Gene 593 (2016) 308–314

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Sequence analyses and chromosomal distribution of the *Tc1/Mariner* element in Parodontidae fish (Teleostei: Characiformes)

Michelle Orane Schemberger [a,b,*], Viviane Nogaroto [b], Mara Cristina Almeida [b], Roberto Ferreira Artoni [b], Guilherme Targino Valente [c], Cesar Martins [d], Orlando Moreira-Filho [e], Marta Margarete Cestari [a], Marcelo Ricardo Vicari [a,b]

[a] Department of Genetics, Federal University of Paraná, Curitiba, Paraná, Brazil
[b] Department of Structural and Molecular Biology and Genetics, State University of Ponta Grossa, Ponta Grossa, Paraná, Brazil
[c] Department of Bioprocess and Biotechnology, São Paulo State University, Botucatu, São Paulo, Brazil
[d] Department of Morphology, Institute of Biosciences, São Paulo State University, São Paulo, Brazil
[e] Department of Genetics and Evolution, Federal University of São Carlos, São Carlos, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 18 February 2016
Received in revised form 31 July 2016
Accepted 20 August 2016
Available online 22 August 2016

Keywords:
Cytogenetic
Repetitive DNA
Deterioration
W chromosome
Bioinformatics
Molecular analyses

ABSTRACT

Transposable elements are able to move along eukaryotic genomes. They are divided into two classes according to their transposition intermediate: RNA (class I or retrotransposons) or DNA (class II or DNA transposons). Most of these sequences are inactive or non-autonomous in eukaryotic genomes. Inactivate transposons can accumulate mutations at neutral rates until losing their molecular identity. They may either be eliminated from the genome or take on different molecular functions. Transposable elements may also participate in the differentiation of sex chromosomes. Therefore, the structural variations and nucleotide similarity of *Tc1/Mariner* sequences were analyzed along with their potential participation in the differentiation processes of sex chromosomes in the genomes of Parodontidae fish. All Parodontidae species presented non-autonomous copies of *Tc1/Mariner* with structural variation, different levels of deterioration (genetic distance), and variations in insertion and deletion patterns. The physical mapping of *Tc1/Mariner* on chromosomes revealed dispersed signals in euchromatins, with small accumulations in terminal regions and in the sex chromosomes. The gene dosage ratios indicated copy number variations of *Tc1/Mariner* among the genomes and high transposase open reading frame deterioration in *Parodon hilarii* and *Parodon pongoensis* genomes. This transposon presented transcriptional activity in gonads, but there was no significant difference between sexes. This may indicate non-functional protein expression or may correspond to DNA binding proteins derived from *Tc1/Mariner*. Thus, our results show *Tc1/Mariner* inactivation along with a diversity in Parodontidae genomes and its participation in the differentiation of the W sex chromosome.

1. Introduction

Transposable elements (TEs) are DNA sequences capable of moving from one genomic location to another. They form a significant proportion of eukaryotic genomes (Kidwell, 2002; Wicker et al., 2007) due to their abilities to increase the number of copies in the absence of selection. TEs have high mutagenic potential, but host genomes evolved epigenetic mechanisms that suppress their activity (Slotkin and Martienssen, 2007). However, several studies have reported the importance of co-opted TEs for genomic functions and evolution (Feschotte and Pritham, 2007; Feschotte, 2008; de Souza et al., 2013).

TEs display extreme diversity and are classified as Class I (retrotransposons that move by an intermediate RNA) or Class II (DNA transposons which move directly into the genome) (Wicker et al., 2007). Both classes exist as autonomous and non-autonomous elements. The autonomous copies encode all enzymes necessary for transposition while non-autonomous have no coding capacity and thus, their mobility depends on enzymes produced by autonomous elements (Wicker et al., 2007). Class II elements can incorporate mutations

Abbreviations: cDNA, complementary DNA; CTAB, cetyltrimethylammonium bromide; DDD, aspartate/aspartate/aspartate; DDE, aspartate/aspartate/glutamate; dsRNA, double-stranded RNA; FISH, fluorescence in situ hybridization; FSR, female specific region; GDR, gene dosage ratios; HLD, homeodomain-like-domain; HTH, helix-turn-helix; miRNA, micro RNA; MITEs, miniature inverted repeat transposable elements; NCBI, national center for biotechnology information; ORFs, open reading frames; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction; siRNA, small interfering RNA; TEs, transposable elements; TFBSs, transcription factor binding sites; TIRs, inverted terminal repeats.

* Corresponding author at: Programa de Pós-Graduação em Biologia Evolutiva, Universidade Estadual de Ponta Grossa, Av. Carlos Cavalcanti, 4748, 84030-900 Ponta Grossa, Paraná, Brazil.
E-mail address: miorane@hotmail.com (M.O. Schemberger).

http://dx.doi.org/10.1016/j.gene.2016.08.034
0378-1119/

which in turn are believed to be related to the formation of miniature inverted repeat transposable elements (MITEs), which are small (generally <600 bp in length) and non-autonomous elements (lack coding sequence) characterized by conserved terminal inverted repeats (Jiang et al., 2004; Fernández-Medina et al., 2012).

Tc1/Mariner superfamily is a Class II element and is present in several animal and plant taxa. These elements are about 1000–5000 bp in length and contain a single gene encoding transposase enzyme. Their common characteristics include inverted terminal repeats (TIRs) flanked by TA nucleotides, catalysis domain (DDE - aspartate/aspartate/glutamate or DDD - aspartate/aspartate/aspartate) and DNA-binding domains (Plasterk et al., 1999).

