Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 1077–1081, February 1996
Genetics

DANA elements: A family of composite, tRNA-derived short interspersed DNA elements associated with mutational activities in zebrafish (*Danio rerio*)

(evolution/microsatellite/retroposon-like element/polymorphism/repetitive DNA)

ZSUZSANNA IZSVÁK*†, ZOLTÁN IVICS*, DAVID GARCIA-ESTEFANIA*, SCOTT C. FAHRENKRUG*, AND PERRY B. HACKETT‡§

*Department of Genetics and Cell Biology and ‡Institute of Human Genetics, University of Minnesota, St. Paul, MN 55108

Communicated by Allan Campbell, Stanford University, Stanford, CA, September 21, 1995 (received for review May 30, 1995)

ABSTRACT **DANA is the first SINE isolated from zebrafish (*Danio rerio*) exhibiting all the hallmarks of these tRNA-derived elements. DANA is unique in its clearly defined substructure of distinct cassettes. In contrast to generic SINE elements, DANA appears to have been assembled by insertions of short sequences into a progenitor, tRNA-derived element. Once associated with each other, these subunits were amplified as a new transposable element with such a remarkable success that DANA-related sequences comprise ≈10% of the modern zebrafish genome. At least some of the sequences comprised by the full-length element were capable of movement, forming a new group of mobile, composite transposons, one of which caused an insertional mutation in the zebrafish *no tail* gene. Being present only in the genus *Danio*, and estimated to be as old as the genus itself, DANA may have played a role in *Danio* speciation by massive amplification and genome-wide dispersion. There are extensive DNA polymorphisms between zebrafish populations and strains detected by PCR amplification using primers specific to DANA, suggesting that the DANA element will be useful as a molecular tool for genetic and phylogenetic analyses.**

SINEs (*s*hort *in*terspersed *e*lements) are typically nonviral retrotransposable DNA sequences that use RNA as an intermediate of transposition and are prevalent components of eukaryotic genomes (reviewed in refs. 1–4). SINEs range in length from about 75 to 500 bp and are highly repeated (10^3 to 5×10^5 copies). The best characterized SINE is the *Alu* element, a derivative of 7SL RNA, in primates (2). Unlike the *Alu* elements, most SINEs are derived from tRNAs and are often regarded as tRNA pseudogenes (1, 5). SINEs are transcribed by RNA polymerase III (Pol III); the transcripts can then be reverse transcribed and the DNA products can be inserted back into the genome.

Some SINEs have the ability to fuse with other sequences to form composite transposable elements (1–4). For example, the human *Alu* element has a dimeric or, sometimes, tetrameric structure composed of related subunits (2, 6), whereas the prosimian *Galago* type II element represents a heterologous composite retroposon that resulted from a fusion of two different SINEs (4). In both cases, the "right" halves of these composite elements lack any apparent function, indicating that some sequences may be picked up by a SINE if they lie between the Pol III initiation and termination sites. The hallmarks of retroposons are the presence of a tRNA-related region including a conserved Pol III promoter, an A-rich 3′ end, and flanking direct repeats that are generated at the target site upon insertion. Unlike DNA transposons, SINEs remain fixed in the genome once they have inserted, and therefore provide excellent evolutionary and phylogenetic markers (7). Retroposons represent an extremely successful class of mobile genetic elements as judged by their high copy number; however, most of the copies of SINEs are transcriptionally silent (4). Nevertheless, the abundance and widespread presence of SINEs in eukaryotic genomes have had a major impact on genome structure and function (1–4).

Here, we characterize the first SINE found in zebrafish (*Danio rerio*), which we named DANA (*Dan*io retroposon *A*). DANA has a complex and unique structure with characteristics of a heterologous composite element: a tRNA-derived region is followed by multiple, unrelated sequence blocks. As DANA transposition has not been observed in the laboratory, its mode of transposition (and hence its status as a bona fide retroposon) is indicated only by its sequence and structural similarity to known retroposons. Our findings indicate that DANA-related sequences have been contributing to the shaping of the zebrafish genome. DANA is specifically found in the genus *Danio*, demonstrating the usefulness of SINEs for determining phylogenetic positions within the teleost family Cyprinidae, the subfamily classification of which is unresolved (8). Moreover, DANA can be used to characterize different zebrafish populations, to identify mutations, and as a tool for mapping loci in the zebrafish genome.¶

MATERIALS AND METHODS

Source of Fish DNAs. All fish were purchased in a local pet store, with the following exceptions: the C-32 homozygous diploid zebrafish line was obtained from the University of Oregon, and zebrafish from Singapore, Hong Kong, and Indonesia were obtained from the Cardiovascular Research Center of Massachusetts General Hospital. The "leopard"/C-32 zebrafish pigmentation mutant is from our laboratory. Carp sperm was from the Fish Research Station (Szarvas, Hungary).

