

TRT, a Vertebrate and Protozoan Tc1-Like Transposon: Current Activity and Horizontal Transfer

Hua-Hao Zhang1,2,†, Guo-Yin Li3,†, Xiao-Min Xiong4,†, Min-Jin Han1, Xiao-Gu Zhang2, and Fang-Yin Dai1,*

1State Key Laboratory of Silkworm Genome Biology, Key Laboratory for Sericulture Functional Genomics and Biotechnology of Agricultural Ministry, Southwest University, Chongqing, China
2College of Pharmacy and Life Science, Jiujiang University, Jiujiang, China
3Department of pathology, Hanzhoung Hospital, Hanzhoung city, Shaanxi, China
4Clinical Medical College, Jiujiang University, Jiujiang, China

*Corresponding author: E-mail: fydai@swu.edu.cn.
†These authors contributed equally.

Accepted: August 12, 2016

Abstract

We report a Danio rerio transposon named DrTRT, for D. rerio Transposon Related to Tc1. The complete sequence of the DrTRT transposon is 1,563 base pairs (bp) in length, and its transposase putatively encodes a 338-amino acid protein that harbors a DD37E motif in its catalytic domain. We present evidence based on searches of publicly available genomes that TRT elements commonly occur in vertebrates and protozoa. Phylogenetic and functional domain comparisons confirm that TRT constitutes a new subfamily within the Tc1 family. Hallmark features of having no premature termination codons within the transposase, the presence of all expected functional domains, and its occurrence in the bony fish transcriptome suggest that TRT might have current or recent activity in these species. Further analysis showed that the activity of TRT elements in these species might have arisen about between 4 and 19 Ma. Interestingly, our results also implied that the widespread distribution of TRT among fishes, frogs, and snakes is the result of multiple independent HT events, probably from bony fishes to snakes or frog. Finally, the mechanisms underlying horizontal transfer of TRT elements are discussed.

Key words: Tc1/mariner transposons, DD37E, horizontal transfer.

Introduction

Transposable elements (TEs) are DNA fragments that move from one host genomic site to another. TEs have been reported in almost all organisms, including plants, invertebrates, vertebrates, fungi, and bacteria, and they occupy a substantial fraction of various host genomes (Biémont 2010). TEs that integrate into a new genomic location play important roles in genome architecture as well as in genetic innovation (Feschotte and Pritham 2007).

TEs are generally classified into two classes (Classes I and II) based on their structural organization and mechanism of transposition (Feschotte and Pritham 2007). Class I or RNA elements are transposed via reverse transcription of an RNA intermediate, whereas that Class II or DNA elements transpose directly as DNA, mostly through a so-called “cut and paste” mechanism. The Tc1/mariner superfamilly of TEs is the most widespread class of DNA transposons in nature (Feschotte and Pritham 2007). Tc1/mariner, which was first discovered in Drosophila mauritiana (Haymer and Marsh 1986; Jacobson et al. 1986), is ubiquitous in eukaryotes (Feschotte and Pritham 2007; Liu and Yang 2014). Tc1/mariner elements are generally 1,300–2,400 bp in size and encode a 340-amino acid transposase that is flanked by terminal inverted repeats (TIRs) and a target site duplication (TSD), TA (Lohe et al. 1996). On the basis of variations in the DDE/D signature motif, Tc1/mariner elements are further classified into various monophyletic groups (Shao and Tu 2001). Although some naturally active Tc1/mariner elements have been identified in nematodes and arthropods, most of these are inactive in vertebrates due to the occurrence of stop codons, deletions, or frameshifts. Importantly, active Tc1/mariner transposons may be potentially used as molecular tools for transgenesis and
insertion mutagenesis (Pavlopoulos et al. 2007; Voigt et al. 2016). To date, only some Tc1-like elements (such as To1, To2, Passport, and Tana1) in bony fishes are known to be both intact in their native form and transpositionally active (Koga et al. 2006; Clark et al. 2009; Pujolar et al. 2013; Watanabe et al. 2014).

TEs are ubiquitous in organisms despite extensive evidence that their presence and mobility causes a variety of deleterious mutations (Craig et al. 2002). One explanation is that TEs have the ability to invade a new host by horizontal transfer (HT). A previous study demonstrated that almost all kinds of eukaryotic TEs are capable of HT (Schaack et al. 2010). HT events occurring in vertebrates mainly involve the hAT superfamily of DNA transposons (Pace et al. 2008; Novick et al. 2010; Gilbert et al. 2010; Gilbert et al. 2012; Gilbert et al. 2013).

In the present study, we performed a detail analysis of a Tc1-like transposon in the zebrafish (Danio rerio). Further analysis confirmed that TRT elements constitute a new subfamily within the Tc1 family. We also show that TRT elements are both intact in their native form and functionally active in Pundamilian yererei, Maylandia zebra, and Haplochromis burtoni. Finally, we demonstrate that TRT elements might have invaded into its hosts via multiple HT events.

