Evolution of the *Xenopus piggyBac* Transposon Family TxpB: Domesticated and Untamed Strategies of Transposon Subfamilies

Akira Hikosaka,* Toshihiro Kobayashi,† Yumiko Saito,* and Akira Kawahara*

*Graduate School of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan; and †Graduate School of Biosphere Science, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan

A new family, termed TxpB, of DNA transposons belonging to the *piggyBac* superfamily was found in 3 *Xenopus* species (*Xenopus tropicalis, Xenopus laevis*, and *Xenopus borealis*). Two TxpB subfamilies of Kobuta and Uribo1 were found in all the 3 species, and another subfamily termed Uribo2 was found in *X. tropicalis*. Molecular phylogenetic analyses of their open reading frames (ORFs) revealed that TxpB transposons have been maintained for over 100 Myr. Both the Uribo1 and the Uribo2 ORFs were present as multiple copies in each genome, and some of them were framed by terminal inverted repeat sequences. In contrast, all the Kobuta ORFs were present as a single copy in each genome and exhibited high evolutionary conservation, suggesting domestication of Kobuta genes by the host. Genomic insertion polymorphisms of the Uribo1 and Uribo2 transposons (nonautonomous type) were observed in a single species of *X. tropicalis*, indicating recent transposition events. Transfection experiments in cell culture revealed that an expression vector construct for the intact Uribo2 ORF caused precise excision of a nonautonomous Uribo2 element from the target vector construct but that for the Kobuta ORF did not. The present results support our viewpoint that some Uribo2 members are naturally active autonomous transposons, whereas Kobuta members may be domesticated by hosts.

Introduction

Since the discovery of transposons by McClintock, a variety of them have been shown to be major components of eukaryotic genomes (for a review, see Kidwell 2005). It is believed that transposons invade a genome and subsequently spread throughout it during evolution. The “selfish” mobility of transposons is harmful to the host; hence, they are eliminated or inactivated by the host through natural selection. Even harmless transposons lose the activity eventually because of the absence of conservative selection for them. Thus, in general, transposons have a short life span in a host and they subsequently become fossils in the genome. However, if a transposon moves horizontally to a different species, it could survive further in the new host. In vertebrates, only one DNA transposon *Tol2* has been found to naturally occur in the fish medaka (Koga et al. 1996) and its recent invasion of the host genome has been suggested (Koga et al. 2000). No naturally occurring active DNA transposon has been found in tetrapods thus far. Sleeping Beauty (Ivics et al. 1997) and Frog Prince (Miskey et al. 2003) are 2 active DNA transposons that have been reconstructed from their fossils. A strategy of transposons to survive in a host is to reproduce autonomous copies continuously. This results in the accumulation of abundant nonautonomous or inactive copies along with a few autonomous copies in the genome. An alternative strategy of transposon survival in a host organism is so called “molecular domestication.” Transposons are rarely domesticated in hosts if the transposon is modified by mutations that are beneficial for the host (for a review, see Wolff 2006). In this case, transposons can be conserved by natural selection without the continuous amplification. Transposons living longer in the same host species are believed to have adopted either one of these 2 strategies (untamed or domestication strategy).

Previously, we have reported a new family of miniature inverted-repeat-repeat transposable elements (MITEs) termed as Xmix (Hikosaka et al. 2000; Hikosaka and Kawahara 2004) in *Xenopus*. In order to find the transposases (Tpases) responsible for the mobility of Xmix, we focused on the *piggyBac* superfamily of transposons. This superfamily is known to utilize the common target sequence TTAA (Robertson 2002), which is also utilized for transposition of the T2 family MITEs, including Xmix (Ünsal and Morgan 1995; Izsák et al. 1999; Hikosaka et al. 2000). The *piggyBac* superfamily members were first discovered in *Trichoplusia ni* (Cary et al. 1989; Fraser et al. 1996) and subsequently in a few other species (Handler and McCombs 2000; Sarkar et al. 2003). There are a few reports regarding the composite transposon family Tx with a targeting preference for TTAA in *Xenopus laevis* (Garrett and Carroll 1986; Garrett et al. 1989). This family is believed to belong to the *piggyBac* superfamily (Robertson 2002). Sarkar et al. (2003) discovered certain partial sequences of *piggyBac*-like Tpases in *X. laevis* although a full-length Tpase gene has not been detected thus far.