Usually, TEs in genomes are present as non-autonomous copies, generated by a degradation process (Fernández-Medina et al., 2012). When TEs are inactivated they accumulate mutations leading to the loss of their identity. Hence, the life cycle of a TE consists of the invasion of a host genome, an increase in its number of copies followed by an inactivation and finally, its elimination (Kidwell and Lisch, 2001). Mutated TEs can behave as "neutral sequences" in genomes and can pass through the evolutionary process by "co-option", "exaptation" or "molecular domestication", losing its characteristics of a TE and acquiring new genomic functions which favor the host (Sinzelle et al., 2009). In this respect, studies have reported the importance of TEs in post-transcriptional and translation regulation, in the origin of new genes and noncoding RNAs (Long et al., 2003; Marino-Ramirez et al., 2005; Feschotte, 2008; Herpin et al., 2010; Kapusta et al., 2013). Furthermore, TEs are important for the evolution of heteromorphic sex chromosomes, they are capable of accumulating in nonrecombining genomic regions of proto-sex chromosomes and reducing gene expression through chromatin modification (Charlesworth et al., 2005).

Parodontidae is a group of Neotropical fish with 54 conserved chromosomes. The group is characterized by the presence of diverse sex chromosome systems, including species without heteromorphic sex chromosomes, species with proto-sex chromosomes, species with a ZZ/ZW sex chromosome system, and species with a multiple ZZ/ZW_1W_2 sex chromosome system (Bellafronte et al., 2011; Schemberger et al., 2011). Studies with transposable elements in fish are scarce (de Souza et al., 2013; Herpin et al., 2010; Lowe and Haussler, 2012). In this study, the *Tc1/Mariner* element was isolated, characterized and mapped on chromosomes in order to analyze the sequence structural variations and their potential role in the differentiation processes in the sex chromosomes of Parodontidae fishes.

2. Material and methods

2.1. Species analyzed, chromosome preparation and DNA extraction

Nine species from two genera of Parodontidae were selected for sequence analyses: *Apareiodon piracicabae*, *Apareiodon vittatus*, *Apareiodon* sp., *Apareiodon vladii*, *Apareiodon affinis*, *Apareiodon hasemani*, *Parodon hilarii*, *Parodon pongoensis* and *Parodon nasus* (Table 1). Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) procedure, according to Murray and Thompson (1980). Chromosomal preperations, from *Apareiodon* sp. (ZZ/ZW), *P. hilarii* (ZZ/ZW) and *A. affinis* (ZZ/ZW_1W_2), were obtained from anterior kidney cells using an "air-drying" method (Bertollo et al., 1978). This study was carried out in strict accordance with the recommendations of the Editorial Ethical justification for the use and treatment of fish in research. The protocol was approved by the Committee on the Ethics of Animal Usage of the State University of Ponta Grossa and Brazilian law (Permit Number CEUA 013/2014).

2.2. PCR amplification, DNA cloning and sequencing

A single primer was designed, using the *Tc1/Mariner* sequence of *Takifugu rubripes* (Smit, 2002) (GenBank accession number AJ620741.1), to perform PCR amplification on all Parodontidae samples (Table S1a). PCR consisted of: 0.4 μM of the primer, 70 ng of the genomic DNA, 0.16 mM deoxynucleotide triphosphates (dNTPs), 1× polymerase reaction buffer, 2.5 mM $MgCl_2$ and 1 U of Taq DNA polymerase (Biotools). Cycling conditions were performed as follows: 5 min at 94 °C, 35 cycles of 1 min at 94 °C, 1 min at 64 °C and 1 min at 72 °C and a final elongation with 5 min at 72 °C. The obtained DNA fragments were purified with ExoSAP (GE Healthcare), inserted into the plasmid vector from the pTZ57R/T (InsTAclone PCR Cloning Kit – Life Technologies) and cloned into DH5α *Escherichia coli* competent cells. The obtained clones were sequenced using an ABI-PRISM Genetic Analyzer (Carlsbad, CA).

2.3. Tc1/Mariner molecular characterization

All DNA sequences obtained were submitted to the CENSOR online software (Jurka et al., 2005), RepeatMasker (http://www.repeatmasker.org) and Blastn searches at the National Center for Biotechnology Information (NCBI) to check for similarities of the *Tc1/Mariner* element. The sequences were deposited into the GenBank database (Genbank Accession Numbers: Table 1). These sequences were aligned

Table 1
Parodontidae species analyzed, their sampling locality, *Tc1/Mariner* sequence characterized and Genbank Accession Numbers.

Species	Sex	River (state)	Hydrographic basin	GPS localization	Sequence name (*Tc1/Mariner*)	Size (bp)	Genbank accession numbers
A. piracicabae	Male	Piumhi river (MG)	São Francisco	−20°31′55″ and −46°02′42″	1a	1218	KP860108
A. piracicabae	Female	Piumhi river (MG)	São Francisco	−20°31′55″ and −46°02′42″	2a; 3a	1221; 1221	KP860109; KP860110
A. vittatus	Male	Jordão river (PR)	Iguaçu	−25°42′31″ and −51°53′53″	1b; 2b	1221, 1222	KP860111; KP860112
A. vittatus	Female	Jordão river (PR)	Iguaçu	−25°42′31″ and −51°53′53″	3b; 4b	1221; 1221	KP860113; KP860114
Apareiodon sp.	Male	Verde river (PR)	Alto Paraná	−25°04′35″ and −50°04′03″	1c; 2c	1221; 1215	KP860115; KP860116
Apareiodon sp.	Female	Verde river (PR)	Alto Paraná	−25°04′35″ and −50°04′03″	3c; 4c	1221; 1221	KP860117; KP860118
A. vladii	Female	Piquiri river (PR)	Alto Paraná	−25°01′40″ and −52°27′32″	1d; 2d	1218; 1221	KP860119; KP860120
A. hasemani	Male	São Francisco river (MG)	São Francisco	−17°21′17″ and −44°57′18″	1e; 2e	1229; 1221	KP860121; KP860122
A. hasemani	Female	São Francisco river (MG)	São Francisco	−17°21′17″ and −44°57′18″	3e	1221	KP860123
A. affinis	Male	Passa-Cinco river (SP)	Alto Paraná	−22°25′26″ and −47°41′56″	1f; 2f	1218; 1233	KP860124; KP860125
A. affinis	Female	Passa-Cinco river (SP)	Alto Paraná	−22°25′26″ and −47°41′56″	3f	1222	KP860126
P. nasus	Male	Paraguai river (MT)	Paraguai	−15°34′40″ and −56°09′58″	1g; 2g	1221; 1221	KP860127; KP860128
P. nasus	Female	Paraguai river (MT)	Paraguai	−15°34′40″ and −56°09′58″	3g; 4g	1221; 1220	KP860129; KP860130
P. hilarii	Male	Córrego do Porta (MG)	São Francisco	−17°21′17″ and −44°57′15″	1h; 2h	475; 554	KP860131; KP860132
P. hilarii	Female	Córrego do Porta (MG)	São Francisco	−17°21′17″ and −44°57′15″	3h; 4h	538; 232	KP860133; KP860134
P. pongoensis	Female	Taquaralzinho river (MT)	Araguaia	−15°53′28″ and −52°14′56″	1i; 2i	456; 456	KP860135; KP860136

