ENVIRONMENTAL BIOTECHNOLOGY 6 (1) 2010, 24-36

Properties of mobile elements of genomes and their application in biotechnology

Anita Poćwierz-Kotus, Roman Wenne

Laboratory of Genetics and Marine Biotechnology, Institute of Oceanology, Polish Academy of Sciences,
Powstańców Warszawy 55, 81-712 Sopot, Poland
Corresponding author: Anita Poćwierz-Kotus, E-mail: anita@iopan.gda.pl

ABBREVIATIONS

DTR	direct terminal repeat
HT	horizontal transfer
IS	insertion sequence
ITR	inverted terminal repeat
LINE	long interspersed nuclear element
LTR	long terminal repeat
ORF	open reading frame
SINE	short interspersed nuclear element
TE	transposable element; transposon
Tn	prokaryotic transposon

1. CLASSIFICATION OF TRANSPOSABLE ELEMENTS (TES)

Transposons (transposable elements, TEs) are DNA fragments capable of transposition, i.e., changing their location in the genome. They can spread inside single genomes, between genomes within a population and between species in a process called horizontal transfer. The range of their occurrence, with few exceptions (Gardner et al. 2002), includes all known species.

Transposons were first discovered in maize genome by Barbara McClintock (McClintock 1950, 1953, 1956). McClintock established that "controlling elements" are responsible for pigmentation patterns of corn seeds. Studies on the molecular basics of gene expression in lactose and galactose operons in *Escherichia coli* led to the discovery of insertion sequences (IS) which are the transposons simplest in structure (Hirsch et al. 1972). The enormous diversity of structural features and transposition mechanisms of TEs required a classification system.

1.1. Prokaryotic insertion sequences (ISs)

Insertion sequences (ISs) are short DNA fragments (about 1000 bp) that are capable of translocation in a genome, often together with neighbouring genes. Their presence has been noted in most studied genomes of bacteria and archaea. Because of their high frequency in genomes, it is thought, that ISs play an important evolutionary role by promoting gene inactivation and genome plasticity (Touchon and Rocha 2007).

Most bacterial insertion sequences consist of two characteristic elements. Inverted terminal repeats (ITR) which are sequences 10 to 40 bp long that are located on both ends of the element but oriented in opposite directions. The

second is a region that encodes an enzyme taking an active part in the transposition, a transposase.

In the structure of the transposase enzyme, there is a conserved motif, which consists of three amino acids: DDE or DDD, critical to the process of transposition (Siguier et al. 2006). IS elements are capable of moving into different acceptor sites in the chromosome. But in the genomes of some bacteria, hot spots have been found, where insertion of the IS is more likely (Tobes and Pareja 2006). Insertion sequences are classified into 20 families (Siguier et al. 2006) based on the similarity of transposases, ITR elements and conserved catalytic sites.

During the transposition process, the acceptor site is duplicated outside of each of the ITRs. A sequence 4 to 12 bp long is copied, and the sequence which is to be transposed is placed between these duplicated fragments. The transposase binds to inverted terminal repeats (ITR) on the ends of insertion sequences and cleaves the DNA, as a result sticky ends arise both in acceptor and donor sites (Filee et al. 2007).

1.2. Prokaryotic transposons (Tn)

Tn elements, in contrast to insertion sequences, contain a gene which codes transposase and genes which encode additional functions: resistance to antibiotics, resistance to heavy metals, or conferring specific metabolic abilities, e.g. gene encoding β-galactosidase (Brown and Evans 1991; L'Abee-Lund and Sorum 2000). The length of prokaryotic transposons may vary significantly from about 2000 bp (Tn9 transposon) to 20000 bp (Tn4 transposon).

Many transposons consist of two insertion sequences responsible for transposition, between which there is a DNA fragment responsible for a phenotype feature, but this is not needed in the process of transposition. Some insertion sequences exist as independent elements. Others, despite having the ability of independent transposition, are always a part of compound transposons (Top and Springael 2003). Transposing functionality may be carried out by only one of the IS flanking a transposon (e.g. Tn5 transposon), or by both IS (Tn9). Where both ISs are functional, the combined transposition activities of these ISs improve the effectiveness of transposition of the whole transposon (Tan 1999).

1.3. Class I of the eukaryotic transposons – retrotransposons

In eukaryotic organisms, transposable elements are classified by structure and mechanism of transposition into two classes (Finnegan 1989, 1992a). Class I includes retrotransposons that translocate by inverted transcription through RNA intermediate. Class II transposons, move directly from DNA to DNA and use a mechanism called 'cut-and-paste' in the process of transposition.

Retrotransposons are amplified during transposition and therefore remain in their original place in the genome - only their copy is inserted into a new spot. This mechanism generates enormous potential to increase the number of their copies (Kidwell 2002). Retrotransposons are transcribed and then, because they encode reverse transcriptase, the genetic information is transcribed from RNA to DNA.

