MOLECULAR AND CELLULAR BIOLOGY, Mar. 1986, p. 933–941
0270-7306/86/030933-09$02.00/0

Vol. 6, No. 3

Tx1: a Transposable Element from *Xenopus laevis* with Some Unusual Properties

JAMES E. GARRETT AND DANA CARROLL*

Department of Cellular, Viral, and Molecular Biology, University of Utah Medical School, Salt Lake City, Utah 84132

Received 15 October 1985/Accepted 10 December 1985

A family of transposable genetic elements in the genome of the frog, *Xenopus laevis*, is described. They are designated Tx1. Transposability of the elements was deduced by characterization of a chromosomal locus which is polymorphic for the presence or absence of a Tx1 element. Nucleotide sequence analysis suggested that Tx1 elements show target site specificity, as they are inserted at the pentanucleotide TTTAA in all four cases that were examined. The elements appear to have 19-base-pair (bp) inverted terminal repeats, and they are flanked by 4-bp target duplications (TTAA), although the possibility that they do not create target site duplications is discussed. Tx1 elements have several unusual characteristics: the central portion of each element is comprised of a variable number of two types of 393-bp repeating units; the rightmost 1,000 bp of the element contains separate regions potentially capable of forming bends, left-handed Z-form DNA, and alternative stem-loop structures. Comparisons among single frogs suggest that germ line transposition is relatively infrequent and that variations in numbers of internal repeats accumulate quite slowly at any locus.

Transposable genetic elements have been discovered in a variety of organisms. They represent a diverse collection of DNA sequence elements, although some features are held in common. These include the presence of inverted terminal repeats and the apparent generation of duplications of short target sequences. No function has been firmly established for any eucaryotic transposable element, although the consequences of transposition have been elucidated in several cases. Mobile elements can cause mutations by insertion and have been associated with subsequent chromosomal rearrangements (for a review, see reference 39).

While details of transposition processes are still lacking, the known eucaryotic elements can be classified generally on the basis of their apparent mechanism. Some mobile sequences appear to be integrated reverse transcripts of cellular RNAs, like the L1 family in the mouse (34), small nuclear RNA pseudogenes in several mammalian species (12), and the rarer processed pseudogenes of known mRNAs like globin (28) and tubulin (22). These elements do not encode any obvious functions which would facilitate their transposition and can be characterized as unintentional transposons. The retroviruslike elements (termed retroposons), including the copia-like elements of *Drosophila melanogaster* (36) and the yeast Ty elements (33), also move via RNA intermediates (4) but have discrete ends and apparently encode functions needed for transposition. They are characterized by long terminal repeats, full-length RNA transcripts, and sequence homologies with known retroviruses (10, 40). Another broad category includes elements, like P elements of *D. melanogaster* (29) and Ac elements in maize (14), which support their own transposition but which probably do not involve RNA intermediates. It seems likely that they move by DNA-based mechanisms similar to ones proposed for bacterial transposable elements (39).

Most of the transposable elements of this latter class identified to date reside in genomes of procaryotes, lower eucaryotes, and plants. We report here evidence for a family of transposable elements in the genome of the vertebrate *Xenopus laevis*. We call these elements Tx1 (transposon from *X. laevis* 1). They are discrete but not retroviruslike, so they fall into the category with P and Ac elements. Tx1 elements share some structural features with other mobile sequences, but have unique features as well.

MATERIALS AND METHODS

DNA preparation. *X. laevis* DNA was prepared from blood by previously described methods (6). Plasmid and phage λ (Charon 4) DNAs were prepared by standard methods (3, 11, 23).

DNA analysis. Restriction enzymes were purchased from commercial suppliers and used according to their recommendations. Electrophoresis in agarose or acrylamide gels was performed by standard procedures (23). Southern blot hybridization was done using the dextran sulfate-accelerated procedure (42), and plasmid DNA was nick translated to high specific activity (>10^8 cpm/μg). Routinely, the hybridization solution contained 50% formamide–1 × 10^6 to 2 × 10^6 cpm of the appropriate probe per ml, and hybridization was carried out at 42°C for 24 h.

A *X. laevis* genomic library (in the λ vector Charon 4) (41) was screened by an in situ plaque hybridization technique (1). Subcloning of restriction fragments was by standard methods (23).

DNA sequence analysis. Sequence determinations were performed by the chemical degradation method (25) and the dideoxy chain-terminator method (37). For sequencing by the chemical degradation method, restriction fragments were end labeled with ^{32}P using T4 polynucleotide kinase or avian myeloblastosis virus reverse transcriptase (23). For sequencing by the chain-terminator method, restriction fragments were subcloned into pEMBL vectors (13) to allow the generation of single-stranded templates. To sequence long inserts, ordered deletion series were generated with exonuclease III and S1 nuclease by the method of Henikoff (19). [α-^{35}S]deoxynucleoside triphosphates were used as the label in the sequencing reactions (2), and the sequencing gels were

* Corresponding author.

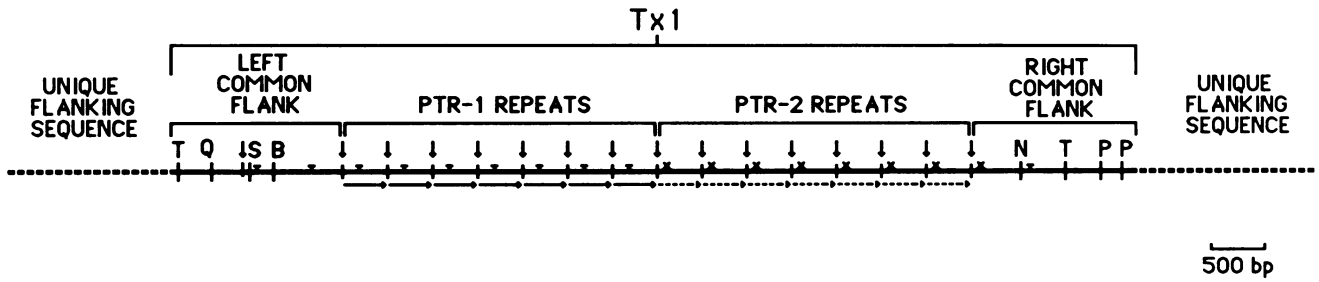

FIG. 1. Consensus restriction enzyme map of a Tx1 element. The map was derived from eight genomic clones, selected by hybridization to PTR-2 sequences, that contained complete Tx1 elements (9). Individual Tx1 elements vary in the number of both types of PTR repeats; the map is drawn with the average number of seven repeats of each type. The ends of the Tx1 elements were determined by S1 nuclease digestion of heteroduplexed Tx1 elements from different genomic clones (data not shown). Outside the regions of the Tx1 elements, the genomic clones all differed. Abbreviations: X, *Acc*I; ↓, *Hin*dIII; ▼, *Sau*961; B, *Bcl*I; N, *Nde*I; P, *Pst*I; Q, *Taq*I; S, *Stu*I; T, *Tha*I.

dried onto glass plates to improve resolution (17). DNA sequences were analyzed with the BIONET sequence analysis programs (Intelligenetics, Stanford, Calif.).