using MUSCLE (Edgar, 2004); Open Reading Frames (ORFs) and amino acid residues were obtained using Geneious 4.8.5 (Drummond et al., 2009). The protein sequences were submitted to Pfam (Finn et al., 2010) to look for protein domains. The MUST software (http://csbl1.bmb.uga.edu/ffzhou/MUST/) was used for the identification of MITEs (Chen et al., 2009).

2.4. Sequence similarity (Plotcon)

The DNA sequence alignments (excluding primer sequence and divided into regions) of the *Tc1/Mariner* were submitted to EMBOSS Plotcon (http://emboss.bioinformatics.nl/cgi-bin/emboss/plotcon), which gives a graphic representation of similarities among sequences in order to characterize their degree and pattern of deterioration. The W*size* (Window Size) value stipulated for this study was 5.

2.5. Genetic distances and statistics

In order to complement the Plotcon results, quantitative analysis for the number of base substitutions per site between sequences (excluding primer sequence and divided into regions) were calculated using MEGA 5.0 (Tamura et al., 2011). Analyses were conducted using the Kimura 2-parameter model and the rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The distance values were submitted to Action statistical program developed under the R platform (http://www.portalaction.com.br/). Non-parametric Kruskal-Wallis and ad-hoc Tukey tests were used, to compare different regions of the *Tc1/Mariner* sequences. Differences were considered significant, with a P value < 0.05.

2.6. Fluorescence in situ hybridization (FISH)

FISH was performed in *Apareiodon* sp., *P. hilarii* e *A. affinis* using the *Tc1/Mariner* sequence probe of *Apareiodon* sp. labeled with digoxigenin11-dUTP (Roche Applied Science, Mannheim, Germany) through nick translation. The FISH procedure followed the protocol described by Pinkel et al. (1986), under high stringency conditions. The signal detection was performed using an anti-digoxigenin antibody conjugated to rhodamine (Roche Applied Science). The chromosomes were counterstained with 0.2 μg/mL of 4′, 6-diamidino-2-phenylindole (DAPI) in the Vectashield mounting medium (Vector, Burlingame, CA) and analyzed using a Zeiss AxioCam MRm Capture Monochrome CCD sensor and 1.4 megapixel resolution.

2.7. Copy number analysis

The internal region of *A. piracicabae Tc1/Mariner* degenerated sequence (obtained in this study) was selected for the construction of qPCR primers (Table S1b), due to the highest similarity with the *Tc1/mariner* of *Takifugu rubripes*. The copy number estimations by qPCR, of all the Parodontidae gDNA (male and female), were performed using the gene dose ratio by a Ct method of relative quantification (Nguyen et al., 2013). Gene dosage ratios (GDR) of the target genes were compared with a single-copy autosomal gene β-actin (Table S1c). If the target gene is autosomal, its copy number ratio compared with β-actin is expected to be 1:1. qPCR was carried out in a Stratagene MxPro3005P thermocycler. The target and reference genes were analyzed simultaneously in duplicates of two independent samples (1 male and 1 female of each species). The cycling conditions were 94 °C for 10 min; 40 cycles of 94 °C for 15 s, and 60 °C for 30s, 72 °C for 30s. Specificity of the PCR products was confirmed by analysis of the dissociation curve.

2.8. RNA extraction and expression analysis

Total RNA of *Apareiodon* sp. gonad tissue (chosen for this analysis because it has heteromorphic ZZ/ZW sex chromosome system) was isolated using TRIZOL (Invitrogen), according to the manufacturer's instructions. The total RNA was submitted to First Strand cDNA Synthesis *Kit* (GE Healthcare), as recommended by the manufacturer's protocol. The cDNAs were amplified in duplicate PCRs using the SYBR Green master mix (Roche) and 8 μM of each primer of *Tc1/Mariner* of the *A. piracicabae* (Table S1b) with a total volume of 20 μL. The β-actin (Table S1c) was used as an internal control. The PCR was performed in the Stratagene MxPro3005P thermocycler with the following cycling conditions: 10 min at 94 °C; 40 cycles of 15 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C; and ended with a dissociation curve. The threshold cycle (Ct) was measured and there was a relative change in the expression level of one specific gene of $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001). A statistical analysis was performed using a Student's *t*-test ($P < 0.05$), with the Action statistical software developed under the R platform (http://www.portalaction.com.br/).

