Extrachromosomal circular copies of the transposon Tc1

Anthony D.Radice⁺ and Scott W.Emmons*

Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Ave, Bronx, NY 10461, USA

Received January 25, 1993; Revised and Accepted April 29, 1993

ABSTRACT

The 1.6 kb Tc1 transposable element of *Caenorhabditis elegans* undergoes excision and transposition in the germline. In somatic tissue it is excised at high frequency. Extrachromosomal linear and circular copies of Tc1 have been identified that are likely to be products of somatic and germline excision. In the present study, we have determined the sequences of the sites of circularization in circular extrachromosomal Tc1 molecules. DNA molecules containing these sites were cloned after PCR amplification with primers directed outward from within Tc1. Sequences were obtained with two complete Tc1 ends and one or more intervening copies of the TA dinucleotide, with one complete end and one deleted end, and with two deleted ends. The 24 clones had different structures, indicating the pool of molecules serving as PCR templates was heterogeneous. The predominant circular junction had one or more nucleotides deleted from at least one transposon end. Such a molecule without two complete ends might not be expected to serve as a transposition intermediate. Hence, some extrachromosomal circular Tc1 molecules may result from a deadend excision pathway.

INTRODUCTION

The genome of the nematode *Caenorhabditis elegans* contains a family of active 1.6 kb transposable elements called Tc1 elements (reviewed in 1). Similar elements are present in genomes of several animal phyla (2). Tc1 elements have 54 bp inverted terminal repeats, and contain a single open reading frame that encodes a DNA-binding protein (3,4). They have an overall structure similar to IS elements in bacteria and certain transposable elements in eukaryotes (such as P elements in *Drosophila* and Ac and En/Spm elements in maize). These elements are thought to undergo transposition by a conservative mechanism involving excision of the donor element followed by reinsertion elsewhere in the genome (see 5 for reviews).

Tc1 elements are excised in both somatic cells and germ cells (6,7,8). Excision appears to generate a double-strand break in

the chromosome at the former site of insertion. This can undergo templated repair, taking advantage of sequences on the chromosomal homolog, on the sister chromatid, or, in transgenic animals, from a transgene (9,10). These observations are consistent with a conservative replication mechanism for Tc1, in which a donor element is excised, inserts at a new genomic site without replication, and a new copy of the element is recreated at the donor site by the templated repair mechanism (9).

Consistent with such a conservative transposition mechanism, extrachromosomal Tc1 elements have been detected (11,12). The predominant species was a linear monomer of Tc1, but relaxed and supercoiled circular monomers were also detected. Any or all of these species might be transposition intermediates, competent to reinsert into the genome.

A competent extrachromosomal transposition intermediate would be expected to contain all of the transposon sequence. The *C.elegans* genome contains very few defective Tc1 elements (13,14), and hence most or all transposition events must generate a complete element. In order to determine whether extrachromosomal circular Tc1 molecules consisted of complete, circularized copies of Tc1, we cloned PCR products designed to contain the site of circularization. We found that these molecules in most, if not all, instances were deleted for one or more nucleotides at at least one of the two ligated transposon ends.

MATERIALS AND METHODS

Nematodes

C.elegans wild type was the Bristol N2 strain (15). The Bergerac BO strain originated from an independent wild isolate (16). RW7406 was a hybrid Bristol/Bergerac BO strain constructed in a study of Tc1 mutators (17); it retains Tc1 germline activity. TR679 was the 'high hopper' strain in which Tc1 has elevated germline activity (18). Nematodes were maintained following the general methods described (19).

Isolation of extrachromosomal Tc1 circles

DNA was isolated with an Applied Biosystem 340A Nucleic Acid Extractor following the manufacturer's instructions. β -mercaptoethanol (1%) was added to worms in phosphate buffer

* To whom correspondence should be addressed

⁺ Present address: Department of Virology and Parasitology, Lindsley F.Kimball Research Institute, The New York Blood Center, 310 E. 67th Street, New York, NY 10021, USA

before the worms were added to lysis buffer. For isolation of extrachromosomal circles five to ten micrograms of genomic DNA was fractionated on a 1% agarose preparative gel. Fractions expected to contain 1.6 kb circles (both relaxed and supercoiled) were identified by ethidium bromide staining of adjacent lanes containing circular and linear plasmid standards. DNA was isolated from appropriate fractions by electroelution.

