Characterization of a Tc1-like transposable element in zebrafish (Danio rerio)

Received: 7 September 1994 / Accepted: 13 December 1994

Abstract We have characterized Tdrl, a family of Tc1-like transposable elements found in the genome of zebrafish (Danio rerio). The copy number and distribution of the sequence in the zebrafish genome have been determined, and by these criteria Tdrl can be classified as a moderately repetitive, interspersed element. Examination of the sequences and structures of several copies of Tdrl revealed that a particular deletion derivative, 1,250 bp long, of the transposon has been amplified to become the dominant form of Tdrl. The deletion in these elements encompasses sequences encoding the N-terminal portion of the putative Tdrl transposase. Sequences corresponding to the deleted region were also detected, and thus allowed prediction of the nucleotide sequence of a hypothetical full-length element. Well conserved segments of Tc1-like transposons were found in the flanking regions of known fish genes, suggesting that these elements have a long evolutionary history in piscine genomes. Tdrl elements have long, 208 bp inverted repeats, with a short DNA motif repeated four times at the termini of the inverted repeats. Although different from that of the prototype C. elegans transposon Tc1, this inverted repeat structure is shared by transposable elements from salmonid fish species and two Drosophila species. We propose that these transposons form a subgroup within the Tc1-like family. Comparison of Tc1-like transposons supports the hypothesis that the transposase genes and their flanking sequences have been shaped by independent evolutionary constraints. Although Tc1-like sequences are present in the genomes of several strains of zebrafish and in salmonid fishes, these sequences are not conserved in the genus Danio, thus raising the possibility that these elements can be exploited for gene tagging and genome mapping.

Key words Transposable element • Evolution
Zebrafish • Transposase • Inverted repeats

Introduction

The zebrafish (Danio rerio) is a superb model animal for studies of the molecular genetics of vertebrate development (Driever et al. 1994; Nüsslein-Volhard 1994; Rossant and Hopkins 1992). However, two methods of investigating the roles that genes play in development are lacking in zebrafish: a method of randomly tagging genes and of establishing transgenic fish lines that continue to express transgenic DNA over multiple generations. Hence, there is no method currently available for efficient integration of intact copies of foreign DNA into the zebrafish genome. Development of such a technology would not only be beneficial for the production of transgenic fish, but would also be useful for DNA-based insertional mutagenesis and gene tagging, which could be applied to generate and identify developmental mutants. A potential solution to the above problems would be to harness transposable elements as vectors for carrying specific DNA sequences into chromosomes for the purpose of mutagenesis, gene tagging and genetic engineering.

Since their discovery, repetitive elements with the characteristics of transposons have been found in the genomes of all organisms in which they were sought. Transposable elements are usually divided into two major groups according to their mechanism of transposition (Finnegan 1989). The first group is composed of DNA-based transposable elements that transpose via a DNA intermediate. The second group, the retroelements, move
by reverse transcription of an RNA intermediate. DNA transposons are common in bacteria, invertebrates and plants, whereas in vertebrates retrotransposons are more abundant (Henikoff 1992). In eukaryotes, DNA-based transposable elements typically range from 1 to 10 kb in length, have short inverted terminal repeats, and have sequences that may encode a transposase required for the transposition process. Elements of this sort include the P elements of Drosophila melanogaster, the Ac and Spm elements of maize and the Tcl transposon of Caenorhabditis elegans. Transposons exist in active and inactive forms (Hartl et al. 1992). Active transposable elements encode genes required in trans, as well as flanking sequences required in cis for mobility. Active elements are relatively rare, since they destabilize the genomes of their hosts by insertion mutagenesis. More common are inactive elements which often contain mutations in the sequence encoding the trans-acting transposase, thereby rendering these elements unable to transpose on their own. However, these inactive elements may still retain cis-active sequences needed for recognition by the transposase, and therefore can be mobilized by transposase expressed from an active element. Here, we use the term transposable element for sequences that have the hallmarks of either active or inactive transposons.