3. Results

3.1. Tc1/Mariner molecular characterization

Twenty three clones containing *Tc1/Mariner* fragments ~1200 bp were amplified in Parodontidae genomes (sequence names in the Table 1), with the exception of *P. hilarii* and *P. pongoensis.* In the case of these last two, the six clones presented fragments of ~600 bp (sequence names in Table 1). CENSOR and RepeatMasker analysis revealed high similarity with non-autonomous *Tc1/Mariner* of *Takifugu rubripes* (~91.5%) and *Danio rerio* (~93%) respectively, with the exception of *P. hilarii* and *P. pongoensis* that showed low similarity values. The lower similarity of these sequences in the two species is due to higher accumulation of mutations when compared to other parodontids (Fig. S1). The domain protein identifications by Pfam showed different levels of *Tc1/Mariner* degradation (Fig. 1).

The TIRs of these elements had 27 bp or less in some species (Fig. S2). A comparison of terminal repeats excluding primer sequences (21 bp) showed that *P. pongoensis* and *P. hilarii* presented the highest deterioration levels in this region (Fig. S2).

The ORF related to the transposase (~346 amino acid residues) contained three domains, the homeodomain-like-domain (HLD), helix-turn-helix (HTH) and DDE endonuclease (Fig. S3A,B), the last one being truncated in all the sequences (Fig. 1). The distance between the second D (179 position) and E (212 position) of DDE endonuclease was 33 amino acid residues (Fig. S3B). Protein domains also exhibited missense mutations (Fig. S3). The 3e sequence had an amino acid substitution in the codon corresponding to the second D (Fig. S3B).

Some sequences revealed more than one ORF due to the presence of premature stop codons, in which one ORF can encode HLD and HTH and the other one can encode the endonuclease DDE. The first ORF of the sequences 2b, 2e and 2f potentially encoded a protein with both intact HLD and HTH domains (Fig. 1). *P. hilarii* and *P. pongoensis Tc1/Mariner* sequences did not present valid ORFs, and when subject to MUST software these sequences were identified as MITEs (similarity values: >0.57; except 3 h sequence) corroborating with results of Fig. S1.

3.2. Tc1/Mariner similarity and genetic distance

Plotcon similarity results demonstrated several *Tc1/Mariner* regions with insertions, substitutions and deletions. There were low similarity from 24 to 26 bp position of 5′-TIR in Fig. S2 that corresponded from the 3 to 5 bp in Fig. S4A (primer sequence was not considered), and the six initials positions of the 3′-TIR (Fig. S2) that corresponding the same sequence of the Fig. S4B (primer sequence was not considered), showing similarity decreasing from the 2 to 4 bp.

In the *Tc1/Mariner* complete sequences graph, there was low similarity ~400–1150 bp regions (highlighted in red in the Fig. S4C). This occurred due to deletions in this region observed in *P. hilarii* and *P.*