PCR amplification of junction sequences

The four primers used in PCR amplification of junction sequences in Tc1 circles had the following sequences (numbers indicate the positions within Tc1 of the terminal nucleotides, see Figure 1): SWE-6: 924-CACTTCCATGGGCCCCACTAA-943; SWE-24: 1211-CAGATGCCAAATTCAACCA-1229; SWE-18C: 97-TCGCTTTATGCACACGGA-80; SWE 15A: 666-TCCG-TGTAGTCCTGCTT-650. Amplification was with one of three pairs of primers SWE-6 and SWE-15A, SWE-15A and SWE-24, or SWE-18C and SWE-24. PCR reactions contained 1–2 μ g of DNA template in PCR buffer (50mM KCl; 10 mM Tris-Cl, pH 8.3; 1.5 mM MgCl₂; 0.01% gelatin); 50–100 pmole of each primer; 100 mM each deoxynucleotide triphosphate; and 2.5 units AmpliTaq (Perkin-Elmer/Cetus). Amplification reactions, carried out in a Perkin-Elmer/Cetus DNA Thermal Cycler, consisted of 30 cycles of the following stepped cycle program: 94°, 1 min; T_m, 1 min; 72°, 2 min. T_m was estimated, from the formula $4(G+C)+2(A+T) = ^\circ\text{C}$, to be 52° for the SWE-18C/SWE-24 and SWE-6/SWE-15A pairs, and 54° for the SWE-24/SWE-15A pair.

Subcloning of PCR products and DNA sequencing

Single-stranded ends of the PCR products were made double-stranded by addition to the reaction mixture of 5–10 units of Klenow polymerase (New England Biolabs), followed by incubation for 30 min at room temperature. DNA fragments were cloned into the HincII site of pBluescript II (KS+) (Stratagene) either directly, or after gel fractionation and electroelution of amplification products of a particular size. DNA sequencing was by the dideoxynucleotide method using Sequenase version 2.0 (U.S. Biochemical) and one of two primers oriented across the expected Tc1 junction sequence: either SWE-18C mentioned above, or SWE-38 (1475-AATCTGTTGAACATTTT-GATG-1496). All clones with identical sequences (Tables 1 and 2) were from different PCR reactions, ensuring the independence of all the clones sequenced.

RESULTS

Isolation of clones containing joined Tc1 ends

We used a PCR strategy to amplify the junction sequence of extrachromosomal circular Tc1 molecules. The PCR primers employed annealed within Tc1 and primed DNA synthesis towards the ends of the element (Figure 1). Thus, only molecules containing nearby or adjacent right and left ends of Tc1 could serve as amplification templates. Figure 2 shows the results of one such amplification on DNA templates from two *C. elegans* strains: Bergerac BO, which contains over 300 genomic Tc1 elements; and TR679, which contains about 80 genomic Tc1 elements. In these strains, Tc1 undergoes excision in both somatic and germline cells (6,7,8,18,20). Total genomic DNA was used as the template with two sets of primers (SWE-6 and SWE-15A, Figure 2, lanes 2 and 3; and SWE-24 and SWE-15A, Figure 2, lanes 4 and 5). The expected size of the PCR products was 1354

Figure 1. Hypothetical diagram of a circularized Tc1 molecule with the location of the oligonucleotide primers used in PCR reactions to amplify junction sequences.

Figure 2. Products of PCR amplification of DNA from various strains were fractionated on a 1% agarose gel and stained with ethidium bromide. The primers for lanes 2 and 3 were SWE-6 and SWE-15A and lanes 3 and 4 were SWE-24 and SWE-15A; amplification of a circularized Tc1 element with the first two primers is expected to give a 1354 bp product and with the second two primers is expected to give a 1077 bp product.

bp for SWE-6 and SWE-15A and 1065 bp for SWE-24 and SWE15A when amplifying a template consisting of circularized Tc1. Amplification products of this expected size were generated as well as additional products of lower and higher molecular weight.