DNA Preparation, Blotting, and Hybridization. Preparation of high molecular weight DNA and blottings were done essentially as described (9). An α-^{32}P-labeled, ≈140-bp fragment of DANA served as probe in DNA hybridizations (Fig. 1). Low-stringency hybridizations were done in 30% (vol/vol) formamide, and the filter was washed in 0.5× standard saline citrate (SSC)/0.1% (wt/vol) SDS at 42°C for 7 hr.

PCRs and Sequence Analysis. PCRs were done as described (10), with the following oligonucleotide primers: Dana-1,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Pol III, RNA polymerase III; Myr, million years.

†Permanent address: Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary.
§To whom reprint requests should be addressed at: Department of Genetics and Cell Biology, University of Minnesota, 1445 Gortner Avenue, St. Paul, MN 55108-1095.

¶Two representative DANA sequences have been deposited in the GenBank database (accession nos. L42294 and L42295).

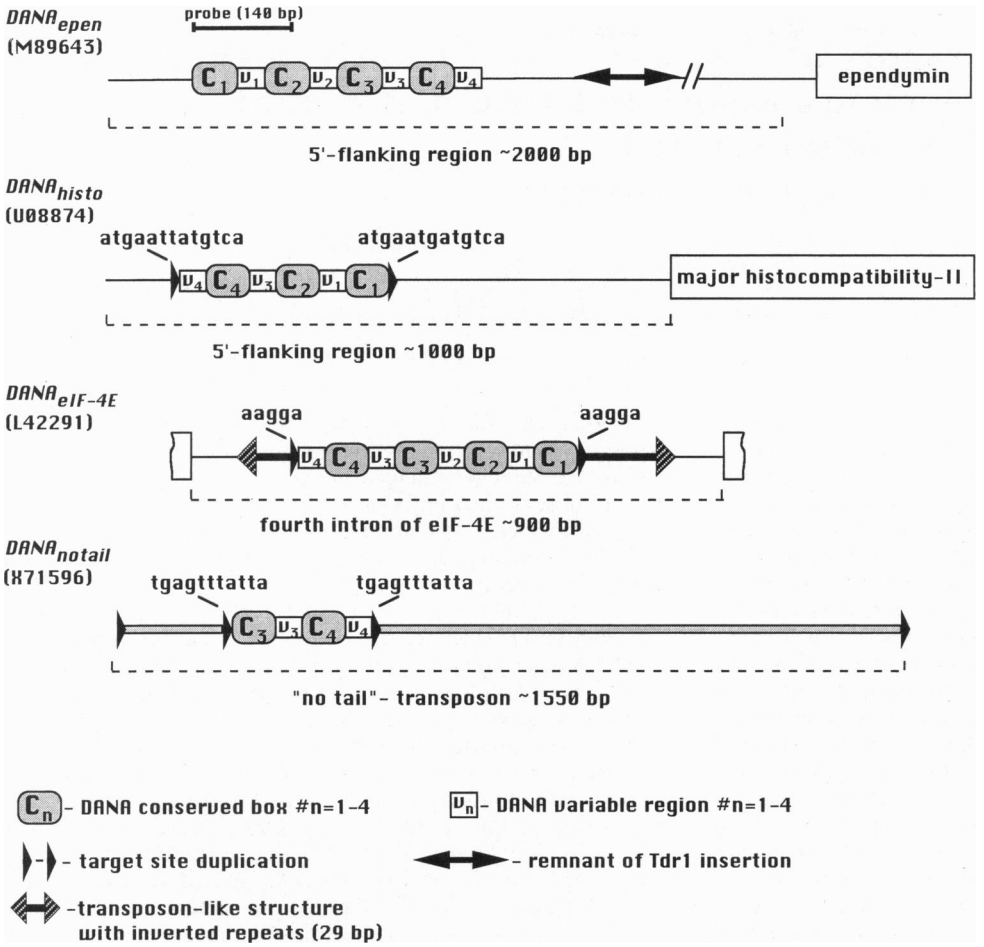

FIG. 1. DANA is found in close proximity to various zebrafish genes. The conservation, positions, and orientations of DANA sequences relative to structural genes are represented schematically. Shaded boxes in the DANA sequences indicate the four conserved regions (C_{1-4}); open boxes of smaller size represent the variable regions of DANA (v_{1-4}). Sequences of bracketing duplications (imperfect in $DANA_{histo}$) are shown above the DANA elements. The fragment used as a hybridization probe is shown above the $DANA_{epen}$ locus. Transcription of genes is from left to right.