RESULTS

Tx1 element structure. The elements we now call Tx1 were described previously as members of a family of dispersed, repetitive sequences from the *X. laevis* genome, which we then called paired tandemly repeat (PTR) elements (9). The central portions of these elements consist of variable numbers of two different types of 393-base-pair (bp) repeating units called PTRs (Fig. 1). The average number of repeats is seven for both PTR-1 and PTR-2, but elements with as few as 1 or as many as 15 of either repeat have been identified.

All Tx1 elements share roughly 1,000 bp of sequence to the left of PTR-1 (left common flank [LCF]) and another 1,000 bp to the right of the PTR-2 cluster (right common flank [RCF]) (Fig. 1). S1 nuclease protection experiments showed that the ends of all the elements are discrete. Beyond these shared sequences, the context of each element seemed to be unique. There are approximately 750 Tx1 elements in the *X. laevis* genome, and in situ hybridization experiments (D. C. Jamrich and M. Jamrich, unpublished data) have shown them to be distributed to all chromosome arms. The discrete ends and dispersed locations of the elements suggest that they might be transposable.

Tx1 element polymorphism at a specific locus. Because there are so many Tx1 elements, it is not possible to examine individual loci using Tx1 sequences as hybridization probes. To analyze specific chromosomal sites for Tx1 insertion polymorphism, we identified and subcloned unique sequences adjacent to the elements in our original lambda clones (9). These were used as probes in Southern blot hybridizations to digested DNAs from individual frogs.

One such locus is diagrammed in Fig. 2. The 1,100-bp *Hin*dIII-*Bgl*II fragment to the right of the Tx1 element in lambda clone B4 was subcloned to yield the plasmid pM11. When this was used to probe Southern blots of DNAs from

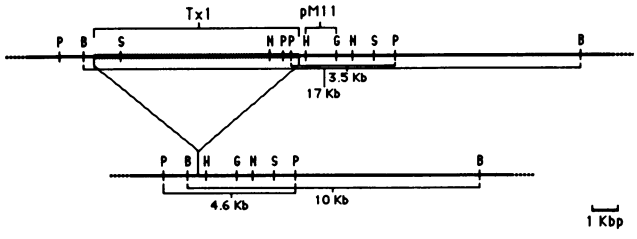

FIG. 2. Restriction map of the genomic locus represented by the B4 clone. The upper map shows the locus containing a Tx1 element; the 1,100-bp *Hin*dIII-*Bgl*II fragment that was subcloned to give the plasmid pM11 is indicated. The lower map diagrams the positions of restriction sites at the same locus but that lack the Tx1 element. The lengths of fragments expected in *Bam*HI and *Pst*I digests are shown for both versions. Restriction sites are abbreviated as follows: B, *Bam*HI; G, *Bgl*II; H, *Hin*dIII; N, *Nde*I; P, *Pst*I; S, *Stu*I.

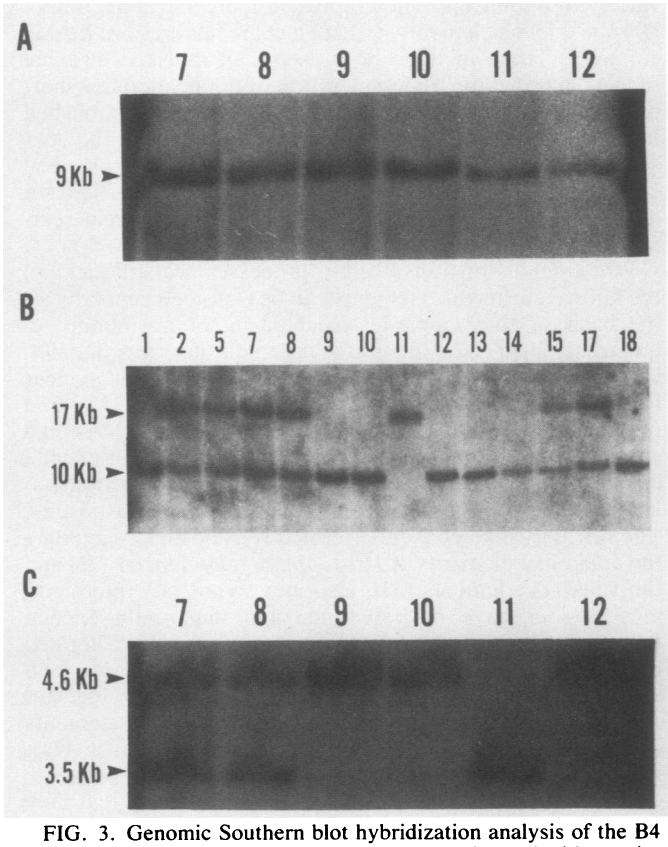

FIG. 3. Genomic Southern blot hybridization analysis of the B4 locus. DNA (10 μg) from individual frogs was digested with restriction enzyme(s), fractionated on agarose gels, transferred to nitrocellulose membranes, and then hybridized with nick-translated pM11. The numbers above the lanes indicate the animals from which DNA was obtained, and is consistent among the three autoradiographs in panels A, B, and C. The digests were *Hin*dIII-*Bam*HI (A), *Bam*HI (B), and *Pst*I (C). The sizes of hybridizing bands are indicated.