Altogether thirty independent PCR products were cloned and sequenced. In some experiments amplification products of the expected size were excised from a gel such as the one shown in Figure 2 and cloned, whereas in other experiments DNA molecules in the amplification reaction were cloned directly without prior fractionation. Six of the thirty cloned fragments

Figure 3. Southern analysis of the products of PCR amplification of gel purified DNA fractions. Total genomic DNA from strain TR679 was fractionated by mobility on a 0.8% agarose gel. The lane containing fractionated DNA was sliced into 12 fractions and the DNA was recovered from each fraction by electroelution. Based on cofractionation of undigested plasmid molecules and estimation from standard curves of plasmids, fraction 11 was expected to contain 1.6 kb supercoiled molecules and fractions 7 and 8 were expected to contain relaxed circles. Fractions 5 and 6 fractionated the same to fractions containing relaxed circular plasmids of approximately 3.0 kb; these fractions may contain circular Tc1 multimers. DNA from each fraction was amplified in PCR reactions with SWE-6 and SWE-15A primers. The expected product is 1354 bp. The products of the PCR reaction were fractionated on a 1% agarose gel, transferred to Nytran membrane and hybridized to a Tc1 probe. The autoradiograph was overexposed to show the fainter bands; this accounts for its different appearance from Figure 2.

contained a single Tc1 end joined to an unidentifiable DNA sequence. These clones could have resulted from mispriming by one of the primers in genomic DNA flanking an inserted Tc1 element; therefore they are not considered further. The remaining 24 clones had 22 distinct structures involving a head-to-head joining of sequences from the left and right ends of Tc1, as expected.

Thirteen clones were isolated after PCR amplification of templates enriched for extrachromosomal DNA by gel fractionation. The result of one such gel, in which DNA was isolated and analyzed from all fractions across the entire gel, is shown in Figure 3. DNA from each fraction was used as a PCR template with the primers SWE-6 and SWE-15A, and the amplified products were analyzed by Southern hybridization with a Tc1 probe. A 1.35 kb amplification product was generated as expected from fractions expected (based on size standards) to contain 1.6 kb relaxed circles (fractions 7 and 8) and 1.6 kb supercoiled molecules (fraction 11). Fractions containing DNA of slightly lower mobility (fractions 5 and 6) may contain multimer extrachromosomal molecules.

Unexpectedly, an amplification product was also generated from fractions containing high molecular weight chromosomal DNA. We consider four possible explanations for this observation: 1. the high molecular weight DNA might be

Table 1. Isolation and structure of junction fragments

Strain	Source of DNA ¹	Clone	Structure ²	Cloning Method ³
Bergerac BO	Genomic	B3	16,46	Shotgun
		B10	0,[TA] ₃ ,0	Shotgun
		B13	11,19	Shotgun
		B14	0,TA ₀	GP
		B15	0,360	GP
		B16	0,259	GP
	Gel fractions	A4	17,0	GP
		A5	1,12	GP
		A6	18,1	GP
		A7	18,2	GP
		A8	0,9	GP
		A9	13,1	GP
		A10	0,7	GP
RW7406	Genomic	RW2	0,TA _N ,203,TA ₀	GP
TR679	Genomic	T8	0,[TA] ₃ ,0	Shotgun
		T12	0,[TA] ₃ ,0	GP
		T23	0,[TA] ₆ ,0	GP
		T25	30,9	Shotgun
	Gel fractions	T4	0,3	GP
		T5	8,0	GP
		T10	2,0	GP
		T18	16,1	GP
Bristol	Gel fractions	N22	2,1	GP
		N28	26,45	GP

Notes: 1. The PCR template was either whole genomic DNA, or DNA eluted from gel fractions expected to contain 1.6 kb circles. 2. The first number is the number of Tc1 nucleotides deleted on the left side of the junction (from the right end of Tc1); the last number is the number of Tc1 nucleotides deleted on the right side of the junction (from the left end of Tc1). Nucleotides between the joined Tc1 sequences, if present, are shown between these numbers. 3. After amplification, PCR products were either cloned directly (Shotgun), or were fractionated on agarose to obtain amplification products of the expected size, which were eluted from the gel and cloned (GP = gel purified).

contaminated with extrachromosomal Tc1 circular molecules; 2. integrated Tc1 copies might result in a PCR artifact; 3. amplification might be due to tandem Tc1 dimers in *C. elegans* germline DNA; 4. amplification might be due to tandem Tc1 dimers or circles integrated into somatic DNA but not present in germline DNA. In a reconstruction experiment we showed that extrachromosomal circular plasmid molecules separated completely from high molecular weight DNA, suggesting that explanation 1 is unlikely. Explanations 2 and 3 also appear unlikely in view of the structure of PCR products amplified from unfractionated genomic DNA, as discussed further below. We have no evidence for or against explanation 4.