5′-GGCGACRCAGTGGCGCAGTRGG-3′; Dana-2, 5′-TTTTCTTTTTGGCTTAGTCCC-3′; TdrA, 5′-TCCATCAGACCACAGGACAT-3′; and FTC-12, 5′-TGTCAGGAGGAATGGGCCAAAATTC-3′. Inter-PCR was performed in a total volume of 50 μl, with a 6-min initial denaturation step at 94°C followed by 20 cycles of 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C. To each reaction mixture, 10 μCi (370 kBq) of [α-^{32}P]dCTP (6000 Ci/mmol; Amersham) was added. The amplification products were passed through a Sephadex G-50 column (Pharmacia), and 1/10 of the product mixture was separated in a 5% polyacrylamide gel in 1× TBE buffer (90 mM Tris/90 mM boric acid/2 mM EDTA). DNA sequence analysis was carried out with the GCG sequence analysis package (University of Wisconsin Genetics Computer Group). Two representative DANA sequences have been deposited in GenBank (L42294 and L42295). Computer searches were performed on GenBank release 88.0 and EMBL release 42.0.

RESULTS

DANA Is a Repetitive Retroelement Associated with Various Zebrafish Genes. We recently described a Tc*1*-like DNA-transposon, Tdr*1*, the fragments of which can be found in the neighborhood of several fish genes, including the ependymin genes in Atlantic salmon and zebrafish (10). Upstream of the Tdr*1* remnants in the 5′ flanking region of the zebrafish ependymin gene we discovered an additional, transposon-like DNA sequence. Sequences in the fourth intron of the zebrafish eIF-4E gene were also found to be similar to the 5′ region of the ependymin gene, but unrelated to Tdr*1*. Based on the presence of the same sequence in two separate genes, we searched the databanks and identified two additional copies of this sequence. One copy is in the 5′ region of the histocompatibility protein type II gene and the other is a part of a transposon-like insertion element that caused a spontaneous mutation of the *no tail* gene in zebrafish (11). These four loci represent five percent of the total number of zebrafish genes deposited in the databanks, suggesting that the sequences are members of a repetitive family. Most of these elements are framed by short (5- to 13-bp) direct repeats (Fig. 1), indicative of duplication of sequences at the insertion site. A nucleic acid alignment of these four genomic copies allowed us to design PCR primers encompassing the putative 5′ and 3′ sequences of the element (Fig. 2), which were used to amplify other members of this family in zebrafish.

DANA Shows All the Characteristic Features of a tRNA-Derived Retroposon and Has a Unique, Complex Structure. Fifteen additional copies of DANA were amplified and cloned from zebrafish genomic DNA and from a cDNA library. Together with the databank entries, 19 copies were analyzed and used to generate the consensus DANA sequence shown in Fig. 2. Seventeen of the 19 elements are ≈350 bp long and show a complex structure consisting of four conserved regions (C_{1-4}) alternating with four variable regions (v_{1-4}), as shown in Figs. 1 and 2. The nucleotide sequence of C_1 shows approximately 65–67% similarity to a variety of tRNAs, with the best matches being to $tRNA^{Met}$ and $tRNA^{Ala}$. In addition, C_1 exhibits the consensus A and B boxes of a Pol III promoter (Fig. 2). The tRNA-related region and the conserved promoter elements indicate that DANA belongs to the family of retroposons. The origins of C_2 and C_3 are not known; there are no significantly similar sequences in the databases. C_4 has a size similar to that of C_1 and contains Pol III promoter-like sequences suggesting that it is derived from a gene transcribed