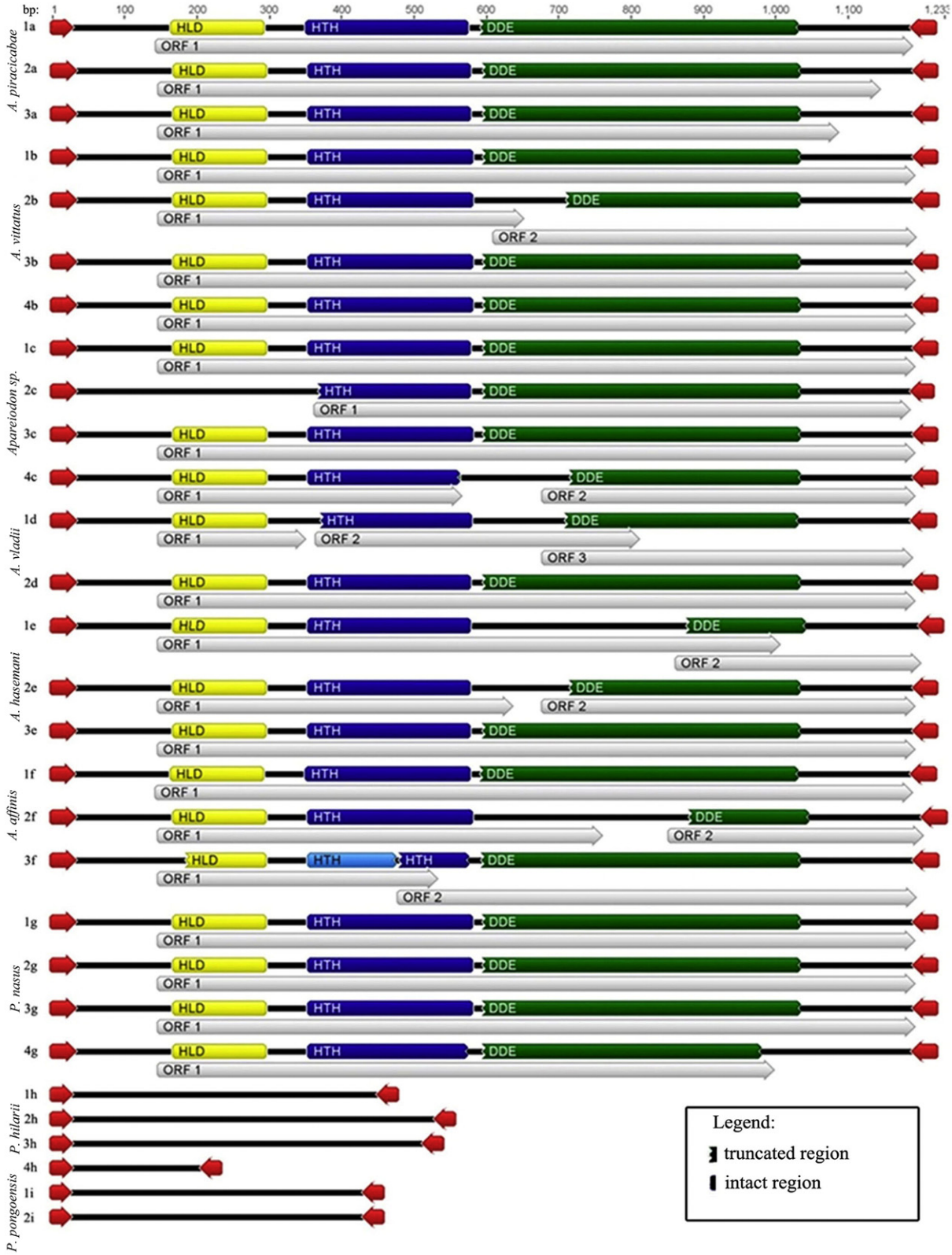

Fig. 1. Molecular characterization of *Tc1/Mariner* sequences of Parodontidae species. The red represent TIRs regions; in yellow *homeodomain-like-domain* (HLD) domain; in blue *helix-turn-helix* (HTH) domain; in green endonuclease DDE (DDE) domain. Figure performed using Geneious 4.8.5 software (Drummond et al., 2009).

pongoensis sequences (MITEs sequences). The ORF alignment presented punctual variations (highlighted in red in the Fig. S4D) and the domain that presented regions with less similarity value was endonuclease DDE (highlighted in red in the Fig. S4D). Sequences with high degradation levels presented greater heterogeneity than the less degraded sequences (Fig. S4 E,F).

Kimura 2-parameter test demonstrated that *Tc1/Mariner* 5′-TIR and 3′-TIR had less genetic distance than internal region (significant by

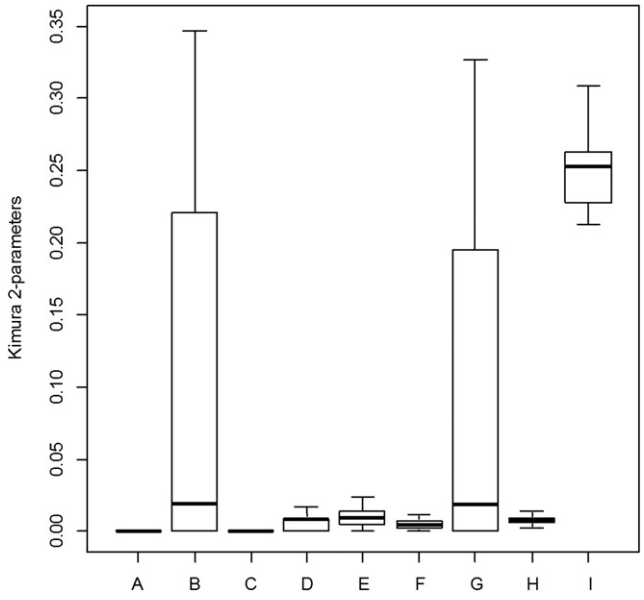

Fig. 2. Number of base substitutions per site (genetic distance) between Parodontidae *Tc1/Mariner* sequences. For this analysis *Tc1/Mariners* were divided into seven regions in order to obtain a detailed analysis of each sequence region. A – 5'TIR; B – Internal region without TIRs; C – 3'TIR ; D – HLD domain; E – HTH domain; F – endonuclease DDE domain; G – all complete sequences; H – complete sequences with low degradation, not including *P. hilarii* and *P. pongoensis* sequences; I – complete sequences with high degradation, only *P. hilarii* e *P. pongoensis* sequences were included (Kruskal-Wallis test Ad hoc Tukey test: significant comparisons with $p < 0.05$ value for H-I. The B-A; B-C; D-E; E-F presented significant differences $p < 0.05$ only for the Kruskal-Wallis).

Kruskal-Wallis test: $p < 0.05$), however, the small sequence length of terminal regions were analyzed (Fig. 2A,B,C). The range of the divergence between the sequences corresponding to the HTH domain was greater than the HLD domain (only significant by Kruskal-Wallis test: $p < 0.05$) (Fig. 2D,E). Significant differences ($p < 0.05$) for the same test were obtained between the HTH and the endonuclease DDE domains (Fig. 2E,F).

Tc1/Mariner complete sequence analysis presented a broad range distance of Kimura 2-parameter (Fig. 2G). However, point comparisons between several complete sequences of different species, pairwise analysis, had distances equal to 0 (highlighted in red in the Table S2). When *P. hilarii* and *P. pongoensis* sequences (with a high level of degradation) were removed from the sample, the difference between the distances decreased considerably with few variations around zero (Fig. 2H). *P. hilarii* and *P. pongoensis* sequences presented higher distance values (Fig. 2I) than low levels of degradation sequences (Fig. 2H), with significant differences (Kruskal-Wallis test: $p < 0.05$; Tukey test: $p = 7,24451E-06$).

3.3. Tc1-Mariner physical chromosome mapping, copy number and expression analysis

The pattern of the *Tc1/Mariner* distribution in *Apareiodon* sp. and *P. hilarii* (ZZ/ZW heteromorphic sex chromosome system) and *A. affinis* (ZZ/ZW_1W_2 heteromorphic sex chromosome system) was dispersed in euchromatic regions, with minor terminal accumulations (Fig. 3). Furthermore, interstitial *Tc1/Mariner* sites on W chromosome of *Apareiodon* sp. (Fig. 3A) and *P.hilarii* (Fig. 3B) and terminal *Tc1/Mariner* sites on the Z chromosome of *A. affinis* (Fig. 3C) were mapped. GDR demonstrated that *P. hilarii* (male: 1.07 and female: 0.55) and *P. pongoensis* (female: 21.48) had a smaller number of copies for the *Tc1/Mariner* with internal region than other species, which ranged from 4513.403355 to 39,649.08632 (Table S3). Expression analysis demonstrated transcriptional activity in gonadal tissue of *Apareiodon* sp. (Fig. S5, Table S4) without significant differences between the sexes ($p = 0.27$).