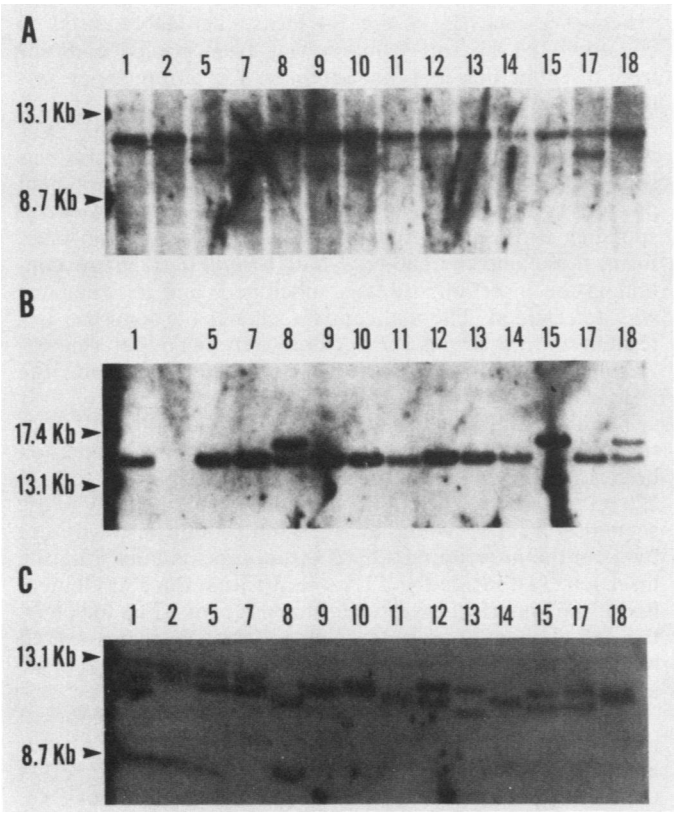

FIG. 4. Genomic Southern blot analysis of Tx1-containing genomic loci A1, B1, and C4. DNA from individual frogs was digested with restriction enzyme(s), size fractionated on agarose gels, transferred to nitrocellulose filters, and then hybridized to nick-translated probes derived from single-copy sequences located adjacent to cloned Tx1 elements (9). The numbers above the lanes are as described in the legend to Fig. 3. (A) DNAs were digested with *Eco*RI-*Bgl*II. The probe was a 1,400-bp *Pst*I fragment located to the right of the Tx1 element in the genomic clone A1. (B) DNAs were digested with *Eco*RI. The probe was a 1,600-bp *Hin*dIII-*Eco*RI fragment located to the left of the Tx1 element in the genomic clone B1. (C) DNAs were digested with *Bgl*II. The probe was a 1,900-bp *Eco*RI fragment located to the right of the Tx1 element in the genomic clone C4.

single frogs digested with both *Hin*dIII and *Bam*HI, it hybridized only to a 9-kilobase (kb) fragment of single-copy intensity (Fig. 3A). This digest unlinks the probe sequences from the Tx1 element and measures the distance to the nearest *Bam*HI site to the right of the probe. Digestion with *Bam*HI alone yields fragments spanning both the region of the probe and the site of insertion. Hybridization of pM11 to such digests showed bands at 17 kb, which is the expected size when a Tx1 element is present, and 10 kb, which is expected if the element is absent (Fig. 3B). Some individual frogs were homozygous for the 17- or the 10-kb band, and heterozygotes showing both bands were found.

Digests with other restriction enzymes (*Bgl*II, *Eco*RI) which generate fragments that span the probe and the insertion site gave comparable results: two bands that were different by 7 kb were found (data not shown). Judgments concerning particular frogs were entirely consistent with the *Bam*HI digests. Of 14 frogs examined, 1 was homozygous $Tx1^+$, 7 were homozygous $Tx1^-$, and 6 were heterozygous (at the B4 locus).

Enzymes that cut within Tx1 were also used to characterize the B4 locus in these frogs. *Pst*I cleaved within RCF and yielded a 3,500-bp fragment that hybridized to pM11 when Tx1 was present; when Tx1 was absent, the distance between the flanking *Pst*I sites was 4,600 bp (Fig. 2). The expected bands were indeed seen, and their occurrence in individual frogs was just as predicted based on the spanning digests (Fig. 3C). Equivalent results were obtained with three other enzymes that cut both inside and outside the element: *Stu*I, *Acc*I, and *Nde*I (data not shown).

The frogs used in these studies were from several different sources. Many were caught in the wild at different periods of time. We presume that they were not closely related but were representative of the *X. laevis* population. Our analysis indicates that both the $Tx1^+$ and $Tx1^-$ alleles of the B4 locus depicted in Fig. 2 are spread throughout the population of that species.

Analysis of other specific loci. Similar analyses were done with single-copy probes from sequences flanking three other Tx1 elements, corresponding to the lambda clones A1, B1, and C4 (9). In all three cases, when spanning digests were examined, one or two bands of single-copy intensity were observed in the DNA from each single frog (Fig. 4). Unlike the situation at the B4 locus, the bands differed from each other by one or a few multiples of 400 bp, rather than 7 kb,

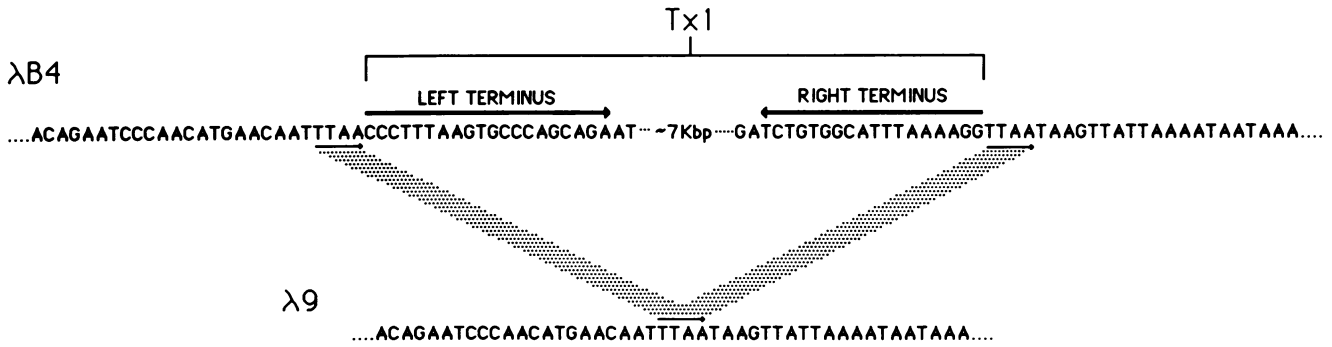

FIG. 5. Comparison of sequences of inserted (λB4) and uninserted (λ9) alleles of the B4 locus. Restriction fragments spanning the left and right termini of the Tx1 element on the λB4 clone were sequenced by the method of Maxam and Gilbert (25). A restriction fragment containing the Tx1 insertion site on the genomic clone of the unoccupied B4 locus (λ9) was subcloned and sequenced by the dideoxy chain-terminator method. Large arrows mark the 19-bp inverted terminal repeats of the Tx1 element. The smaller arrows mark the TTAA direct repeats flanking the Tx1 element in λB4, one copy of which was found at the corresponding location in the λ9 sequence. A total of 220 bp of sequence flanking the insertion site were compared and found to match exactly. Kbp, kilobase pairs.