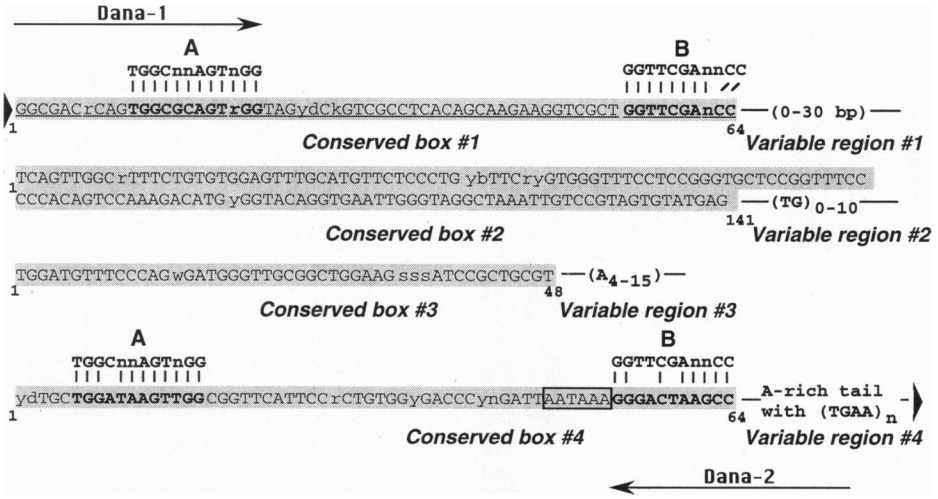

FIG. 2. DANA is a tRNA-derived SINE with a unique, complex structure. The consensus DANA sequence is shown with the conserved and variable boxes indicated below the nucleotide sequence. Because the DANA sequence contains variable regions, each conserved region is numbered from position 1. The tRNA-related sequence is underlined, with the A and B boxes of the Poly III promoter and their conservation shown. PCR primers are indicated by arrows above and below the sequence. The putative poly(A) signal is framed. Target-site duplications are marked by filled arrowheads.

by Pol III. However, these promoter sequences are more divergent from the consensus (Fig. 2), and C_4 does not show any obvious relationship to any tRNA or to C_1. Unlike C_1, C_4 contains a consensus poly(A) signal near its 3′ end that is followed by an A-rich, polypurine sequence. Presence of Pol III promoter elements and an A-rich sequence at the 3′ end suggest that C_4 might have been an independent SINE that has fused with the 3′ end of another Pol III transcriptional unit.

The sequences between the conserved boxes are hypervariable, primarily in length of simple mono-, di-, or tetranucleotide repeats; we did not encounter any two DANA elements with the same sequence in these positions. v_1 is juxtaposed to the B box of the promoter in C_1 (Fig. 2); it does not bear resemblance to any recognizable RNA or DNA motif. v_2 generally consists of a moderately long tract of a microsatellite-like dinucleotide repeat, $(TG)_n$, occasionally punctuated by other bases, mostly A. This particular type of dinucleotide repeat is commonly referred to as a CA repeat and constitutes a major fraction of repetitive DNA in some eukaryotic genomes. However, we have found exceptions; for instance, v_2 is a tandem repeat of a 15-bp sequence in $DANA_{epen}$. v_3 is an internal poly(A) tract, usually found at the 3′ ends of SINEs, whereas v_4 is always an A-rich polypurine tract followed by a repeat of a short DNA motif, frequently $(TGAA)_n$. Interestingly, the TGAA motif can also be found in some v_2 sequences, thus raising the possibility that these short, repeated sequences were generated by the insertion of a C_3-v_3-C_4-v_4 sequence into the 3′ end of another transpositional unit.

Most SINEs have a simple structure, consisting of a tRNA-related region followed by a tRNA-unrelated sequence and an A-rich tail (2). Composite elements usually do not contain more than two recognizable parts. Thus, DANA is an atypical SINE because it is often composed of four distinct cassettes.