Eleven clones were isolated after PCR amplification of unfractionated genomic DNA. The analysis of gel fractions described above suggests that such clones might have been derived from amplification of sequences in the high-molecular-weight DNA fraction, and could therefore have resulted either from a PCR artifact generated by the large number of genomic Tc1 elements, or from amplification of integrated tandem Tc1 dimers. Since eight out of the eleven clones have a precise Tc1 end defining at least one side of the junction, it appears that a Tc1-end-specific activity has been involved in their generation (the remaining three clones have terminal deletions from 9 to 46 base pairs). We believe they are therefore unlikely to be a PCR artifact.

Further, five of the eleven clones have a structure inconsistent with an origin from integrated tandem Tc1 dimers. Clones B3, B13, B15, B16, and T25 (see Tables 1 and 2), isolated from amplification of unfractionated genomic DNA of Bergerac BO and TR679, contained one incomplete Tc1 end that lacked an

Table 2. Sequence of junction in PCR clones.

Conceptual head-to-head joining of Tc1 ends (0,0)			...TATCTTTTGGCCAGCACTG	CAGTGCTGGCCAAAAAGATA...
			right transposon end	left transposon end
Molecules with two complete ends (0,N,0)				
B10	0,[TA] ₃ ,0	TATCTTTTGGCCAGCACTG	TATATA	CAGTGCTGGCCAAAAAGATA
B14	0,TA,0	TATCTTTTGGCCAGCACTG	TA	CAGTGCTGGCCAAAAAGATA
T8	0,[TA] ₃ ,0	TATCTTTTGGCCAGCACTG	TATATA	CAGTGCTGGCCAAAAAGATA
T12	0,[TA] ₃ ,0	TATCTTTTGGCCAGCACTG	TATATA	CAGTGCTGGCCAAAAAGATA
T23	0,[TA] ₆ ,0	TATCTTTTGGCCAGCACTG	TATATATATATA	CAGTGCTGGCCAAAAAGATA
RW2	0,TA,N ₂₀₃ ,TA,0	TATCTTTTGGCCAGCACTG	TA,N ₂₀₃ ,TA	CAGTGCTGGCCAAAAAGATA
Molecules with 1 complete end (0,N) or (N,0)				
A4	17,0	TATCTTTTGGCCAGCACTG		ATA
A8	0,9	TATCTTTTGG		CAGTGCTGGCCAAAAAGATA
A10	0,7	TATCTTTTGGCC		CAGTGCTGGCCAAAAAGATA
T4	0,3	TATCTTTTGGCCAGCA		CAGTGCTGGCCAAAAAGATA
T5	8,0	TATCTTTTGGCCAGCACTG		GCCAAAAAGATA
T10	2,0	TATCTTTTGGCCAGCACTG		GTGCTGGCCAAAAAGATA
B15	0,360	GGAAGC(1250)		CAGTGCTGGCCAAAAAGATA
B16	0,259	GCATAAT(1351)		CAGTGCTGGCCAAAAAGATA
Molecules without a complete end (N,M)				
B3	46,16	AGAAAAA(1564)		(17)GATATCACTTTTAA ^{TT}
B13	11,19	AAGTGGAT(20)		AAAAAGATA
A5	1,13	TATCTTT		GTGCTGGCCAAAAAGATA
A6	19,1	TATCTTTTGGCCAGCACT		A
A7	19,2	TATCTTTTGGCCAGCAC		A
A9	14,1	TATCTTTTGGCCAGCACT		AAGATA
T18	17,1	TATCTTTTGGCCAGCACT		ATA
T25	30,9	TATCTTTTGG		(31)TTTTTTGTGT
N22	2,1	TATCTTTTGGCCAGCACT		GTGCTGGCCAAAAAGATA
N28	27,45	GAAAAAG(1565)		(27)TTTGGTTTTTT

The structure of each clone named in the first column is summarized in the second column, following the convention defined in Note 2, Table 1. In the DNA sequences, numbers in parentheses give the position of the adjacent nucleotide in the Tc1 sequence of Rosenzweig *et al.* (1983). Underlined nucleotides differ from the Tc1 sequence of Rosenzweig *et al.* (1983).