Materials and Methods
Identification and Copy Number Determination of TRT

TBLASTN analysis (Altschul et al. 1990) of the zebrafish genome was performed using the amino acid sequences of Bombyx mori Bmmar1 (U47917) as query, which in turn detected a Tc1/mariner transposon that encoded a DD37E motif-harboring polypeptide. We named these elements DrTRT (an abbreviation for Danio rerio Transposon Related to Tc1) because of its several similarities to transposons of the Tc1 family, as described in Results part. To determine the distribution of DrTRT, its transposase sequence was used as query to search against 1,176 species genomes, including 437 fungus, 216 invertebrates, 226 vertebrates, 150 plants, and 147 protozoans that are available at National Center for Biotechnology Information (NCBI). This transposon was determined to exist in one species when the unique DD37E motif was detected in the catalytic domain of one transposon.

The whole-genome shotgun sequences of all studied species were downloaded from NCBI. The consensus sequences of TRT were reconstructed using a multiple alignment of full-length copies in each genome using DAMBE (Xia and Xie 2001). Then, these respective consensus sequences were employed to mask each host genome to estimate copy number. All BLAST hits with more than 100 bp in size and 80% identity were used to calculate copy number.

Sequence and Phylogenetic Analyses

Inverted repeats were manually searched by the FastPCR software (Kalendar et al. 2014). The potential open reading frame of TRT used in the present study was analyzed using GENSCAN (http://genes.mit.edu/GENSCAN.html) or getorf in the EMBOSS-6.3.1 package (Rice et al. 2000), with default parameters. A multiple alignment of these elements was created by MUSCLE (Edgar 2004), with default parameters. The secondary structure of transposase was predicted by PSIPRED (McGuffin et al. 2000), with default parameters. Putative nuclear localization signal (NLS) motifs were predicted using PSORT II Prediction as provided in the PSORT WWW server (http://psort.nibb.ac.jp/). Each pairwise identity was calculated by Bioedit (Hall 1999) after all ambiguous and gapped sites were removed.

Sequences of Recombination-activating gene 1 (RAG1) were used in the comparison with transposon distance, with the purpose of testing HT hypothesis. Their accession numbers were listed in supplementary table S4, Supplementary Material Online. Multiple alignments of RAG1 and TRT were created using MUSCLE (Edgar 2004). Then, comparison distances of RAG1 and TRT were calculated using MEGA 4 (Tamura et al. 2007; pairwise deletion, maximum composite likelihood) based on two aligned files (supplementary dataset S1 and S2, Supplementary Material Online).

Transposase sequences of Tc1, mariner, and maT were downloaded from GenBank. The accession numbers are as follows: Anopheles albimanus Quetzal (L76231), Pleuronectes platessa PplTc1 (AJ303069), Haemonchus contortus HcTc1 (AF999908), Rana pipiens RpTc1 (BK001476), Drosophila virilis Paris (Z49253), Fusarium oxysporum Impala (AF282722), Ceratits capitata Ccmar1 (U40493), Homo sapiens Hsmar1 (U52077), H. sapiens Hsmar2 (U49974), Drosophila mauritiana Dmmar1 (X78906), Drosophila simulans Dsmar1 (X89927), Chrysoperla plumibunda Cpmar1 (U11650), Girardia tigrina Dtmar1 (X71979), Adineta vaga Avmar1 (DQ138246), Apis mellifera Ammar1 (AY155490), Drosophila mauritiana Mos1 (AEZ51500), Forficula auricularia Famar1 (AAP51098), Shigella sonnei 15630 (X05955), Bombyx mori Bmmar6 (U47917), B. mori Bmmar6 (AF461149), Caenorhabditis briggsae Cbmmar1 (AC084526), C. briggsae Cbmmar4 (AC084524), C. briggsae Cbmmar5 (AC084578), C. elegans CemaT1 (U41268), and Philodina rosea PrD37E (DQ138288). PrD37D (P. roseola), Tc1 (C. elegans), Tc1-1_Pm (Petromyzon marinus), Tc1-1_Ic (Latimeria chalumnae), Tc1-1_Dr (D. rerio), and Tc1-1_Ory (Oryzias latipes) were downloaded from Repbase (Jurka et al. 2005). A multiple alignment of full-length transposase sequences of the above transposons and all TRT was created by MUSCLE with
default parameters (Edgar 2004). Then, the appropriate amino acid (aa) substitution model was selected using the Akaike information criterion in the ProtTest 3 server (Darnilla et al. 2011). The best-suited aa substitution model for these data was the WAG model. Phylogenetic trees were then built using the MRBAYES 3.1.2 software (Ronquist and Huelsenbeck 2003) until the values of the average SD of split frequencies were stably below 0.01.