4. Discussion

The largest group of eukaryotic Class II transposons is composed of members related to the mariner and the Tc1 families (Benjamin et al., 2007). However, these sequences are usually inactivated, including those in fish genomes (Muñoz-López and García-Pérez, 2010; Izsvfik et al., 1995). The *Tc1/Mariner* Parodontidae element was characterized by the presence of the HLD, HTH and DDE domains in the ORF corresponding to transposase. However, at least one domain is truncated in all analyzed sequences, indicating non-autonomous transposons. HLD and HTH domains are non-truncated in most of the sequences analyzed, due to the low mutation rates or synonym mutations. The HTH and HLD domains are necessary for recognition and binding in the TIRs, and the DDE signature-sequence is the catalytic domain responsible for site-specific cleavage and junction in the transposition processes, essential for the transposase's activity (Craig, 1995; Plasterk et al., 1999).

Although only Parodontidae *Tc1/Mariner* non-autonomous sequences have been identified, there is transcriptional activity in the gonads of these degenerated sequences. This may indicate truncated protein synthesis without utility in genomes or new protein synthesis with another function because the presence of ORFs corresponding to proteins with ligation domains derived from *Tc1/Mariner* of Parodontidae were detected. Studies on molecular co-option of transposases-derived (originated from mutated transposon) showed evidence that these are more likely to create the new DNA-binding proteins, which are involved in the cellular recombination cycle control and

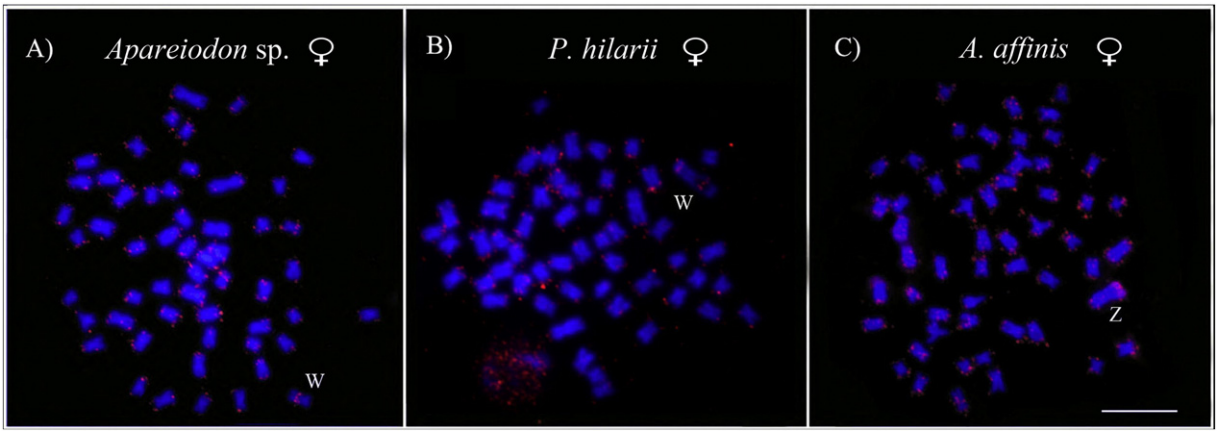

Fig. 3. Metaphases of females of *Apareiodon* sp. (A) *P. hilarii* (B) and *A. affinis* (C) subjected to FISH using de *Tc1/Mariner* probe (red) obtained from a *Apareiodon* sp. clone (3c sequence). Scale bar = 5 μm.

other functions related to chromatin modification (Feschotte and Pritham, 2007; Feschotte, 2008). In fact, DNA binding proteins Pax have highly conserved function across the animal kingdom, including eye development and cephalization. However, they acquired binding domain from an ancestral *Tc1/Mariner's* transposase after the emergence of metazoans, about one billion years ago (Breitling and Gerber, 2000). The *Tc1/Mariner* HLH domain identified in this study is commonly found in transcription factors (Gehring, 1992; Gehring, 1993). Furthermore, some sequences presented one ORF with intact HLD and HTH that may be the source for new protein origins. Studies also indicate that RNAs produced by TE can be recruited to generate siRNAs (Feschotte, 2008). In other organism complete transcription of the *Tc1/Mariner* elements or MITEs can form hairpins due to the complementarity of the TIRs and form a region of dsRNA, which can be processed by the enzymatic machinery to form short siRNAs that can, in turn, silence these elements (Piriyapongsa and Jordan, 2008). Moreover, the miRNA originating from transposons are evolutionarily new regulators involved in the control of endogenous genes (Lohe and Hartl, 1996).

The similarity analyses demonstrate that Parodontidae *Tc1/Mariner* sequences incorporated different errors and *P. pongoensis* and *P. hilarii* sequences presented the highest rates of deterioration. These sequences have a wide diversity, due to the fact that Class II elements tend to incorporate internal deletions and nucleotide substitutions that are related to the formation of MITEs (Fernández-Medina et al., 2012). This deterioration process is expected in the genome during TE deterioration, both in the structure and in the nucleotide composition. Different classes and orders of TEs can incorporate errors in different manners, according to their transposition mechanism (Fernández-Medina et al., 2012). The Parodontidae *Tc1/Mariner* element presented a large rage of Kimura 2-parameter distance when all the sequences were analyzed (Fig. 2B,H) and there was high deletion in the region between ~400–1150 bp (Fig. S4), which corresponds to the HTH and DDE domains when comparing all the sequences. Moreover, when comparing the sequences present in ORFs, the DDE contains regions with larger difference than the HTH and HLD domains (Fig. S4D). In addition, Kimura 2-parameter model comparisons indicated a high range of distance when comparing HTH against HLD and DDE domains (Fig. 2 D,E,F), but these were only significant for the Kruskall-Wallis test. The MITEs sequences presented higher distance values in comparison with the sequences that have the ORF corresponding to transposase (Fig. 2 I,J). Thus, the *Tc1/Mariner* results revealed degenerated sequences in Parodontidae species. The loss of *Tc1/Mariner's* identity in Parodontidae and the occurrence of some sequences in pairwise analysis with a genetic distance of 0, present in different species, indicated sequence similarity even though copies of TEs were not completely intact (Table S2). This is probably due to the fact that they are found in recently diverged species (Bellafronte et al., 2013) or the *Tc1/Mariner* element had been recruited for new functions in the Parodontidae genome, which is reinforced by euchromatic dispersed location.