DANA Can Form New, Mobile Transposons by Fusion with Other Transposon-Like Structures. The association of DANA with other repetitive sequences is exemplified by $DANA_{eIF\text{-}4E}$, which was inserted into a DNA sequence with 29-bp terminal inverted repeats (Fig. 1). This motif exists at several other genomic locations, suggesting repetitive characteristics (unpublished results). Incomplete DANA elements were also found, and some of the deletion breakpoints coincide with the boundaries of the conserved or the variable regions described above. For instance, the C_3 box is missing from $DANA_{histo}$ (Fig. 1). Another example is the $DANA_{ntl}$ sequence, which is severely truncated at the 5′ region and consists of only C_3-v_3-C_4-v_4 (Fig. 1). The presence of 11-bp direct repeats flanking the boundaries of the incomplete $DANA_{ntl}$ sequence suggests that the truncated element was generated by retroposition. Apparently, $DANA_{ntl}$ transposed into another sequence, also framed by direct repeats (11). Although the recipient sequence has not yet been classified, the entire 1550-bp sequence (Fig. 1) is a member of a family of middle repetitive elements in the zebrafish genome (11), which inserted into the *no tail* gene to inactivate the coding region. The *no tail* insertion is one of only three characterized mutations recovered from zebrafish (11). These findings underscore the flexible nature of SINE elements in transposition and their capacity to contribute to genetic variability.

DANA Is a *Danio*-Specific SINE. The presence of a given SINE is usually restricted to relatively few related species, but in a permissive genetic environment transposition is efficient. As a result, genomes accumulate large numbers of specific elements during evolution. The copy number of DANA in the zebrafish genome was determined by probing, with the fragment shown in Fig. 1, a blot of serial dilutions of genomic DNA and a plasmid containing a cloned DANA element (data not shown). We estimate that there are about $4–5 \times 10^5$ copies of the DANA element per haploid genome. Assuming that the size of the zebrafish haploid genome is 1.6×10^9 bp (10) and that the majority of DANA elements are about 350 bp, these mobile elements comprise ≈10% of the zebrafish genome.

In order to determine the distribution of DANA among related fish species from the order Cypriniformes, the PCR primers shown in Fig. 2 were used to amplify potential DANA-related elements. Probing PCR amplification products with the DANA-specific fragment revealed positive signals in DNA samples from zebrafish as well as in pearl danio (*Danio albolineatus*) and giant danio (*Danio aequipinnatus*), the closest relatives of zebrafish (12) (Fig. 3). The major elements were full-length units of about 350 bp, although products of slightly smaller and larger sizes were also found, but not characterized. No DANA-hybridizing products were detected from fish outside the genus *Danio*. To corroborate these findings, total genomic DNAs from the same fish species were subjected to Southern hybridization using the DANA probe. Again, significant hybridization signals were obtained only with DNA samples from the three *Danio* species (data not shown), confirming that this repetitive element was amplified specifically in the lineage leading to the genus *Danio*.

On average, the 19 DANA copies we analyzed have an 86% similarity, ignoring large deletions and excluding the variable regions. An average sequence divergence of 14% suggests that the major amplification of DANA probably occurred approximately 25–30 million years (Myr) ago, assuming that the substitution rate of 0.5%/Myr used for pseudogenes (13) is valid for retroposons (1).

DNA Polymorphisms Between Different Populations and Strains of Zebrafish Detected by a DANA-Specific Primer in PCR. The human *Alu* element has been used to detect polymorphic DNA fragments (14) by a method known as

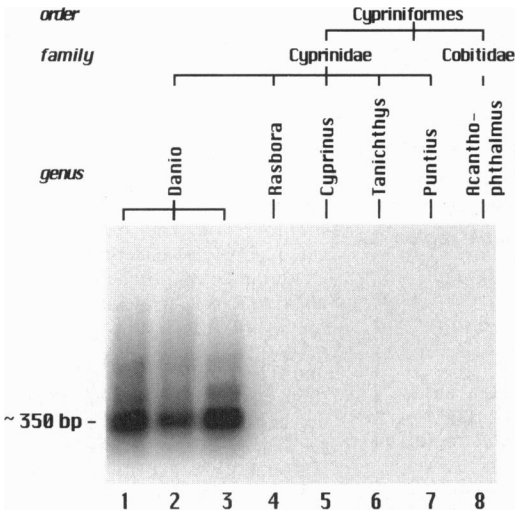

FIG. 3. DANA is *Danio*-specific. Southern blot of PCR products obtained with primers Dana-1 and Dana-2 (Fig. 2) was hybridized with a DANA probe (Fig. 1). Genomic DNA samples from the following species were analyzed: zebrafish (*Danio rerio*) (lane 1), pearl danio (*Danio albolineatus*) (lane 2), giant danio (*Danio aequipinnatus*) (lane 3), rasbora (*Rasbora heteromorpha*) (lane 4), common carp (*Cyprinus carpio*) (lane 5), white cloud mountainfish (*Tanichthys albonubes*) (lane 6), tiger barb (*Puntius tetrazona*) (lane 7), and striped loach (*Acanthophthalmus kuhli*) (lane 8). The phylogenetic classification of these fish species is given at the top.