Results

Identification and Characteristics of DrTRT in the Zebrafish Genome

Using the deduced protein sequence of Bmmar1, we retrieved one sequence (location: CZQB01060324 6427 7413 +) that showed significant similarity (E values: 3e-13) in a TBLASTN search against the zebrafish genome in NCBI. In contrast to the DD37D signature encoded by Bmmar1, the spacing of the DDE motif within the catalytic domain of DrTRT was DD37E (fig. 1). Although DrTRT (named Mariner-14_DR in Repbase) has been deposited in Repbase (http://www.girinst.org/repbase/), knowledge and evolutionary history of DrTRT remain largely unknown. Therefore, DrTRT was utilized in subsequent analyses of Tc1/mariner elements in the present study. The nucleotide sequence of DrTRT was used as query for BLASTN analysis of the zebrafish genome. All obtained significant hits were extracted with 1,000-bp flanking sequences using our Perl script, and these were aligned to determine their boundaries. The consensus sequence of DrTRT was reconstructed, which was determined to be 1,563 bp in length (fig. 1). The transposase was flanked by TIRs that were 38 bp in length (fig. 1). DrTRT also contained a 23-bp 5’-subterminal inverted repeat (SIR) (TGTCATAATTAGGCAACT T) that pairs with a reverse complementary 3’-SIR (AAGTTCGCC TAAATATAGCA) (fig. 1). A previous study suggested that the subterminal region of DrTRT plays critical structural or functional roles during transposition (Tu, 2000). DrTRT was detected as about 44 copies (table 1). Further analyses showed that 50% (22/44) of DrTRT were verified full-length elements that showed >95% sequence identity to the consensus sequence (supplementary table S1, Supplementary Material Online).

The coding domain of the transposase was 1,014 bp (338 amino acids) in length. Several conserved motifs that were characteristics of Tc1-like transposons (Plasterk et al. 1999) were also observed in the DrTRT sequence (fig. 1). First, there were two helix-turn-helix (HTH) motifs at the N-terminal of the transposase, and each motif consisted of three α-helices. Second, a GRKK putative AT hook-like motif was present between the two HTH motifs. Third, two bipartite NLSs were identified at the N-terminal of the transposase. Finally, a catalytic triad DDE motif within their catalytic domain, with 37 amino acids between the second aspartic acid (D) and glutamic acid (E).

Distribution of TRT in Other Sequenced Species

To determine the species distribution of TRT, a TBLASTN search against 1,176 species (437 fungus, 216 invertebrates, 226 vertebrates, 150 plants, and 147 protozoans) whose complete or nearly complete genome sequences are available at NCBI was performed using the protein encoded by DrTRT as query. Significant hits encoding the conserved DD37E motifs were detected in bony fishes, clawed frog, snakes, protozoans, and fungi (table 1 and supplementary table S2, Supplementary Material Online). We should notice that TRT copies may reside within unassembled portions of a genome because many genomes available in public databases are not truly complete in the sense that they are highly fragmented and transposons are more likely to be interrupted by gaps in the sequence. Their copy number per genome extensively varied among species, changing from 1 copy in the fungi Cunninghamella bertholletiae to 435 copies in the Atlantic salmon Salmo salar (table 1), thereby suggesting that these transposons underwent species-specific proliferation in their host genomes. Interestingly, the transposases of several full-length copies of TRT that were identified in bony fishes and fungi did not harbor internal stop codons or frameshift mutations and presented all expected functional domains, as well as an intact TIR (fig. 2 and supplementary table S1, Supplementary Material Online), suggesting that TRT might have current or recent activity in these host genomes. For example, TRT identified in the bony fish P. yererei had 257 copies, and almost all copies (201/257) were present in full length (table 1 and supplementary table S1, Supplementary Material Online and fig. 2). Meanwhile, all complete copies revealed a strikingly high level of sequence identity (97–99% pairwise nucleotide identity) to the entire length of TRT elements, of which 84/201 encoded intact transposases (fig. 2).

To further investigate the activity of TRT, the nucleotide sequences of the transposases were used as a BLASTN query against the species Transcriptome Shotgun Assembly (TSA) database. Interestingly, when the putative complete transposase sequences were aligned against the full transcriptome database of the bony fishes M. zebra and H. burtoni, the references were completely covered (100% in length) (fig. 2). Our results also showed that the active TRT elements might be restricted to members of Haplochromini (P. yererei, M. zebra, and H. burtoni). The bony fishes P. yererei, M. zebra, and H. burtoni shared the last common ancestor about 4 Ma (Hedges et al. 2006) and they were diverged from the lyretail cichlid Neolamprologus brichardi (no active TRT was found in this species) about 19.6 Ma, thereby suggesting that the activity of TRT elements in these species might have begun at about between 4 and 19 Ma (fig. 2).
Despite the observation that transposase sequences have consensus sequences with similar lengths, those of TIRs significantly differed between metazoans and fungi (table 1). Generally, TRT elements in metazoans except for the clawed frog had a short TIR sequence (26–44 bp) at their termini. TRT elements in fungi was flanked by much longer TIRs (generally >100 bp) compared to those of metazoan TRT elements. TIRs played important roles (e.g., providing a cleavage signal sequence and binding site for transposases) in the transposition of Tc1/mariner transposons (Brillet et al. 2007). Tc1/mariner transposons could be divided into different groups based on variations in their TIRs, including length and motif content (Brillet et al. 2007). Therefore, it is important to analyze the functional domain of these length-variable TIRs during