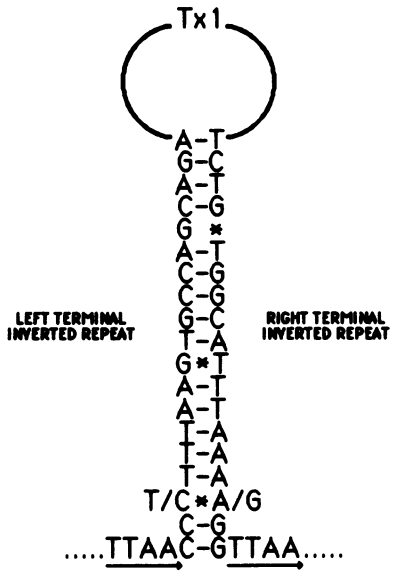

FIG. 6. Tx1 element inverted terminal repeats. The repeats are depicted with the left repeat base-paired to the right repeat. There are two mismatches and one insertion or deletion in the 19-bp sequence. The third nucleotide into the inverted terminal repeat varies between Tx1 elements. The TTAA direct repeats flanking the inverted terminal repeats could also be construed as part of the inverted repeats.

and (with one exception) all were in the range of sizes expected for the particular locus with a resident Tx1 element. We interpret the polymorphism that was observed as variations in the number of PTR-1 and PTR-2 repeats in the Tx1 elements at the same site in different chromosomes. We conclude that all 28 chromosome sets examined contain elements at all three sites.

The elements at the A1 and B1 loci showed a relatively low degree of polymorphism, while that at C4 was highly polymorphic. The Tx1 element at the C4 locus had an unusual structure. Both in the original lambda clone and in the chromosomal versions that we analyzed, this element lacked RCF sequences. We suggest that the deletion of one terminus may have stranded the Tx1 element at this site, rendering it unable to transpose or excise and that the observed variants may have accumulated over a very long period of time. Tx1 elements at the other sites may be more recent arrivals and, consequently, have a lower degree of divergence from their founder elements.

Sequence analysis of the B4 locus. The DNA used to construct the Charon 4 library was from pooled embryos from three matings and thus represents 12 chromosome sets (43). We screened digests of the DNA used for cloning with pM11 and found that $Tx1^-$ B4 loci outnumbered the $Tx1^+$ version in this pool (data not shown). The library was screened with pM11 used as the probe, and a recombinant phage was isolated (called λ9) which had an insert corresponding to the uninterrupted locus in its restriction map. From this clone the 580-bp *Hin*dIII-*Bam*HI fragment containing the insertion site was subcloned, and its sequence was determined. The nucleotide sequences around the Tx1 termini in the original B4 clone were also determined. Results of the comparison of these two clones are shown in Fig. 5.

The $Tx1^+$ B4 sequence corresponded to a simple interruption of the $Tx1^-$ λ9 sequence. Flanking this interruption on both sides, the two sequences were identical over the total of 220 bp compared. The uninterrupted site had one copy of the sequence TTAA, while this tetranucleotide was present twice in the interrupted site as direct repeats bordering the insertion. Just inside the TTAA sequences, the Tx1 element had 19-bp imperfect inverted terminal repeats (Fig. 6). Thus, the Tx1 element appears to have two characteristics shared by most transposable elements: terminal inverted repeats and flanking duplications of target sequences.

It should be pointed out that we do not know the historical relationship between the $Tx1^+$ and $Tx1^-$ B4 loci. All our evidence suggests that they are allelic. We argue below that the unoccupied version is a preinsertion site. It is possible, however, that it resulted from the excision of a Tx1 element. In either case, it is clear that Tx1 is capable of movement as a discrete entity within the *X. laevis* genome.

Sequences of additional Tx1 termini. The nucleotide sequences of junction regions, where Tx1 meets the surrounding sequence, were determined for three additional clones. Fragments covering the left and right junctions were subcloned from lambda clones A1, B1, and B10 (9). Portions of their sequences are presented in Fig. 7, along with that of clone B4.

The ends of all four elements are essentially completely homologous, out to and including the 19-bp inverted repeat. Like the B4 element, all are flanked by direct repeats of the sequence TTAA. In addition, every element is preceded on the left by another T residue. This feature creates ambiguity in the interpretation of the element and target sequences.

Most transposable elements create a short target site duplication on insertion (7), and some elements show a relatively high degree of target specificity. For example, the

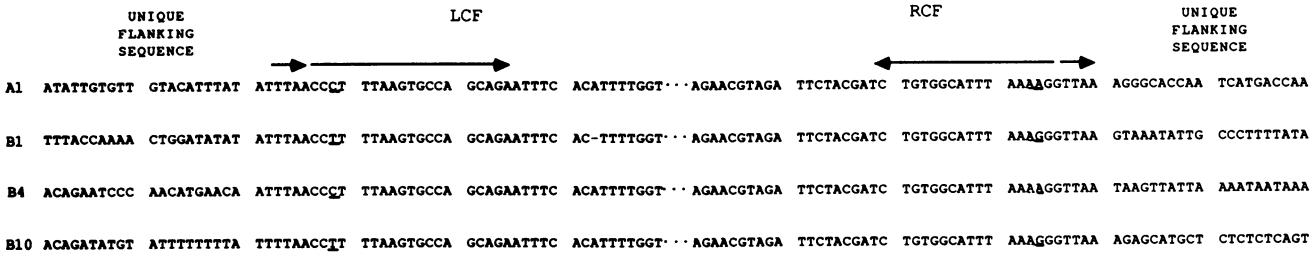

FIG. 7. Sequence analysis of the regions surrounding Tx1 element termini. The sequences of both left and right terminal regions were determined for the Tx1 genomic clones A1, B1, B4, and B10 (9). Restriction fragments covering the termini of the cloned Tx1 elements were isolated and sequenced by the method of Maxam and Gilbert (25). In most instances, one strand was sequenced. Various lengths of sequence were determined; only the regions immediately surrounding the Tx1 termini are shown. The arrows mark the TTAA direct repeats and the LCF and RCF inverted terminal repeats. The third nucleotides into the inverted repeats, which vary from clone to clone, are underlined.

gypsy element of *D. melanogaster* appears to insert specifically at the sequence TACATA, duplicating the first four of those nucleotides (16). By analogy, Tx1 may recognize the sequence TTTAA and duplicate TTAA on insertion. Alternatively, the ends of the element could be AA. . .TT, and an insertion into the specific target site TTTAA by a blunt-end cut between T and A in the target would give rise to the apparent TTAA direct repeats without actually generating a duplication of the original target sequence.