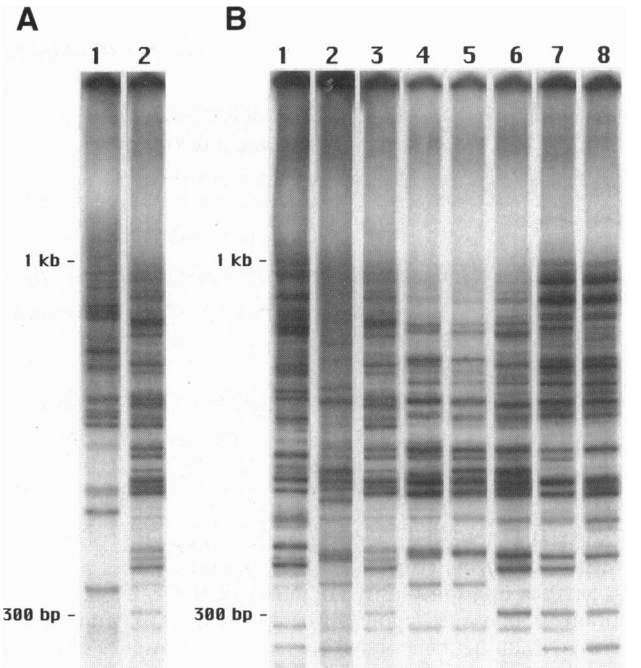

FIG. 4. DNA polymorphisms generated by DANA-specific primers in PCR. ^{32}P-labeled PCR amplification products were separated in polyacrylamide gels. (*A*) Effect of the specificity and combination of PCR primers on the number of DNA fragments amplified from Singapore zebrafish DNA. Lanes: 1, Dana-1 primer only; 2, Dana-1 supplemented with two Tdr*1*-specific primers, TdrA and FTC-12. (*B*) Polymorphic patterns generated by DANA-PCR using the same primer combination as in lane 2 of *A*, on DNA samples from different zebrafish populations. Lanes: 1, Hong Kong; 2, Indonesia; 3, Singapore; 4, wild-type 1 (United States); 5, wild-type 2 (United States); 6, "gold" zebrafish; 7, clonal line C-32; 8, "leopard"/C-32.

SINE-PCR. Essentially, single-copy genomic DNA flanked by *Alu* repeats is amplified by PCR using *Alu* primers to produce polymorphic fragments that are inherited in a Medelian fashion. We exploited the DANA sequences in a similar way in zebrafish, employing a single PCR primer complementary to one of the conserved regions of DANA (Dana-1 in Fig. 2) to amplify sequences between various pairs of DANA elements. Lane 1 in Fig. 4*A* shows an autoradiogram of radiolabeled PCR products from Singapore zebrafish separated in a polyacrylamide gel. Many bands between approximately 300 and 1000 bp are visible in the gel, suggesting that several of the DANA elements are sufficiently close to one another for efficient amplification. We observed DNA polymorphisms between zebrafish populations using a single DANA-specific PCR primer (data not shown). However, by adding primers specific for Tdr*1*, a DNA transposon represented by 1000 interspersed copies in the zebrafish genome (10), we were able to improve resolution of more marker bands because the distance between the heterologous primers was decreased (Fig. 4*A*, lane 2). The utility of the procedure is demonstrated in Fig. 4*B* by the distinctive patterns of PCR products obtained from genomic DNA samples of zebrafish populations from different geographical sources and from different laboratory strains. We detected a number of polymorphic bands between zebrafish populations from Hong Kong, Indonesia, Singapore, and the United States, demonstrating genomic rearrangements presumably generated by microsatellite variation and/or transposon insertion. Fewer polymorphic bands were detected between laboratory strains of zebrafish (Fig. 4*B*, lanes 5–7), and fewer yet between individuals from the same laboratory populations (lanes 4 and 5 and lanes 7 and 8), suggesting that these repeated elements are suitable for the characterization of different zebrafish populations as well as for gene mapping.

DISCUSSION

The zebrafish DANA element is an exceptional SINE in terms of its distinct substructure, consisting of four cassettes of alternating conserved and hypervariable sequences. Retroposons which have transposed into one another either become "dead-ends" in evolution (1, 15) or, rarely, they may be amplified further as a new unit, as long as the new sequence does not contain a Pol III terminator (3, 4).