We discovered a new family of DNA transposons belonging to the *piggyBac* superfamily termed TxpB in 3 *Xenopus* species (*Xenopus tropicalis, X. laevis*, and *Xenopus borealis*). The members of this family were classified into the following 3 subfamilies: Kobuta, Uribo1, and Uribo2. Since the first divergence of these species (ca. 100 MYA), certain members of each subfamily were found to retain well-conserved open reading frames (ORFs). To elucidate the strategies adopted by these transposons for their long survival, we characterized them regarding copy numbers, structural conservations, and Tpase activity. These analyses suggested that some Uribo2 members encode active Tpases, whereas all the Kobuta members might be domesticated by hosts.

Materials and Methods

Database Search

Expressed sequence tag (EST) data for *X. laevis* was obtained using TBLASTN search from the National Center for Biotechnology Information (NCBI) Blast web site (http://www.ncbi.nlm.nih.gov/BLAST/). *Trichoplusia*
(GenBank accession number: AAA87375) and human (BAB71379) piggyBac Tpase protein sequences were used as queries. The hit sequences were used to design polymerase chain reaction (PCR) primers to amplify the Urib01, Xla-Kobuta, and Xbo-Kobuta Tpase ORFs (see the following section of PCR and cloning). The unmasked genome sequence data (release 4.1) for *X. tropicalis* was downloaded from the Department of Energy (DOE) Joint Genome Institute (JGI) web site (http://genome.jgi-psf.org/Xentr4/Xentr4.home.html) and searched by the NCBI Blast program (Altschul et al. 1997) using cloned Urib01 and Xla-Kobuta ORF sequences (see Results) as queries. The hit sequences were used to design PCR primers for amplifying the Xtr-Kobuta and Xtr-Urib02 Tpase ORFs.

**PCR and Cloning**

*Xenopus laevis* and *X. borealis* were obtained as mentioned previously (Hikosaka et al. 2000). *Xenopus tropicalis* strains were provided by the National Bioresource Project, Japan. DNA was extracted from animals as described elsewhere (Hikosaka et al. 2000).

For usual PCR conditions, 100 ng genomic DNA was used. Short PCR (≤3 kb) was performed using KOD Plus polymerase (Toyobo, Osaka, Japan) with the following thermal cycling parameters: 94°C, 2 min (94°C, 30 s; 50–60°C, 15 s; and 68°C, 1 s to 4 min); 30× cycles; and 68°C, 5 min. Long PCR (>3 kb) was performed using KOD Dash polymerase (Toyobo) with the following thermal cycling parameters: 94°C, 2 min (94°C, 30 s; 60°C, 2 s; and 74°C, 3 min); 30× cycles; and 74°C, 5 min. The following PCR conditions as to the primer set and annealing temperature were used for amplification of the Tpase ORFs: for Urib01, GCCGTCTAGAGATTTTTTCTGGCA/CAATGGCCTAAGTCTGAGAG/AACAGTGCTAC, Uribo2 (GCTGAAGAAGCTGCTGCACG/ACCTGCAGATTTCACTGTTAC, and Kobuta (GCACTGGAAAACCTGATAGGG and CTACTGCATCTACATTGAG). The PCR products were cloned into the EcoRV site of the pBlueScript II and used for in vitro transcription with T3 RNA polymerase using a digoxigenin-labeling mix (Roche Applied Science, Mannheim, Germany). Southern blot hybridization was performed as described elsewhere (Hikosaka and Kawahara 2004) with minor modifications (hybridization was performed at 50°C and followed by posthybridization washing at 60°C in a 0.5× saline sodium citrate (SSC)–0.1% sodium dodecyl sulfate solution).