Comparison of sequences within these terminal regions of the Tx1 elements showed less than 1% overall variation (data not shown). However, there was one consistent and curious variation in the third nucleotide of the 19-bp inverted terminal repeat on both ends of the element (Fig. 7). This position was occupied by either C or T in the left inverted repeat and by either A or G on the right, but it varied such that there was a mismatch between these positions when the ends of a given element were compared.

Apart from the inverted terminal repeats, the right and left ends of Tx1 had no obvious relationship with each other. There were no outstanding features of the genomic sequences surrounding the elements, and the 580 bp sequenced from the uninterrupted B4 locus had multiple translational stop codons in all six potential reading frames.

Sequence of a complete Tx1 element. Nucleotide sequences of several PTR-1 and PTR-2 repeats have been determined previously, and consensus sequences have been reported (9). To complete the analysis of a whole Tx1 element, the example in the λB4 clone was chosen. Its structure appeared to be representative of the other cloned elements, and its polymorphism in frog populations suggested that it might have undergone a relatively recent transposition event. Restriction fragments containing LCF and RCF regions from the λB4 clone were subcloned into pEMBL vectors (13). Deletion series were generated with exonuclease III and S1 nuclease (19), and the sequences were determined (Fig. 8 and 9).

Much of LCF was composed of diverged relatives of the PTR-1 sequence. The divergence increases as one reads leftward toward the end of Tx1 (Fig. 8). This terminal deterioration of the sequence is characteristic of other clusters of tandem repeats, e.g., the 15-bp internal repeats in the spacer of *X. laevis* 5S DNA (15). Approximately 400 bp at the left terminus of Tx1 had no discernible relationship to PTR-1 or to any other sequences in the element, apart from the 19-bp inverted terminal repeat. This region had no long open reading frames or other outstanding features.

On the right side of the element, the transition between PTR-2 repeats and RCF sequence was abrupt (Fig. 9). Reading rightward from the last *Hin*dIII site of the PTR-2 repeats, there are 148 bp of exact homology with the PTR-2 consensus, beyond which the remaining 998 bp of the element show no relationship to PTR sequences.

There were several open reading frames in the RCF sequence. The longest of these, reading from nucleotides 1092 to 597, could encode 165 amino acids. There were no transcription initiation signals preceding this open reading frame, but it was oriented such that transcription originating outside Tx1 could proceed across it. Some of the smaller open reading frames could conceivably be spliced into longer continuous, translatable sequences. We found no evidence for transcripts of Tx1 in a variety of frog tissues and developmental stages (21; unpublished data), but very low levels could have escaped detection.

The RCF sequence contained three other regions of particular interest. Between nucleotides 350 and 450 there were a number of short clusters of A residues spaced about one helix turn apart. Restriction fragments containing this region had abnormally low electrophoretic mobility in polyacrylamide gels (data not shown). Similar sequences in other DNAs have been shown to produce kinks or bends which are responsible for electrophoretic retardation and other abnormal physical attributes (18, 24, 44). At least one example exists in which alteration of such sequences affects expression of a neighboring gene (5).

Second, between nucleotides 560 and 755 there were several extended stretches comprised mostly of alternating purine and pyrimidine residues which were dominated by G and T. These RCF sequences did not hybridize at normal stringencies with the thousands of dispersed clusters of $(CA)_n$:$(GT)_n$ repeats in the *X. laevis* genome (26). Nonetheless, these alternating purine-pyrimidine sequences are long enough potentially to adopt the Z form of DNA (32).

Finally, a number of direct and inverted repeats were found within the terminal 85 bp of RCF. There are two mutually exclusive secondary structures that can be proposed for one strand of this sequence; these structures have either one or two stem loops (Fig. 10). It may be significant that this region is positioned near the beginning of the longest open reading frame in Tx1. Alternative secondary structures have been shown to participate in regulation of bacterial operon expression (45). Whether such structures are ever formed in Tx1 sequences in their natural setting is not known.

DISCUSSION

Tx1 elements comprise a highly homogeneous family of dispersed, repetitious sequences in the *X. laevis* genome. Analysis of a specific genomic locus indicates that Tx1 is capable of transposition as a discrete entity. Furthermore, the element shares nucleotide sequence features with transposable elements in other organisms. It has inverted terminal repeats 19 bp in length, including three mismatches. This is similar in length and degree of homology to termini of other eucaryotic and procaryotic transposons (35, 36, 41).

Nearly all known transposable elements are found to be bordered by direct repeats of a short sequence that was present only once at the target site prior to insertion. It is thought that this target duplication arises from the repair of staggered single-strand breaks introduced into the recipient DNA during the insertion event (38). For most transposons, the size of the target duplication is constant, but its sequence is variable. Some elements, however, display a strong preference for particular target sequences. For example, the *D. melanogaster* element, gypsy, inserts specifically into the sequence TACATA, generating a duplication of TACA (16).

The target site situation for Tx1 is somewhat ambiguous. We believe that the unoccupied B4 locus which we sequenced represents a preinsertion site, not a postexcision site. Comparison of the four Tx1 loci that were characterized indicates that elements completely spread throughout the population are more internally polymorphic. The fact that the element at the B4 locus is identical in the restriction map of all frogs that carry it suggests that it is a relatively recent arrival and is in the process of spreading through the frog population.

If this is true, then the Tx1 element at B4 is indeed flanked by direct repeats of a short sequence found only once at the unoccupied target. Precisely the same 4-bp sequence (TTAA) was found on both ends of all four Tx1 elements that were examined. Reconstruction of the target sites, using the B4 locus as a model, indicates that all the elements inserted