![Diagram](image)

**FIG. 1.**—Nucleotide sequence and amino acid translation of DrTRT. Its TIRs and SIR are underlined in black and red, respectively. The stop codon is indicated by an asterisk. The six α-helices at the N-terminal of the transposase are indicated by blue letters. A GRKK putative AT-hook like motif is shown by a green letter. Two bipartite NLS at the N-terminal of the transposase are underlined in blue. The catalytic triad DD37E motif within the catalytic domain is indicated by red letters.
Table 1
Characteristics of TRT and Tc1 transposons identified in this study

| Species                                      | TRTs (bp) | Length (bp) | Tpase (aa) | Number of copies | Representatives                  |
|----------------------------------------------|-----------|-------------|------------|------------------|----------------------------------|
| **Vertebrates**                              |           |             |            |                  |                                  |
| Amphilophus citrinellus                      | 2994      | 1561        | 263        | 48               | emb:CCE01000287.1                 |
| Anguilla japonica                            | 2995      | 1444        | 217        | 6                | gb:AVP01016937.1                 |
| Anoploparma timbria                          | 2996      | 1045        | 27         | 27               | gb:AWG010162360.1                |
| Astyanax mexicanus                           | 2997      | 1549        | 202        | 27               | gb:APW010059627.1                |
| Boleophthalmus pectinirostris                | 2998      | 1592        | 265        | 19               | gb:JACK01022919.1                |
| Cynoglossus semilaevis                        | 2999      | 1552        | 27         | 150              | gb:AGRL01027172.1                |
| Cyprinodon variegatus                        | 3000      | 1570        | 338        | 254              | gb:JPKM01031331.1                |
| Danio rerio                                  | 3001      | 1563        | 338        | 44               | chr2 13613290 13614851           |
| Dicentrarchus labrax                         | 3002      | 1282        | 317        | 12               | emb:CAK01017223.1                |
| Esox lucius                                  | 3003      | 1561        | 27         | 323              | gb:AZR01046952.1                 |
| Haplochlamis burtoni                         | 3004      | 1563        | 27         | 338              | gb:AFN01059715.1                 |
| Larimichthys crocea                         | 3005      | 1574        | 335        | 13               | gb:JPKY01009954.1                |
| Maylandia zebra                              | 3006      | 1563        | 27         | 338              | gb:AGTA01022885.1                |
| Neolamprologin brichardi                     | 3007      | 885         | 294        | 22               | gb:AFNY01026505.1                |
| Nothobranchius furzeri                       | 3008      | 1011        | 336        | 6                | gb:ABLO01004310.1                |
| Orzyas latipes                               | 3009      | 1513        | 231        | 10               | dbj:BAA0104056582.1              |
| Pampus argenteus                             | 3010      | 982         | 224        | 46               | gb:JHEK01284027.1                |
| Periplophamus magnuspinnatus                 | 3011      | 1559        | 22         | 338              | gb:JACL01047866.1                |
| Pseudoleuronectes yokohamae                  | 3012      | 1398        | 250        | 31               | dbj:BDB01014196.1                |
| Penumilus yerere                             | 3013      | 1563        | 27         | 338              | gb:AFNY01048038.1                |
| Salmo salar                                  | 3014      | 1561        | 27         | 338              | gb:AGKD04000285.1                |
| Scartelos strophous                           | 3015      | 1562        | 27         | 264              | gb:JACN01055086.1                |
| Stegastes partitus                           | 3016      | 1562        | 27         | 338              | gb:JMKM01022513.1                |
| Takifugu flavius                             | 3017      | 1552        | 27         | 338              | gb:AOOT01008521.1                |
| Takifugu rubripes                            | 3018      | 1551        | 27         | 338              | gb:scaffold 12 809695810893       |
| Xiphophorus maculatus                        | 3019      | 1556        | 27         | 250              | gb:AGAJ01016295.1                |
| **Amphibians (clawed frog)**                 |           |             |            |                  |                                  |
| Xenopus (Silurana) tropicalis                | 3020      | 1573        | 202        | 326              | gb:ABLO01022242.1                |
| **Lepidosauria (snakes)**                    |           |             |            |                  |                                  |
| Crotalus mitchelli pyrrhus                   | 3021      | 1562        | 27         | 338              | gb:JPMF01075523.1                |
| Ophiophagus hannah                           | 3022      | 1561        | 26         | 234              | gb:AZIM01005391.1                |
| Python bivittatus                            | 3023      | 1729        | 27         | 320              | gb:AEQU01214040.1                |
| **Protozoa**                                 |           |             |            |                  |                                  |
| Perkinsus marinus                            | 3024      | 1229        | 289        | 48               | gb:AXA01001025.1                 |
| **Fungi (Mucorales)**                        |           |             |            |                  |                                  |
| Cokeromyces recurvatus                       | 3025      | 1672        | 218        | 286              | gb:JNHE01000254.1                |
| Cunninghamella bertholletiae                 | 3026      | 1488        | 126        | 343              | gb:GNEG01000333.1                |
| Lichtheima corymbifera                       | 3027      | 1368        | 59         | 303              | gb:JNEE01009092.1                |
| Mucor cirrhioides                            | 3028      | 1579        | 343        | 53                | gb:DMJN01013556.1                |
| Mucor racemosus                              | 3029      | 1647        | 38         | 343              | gb:JNEI01009390.1                |
| Mucor ramannianus                            | 3030      | 1647        | 216        | 342              | gb:JNFE01002670.1                |
| Mucor velutinizens                           | 3031      | 1646        | 33         | 343              | gb:JNDO01015981.1                |
| Rhizopus delemar                             | 3032      | 1671        | 211        | 343              | gb:AAAC02000027.1                |
| Rhizopus oryazae                             | 3033      | 1672        | 211        | 343              | gb:JNDV01011981.1                |
| Rhizopus microsporus                         | 3034      | 861         | 195        | 24               | gb:JNEJ01004884.1                |

