D. THIERRY-MIEG and J. MANCILLAS, 1993 Cloning by insertional mutagenesis of a cDNA encoding *Caenorhabditis elegans* kinesin heavy chain. Proc. Natl. Acad. Sci. USA **90:** 9181–9185.

PERIERA A., H. CUYPERS, A. GIERL, Z. SCHWARTZ-SOMMER and H. SAEDLER, 1986 Molecular analysis of the *En/Spm* transposable element system of *zea mays*. EMBO J. **5:** 835–841.

POHLMAN, R. F., N. V. FEDEROFF and J. MESSING, 1984 The nucleotide sequence of the maize controlling element *Activator*. Cell **37:** 635–643.

RIO, D. C., and G. M. RUBIN, 1988 Identification and purification of a *Drosophila* protein that binds to the terminal 31-base-pair inverted repeats of the *P* transposable element. Proc. Natl. Acad. Sci. USA **85:** 8929–8933.

ROBERTSON, H. M., 1993 The mariner transposable element is widespread in insects. Nature **362:** 241–245.

ROSENZWEIG, B., L. LIAO and D. HIRSH, 1983 Sequence of the *C. elegans* transposable element Tc1. Nudeic Acids Res. **11:** 4201–4209.

RUSHFORTH, A. M., B. SAARI and P. ANDERSON, 1993 Site-selected insertion of the transposon Tc1 into a *Caenorhabditis elegans* myosin light chain gene. Mol. Cell. Biol. **13:** 902–910.

RUVOLO, V., J. E. HILL and A. LEVITT, 1992 The Tc2 transposon of *Caenorhabditis elegans* has the structure of a self-regulated element. DNA Cell Biol. **11:** 111–122.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SAVAGE, C., and M. CHALFIE, 1989 *mec-7* is a β-tubulin gene required for the production of 15-protofilament microtubules in *Caenorhabditis elegans*. Genes Dev. **3:** 870–881.

SMITH, D., J. WOHLGEMUTH, B. R. CALVI, I. FRANKLIN and W. M. GELBART, 1993 *hobo* enhancer trapping mutagenesis in Drosophila reveals an insertion specificity different from *P* elements. Genetics **135:** 1063–1076.

UPADHYAYA, K. C., H. SOMMER, E. DREBBERS and H. SAEDLER, 1985 The paramutagenic line *niv*-44 has a 5 kb insert, Tam2, in the chalcone synthase gene of *Antirrhinum majus*. Mol. Gen. Genet. **199:** 201–207.

VAN LUENEN, H. G. A. M., S. D. COLLOMS and R. H. A. PLASTERK, 1993 Mobilization of quiet, endogenous Tc3 transposons of *Caenorhabditis elegans* by forced expression of Tc3 transposase. EMBO J. **12:** 2513–2520.

VODKIN, L. O., P. R. RHODES and R. B. GOLDBERG, 1983 A lectin gene insertion has the structural features of a transposable element. Cell **34:** 1023–1031.

VOS, J. C., H. G. A. M. VAN LUENEN and R. H. A. PLASTERK, 1993 Characterization of the *Caenorhabditis elegans* Tc1 transposase *in vivo* and *in vitro*. Genes Dev. **7:** 1244–1253.

WATERSTON, R. H., J. N. THOMSON and S. BRENNER, 1980 Mutants with altered muscle structure in *Caenorhabditis elegans*. Dev. Biol. **77:** 271–302.

YUAN, J., M. FINNEY, N. TSUNG and H. R. HORVITZ, 1991 Tc4, a *Caenorhabditis elegans* transposable element with an unusual foldback structure. Proc. Natl. Acad. Sci. USA **88:** 3334–3338.

ZAGURSKY, R., N. BAUMEISTER, N. LOMAX and M. BERGMAN, 1985 Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Tech. **2:** 89–94.

Communicating editor: R. K. HERMAN