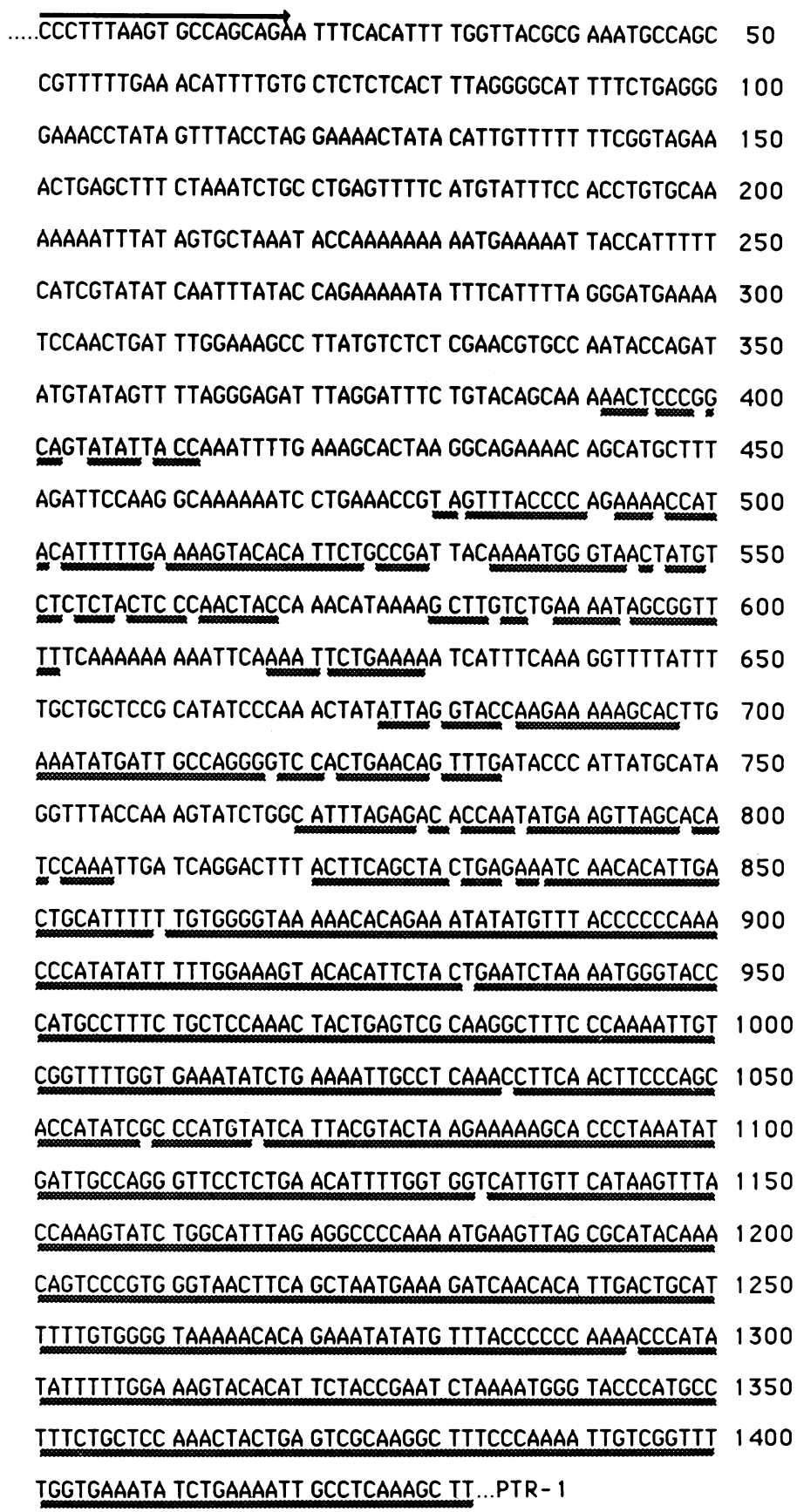

FIG. 8. Nucleotide sequence of the Tx1 LCF region. A restriction fragment covering this region (from the genomic clone B4) was subcloned into a pEMBL vector, an ordered series of deletions was generated with exonuclease III and S1 nuclease, and the sequence was determined by the dideoxy chain-terminator method. Both strands were completely sequenced. Nucleotide 1 is the first nucleotide of the 19-bp inverted terminal repeat, which is indicated by an arrow. The last six nucleotides shown (1427 to 1432) constitute the first *Hin*dIII site of the regular PTR-1 repeats. The underlined sequence shows regions of homology between LCF and PTR-1 sequences.