EcoRV site present in the Tc1 inverted terminal repeat 20 nucleotides from the end. Since none of the five had the same structure, they must have been derived from a fairly large heterogeneous population of such molecules. Yet *C. elegans* germline DNA of the Bristol and Bergerac BO strains has been shown to contain only a single Tc1 element lacking a terminal EcoRV site (14) (strain TR679 is a hybrid strain derived by crossing Bristol and Bergerac BO (18)). This variant element resulted from insertion of another transposon (Tc2) within a Tc1 element, and does not have the structure of the molecules isolated here (14). We conclude that these 5 clones did not arise by amplification of germline tandem Tc1 dimers. If they resulted from amplification of Tc1 sequences integrated into high molecular weight DNA, they must derive from structures unique to somatic DNA.

Structure of joined Tc1 ends

The origin and structure of all the sequenced PCR products are given in Table 1, and the corresponding DNA sequences are given in Table 2. Among the 24 independent clones, only 3 had the same structure. This indicates that the pool of molecules serving as PCR templates was large and heterogeneous. Fourteen of the 24 clones contained at least one intact Tc1 end, and 6 additional clones had at least one end with only 1 or 2 nucleotides removed. This supports the inference that the Tc1 transposase was involved in their generation, and argues strongly against the possibility of a PCR artifact. Five of the clones had two intact transposon ends and from one to six intervening TA dinucleotides. Since only a small number of clones were sequenced from some strains, we cannot determine whether the types of sequences differed between strains.

DISCUSSION

We present the structure of 24 PCR amplification products generated by primers directed outwards from within the Tc1 transposon. Twenty of the 24 products contained an end-to-end junction of Tc1 sequences at or within 2 nucleotides of the end of Tc1. This structure argues strongly that recognition of the end by a Tc1-specific transposase was involved in joining the Tc1 sequences, and against the possibility that they have artifactual origin from two template molecules. Thirteen of the 24 products were amplified from gel-fractionated templates, providing direct evidence that they are from extrachromosomal DNA molecules. A variety of explanations for products amplified from unfractionated genomic DNA are considered.

In all the strains tested, Tc1 elements undergo excision in somatic tissues at high frequency. This is the predominant Tc1 activity, exceeding germline activity by several orders of magnitude (20). The molecules sequenced here are therefore likely to have resulted from amplification of extrachromosomal Tc1 molecules generated by somatic excision events.

The data indicate that circular extrachromosomal Tc1 molecules have a heterogeneous structure, usually, if not always, with nucleotides deleted from one or both ligated transposon ends. Molecules with two intact Tc1 ends might serve as competent transposition intermediates. However, the 5 clones with this structure isolated in these experiments (B10, B14, T8, T12 and T23) were all derived from amplification of unfractionated genomic DNA. They are therefore not unambiguously derived from extrachromosomal sequences. In contrast, all of the clones that were isolated after amplification of purified extrachromosomal DNA had nucleotides deleted from one or both

inverted repeats. In addition, there were two instances where one entire inverted repeat was deleted.

Do deleted extrachromosomal circles arise as a result of imprecise excision?

One common feature of the five molecules with two complete ends is the presence of TA dinucleotide(s) at the junction. Previous studies of Tc1 excision sites in the gene *unc-54* indicated Tc1 excision is most often imprecise; in fact, only one of 20 excision sites sequenced was precise (21). In addition, the insertion sites for Tc1 seem to be specific with an unambiguous preference to insert at a TA dinucleotide. Also, examination of chromosomal products of somatic excision indicated eight out of eleven sites were precise, and the remaining three sites contained four extra nucleotides at the site of excision, indicating imprecise excision (22). It may be possible the TA junction sequences we identified between two complete Tc1 ends arose either by imprecise excision or through duplication of the termini of the element prior to religation.