Transposition by biochemical techniques, as well as compare their structural and functional parts with other transposons of the Tc1/mariner superfamily. Phylogenetic analysis of transposons based on their full length indicated that the TRT elements of metazoans and fungi could be classified into two independent groups (fig. 3). Interestingly, phylogeny also suggested that TRT elements from fungi and reported Tc1 elements showed much closer relationship when compared with TRT elements from metazoans, implying that they might be derived from a more recent ancestor. Indeed, the C-terminal

2998 Genome Biol. Evol. 8(9):2994–3005. doi:10.1093/gbe/eww213 Advance Access publication September 25, 2016
domains of TRT elements from fungi and reported Tc1 elements were very conserved (supplementary fig. S1, Supplementary Material Online). The above results suggest that the TRT elements from fungi are members of Tc1. Therefore, it is not appropriate to call these elements TRT. In order to solve this, all elements from fungi were renamed Tc1.

Although more than 437 fungal genomes have been sequenced, Tc1 elements were only detected in fungi belonging
to the Mucoromycotina (fig. 4). Tc1 elements were also extensively distributed in Mucoromycotina, although only 24 the Mucoromycotina species have been completely sequenced. Although the Dikarya embrace two large phyla (the Ascomycota and the Basidiomycota) that account for the vast majority of fungi and most (379/437) of the sequenced fungi, no Tc1 elements were detected in this taxonomic cluster. We should notice that sampling bias of the available databases may influence our level of detection for Tc1 elements in fungi. Sequencing of additional fungal genomes may facilitate in the identification of Tc1 elements. Such expansion will also facilitate in the elucidation of the evolutionary dynamics or patterns of Tc1 elements in a wide range of fungal species. Several different strains of fungal species have been completely sequenced, and thus we investigated the presence/absence of polymorphisms in the Tc1 elements of these strains. The fungi *Lichtheimia corymbifera* and *Cunninghamella bertholletiae* harbored polymorphisms in their TRT elements (supplementary fig. S2, Supplementary Material Online), thereby suggesting that Tc1 elements recently invaded their host and currently have not undergone fixation. Alternatively, Tc1 polymorphisms might result from cut-and-paste excision of the element.

**Comparison of TRT Elements with Other Known Tc1/Mariner Transposons**

To determine the relationship between TRT elements and other members of the Tc1/mariner superfamily, the transposase sequences of each TRT element were aligned to 11 mariner, 12 Tc1, and 7 maT elements. The reported DD37E transposons (*ItmD37E*) were also included in the analysis. The consensus amino acid sequence of the element (I5630) from *Shigella sonnei* was used as out-group. The phylogenetic tree showed that TRT elements were more closely related to Tc1 than to mariner and maT (fig. 3). The bootstrap values obtained using two different methods were all >99%. Our results also showed that TRT elements and the reported DD37E transposons formed independent clades, thereby suggesting that these belong to two different subfamilies of the Tc1/mariner superfamily.

Next, we compared the functional motifs in the catalytic domain of Tc1, mariner, maT, *ItmD37E*, and TRT elements. The results showed that three conserved motifs in the catalytic domain could separate mariner (TXDE, HDN, and SPDLAPS/T/H/DY) from Tc1 (W/FXS/TDE, QND and SPDNLPIE). The three conserved functional motifs of TRT elements included (W/F)TDE, Q/H/DN, and SPD/H/LNPIE (supplementary fig. S3, Supplementary Material Online), which was similar to those of Tc1. However, the *ItmD37E* conserved functional motifs (MDDE, PDLA, and V/CPOA/FRPIE) clearly differed from those of other members of the Tc1/mariner superfamily, thereby suggesting that *ItmD37E* could be organized into a distinct family of the Tc1/mariner superfamily.