The dynamics of TE expansion across genomes varies depending on the species and element type considered (Montiel et al., 2012). Physical mapping of the *Tc1/Mariner* Parodontidae element showed a dispersed distribution, accumulation in the female specific region (FSR) of the W chromosomes and accumulation on the terminal Z chromosome of *A. affinis*. The GDR value was ~2-fold higher in *P. hilarii* male compared to the female because the primer of qPCR analyses binds in the internal region and these sequences presented high deterioration in this region (MITEs sequences). Then, W of *P. hilarii* probably has highly degraded sequences for this TE. Ancient TEs tend to accumulate in euchromatic regions associated with genes, while new and more deleterious elements tend to colonize the heterochromatin (Santiago et al., 2002). In this respect, depending on the location where TEs are inserted and the deterioration pattern, they can be beneficial to the host genome.

The accumulation of *Tc1/Mariner* on the Parodontidae W chromosomes could have been caused by a lack of recombination between sex chromosomes. Other repetitive sequences, such as TE Helitron and sat1Wp (Schemberger et al., 2014) and (GATA)n (Ziemniczak et al., 2014), have reported chromosomes' sexual evolution in Parodontidae. In this context, early stages of sex chromosome evolution show a lack of recombination between proto-sex chromosomes. This may lead to genetic degeneration of the heterozygous sex (Charlesworth et al., 2005). Hence, there is loss in the adaptation level of genes on the Y or W chromosomes and the population genetic processes cause a build-up of both satellite-DNA sequences and transposable elements in genomic regions, where crossing-over is reduced or absent. Furthermore, TEs can influence the course of sex determination cascade by donating new regulatory elements. For example, in the medaka fish a duplicate of the Dmrt1 autosomal transcription factor gene, DMY (present on chromosome Y), has become the sex master regulator. This has been established by the insertion of a TE into its regulatory region (Herpin et al., 2010).

Thus, the *Tc1/Mariner* sequences are an interesting issue to be investigated in Parodontidae genomes. The characterization and location of TEs open new perspective for a better understanding of the composition, diversification and role origin of sex chromosomes in Parodontidae. The genomes of eukaryotes are complex with respect to regulatory networks. Therefore, TEs are fundamental in understanding TFBSs (transcription factor binding sites) and protein origin, evolution and activity, since they are entities of great abundance and with wide genomic diversity.

Conflict of interest

The authors declare that they have no conflict of interest.

The following are the supplementary data related to this article.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2016.08.034.

Acknowledgments

The authors are grateful to the Ministério do Meio Ambiente/Instituto Brasileiro do Meio ambiente e dos Recursos Naturais Renováveis (IBAMA/MMA/SISBIO No. 15117). This study was financed by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq No. 303696/2013-3; 473212/2013-7), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – PPG/Genética 40001016006), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP No. 2012/15258-0; 2013/04533-3), Secretaria de Ciência e Tecnologia do Estado do Paraná (SETI No. 378/2012) and Fundação Araucária de Apoio ao Desenvolvimento Científico do Estado do Paraná (Fundação Araucária No. 22.843).

References

Bellafronte, E., Mariguela, T.C., Pereira, L.H.G., Oliveira, C., Moreira-Filho, O., 2013. DNA barcode of Parodontidae species from the La Plata river basin-applying new data to clarify taxonomic problems. Neotrop. Ichthyol. 11 (3), 497–506.

Bellafronte, E., Schemberger, M.O., Moreira-Filho, O., Almeida, M.C., Artoni, R.F., Margarido, V.P., Vicari, M.R., 2011. Chromosomal markers in Parodontidae: an analysis of new and reviewed data with phylogenetic inferences. Rev. Fish Biol. Fish. 12, 559–570.

Benjamin, B., Yves, B., Corinne, A.G., 2007. Assembly of the Tc1 and mariner transposition initiation complexes depends on the origins of their transposase DNA binding domains. Genetica 130 (2), 105–120.

Bertollo, L.A.C., Takahashi, C.S., Moreira-Filho, O., 1978. Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, Erythrinidae). Rev. Bras. Genet. 1 (2), 103–120.

Breitling, R., Gerber, J.K., 2000. Origin of the paired domain. Dev. Genes Evol. 210, 644–650.

Charlesworth, D., Charlesworth, B., Marais, G., 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95 (2), 118–128.

Chen, Y., Zhou, F., Li, G., Xu, Y., 2009. MUST: a system for identification of miniature inverted-repeat transposable elements and applications to *Anabaena variabilis* and *Haloquadratum walsbyi*. Gene 436, 1–7.