Of the 13 clones that were isolated after amplification of purified extrachromosomal DNA, all had at least one complete end with the other end having deleted nucleotides, or two ends with deletions in both. Such molecules would not be expected to successfully reintegrate. They might arise by ligation of linear extrachromosomal molecules after exonucleolytic attack, or by imprecise excision. Possibly extrachromosomal molecules of this structure accumulate because they are unable to reintegrate. Alternatively, competent transposition intermediates might be regenerated utilizing sequence information from the intact end. This might be particularly the case if multimeric extrachromosomal molecules are present.

Linear and circular extrachromosomal transposons containing only transposon sequences or deleted versions of elements have also been detected for Tn10 (23), Mu (24) and copia (25). Although copia is a retrotransposon, the overall structure of the copia junction sequences resembled those found in Tc1. These extrachromosomal molecules contained 1 to 15 nucleotides inserted between the direct repeats. One extrachromosomal circle had a single terminal repeat and a 500 bp deletion (26). It was suggested the origin of two copia circles containing 15 nucleotide insertions may be due to imprecise excision.

ACKNOWLEDGMENTS

We are grateful to the members of the Emmons lab for their advice and comments on the manuscript. We thank Dr Colin Campbell for numerous discussions and advise. This work was supported by NIH grant GM38174. ADR was supported by Training Grant T32CA09060. SWE was the recipient of the Harry Winston Established Fellowship for Research of the American Heart Association.

REFERENCES

- Moerman, D.G., and Waterston, R.H. (1989) DNA, D.E. Berg and M.M. Howe, eds. American Society for Microbiology, pp537–556.
- Radice, A.D., Bugaj, B., Fitch, D.H.A. and Emmons, S.W. PNAS (in press).
- Rosenzweig, B., Liao, L., and Hirsh, D. (1983) NAR 11, 4201–4209.
- Schukkink, R.F. and Plasterk, R.H.A. (1990) NAR 18, 895–900.
- Berg, D.E., and Howe, M.M. (1989) Mobile DNA. Washington, D.C. American Society for Microbiology.
- Emmons, S. W. and Yesner, L. (1984) Cell 36, 599–605.
- Eide, D., and Anderson, P. (1985a) PNAS 82, 1756–1760.
- Moerman, D.G., Benian, G.M., and Waterston, R.H. (1986) PNAS 83, 2579–2583.
- Plasterk, R.H.A. (1991) EMBO J, 10, 1919–1925.
- Plasterk, R.H.A., and Groenen, J.T.M. (1992) EMBO J, 11, 287–290.
- Ruan, K.-S. and Emmons, S. W. (1984) Proc. Natl. Acad. Sci., USA 81, 4018–4022.
- Rose, A.M., and Snutch, T.P. (1984) Nature 311, 485–486.
- Emmons, S.W., Yesner, L., Ruan, K. and Katzenberg, D. (1983) Cell 32, 55–65.
- Levitt, A., and Emmons, S. W. (1989) PNAS 86, 3232–3236.
- Brenner, S. (1974) Genetics 77, 71–94.
- Nigon, V., (1949) Ann. Sci. Nat. Zool., 11, 1–132.
- Mori, I., Moerman, D.G., and Waterston, R.H. (1988) Genetics 120, 397–407.
- Collins, J., Saari, B., and Anderson, P. (1987) Nature 328, 726–728.
- Sulston, J., and Hodgkin, J. (1988) In Wood, W.B. (ed.) Cold Spring Harbor Laboratory Press. pp587–606.
- Eide, D., and Anderson, P. (1985b) Genetics 109, 67–79.
- Eide, D., and Anderson, P. (1988) Mol. Cell Biol. 8, 737–746.
- Ruan, K.-S. and Emmons, S.W. (1987) NAR 15, 6875–6881.
- Benjamin, H.W. and Kleckner, N. (1989) Cell 59, 373–383.
- Sundaresan, V. and Freeling, M. (1987) PNAS 84, 4924–4928.
- Flavell A.J. and Ish-Horowicz, D. (1981) Nature 292, 591–595.
- Flavell A.J. and Ish-Horowicz, D. (1983) Cell 34, 415–419.