**Analysis of Transposons by In Silico Walking and PCR**

Sequences homologous to the 200 bp N- or C-terminals of TxpB ORFs were identified by BlastN search, and the 5 kb sequences upstream from the N-terminals or downstream from the C-terminals were aligned to identify arm regions and terminal inverted repeat (TIR). Using the predicted TIR consensus sequences of CCYTTTTNNTGCCA (see Results), the transposable elements related to Xtr-Urib01 and Urib02 were extracted from the *X. tropicalis* genome database by a regular expression search using Perl. Namely, in order to find the sequences (51–10,000 nt in length) framed by the 14 nt TIR consensus sequences, the following regular expression was used: CC[CT][T][T][TT][AGT][C] [3] [TGCC][AAG][GTC][23,9972]?[TGCC][AAC][GAG]. Hit sequences were extracted along with 4 nt (the length of target site duplication [TSD]) plus additional 30 nt flanking sequences from the database. The flanking sequences were used to remove redundant hits from the total hits. The finally extracted 627 unique loci were analyzed and classified into Urib01 or Urib02. Some transposon sequences were compared using the polydot program of the Emboss package.

**Transposon Excision Assay in Cell Cultures**

An expression vector for Urib02 Tpase (pExp-Urib02) was constructed from the cloned Xtr-Urib02 Tpase molecular phylogenetic analysis. As outgroups within the piggyBac superfamily, 5 Tpase sequences (human HsaPGBD4, ascidian CinPBD1 and CinPB3, moth ThiPB, and mosquito AgaPB8) from the data reported by Sarkar et al. (2003) were used for the analysis. HsaPGBD4, CinPB3, and AgaPB8 are the Tpases most closely resembling the TxpB Tpases. The amino acid sequences were aligned using ClustalW (Thompson et al. 1994). Molecular phylogenetic trees were constructed using Neighbor-Joining methods (Saitou and Nei 1987). The cloned TxpB Tpases were registered as accession numbers AB300551–63.

**Southern Blot Hybridization Analyses**

The probes used for Southern blot hybridization were constructed from the desired regions of the cloned Tpases that have been mentioned in supplementary material 1 (Supplementary Material online) (Xtr-Urib01_PCR, Xtr-Urib02_PCR, Xtr-Urib01_R, and Xtr-Kobuta_PCR_Ni1). To amplify these probe sequences, the following primer sets were used: Urib01 (ACCGAGGAAGCTGCTAGTGA and AACTGCTGCATCACTTAC); Uribo2 (GCTGAAGAAGCTGTCAGCACA and AGCTGCACTACAGCTCCTAT), and Kobuta (GCACTGGAAAACCTGATAGGG and CTACTGCATCTACATTGAG). The PCR products were cloned into the EcoRV site of the pBlueScript II and used for in vitro transcription with T3 RNA polymerase using a digoxigenin-labeling mix (Roche Applied Science, Mannheim, Germany). Southern blot hybridization was performed as described elsewhere (Hikosaka and Kawahara 2004) with minor modifications (hybridization was performed at 50°C and followed by posthybridization washing at 60°C in a 0.5× saline sodium citrate (SSC)–0.1% sodium dodecyl sulfate solution).