There is a growing number of repetitive elements known from fish species (Datta et al. 1988; Kido et al. 1991; Naruse et al. 1992; Winkfein et al. 1988; Wright 1989). In zebrafish, three types of repetitive elements have been reported: satellite-like sequences (Ekker et al. 1992); Alu-elements (He et al. 1992); and recently, a DNA transposon, Tdrl (Radice et al. 1994). Tdrl belongs to the family of Tcl-like transposable genetic elements, named after the prototype found in C. elegans (Emmons et al. 1983). Members of this transposon family share certain features including a length of about 1600 bp, conserved terminal nucleotides (TACAGT) of the inverted repeats and conserved sequence domains in the encoded transposase. Tcl-like sequences, as defined by the homology of their transposable genes, have been found in fungi (Daboussi et al. 1992), protozoa (Doak et al. 1994), and are widespread in animals including nematodes (Collins et al. 1989; Emmons et al. 1983; Prasad et al. 1991), insects (Brezinsky et al. 1990; Caizzi et al. 1993; Franz and Savakis 1991; Harris et al. 1988; Simmonds and Merriman 1994), hagfish (Heierhorst et al. 1989), and teleost fishes such as zebrafish, channel catfish (Ictalurus punctatus; Wilson et al. 1990), rainbow trout (Oncorhynchus mykiss; Radice et al. 1994) and Atlantic salmon (Salmo salar; Goodier and Davidson 1994; Radice et al. 1994).

Transposable elements can be powerful tools for gene tagging and genetic mapping, especially if strains of the same organism contain small numbers of these sequences. Moreover, the transposase of active elements should serve as an excellent source for recombinase-mediated insertion of transgenic DNA (Ivics et al. 1993). Accordingly, we examined the Tdrl-like elements in various strains of zebrafish in order to find an active copy, and to determine whether a line of zebrafish could be found that had a small number of elements, and therefore would be suitable for transposon tagging and mutagenesis studies.

Here we present our characterization of several copies of Tdrl, the first DNA transposon characterized in zebrafish. All Tdrl elements we have analyzed exhibit variability in size and sequence, suggesting that Tdrl is not a recent invader of the zebrafish genome. The majority of Tdrl elements appear to be descendants of a single progenitor transposon with a deletion that removed the N-terminal region of the putative transposase. However, because sequences complementary to this missing portion can also be found in the genome, and a homologous element exists in salmon, the sequence of a hypothetical transposase gene can be predicted. A byproduct of this investigation is a preliminary phylogenetic analysis of sequences and structures of several Tcl-like transposons, which suggests that the Tdrl element may be a hybrid of two types of transposons.

**Materials and methods**

Source of fish DNAs

Wild-type zebrafish, "blue" danios, "giant" danios, "gold" danios and "leopard" danios were purchased in a local pet store. The C-32 homozygous diploid zebrafish line was obtained from the University of Oregon. Zebrafish from Singapore, Hong Kong and Indonesia were obtained from the Massachusetts General Hospital, Cardiovascular Research Center. The zebrafish pigmentation mutant is from our laboratory. Carp sperm was obtained from the Fish Research Station, Szarvas, Hungary. Chinook salmon DNA was prepared from the CHSE cell line, and northern pike DNA was extracted from the PG cell line.

DNA preparation, blotting and hybridization

High molecular weight DNA was prepared from adult zebrafish, essentially as described (Ivics et al. 1993). Tissues were homogenized in 200 mM TRIS-HCl pH 8.0, 50 mM EDTA, 150 mM NaCl, 0.5% (w/v) SDS and 500 µg/ml Proteinase K overnight at 55°C. Samples were then extracted once with phenol-chloroform, once with chloroform and precipitated with 2 vol ethanol. DNA samples were prepared for dot-blotting by the addition of 1 vol 0.8 M NaOH, incubated at room temperature for 10 min, then 2 vol of ice-cold 2 M ammonium acetate was added, and the DNA samples were blotted to a Nytran (Schleicher & Schuell) nylon membrane with the help of a dot blot apparatus (BRL). Blots were then baked at 80°C for 2 h and irradiated with ultraviolet light for 30 s. Southern transfer was done according to standard procedures (Sambrook et al. 1989). An [α-32P]-labeled, approximately 700 bp internal fragment of Tdrl served as probe in DNA hybridizations (Fig. 5A). This fragment was subcloned from the Tdrl transposon isolated by Radice et al. (1994). High stringency hybridizations were done according to standard methods (Sambrook et al. 1989) in 50% (v/v) formamide, while low stringency hybridizations were in 30% (v/v) formamide. Stringent washings of the filters were done in 0.1×SSC, 0.1% (w/v) SDS at 65°C, whereas low-stringency washings were in 0.5×SSC, 0.1% (w/v) SDS at 42°C for 7 h. Autoradiograms were scanned with an AMBIS densitometry system.