Two of the conserved boxes of DANA carry Pol III promoter motifs. In composite elements, the promoter most closely matching a consensus Pol III sequence is located in the upstream part of the SINE, thereby occluding weaker promoters in downstream sequences. The promoter elements in the 5′ conserved box (C_1) are the better conserved and, thus, presumably provide the transcriptional activity for DANA. The 3′ tail consists of an A-rich region followed by short, repeated DNA motifs, as has been found in some families of artiodactyl retroposons (16). The internal variable region v_2 of DANA is composed of a microsatellite-like dinucleotide repeat. Similar repeats were also reported from certain *Alu* subfamilies (17); from seal, mink and dog SINEs (18); and from the artiodactyl family (15). Association of SINEs with microsatellite-like sequences can be explained by a posttranspositional mechanism similar to the expansion of simple satellite sequences (1).

We view DANA as a composite retroposon that was probably formed by repeated insertions of three short sequences into an ancient tRNA-derived SINE, which then was amplified as a transpositional unit. However, components of the DANA element apparently are capable of transposing by themselves. $DANA_{ntl}$, a DANA-related sequence severely shortened at its 5′ end, fused with another transposon-like structure, presumably as a result of retroposition, to produce a new family of active transposable elements (11). Two possibilities for the generation of the shorter mobile sequences are most likely. First, there might be multiple DANA "master" genes in the zebrafish genome encoding short transcripts which are the ancestors of the different Cv blocks of the DANA element.

Second, there may be only full-length DANA genes, but the transcripts are either processed at their 5′ ends or reverse transcription is often incomplete (1). In contrast, there is no simple mechanism that would explain the precise and specific deletion of C_1-v_1-C_2-v_2 from a complete DANA element to form $DANA_{ntl}$. Resolution of DANA evolution will require further characterization of more elements to establish the "rules of assembly."

SINEs have been isolated from salmonid fish species (7, 19). These salmonid retroposons do not possess the complex structure described for DANA and do not show any significant sequence similarity with DANA. Certain salmonid SINEs, estimated to be <5 Myr old, have been implicated to play a role in salmonid speciation (19). Based on sequence divergence between individual copies, DANA appears to be at least 25 Myr old. Fossil records suggest that the genera *Puntius* and *Rasbora*, close relatives of the genus *Danio*, can both be traced back to the Miocene, also ≈25 Myr ago (8). This coincidence and the restricted presence of DANA only in species belonging to the genus *Danio* are consistent with the speculation that the massive burst(s) of DANA transposition may have been either the cause or the consequence of the speciation process that eventually led to the various *Danio* species. Our results underscore the power of SINEs to establish phylogenetic relationships for such a widely used model organism in developmental biology as the zebrafish. Resolution of such relationships is important if developmental functions are to be addressed in an evolutionary context (12).

DANA is affiliated with the *no tail* insertional mutation, which has a severe effect on development. Besides their potential impact on genome organization, gene structure, and function, SINEs and other repetitive elements—e.g., *copia*-like elements in plants and *gypsy* elements in *Drosophila*—influence gene regulation as well (4, 20–22). The zebrafish ependymin locus appears to be a "transposon trap"; it contains both a DANA retroelement and fragments of a Tdr*1* transposon (10). These two transposon-originated sequences occupy about 30% of the region responsible for the spatial and temporal regulation of ependymin gene expression (23).

Some retroposons are particularly suitable as genetic markers because they have high copy number and an apparently random distribution in the genome, remain fixed in the chromosomes after insertion, and segregate in Mendelian fashion. *Alu*-PCR has been developed to a versatile technique for human genome mapping, fingerprinting, and yeast artificial chromosome cloning (14, 24, 25). The number of polymorphic bands may be increased by the combination of various primers to repetitive sequences in the zebrafish genome such as Tdr*1* transposons (10) and satellite-like repeats (26). Polymorphic fragments can be isolated and become sequence-tagged reference points for physical genome mapping. Here we have shown that transposon-specific PCR primers are capable of detecting multiple polymorphic bands from zebrafish genomic DNAs of different sources. Association of DANA with microsatellite-like repeats of variable length permits the combination of SINE-PCR and microsatellite marker mapping (15). Thus, DANA can be used for screening polymorphisms between fish stocks, identification of natural and induced mutations, detection of transgenic sequences, and genome mapping (4, 27), using only one to three primers. This method will complement that of RAPD (rapid amplification of polymorphic DNA) mapping, in which several hundred PCR primer pairs are used to map zebrafish genes (28).