**Molecular Phylogenetic Analyses**

The data obtained from the Xenopus genomic database and the PCR clone sequences (supplementary materials 1 and 2, Supplementary Material online) were used for the molecular phylogenetic analysis. As outgroups within the piggyBac superfamily, 5 Tpase sequences (human HsaPGBD4, ascidian CinPBD1 and CinPB3, moth ThiPB, and mosquito AgaPB8) from the data reported by Sarkar et al. (2003) were used for the analysis. HsaPGBD4, CinPB3, and AgaPB8 are the Tpases most closely resembling the TxpB Tpases. The amino acid sequences were aligned using ClustalW (Thompson et al. 1994). Molecular phylogenetic trees were constructed using Neighbor-Joining methods (Saitou and Nei 1987). The cloned TxpB Tpases were registered as accession numbers AB300551–63.
(Xtr-Uribo2_PCRIv1b). After PCR of the clone using the T7 primer (TAATACGACTCAGTGAGG) and N-terminal primer (CGGATCCAGATCCCTAGGACAAAAAGGGTTATT: the underlined sequence is the Kozak consensus sequence including translation initiation site), the PCR products were digested with BamHI and SalI and then inserted between the BamHI and XhoI sites of the pcDNA3.1/Hygro (+) vector (Invitrogen, Carlsbad, CA) and cloned. The expression vector for Kobuta “Tpase” (pExp-Kobuta) was also constructed in a similar manner (a primer set of T7 and CGGATCCAGATCCCTAGGACAAAAAGGGTTATTATA and a template of the Xtr-Kobuta PCR Ni4a clone were used). Their structures are shown in figure 6A. A target vector was constructed from a nonautonomous Uribo2 transposon by amplifying the U2_#2 insertion site (accession number AB332396) as described above and cloned into the EcoRV site of the pBlueScript II vector (fig. 6B). Transfection experiments were performed using human 293T cells according to the procedure described elsewhere (Saito et al. 2005). The cells were cultured for 2 days, and their DNAs were extracted using a Wizard genomic DNA Purification Kit (Promega, Madison, WI). The excision reaction of the target was detected by PCR using the primer set of TTTTCCCAGTCACGAC-GTT (primer U19) and ATGACCATGATTACGC (primer R20) that flanked the EcoRV site of the pBlueScript II at both sides. The PCR product was isolated by gel electrophoresis, and its sequence was analyzed as mentioned above.

Results

Detection and Cloning of Xenopus piggyBac Tpase Genes

In the X. laevis EST database, 2 types of piggyBac Tpase homologues were found by TBLastN searches using Trichoplusia (accession number: AAB60375) and human (BAB71379) sequences as queries. The corresponding genomic DNA sequences of X. tropicalis, X. laevis, and X. borealis were amplified by PCR using primers designed for sequences at both ends of the Tpase ORFs (see Materials and Methods) and then cloned. We obtained 2 types of ORFs (Uribo1 and Kobuta) from both X. laevis and X. borealis and 1 type of ORF (Uribo1) from X. tropicalis. We refer to these ORFs as Xla-Uribo1 and Xla-Kobuta (X. laevis), Xbo-Uribo1 and Xbo-Kobuta (X. borealis), and Xtr-Uribo1 (X. tropicalis). Differences in the ORFs were observed as gaps and nucleotide substitutions that were characteristic to their respective types (supplementary material 1, Supplementary Material online). We termed the present transposon family as the TxpB family. Sarkar et al. (2003) reported 2 partial ORF sequences (BJ073130 and BJ073435) that exhibited some similarity to Xla-Uribo1. These sequences could be included in the TxpB family.