**Fig. 3.**—Phylogenetic relationships of all TRT and Tc1 elements identified in this study with other members of the Tc1/mariner superfamily based on their transposases. Bootstrap values of <70% are not shown. Three clusters of TRT elements might be involved in HTs were circled up in circles. TRT and Tc1 elements identified in this study were shown using red color.
Evidence for Multiple HTs of TRT

The above phylogeny showed that TRT elements identified in this study could be classified into three major clusters (fig. 3): Cluster 1 includes 3 species (1 frog and 2 bony fishes); Cluster 2 includes 17 bony fishes; Cluster 3 includes 8 species (3 snakes and 5 bony fishes). Phylogenetic analysis also suggested that the host and TRT phylogenies were incongruent. This result might imply that TRT elements have been exposed to multiple episodes of HTs. To further demonstrate this conclusion, pairwise distances between all consensus sequences...
were calculated. Distances of almost all pairwise comparisons were extremely low (average = 0.084; SD = 0.053; range = 0.000–0.313) (fig 5 and supplementary table S3, Supplementary Material Online). Meanwhile, most of species involved in TRT pairwise distances shared last common ancestor more than 110 Ma (supplementary table S3, Supplementary Material Online). Considering these deep divergence time and the strikingly low-pairwise TRT distances seem incompatible with the possibility that these transposons were obtained from parents to offspring. Indeed, for almost all pairwise comparisons (153/167), the distances computed for TRT are much lower than those calculated for RAG1 (average = 0.302; SD = 0.160; range = 0.001–0.698) (supplementary table S3, Supplementary Material Online), which was usually used to infer the HT events of transposons in vertebrates (Gilbert et al. 2010; Gilbert et al. 2012). Two additional lines of evidence ruled out the scenario that TRT elements were vertically inherited from the last common ancestor of these species. First, TRT elements identified in closely related bony fishes showed a higher level of nucleotide sequence divergence than those of bony fishes and snakes (data not shown). Second, the taxonomic distribution of TRT elements was highly discontinuous. Although more than 226 vertebrates have been sequenced, TRT elements were only detected in bony fishes, frog, and snakes. Together, our results showed that the presence of TRT in most of species reported in this study is as a result of HT events. However, for 14 pairs of species, we also noticed that the distance computed between TRT was greater than the distances calculated for RAG1 (supplementary table S3, Supplementary Material Online). Two hypotheses may explain this phenomenon: First, TRT was transferred into the common ancestor of comparison species by HTs and these species got TRT by vertical transfers; Second, TRT was introduced in each species by HTs but the happening time of these HT events and the species diverged was very close. This phenomenon was also observed in other vertebrates (Gilbert et al. 2012).

In conclusion, our findings indicate that vertical and HTs may not be mutually exclusive and may have concurred in the evolution of TRT elements.

**Discussion**

**Description of a Tc1-like Transposon**

The present study identified a Tc1-like transposon in zebrafish. DrTRT presented all the hallmark features of Tc1-like elements, including the existence of a transposase of about 340 amino acids in length, a DDE motif, two HTH motifs in the DNA-binding domains, TiRs, and TSD TA at each end. The spacing of the DDE motif within the catalytic domain of DrTRT was unique, with 37 amino acids (DD37E) separating the second aspartic acid and the glutamic residues. Although ITmD37E has been previously reported (Shao and Tu 2001), phylogenetic analysis and functional domain comparison demonstrated that TRT elements and ITmD37E belong to two distinct subfamilies of the Tc1/mariner superfamily. Phylogenetic analysis also showed that TRT elements belong to the Tc1 family (fig. 3). However, the observed high genetic
distance with the rest of the Tc1 elements (DD34/35E) and the presence of a unique DDE spacing (DD37E) suggest that TRT elements constitute a novel subfamily within the Tc1 family.

The present study observed different species distributions for TRT and ITmD37E elements. TRT elements were detected in the genomes of bony fishes, clawed frog, snakes and protozoans following a search in GenBank. No TRT elements were observed in insects. In contrast, previous studies and our results (data not shown) indicated that ITmD37E extensively occurred in various species, including Arthropoda, Rotifera, flatworms, Hydrozoa, and ciliates (Shao and Tu 2001; Arkhipova and Meselson 2005; Biedler et al. 2007). On the other hand, ITmD37E was not detected in vertebrates.