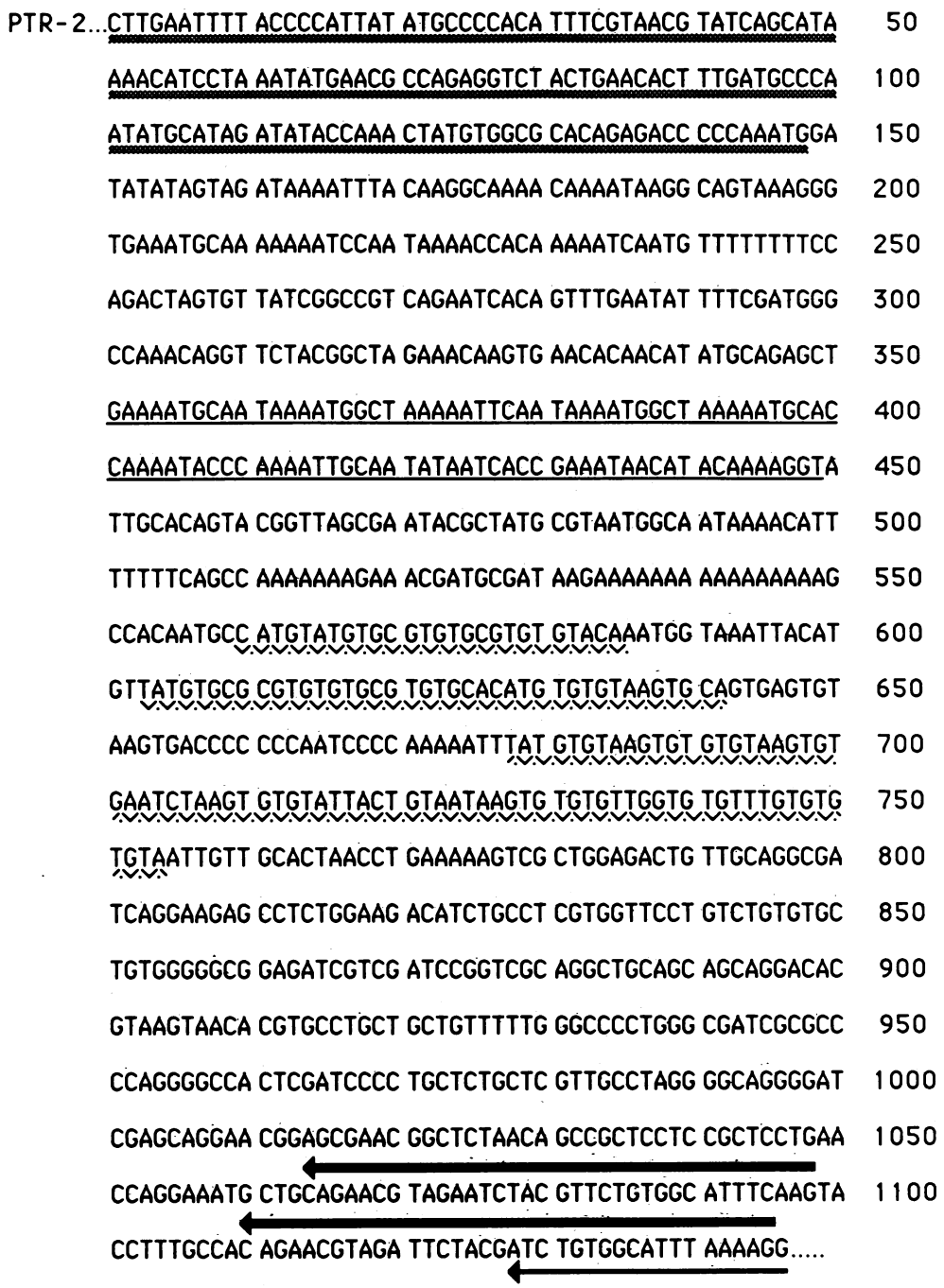

FIG. 9. Nucleotide sequence of the Tx1 RCF region. A restriction fragment covering this region (from the Tx1 genomic clone B4) was subcloned into a pEMBL vector, and its sequence was determined as described in the legend to Fig. 8. Both strands were completely sequenced. Nucleotide 1 is at the middle of the last *Hin*dIII site of the PTR-2 repeats. Nucleotide 1146 marks the right terminus of Tx1. The underlined sequences indicate the region of homology to PTR-2 (positions 1 to 148), the region of kinked DNA (positions 350 to 450), and the regions of extensive alternating purine and pyrimidine residues (positions 560 to 755). The bold arrows show the two direct repeats near the terminus of the RCF region; the thinner arrow marks the inverted terminal repeat.

into the sequence TTTAA. Unfortunately, this means that the ends of the element cannot be defined precisely. Either the element carries a 19-bp inverted terminal repeat and generates a 4-bp target duplication (TTAA) or it carries a 21-bp inverted terminal repeat, including AA on one end and TT on the other, and creates no target duplication. Analogy with other transposable elements suggests that the former is more likely, but there are precedents for blunt-end insertion (8, 27), and a similar ambiguity exists for the nematode element Tc1 (35).

Tx1 also has sequence features which are not common among transposable elements. The most obvious of these is the presence of variable numbers of two types of internal repeats, the PTR-1 and PTR-2 units. No function is currently ascribable to these repeats, but both sequences are present in elements very similar to Tx1 in the related species

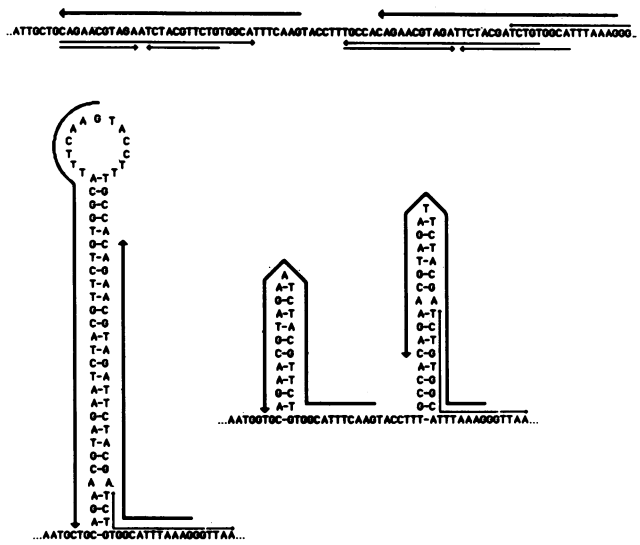

FIG. 10. Direct and inverted repeats at the right terminus of Tx1. The sequence at the right terminus of Tx1 is shown. The smaller arrow above the sequence indicates the location of the inverted terminal repeat; the larger arrows show the two direct repeats. The arrows below the sequence indicate the positions of the inverted repeats. Below, the sequence is drawn in the two potential stem-loop structures that it could assume.

Xenopus borealis and *Xenopus mulleri* (9). It is curious that another family of possible *X. laevis* transposons also carries internal repeats (20). These 1723 elements are abundant, dispersed, and have inverted terminal repeats and flanking 8-bp direct repeats, which is suggestive of a mobile life-style. Length variation among different 1723 elements is due to variable numbers of a 180-bp repeating unit. The FB elements of *D. melanogaster* have variable numbers of tandem repeats, but these are terminal and inverted (30). Other transposon families have members that vary in size, but the variations are due to irregular deletions rather than regular variation in numbers of internal repeats.

Other unusual features of Tx1 elements include the variable sequence, but conserved mismatch, at the third position of the 19-bp inverted terminal repeat. The precise sequences of the ends of transposable elements are important for the transposition process, so this mismatch may be required for or generated by transposition events, even though there is no evidence for direct interaction between the left and right termini. Sequences in the RCF region capable of forming bends and left-handed Z-form helices are intriguing but of unknown significance. Several other transposable elements have been shown to have terminal regions, like those in RCF, that are capable of forming stem loops (31); however, no definite role in transposition has been assigned to them.

A surprising feature of the Tx1 elements is the limited degree of heterogeneity in the numbers of PTR-1 and PTR-2 repeats at a given locus. On growth of lambda phage carrying Tx1 elements, variants are rapidly accumulated due to unequal crossing over between these repeats (9). In contrast, the elements in frogs are entirely homogeneous at the B4 locus. Those at the A1 and B1 loci are ancient enough to have spread through the whole population, yet very few different alleles are present and the overall level of polymorphism is less than 20%. The C4 locus is highly polymorphic, but the damaged element at that site may have been stranded for a much longer time. The low level of variation indicates that very little unequal crossing over is occurring at a given site and that there is very little genetic interaction between Tx1 elements at different loci, because these are known to be quite heterogeneous in their content of PTR repeats (9, 21).

Our characterization of the Tx1 elements has been confined to elucidation of their structure. We have no direct evidence for any consequences, positive or negative, they may have for the frogs or their genetic organization. Clearly, individual frogs can exist quite healthily when homozygous for either the interrupted or the uninterrupted version of the B4 locus. Very similar element families are present in the genomes of *X. borealis* and *X. mulleri* (9), but no cross-hybridizing sequences have been found outside the genus *Xenopus* (unpublished data).

ACKNOWLEDGMENTS

We are grateful to Charles Hussey for assistance in some of the nucleotide sequence determinations and to Igor Dawid for providing the *X. laevis* library and the mixed embryo DNA from which it was derived. We appreciate continuing discussions with Joe Gall, Milan Jamrich, Sue Bromley, Igor Dawid, Glenn Herrick, and Costa Georgopoulos and comments on the manuscript by G. Herrick and C. Georgopoulos. D.C. thanks Joe Gall and colleagues for their hospitality during a sabbatical visit to their laboratory.