Subsequently, we searched the X. tropicalis genome database for sequences related to the TxpB Tpase ORF and found many candidate sequences (supplementary material 2, Supplementary Material online). Most of the sequences were destroyed by indels, frameshifts, and nonsense mutations in their ORFs; however, some sequences appeared to have complete ORFs. These ORFs were classified into the Xtr-Kobuta, Xtr-Uribo1, and Xtr-Uribo2 subfamilies based on similarities among their sequences. Finally, 2 of 122 Xtr-Uribo1-related and 7 of 149 Xtr-Uribo2-related sequences were found to possess complete ORFs of about 1.8 kb (we termed these long ORFs “complete” ORFs). Only one copy was found for Xtr-Kobuta, which possessed a complete ORF. To isolate these ORFs, we designed primers for the unique sequences flanking these ORFs and amplified them by PCR from the genomic DNAs of 3 X. tropicalis lines (Nigerian, Ivory Coast, and Asashima lines). The sequences that were cloned from 2 individuals of the Nigerian line were identical to the Kobuta sequence in the database, whereas those cloned from the other lines exhibited the following polymorphisms: 2 synonymous and 5 nonsynonymous substitutions among the 5 ORFs (supplementary material 1, Supplementary Material online). In contrast, some Xtr-Uribo2 ORFs exhibited fatal mutations such as frameshift or nonsense mutations (data not shown). The diversity of the Uribo1 and Uribo2 members was observed as nucleotide substitutions at 117 positions in 3 complete Xtr-Uribo1 ORFs and at 412 positions in 12 complete Xtr-Uribo2 ORFs (excluding gaps) (supplementary material 1, Supplementary Material online). In addition, amino acid substitutions were observed at 73 positions of the Xtr-Uribo1 ORFs and 155 positions of the Xtr-Uribo2 ORFs. Thus, within a single species of X. tropicalis, sequence diversity of the Uribo1 and Uribo2 ORFs was extremely large but that of the Kobuta ORFs was little, suggesting that Kobuta gene would be, unlike Uribo1 and Uribo2 genes, under the conservative pressure of natural selection.

Molecular Phylogenetic Analyses of the TxpB Family

To elucidate the identity of the TxpB family, we analyzed their phylogenetic relationships among piggyBac superfamily Tpase sequences. As outliers within the piggyBac superfamily, we used 5 Tpases. HsaPGBD4 (Homo sapiens) is the most closely resembling protein to the TxpB Tpases in the GenBank protein database; CinPB3 (ascidian Ciona intestinalis) and AgaPB8 (mosquito Anopheles gambie) are also closely resembling piggyBac Tpases. In addition, TniPB (moth T. ni), the first discovered piggyBac Tpase, and CinPBD1 (C. intestinalis), an example of apparently domesticated Tpase (Sarkar et al. 2003), were also used for the analysis (an alignment of their amino acid sequences is available in supplementary material 3, Supplementary Material online). A molecular phylogenetic tree constructed revealed that the TxpB family was a distinct group within the piggyBac superfamily (fig. 1). Furthermore, the relationships within the TxpB family suggested that the 3 subfamilies of Uribo1, Uribo2, and Kobuta had diverged from a common transposon prior to branching of X. tropicalis from the group of X. laevis and X. borealis. This speciation is believed to have occurred more than 100 MYA (Knochel et al. 1986).

The catalytic DDE motif is known to be one of the characteristics of widespread Tpases. The amino acids of D268, D346, and D447 of T. ni piggyBac Tpase was predicted to be a DDE-like motif (Sarkar et al. 2003). An alignment analysis revealed that this motif was conserved in
Uribo1 and Uribo2 but not in Kobuta (fig. 2). In Kobuta, the first D changed to N in *X. tropicalis* and the third D changed to N in all the 3 *Xenopus* species, suggesting lack or modification of Tpase activity.

Copy Numbers of 3 TxpB Subfamily Tpases in 3 *Xenopus* Species

The present *X. tropicalis* genome database analysis revealed the presence of one “intact” Kobuta Tpase and few intact and many deficient Uribo1 and Uribo2 Tpases. However, the construction of the *X. tropicalis* genome database has not been completed. In addition, no information regarding the other *Xenopus* species genomes is currently available: hence, the copy number of the Kobuta Tpase gene in the genome is unknown. In order to obtain information regarding the copy numbers of the subfamilies, we performed Southern hybridization analysis. Restriction enzyme–digested genomic DNAs of *X. laevis* and 2 *X. tropicalis* lines (Nigerian and Asashima) were analyzed by the hybridization probes specific to each subfamily (fig. 3). The Kobuta probes gave a single band despite the restriction enzymes used, but both the Uribo1 and the Uribo2 probes gave multiple bands. The results suggested that the Kobuta Tpase gene was a single-copy gene, whereas both the Uribo1 and the Uribo2 Tpases genes were multicopy genes.