TRT Elements Are Complete and Potentially Active in Haplochromini

To date, only some Tc1-like elements (such as Tol1, Tol2, Passport, and Tana1) that were identified in bony fishes have been demonstrated to be transpositionally active (Koga et al. 2006; Clark et al. 2009; Pujo lar et al. 2013; Watanabe et al. 2014). Previous studies showed that most Tc1-like elements in fishes are inactive because of stop codons, deletions, or frameshifts within their sequences (Radice et al. 1994; Reed 1999). In contrast, TRT elements were determined to be intact in the genomes of P. yererei, M. zebra, and H. burtoni. These transposases showed no internal stop codons or frameshift mutations, and their TIRs were completely identical, thereby suggesting that TRT elements have current or recent activity in these species. Generally, inactive transposons are expected to rapidly accumulate random mutations not only in the transposase but also in the TIRs, thereby disrupting their complementarity (Pujo lar et al. 2013). Recent activity of TRT elements is also supported by the observations of a very high sequence identity that was observed between full-length copies and consensus sequences, high copy number, and the presence of several elements in the TSA databases (fig. 2). Although intact transposases of TRT or Tc1 elements were also identified in other species, including Salmo salar, D. rerio, and fungi, no complete transposases was detected in their transcriptomes (fig. 2). These findings suggest that TRT elements might not be active in these species. Further analysis suggested that the activity of TRT elements in P. yererei, M. zebra, and H. burtoni might have begun about between 4 and 19 Ma.

Homology-based analysis indicates that TRT elements may have occur as approximately 250 copies in P. yererei, M. zebra, and H. burtoni (table 1). Similar copy numbers were also observed for the other two active Tc1-like elements, Passport (300 copies) and Tana1 (764–1,568 copies) (Leaver 2001; Pujo lar et al. 2013). Furthermore, the detection of Passport and Tana1 in divergent species also suggests that HT played important roles in their dissemination (Leaver 2001; Pujo lar et al. 2013).

Mechanisms and Direction of HT Events Involving TRT Elements

Although we have demonstrated that TRT elements underwent HT events to bony fishes, snakes and frogs, their underlying mechanisms remain unknown. However, it should be noted that DNA viruses are well recognized as potential intermediates for HT of transposons among various animal species (Schaack et al. 2010). Interestingly, it is reported that a snake retroposon integrated into the genome of the taterapox virus, a pox virus that infects West African rodents, by HT (Piskurek and Okada 2007). Therefore, we speculate that pox viruses may be an efficient vector for the HT of TRT elements that were identified in snakes. Parasitism might also facilitate the HT of transposons. Recently, some cases of HT events involving transposons between vertebrate and their parasites were reported (Kuraku et al. 2012; Walsh et al. 2013; Zhang et al. 2014; Suh et al. 2016). In the present study, a high sequence identity of TRT elements was also observed in several bony fishes, indicating that parasitism might have served as an effective delivery system for the HT of TRT elements.

A second important issue that needs to be addressed is the direction of HT. In the present study, the continuous distribution of TRT elements across bony fish genomes and TRT elements from snakes and frog are nested within the bony fishes (fig. 3), thereby suggesting HT events from the bony fishes to the snakes or frog.

Potential Applications

The discovery of a widespread Tc1-like transposon in organisms may have potentially important applications. Molecular tools are being developed for the genetic manipulation of organisms. Active Tc1/mariner transposons could be potentially used as molecular tools for transgenesis and insertion mutagenesis (Pavlopoulos et al. 2007; Voigt et al. 2016). The integrity of the native form (with an intact transposase), the presence of all functional domains, and their occurrence in the M. zebra and H. burtoni transcriptomes are suggestive of current or recent activity of TRT elements. Meanwhile, TRT elements are widespread in organisms. Therefore, we speculate that TRT elements identified in this study might be used as a molecular tool.

Supplementary Materials

Supplementary materials figures S1–S3, dataset S1 and S2, and tables S1–S4 are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The National Natural Science Foundation of China (31560308 to Z.H.H., 31260632 to Z.X.G., and 31401106 to H.M.J.), the Natural Science Foundation of Jiangxi Province (20161BAB214151 to Z.H.H.) and the Hi-Tech Research and
Development (863) Program of China (2013AA102507) supported this study. We also thank the language revision of this manuscript by the LetPub company.

**Literature Cited**

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Arkhipova IR, Meselson M. 2005. Diverse DNA transposons in rotifers of the class Bdelloidea. Proc Natl Acad Sci U S A. 102:11781–11786.

Biedler JK, Shao H, Tu Z. 2007. Evolution and horizontal transfer of a DD37E DNA transposon in mosquitoes. Genetics 177: 2553–2558.

Biémont C. 2010. A brief history of the status of transposable elements: from junk DNA to major players in evolution. Genetics 186:1085–1093.

Brillet B, Bigot Y, Augé-Gouilliou C. 2007. Assembly of the Tc1 and mariner transposition initiation complexes depends on the origins of their transposase DNA binding domains. Genetica 2007 130:105–120.

Clark KJ, Carlson DF, Foster LJ, Fahrenkrug SC. 2009. Passport, a native Tc1 transposon from flatfish, is functionally active in vertebrate cells. Nucleic Acids Res. 37:1239–1247.

Craig NL, Craigie R, Gellert M, Lambowitz AM. 2002. Mobile DNA II. Washington, DC: American Society for Microbiology Press.

Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1161–1165.