This work was supported by Public Health Service research grant GM22232 from the National Institutes of Health. During part of the period of this work, D.C. was an American Cancer Society Scholar in Cancer Research.

LITERATURE CITED

1. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
2. **Biggen, H., T. J. Gibson, and D. F. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
3. **Birnboim, H. C.** 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. **100**:243–255.
4. **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.
5. **Bossi, L., and D. M. Smith.** 1984. Conformational change in the DNA associated with an unusual promoter mutation in a tRNA

operon of *Salmonella*. Cell **39**:643–652.
6. **Brown, D. D., P. C. Wensink, and E. Jordan.** 1971. Purification and some characteristics of 5S DNA from *Xenopus laevis*. Proc. Natl. Acad. Sci. USA **62**:3175–3179.
7. **Calos, M., and J. Miller.** 1980. Transposable elements. Cell **20**:579–595.
8. **Cappello, M., S. M. Cohen, and H. F. Lodish.** 1984. *Dictyostelium* transposable element DIRS-1 preferentially inserts into DIRS-1 sequences. Mol. Cell. Biol. **4**:2207–2213.
9. **Carroll, D., J. E. Garrett, and B. S. Lam.** 1984. Isolated clusters of paired tandemly repeated sequences in the *Xenopus laevis* genome. 1984. Mol. Cell. Biol. **4**:254–259.
10. **Clare, J., and P. J. Farabaugh.** 1985. Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. Proc. Natl. Acad. Sci. USA **82**:2829–2833.
11. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA **62**:1159–1166.
12. **Denison, R., and A. Weiner.** 1982. Human U1 RNA pseudogenes may be generated by both DNA- and RNA-mediated mechanisms. Mol. Cell. Biol. **2**:815–828.
13. **Dente, L., G. Cesareni, and R. Cortese.** 1983. pEMBL: a new family of single stranded plasmids. Nucleic Acids Res. **11**:1645–1655.
14. **Doering, H.-P., and P. Starlinger.** 1984. Barbara McClintock's controlling elements: now at the DNA level. Cell **39**:253–259.
15. **Fedoroff, N. V., and D. D. Brown.** 1978. The nucleotide sequence of oocyte 5S DNA in *Xenopus laevis*. I. The AT-rich spacer. Cell **13**:701–716.
16. **Freund, R., and M. Meselson.** 1984. Long terminal repeat nucleotide sequence and specific insertion of the gypsy transposon. Proc. Natl. Acad. Sci. USA **81**:4632–4636.
17. **Garoff, H., and W. Ansorge.** 1981. Improvements of DNA sequencing gels. Anal. Biochem. **115**:450–457.
18. **Hagerman, P.** 1984. Evidence for the existence of stable curvature of DNA in solution. Proc. Natl. Acad. Sci. USA **81**:4632–4636.
19. **Henikoff, S.** 1984. One-way exonucleolytic digestion to target deletion end-points for DNA sequencing. Gene **28**:331–359.
20. **Kay, B., and I. B. Dawid.** 1983. The 1723 element: a long, homogeneous, highly repeated DNA unit interspersed in the genome of *Xenopus laevis*. J. Mol. Biol. **170**:583–596.
21. **Lam, B. S., and D. Carroll.** 1983. Tandemly repeated DNA sequences from *Xenopus laevis*. II. Dispersed clusters of a 388 bp repeating unit. J. Mol. Biol. **165**:587–597.
22. **Lemischka, I., and P. A. Sharp.** 1982. The sequences of an expressed rat α-tubulin gene and a pseudogene with an inserted repetitive element. Nature (London) **300**:330–335.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Marini, J. C., S. D. Levine, D. M. Crothers, and P. T. Englund.** 1982. Bent helical structure in kinetoplast DNA. Proc. Natl. Acad. Sci. USA **79**:7664–7668.
25. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
26. **Miesfeld, R., M. Krystal, and N. Arnheim.** 1980. A member of a new repeated sequence family which is conserved throughout eukaryotic evolution is found between the human δ and β globin genes. Nucleic Acids Res. **9**:5931–5947.
27. **Murphy, E., and S. Lofdahl.** 1984. Transposition of Tn554 does not generate a target site duplication. Nature (London) **307**:292–294.
28. **Nishioka, Y., A. Leder, and P. Leder.** 1980. Unusual α-globin-like gene that has cleanly lost both globin intervening sequences. Proc. Natl. Acad. Sci. USA **77**:2806–2809.
29. **O'Hare, K., and G. M. Rubin.** 1983. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34**:25–35.
30. **Potter, S. S.** 1982. DNA sequence of a foldback transposable element in *Drosophila*. Nature (London) **297**:201–204.
31. **Rhodes, P. R., and L. O. Vodkin.** 1985. Highly structured sequence homology between an insertion element and the gene in which it resides. Proc. Natl. Acad. Sci. USA **82**:493–497.
32. **Rich, A., A. Nordheim, and A. Wang.** 1984. The chemistry and biology of left-handed Z DNA. Annu. Rev. Biochem. **53**:791–846.
33. **Roeder, G. S., and G. R. Fink.** 1983. Transposable elements in yeast, p. 299–328. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
34. **Rogers, J.** 1983. A straight LINE story. Nature (London) **306**:113–114.
35. **Rosenzweig, B., L. W. Liao, and D. Hirsh.** 1983. Target sequences for the *C. elegans* transposable element Tc1. Nucleic Acids Res. **11**:7137–7140.
36. **Rubin, G. M.** 1983. Dispersed repetitive DNAs in *Drosophila*, p. 329–362. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Shapiro, J. A.** 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc. Natl. Acad. Sci. USA **76**:1933–1937.
39. **Shapiro, J. A. (ed.).** 1983. Mobile genetic elements. Academic Press, Inc., New York.
40. **Varmus, H. E.** 1985. Reverse transcriptase rides again. Nature (London) **314**:583–584.
41. **Vodkin, L. O., P. R. Rhodes, and R. B. Goldberg.** 1983. cA lectin gene insertion has the structural features of a transposable element. Cell **34**:1023–1031.
42. **Wahl, G. M., M. Stern, and G. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulphate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
43. **Wahli, W., and I. B. Dawid.** 1980. Isolation of two closely related vitellogenin genes, including their flanking regions, from a *Xenopus laevis* gene library. Proc. Natl. Acad. Sci. USA **77**:1437–1441.
44. **Wu, H.-M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. Nature (London) **308**:509–513.
45. **Yanofsky, C.** 1981. Attenuation in the control of expression of bacterial operons. Nature (London) **289**:751–758.