Identification of Arms, TIRs, and TSD

An active transposon is considered to have 2 arms (left and right arms located at the 5' and 3' ends of the Tpase ORF, respectively) with TIR and TSD at the ends. We found both the arms and the TIRs for some Xtr-Uribo1 and Xtr-Uribo2 ORFs (supplementary material 4, Supplementary Material online) but not for the Xtr-Kobuta ORF in the genome database. The left and right arms flanking Xtr-Uribo1 and Xtr-Uribo2 ORFs were compared by a dot-plot analysis (fig. 4A). The arms were moderately similar within each subfamily but fairly different between the subfamilies. Therefore, the arms could also be classified into the Uribo1 and Uribo2 subfamilies. Furthermore, a sequence of CCYTTTTNNNTGCCA was predicted as a consensus TIR (supplementary material 4, Supplementary Material online). In most cases, the piggyBac target sequence (TTAA) located adjacent to the TIR sequence. A G nucleotide generally followed the right TIR but not the left TIR. In addition, the ninth position of the TIR sequence was a G in all the Xtr-Uribo1 sequences but not in all the Xtr-Uribo2 sequences. These differences enabled us to distinguish between the left and the right arms as well as between the 2 subfamilies.

In order to profile the TxpB-related transposon population, we searched the *X. tropicalis* genome database for sequences (51–10,000 nt long) framed by the TIR consensus sequence (supplementary material 5, Supplementary Material online). A regular expression search identified 627 unique transposon candidates; of these, 488 (78%) were flanked by TTAA on both sides. *Xenopus tropicalis* TIR sequences were classified into Uribo1 and Uribo2 according to the above-mentioned criterion. The regular TIR combination was observed in 593 sequences (95%), whereas a mismatched TIR combination of Uribo1 and Uribo2 type was observed in 34 sequences. Subsequently, the TIRs were classified into left or right sequences. Most of the sequences (558 sequences, 89%) exhibited the regular left–right orientation identical to those of the Xtr-Uribo1 and Xtr-Uribo2; however, some exhibited a left–left or right–right orientation. Finally, 128 Uribo1 and 319 Uribo2 elements were found to have typical TIR and TSD structures. A majority of these elements did not demonstrate any trace of a Tpase ORF (defined as a nonautonomous transposon);
however, 7 sequences (supplementary material 5, Supplementary Material online) possessed Tpase-derived sequences. A dot-plot analysis was used to compare 15 randomly selected Uribo1 and Uribo2 elements each, and it revealed that the elements were moderately similar within the subfamily but fairly different between the subfamilies (fig. 4B). Furthermore, certain elements were found to contain paired tandem repeat (PTR)-like units and/or ORFs related to non-long terminal repeat (LTR) retrotransposons (supplementary material 5, Supplementary Material online). These profiles are similar to those of the X. laevis Tx transposons (Garrett et al. 1989).

Transposon Insertion Polymorphisms in 3 X. tropicalis Lines

If TxpB Uribo1 or Uribo2 remains active, transposon insertion polymorphism can be detected in a single Xenopus species. From the genome database, 8 insertion sites of non-autonomous Uribo1- or Uribo2-type elements were selected judging from their intactness of TIR (U1_#1 to #4 and U2_#1 to #4; supplementary material 5, Supplementary Material online) and analyzed using PCR with specific primers for each site in 3 frogs of the X. tropicalis Nigerian line, 1 frog of the Ivory Coast line, and 1 frog of the Asashima line. Among them, an Uribo1 insertion site (U1_#1) and an Uribo2 insertion site (U2_#1) showed transposon insertion polymorphisms (fig. 5A). The inserted transposon DNA, expected to be present at the U1_#1 site according to the genome database, was detected in the genomic DNA of 1 Nigerian frog (Ni8). However, it was not detected in DNAs of 2 Nigerian frogs, 1 Ivory Coast frog, and 1 Asashima frog (the PCR products were considerably shorter than 3.7 kb that is the expected length for the U1_#1 site). The insertion at the U2_#1 site was detected in the genomic DNAs of all the Nigerian frogs but not in the Ivory Coast line.
and Asashima frogs (the PCR products were considerably shorter than 4.6 kb that is the expected length for the U2_#1 site). A sequence analysis revealed that the short PCR products completely lacked the transposon insertion present in the database sequences (fig. 5B), and they retained only 1 TTAA target sequence at the prospective insertion site. These insertion polymorphisms suggest recent transposition events for the Uribo1 and Uribo2 transposons.