We thank Tom Sapp for his assistance. Zebrafish from Hong Kong, Singapore, and Indonesia were kindly provided by Drs. D. Stainier and W. Driever. The zebrafish cDNA library was a gift from Drs. Jay White and Martin Petkovich. This work was supported by U.S. Department of Agriculture Grant 92-37205-7842 and National Institutes of Health Grant RO1-RR06625. S.C.F. was supported by National Oceanographic and Atmospheric Administration SeaGrant NA46-RG0101.

1. Weiner, A. M., Deininger, P. L. & Efstratiadis, A. (1986) *Annu. Rev. Biochem.* **55,** 631–661.
2. Deininger, P. L. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 619–636.
3. Rogers, J. (1985) *Int. Rev. Cytol.* **93,** 187–279.
4. Deininger, P. L. & Batzer, M. A. (1993) *Evol. Biol.* **27,** 157–196.
5. Daniels, G. R. & Deininger, P. L. (1985) *Nature (London)* **317,** 819–822.
6. Lee, M., Loomis, G. & Cowan, N. (1984) *Nucleic Acids Res.* **12,** 5823–5836.
7. Murata, S., Takasaki, N., Saitoh, M. & Okada, N. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6995–6999.
8. Winfield, I. J. & Nelson, J. S., eds. (1991) *Cyprinid Fishes* (Chapman & Hall, London), pp. 34–49.
9. Ivics, Z., Izsvák, Zs. & Hackett, P. B. (1993) *Mol. Marine Biol. Biotechnol.* **2,** 162–173.
10. Izsvák, Zs., Ivics, Z. & Hackett, P. B. (1995) *Mol. Gen. Genet.* **247,** 312–322.
11. Schulte-Merker, S., van Eeden, F. J. M., Halpern, M. E., Kimmel, C. B. & Nusslein-Volhard, C. (1994) *Development (Cambridge, U.K.)* **120,** 1009–1015.
12. Meyer, A., Biermann, C. H. & Orti, G. (1993) *Proc. R. Soc. London B* **252,** 231–236.
13. Li, W. H., Luo, C. & Wu, C. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 1–94.
14. Sinnett, D., Deragon, J., Simard, L. R. & Labuda, D. (1990) *Genomics* **7,** 331–334.
15. Kaukinen, J. & Varvio, S. (1992) *Nucleic Acids Res.* **20,** 2955–2958.
16. Watanabe, Y., Tsukada, T., Notake, M., Nakanishi, S. & Numa, S. (1982) *Nucleic Acids Res.* **10,** 1459–1492.
17. Saffer, J. D. & Lerman, M. I. (1983) *Mol. Cell. Biol.* **3,** 960–964.
18. Coltman, D. W. & Wright, J. M. (1994) *Nucleic Acids Res.* **22,** 2726–2730.
19. Kido, Y., Aono, M., Yamaki, T., Matsumoto, K., Murata, S., Saneyoshi, M. & Okada, N. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2326–2330.
20. Brosius, J. (1991) *Science* **251,** 753.
21. White, S. E., Habre, L. F. & Wessler, S. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11792–11796.
22. Finnegan, D. J. (1989) *Trends Genet.* **5,** 103–107.
23. Rinder, H., Bayer, T. A., Gertzen, E. & Hoffmann, W. (1992) *DNA Cell Biol.* **11,** 425–432.
24. Sidhu, M. S., Helen, B. K. & Athwal, R. S. (1992) *Genomics* **14,** 728–732.
25. Soh, J., Mariano, T. M., Bradshaw, G., Donnelly, R. J. & Pestka, S. (1994) *DNA Cell Biol.* **13,** 301–309.
26. He, L., Zhu, Z., Faras, A. J., Guise, K. S., Hackett, P. B. & Kapuscinski, A. R. (1992) *Mol. Marine Biol. Biotechnol.* **1,** 125–135.
27. Zietkiewicz, E., Labuda, M., Sinnett, D., Glorieux, F. H. & Labuda, D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8448–8451.
28. Johnson, S. L., Clare, N. M., Ballinger, E. W. & Postlethwait, J. H. (1993) *Genomics* **19,** 152–156.