Craig, N.L., 1995. Unity in transposition reactions. Science 270 (5234), 253.

de Souza, F.S.J., Franchini, L.F., Rubinstein, M., 2013. Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong? Mol. Biol. Evol. 30 (6), 1239–1251.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., et al., 2009. Geneious v4.8.5. (Available from) http://www.geneious.com.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Fernández-Medina, R.D., Ribeiro, J.M.C., Carareto, C.M.A., Velasque, L., Struchiner, C.J., 2012. Losing identity: structural diversity of transposable elements belonging to different classes in the genome of *Anopheles gambiae*. BMC Genomics 13, 272.
Feschotte, C., 2008. The contribution of transposable elements to the evolution of regulatory networks. Nat. Rev. Genet. 9 (5), 397–405.
Feschotte, C., Pritham, E.J., 2007. DNA transposons and the evolution of eukaryotic genomes. Annu. Rev. Genet. 41, 331–368.
Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., et al., 2010. The Pfam protein families database. Nucleic Acids Res. 38, D211–D222.
Gehring, W.J., 1992. The homeobox in perspective. Trends Biochem. Sci. 17 (8), 277–280.
Gehring, W.J., 1993. Exploring the homeobox. Gene 135 (1–2), 215–221.
Herpin, A., Braasch, I., Kraeussling, M., Schmidt, C., Thoma, E.C., Nakamura, S., Tanaka, M., Schart, M., 2010. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet. 6, 2.
Izsvfik, Z., Ivics, Z., Hackett, P.B., 1995. Characterization of a Tcl-like transposable element in zebrafish (*Danio rerio*). Mol. Gen. Genet. 247, 312–322.
Jiang, N., Bao, Z., Zhang, X., Eddy, S.R., Wessler, S.R., 2004. Pack-MULE transposable elements mediate gene evolution in plants. Nature 431, 569–573.
Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., Walichiewicz, J., 2005. Repbase update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–467.
Kapusta, A., Kronenberg, Z., Lynch, V.J., Zhuo, X., Ramsay, L., Bourque, G., Yandell, M., Feschotte, C., 2013. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. PLoS Genet. 9, 4.
Kidwell, M.G., 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica 115, 49–63.
Kidwell, M.G., Lisch, D.R., 2001. Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution 55, 1–24.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25 (4), 402–408.
Lohe, A.R., Hartl, D.L., 1996. Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Mol. Biol. Evol. 13, 549–555.
Long, M., Betrán, E., Thornton, K., Wang, W., 2003. The origin of new genes: glimpses from the young and old. Nat. Rev. Genet. 4 (11), 865–875.
Lowe, C.B., Haussler, D., 2012. 29 mammalian genomes reveal novel exaptations of mobile elements for likely regulatory functions in the human genome. PLoS One 7 (8), e43128.
Marino-Ramirez, L., Lewis, K.C., Landsman, D., Jordan, I.K., 2005. Transposable elements donate lineage-specific regulatory sequences to host genomes. Cytogenet. Genome. Res. 110, 333–334.
Montiel, E.E., Cabrero, J., Camacho, J.P.M., López-León, M.D., 2012. Gypsy, RTE and mariner transposable elements populate Eyprepocnemis plorans genome. Genetica 140 (7–9), 365–374.
Muñoz-López, M., García-Pérez, J.L., 2010. DNA transposons: nature and applications in genomics. Curr. Genomics 11 (2), 115–128.
Murray, M.G., Thompson, W.F., 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8, 4321–4326.
Nguyen, P., Sýkorová, M., Šíchová, J., Kůta, V., Dalíková, M., Frydrychová, R.Č., Neven, L.G., Sahara, K., Marec, F., 2013. Neo-sex chromosomes and adaptive potential in tortricid pests. Proc. Natl. Acad. Sci. U. S. A. 110 (17), 6931–6936.
Pinkel, D., Straume, T., Gray, J.W., 1986. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc. Natl. Acad. Sci. U. S. A. 83, 2934–2938.
Piriyapongsa, J., Jordan, I.K., 2008. Dual coding of siRNAs and miRNAs by plant transposable elements. RNA 14, 814–821.
Plasterk, R.H.A., Izsvák, Z., Ivics, Z., 1999. Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. 15 (8), 326–332.
Santiago, N., Herráiz, C., Goñi, J.R., Messeguer, X., Casacuberta, J.M., 2002. Genome-wide analysis of the emigrant family of MITEs of Arabidopsis Thaliana. Mol. Biol. Evol. 19 (12), 2285–2293.
Schemberger, M.O., Bellafronte, E., Nogaroto, V., Almeida, M.C., Schühli, G.S., Artoni, R.F., Vicari, M.R., 2011. Differentiation of repetitive DNA sites and sex chromosome systems reveal closely related group in Parodontidae (Actinopterygii: Characiformes). Genetica 139 (11 – 12), 1499–1508.
Schemberger, M.O., Oliveira, J.I.N., Nogaroto, V., Almeida, M.C., Artoni, R.F., Cestari, M.M., Moreira-Filho, O., Vicari, M.R., 2014. Construction and characterization of a repetitive DNA library in Parodontidae (Actinopterygii: Characiformes): a genomic and evolutionary approach to the degeneration of the W sex chromosome. Zebrafish 11 (6), 518–527.
Sinzelle, L., Izsvák, Z., Ivics, Z., 2009. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. Cell. Mol. Life Sci. 66 (6), 1073–1093.
Slotkin, R.K., Martienssen, R., 2007. Transposable elements and the epigenetic regulation of the genome. Nature 8, 272–285.
Smit, A.F.A., 2002. Initial survey of interspersed repeats in *Takifugu rubripes*. Repbase Reports 2 (1), 54.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics gnalysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28 (10), 2731–2739.
Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., et al., 2007. A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. 8, 973–982.
Ziemniczak, K., Traldi, J.B., Nogaroto, V., Almeida, M.C., Artoni, R.F., Moreira-Filho, O., Vicari, M.R., 2014. In situ localization of (GATA)n and (TTAGGG)n repeated DNAs and W sex chromosome differentiation in Parodontidae (Actinopterygii: Characiformes). Cytogenet. Genome Res. 144 (4), 325–332.