Transposon Excision Activity of the Uribo2 Tpase

Tpase expression vectors driven by the cytomegalovirus (CMV) promoter (pExp-Uribo2 and pExp-Kobuta) were constructed for the transfection experiment as shown in figure 6A. A transposon target sequence in the target vector (pTE-Uribo2) was derived from a nonautonomous Uribo2 element at U2_#2 site (fig. 6B and C; supplementary material 5, Supplementary Material online). The excision activity of

![Figure 6](https://academic.oup.com/mbe/article-abstract/24/12/2648/977085)
the Tpase was examined by the cotransfection of 293T cells with the expression and target vectors and detected by PCR using primers specific for the sequences flanking the cloning site (fig. 6B). A short DNA band (ca. 310 bp) was detected in the PCR products of the pExp-Uribo2–cotransfected cells under PCR condition suitable for the amplification of short DNAs (PCR elongation time of 1 s), but not in those of the pExp-Kobuta–cotransfected cells. This short DNA sequence was completely compatible with the transposon-free insertion site sequence, except for a TTAA target site that was retained in the vector (fig. 6C and E). Under PCR condition suitable for the amplification of long DNA, the PCR products from the cotransfected cells as well as those transfected with the target vector alone were observed to contain long DNA (ca. 2.7 kb) corresponding to the intact target sequence (fig. 6D). These results revealed that the Uribo2 protein, but not the Kobuta protein, excised a fraction of the transected target sequences.

Discussion

Tx1C and Tx2C are composite transposable elements found in X. laevis (Garrett and Carroll 1986; Garrett et al. 1989), which have chimeric structures of nonautonomous DNA transposons (Tx1D or Tx2D) containing 393 bp PTRs (PTR1 and PTR2) interrupted by non-LTR retrotransposons (Tx1L or Tx2L). Tx1D and Tx2D are believed to belong to the piggyBac superfamily because they have short TIRs connected to the TSD of TTAA (Robertson 2002). In addition, their consensus TIR sequences are extremely similar to those of the Xtr-Urib1 and Xtr-Urib2 members. Some nonautonomous elements of Xtr-Urib1 or Xtr-Urib2 harbor sequences related to PTRs and/or Tx1L ORFs (supplementary material 5, Supplementary Material online), whereas the arms of the TxPB transposon that bear the Tpase gene or its fragment do not possess these sequences. Taking these similarities and certain other aspects into consideration, we postulated their evolutionary relationships as shown in figure 7. An autonomous DNA transposon (Tx ancestor) invaded the genome of a common ancestor of the Xenopus species. The siblings diverged into 2 major subtypes: the TxP and the TxD ancestors. The TxP siblings retained the Tpase gene, whereas the TxD siblings did not. The nonautonomous TxD members could survive by their replication/transposition driven by the TxP Tpases. The TxD PTR units were inherited from a Tx ancestor or were new appearances in the TxD lineage. The autonomous TxP ancestor diverged into the 2 subfamilies of Uribo1 and Uribo2 that were distinguishably based on the differences in their Tpase and/or TIR sequences. The TxD members also diverged into Uribo1- and Uribo2-like nonautonomous elements, which were distinguishable from nonautonomous TxP-B Uribo1 or Uribo2 elements by the presence of a PTR sequence. It is unclear whether the divergence of the Uribo1 and Uribo2 subfamilies occurred prior to or following the divergence of TxP and TxD. The insertion of a TxL non-LTR retrotransposon into TxD might occur in the common ancestor. One of the autonomous TxP transposons was domesticated by the host and evolved into the Kobuta subfamily.