Dotto BR, et al. 2015. HTT-DB: horizontally transferred transposable elements database. Bioinformatics 31:2915–2917.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Feschotte C, Pritham EJ. 2007. Rampant horizontal transfer of SPIN transposons in squamate reptiles. Mol Biol Evol. 24:1596–1599.

Gilbert C, Hernandez SS, Flores-Benabib J, Smith EN, Feschotte C. 2012. Horizontal transfer of a DD37E DNA transposon in squamate reptiles. Mol Biol Evol. 29:503–515.

Hedges SB, Dudley J, Kumar S. 2006. TimeTree: a public knowledgebase of divergence times among organisms. Bioinformatics 22:2971–2972.

Jacobson JW, Medhora MM, Hartl DL. 1986. Molecular structure of a somatically unstable transposable element in Drosophila. Proc Natl Acad Sci U S A. 83:8684–8688.

Jurka J, et al. 2005. RepBase update, a database of eukaryotic repetitive elements. Cytogenet Genome Res. 110:462–467.

Kalender R, Lee D, Schumian AH. 2014. FastPCR software for PCR, in silico PCR, and oligonucleotide assembly and analysis. Methods Mol Biol. 1116:271–302.

Koga A, Iida A, Hori H, Shimada A, Shima A. 2006. Vertebrate DNA transposon as a natural mutator: the medaka fish Tol2 element contributes to genetic variation without recognizable traces. Mol Biol Evol. 23:1414–1419.

Kuraku S, Qiu H, Meyer A. 2012. Horizontal transfers of Tc1 elements between teleost fishes and their vertebrate parasites, lampreys. Genome Biol Evol. 4:929–936.

Leaver MJ. 2001. A family of Tc1-like transposons from the genomes of fish and frogs: evidence from horizontal transmission. Gene 271: 203–214.

Liu Y, Yang G. 2014. Tc1-like transposable elements in plant genomes. Mob DNA 5:17.

Lohe A, Sullivan D, Hartl D. 1996. Genetic evidence for subunit interactions in the transposase of the transposable element mariner. Genetics 144:1087–1095.

McGuffin LJ, Byson K, Jones DT. 2000. The PspRED protein structure prediction server. Bioinformatics 16:404–405.

Novick P, Smith J, Ray D, Boissinot S. 2010. Independent and parallel lateral transfer of DNA transposons in tetrapod genomes. Gene 449:85–94.

Pace JK 2nd Gilbert C, Clark MS, Feschotte C. 2008. Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. Proc Natl Acad Sci U S A. 105:17023–17028.

Pavlopoulos A, Oehler S, Kapetanaki MG, Savakis C. 2007. The DNA transposon Minos as a tool for transcription and functional genomic analysis in vertebrates and invertebrates. Genome Biol. 8:52.

Piskurek O, Okada N. 2007. Poxviruses as possible vectors for horizontal transfer of retrotransposons from reptiles to mammals. Proc Natl Acad Sci U S A. 104:12046–12051.

Plasterk RH, Izsák Z, Izsák C. 1999. Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. 15:326–332.

Pujolar JM, et al. 2013. Tana1, a new putatively active Tc1-like transposable element in the genome of sturgeons. Mol Phylogenet Evol. 66:223–232.

Radice AD, Bugaj B, Fitch DH, Eommons SW. 1994. Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. Mol Gen Genet. 244:606–612.

Reed KM. 1999. Tc1-like transposable elements in the genome of lake trout (Salvelinus namaycush). Mar Biotechnol. 1:60–67.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16:276–277.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Schaack S, Gilbert C, Feschotte C. 2010. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol Evol. 25:537–546.

Shao H, Tu Z. 2001. Expanding the diversity of the IS630-Tc1-mariner superfamilies: discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons. Genetics 159:1103–1115.

Suh A, et al. 2016. Ancient horizontal transfers of retrotransposons between birds and ancestors of human pathogenic nematodes. Nat Commun. 7:11396.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Tu Z. 2000. Molecular and evolutionary analysis of two divergent subfamilies of a novel miniature inverted repeat transposable element in the yellow fever mosquito, Aedes aegypti. Mol Biol Evol. 17:1313–1325.

Voigt F, et al. 2016. Sleeping Beauty transposase structure allows rational design of hyperactive variants for genetic engineering. Nat Commun. 7:11126. doi: 10.1038/ncomms11126.

Walsh AM, Kortchak RD, Gardner MG, Bertozzi T, Adelson DL. 2013. Widespread horizontal transfer of retrotransposons. Proc Natl Acad Sci U S A. 110:1012–1016.
Watanabe K, et al. 2014. Spontaneous germline excision of Tol1, a DNA-based transposable element naturally occurring in the medaka fish genome. Genome 57:193–199.
Xia X, Xie Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. J Hered. 92:371–373.
Zhang HH, Feschotte C, Han MJ, Zhang Z. 2014. Recurrent horizontal transfers of Chapaev transposons in diverse invertebrate and vertebrate animals. Genome Biol Evol. 6:1375–1386.

Associate editor: Esther Betran