In the present study, we demonstrated that the Xtr-Urib2 Tpase has excision activity toward the target transposon, although there is no evidence for the integration of the excised target into the genome thus far. The accurate excision reaction, which left only a TTAA target site, may minimize damage to the genome and thereby contribute to the long lifetime of this subfamily. The insertion polymorphisms of the nonautonomous Xtr-Urib1 and Xtr-Urib2 elements that were observed in the X. tropicalis lines reveal that their transpositions had occurred recently. No direct evidence for the presence of naturally active DNA transposons in tetrapods has been obtained thus far. Therefore, Xtr-Urib2 is a probable candidate for the first naturally active transposon in tetrapods. Recently, Ray et al. (2007) reported evidence for recent transposition of a transposon family, Myotis-nhAT, in the bat genus Myotis. Although the activity of a putative Tpase of this family has not been proved, Myotis-nhAT could be another candidate for naturally active transposon in tetrapods.

In contrast, Kobuta might be domesticated by the hosts. Evidences are accumulating for the roles of
domesticated transposons in the evolution of organisms (Vollf 2006). Sarkar et al. (2003) suggested the possibility that some piggyBac superfamily Tpases might be independently domesticated by the host organisms. In the present study, we obtained some evidence suggesting domestication of the Kobuta Tpase, namely, 1) the gene exists as a single copy in host genomes and 2) its ORF is conserved over 100 Myr with multiple single amino acid substitutions and a few gaps but without frameshift or nonsense mutation. This high conservation of the single-copy gene could not be explained without purifying natural selection. 3) We found sequences almost identical to the Kobuta Tpases in the EST databases of X. laevis and X. tropicalis (data not shown), suggesting that the Kobuta Tpase genes are expressed in the hosts. In addition, 4) Kobuta Tpases have changed the DDE-like (DDD) motif to DDN or NDN. This motif is believed to be important for Tpase activity of piggyBac Tpases. Although D to N substitution may not necessarily lead to the total loss of Tpase-related activities, the substitution could modify the original Tpase function (Liu et al. 2007; Miskey et al. 2007). For instances, the Kobuta protein might contribute to the host life through its DNA-binding activity or DNA-recombination activity taking over from the ancestral Tpase. As a special function, it might inactivate the Uribo Tpases through heterodimerization, so the active Tpase genes might be not so harmful in the host and thereby survive for a long time during Xenopus evolution.

Supplementary Material

An alignment of the TxpB ORF sequences (supplementary material 1), location of the TxpB ORF-related sequences found in the X. tropicalis genome database (supplementary material 2), an alignment of TxpB and other piggyBac Tpase amino acid sequences (supplementary material 3), left and right arms of the Xtr-Urib01 and Xtr-Urib02 ORFs (supplementary material 4), and location of the Xtr-Urib01 and Xtr-Urib02 transposon candidates (supplementary material 5) are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We wish to express our gratitude to the National Bioresource Project, particularly to Dr Yaoita, Dr Kashiwagi (Center for Amphibian Biology, Hiroshima University), Dr Asashima, and Dr Takahashi (Graduate School of Arts and Sciences, The University of Tokyo) for providing the X. tropicalis strains and to Dr Tochinai (Graduate School of Arts, Sports, Science, and Technology (No. 15770001)).

Literature Cited


Fraser MJ, Ciszczon T, Elick T, Bauser C. 1996. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. Insect Mol Biol. 5:141–151.


Norihiro Okada, Associate Editor

Accepted September 6, 2007