Retrotransposons constitute a large fraction of genomes, and great diversity can be found in fish and other vertebrate genomes. The smallest of all the known vertebrate genomes (fish *Takifugu rubripes* and *Tetraodon nigroviridis*) consist of only about 400 million bp. These genomes include fewer copies but, surprisingly, more clades of retrotransposons, than found in much bigger (about $3.2 \cdot 10^9$ bp) mammalian genomes (Sharov 2006) such as human or mouse (Fisher et al. 2004; Volff et al. 2003).

In the retrotransposons class there are two subclasses. The division criteria are structural features: the first subclass includes elements with long terminal repeats, called LTR-retrotransposons, the second class, called non-LTR retrotransposons, includes elements, which have no LTRs (Dalle Nogare et al. 2002). Some scientists include Penelope-like elements as a third subclass in the retrotransposon class (Volff et al. 2001).

1.3.1. LTR-retrotransposons

LTR-retrotransposons are widely represented in many plant and animal genomes, including chordates, as with the Cigr-1 element in sea squirt *Ciona intestinalis* (Simmen and Bird 2000), the Gmr1 element in cod *Gadus morhua* and the Abr1 in sturgeon *Acipenser baeri* (Butler et al. 2001).

The LTR-retrotransposons subclass includes four distinct groups: Ty1-copia, BEL, DIRS and Ty3-gypsy (Kidwell 2002). Those groups were distinguished on the basis of comparison of amino-acid sequences of the reverse transcriptase. Differences in sequence result in differences in the arrangement of the domains of the protein.

LTR-retrotransposons are commonly found in plant genomes (Park et al. 2007). A particularly wide range of distribution characterizes elements from the subclass Ty1-copia. Their presence was notified in algae, bryophytes, gymnosperm and angiosperms (Alix et al. 2005). The number of their copies is different depending on species – e.g., from several hundreds in *Arabidopsis thaliana* (Terol et al. 2001) and *Oryza sativa* (Vincient and Schulman 2002) up to 50-80% of the whole genome of *Zea mays* (Sanmiguel and Bennetzen 1998).

The length of elements from the Ty3-gypsy subclass in *Drosophila* genome varies from 5 to 14 kbp. In *Dictyostelium* the element DIRS-1 from subclass DIRS is about 5 kbp long, whereas Cer7 element from BEL group in *Caenorhabditis elegans* varies from 3 to 20 kbp.

Horizontal transfer is not well documented for LTR-retrotransposons. But, results of phylogenetic analysis of sequences of elements from Ty1-copia group from Phelipanche and Orobanche, may show that transfer occurred between genomes of those groups of plants (Park et al. 2007).

The LTR-retrotransposons show similarity in structure to retroviruses which encode similar proteins. In the sequence of both retroviruses and LTR-retrotransposons, there is a gene *gag*, which in retrotransposons encodes a protein similar to the capsid proteins of retroviruses, and a *pol* gene encoding a polyprotein. Its activity covers functions of reverse transcriptase, ribonuclease, integrase and protease, thereby enabling the process of reverse transciption and integration of the new copy.

What differentiates retroviruses and LTR-retrotransposons is the presence of the env gene in retroviruses. This encodes capsid proteins, which enable the translocation of the virus from cell to cell. LTR-retrotransposons do not have that gene or have only fragments of it (Kazazian 2004). The length of LTR-retrotransposons differs widely from 3 to 20 kbp.

1.3.2. Non-LTR retrotransposons

The subclass of non-LTR retrotransposons includes LINE (long interspersed nuclear elements) and SINE (short interspersed nuclear elements) elements (Weiner 2002). The characteristic feature for elements included in that subclass, beside lack of repeats, is a polyadenylation sequence at the 3' end (Kazazian 2004).

The lengths of LINE and SINE elements vary from several kbp (LINES) to less than 500 bp (SINES). LINES are autonomous retroelements, whereas the functioning of SINES depends on proteins encoded by LINES.

The occurrence of LINE elements is common and their distribution in genomes may be significant - the number of their copies varies from several thousands up to a million per haploid genome. It is estimated that the haploid human genome contains about 850 thousand copies of LINES, about 21% of the whole genome (Sugano et al. 2006). The abundance of LINE elements does not indicate a high activity in the genome. It is estimated that less than 0.2% of spontaneous mutations in human genome is caused by the insertion of LINE elements. Moreover, although all mammals have a similar level of insertions, the activity of individual groups of retrotransposons varies significantly, i.e. in mice about 10% of spontaneous mutations are caused by LTR retrotransposons (Eickbusch and Furano 2002).

Elements from this group are also common in fish, i.e. Rex1 element is present in the genomes of many teleost fish species (Volff et al. 2000). In *Danio rerio*, over 30 separate lineages of LINE elements were found, which is about 10% of its genome (Furano et al. 2004).

No LINES or Ty3-gypsy elements were found in genomes of Bdelloidae, which belongs to the phylum of rotifers (Rotifera). This lack of LINE elements, in contrast to their ubiquity in other examined species of both vertebrates and invertebrates, is connected with the agamous reproduction of Rotifera. Retrotransposons, as nuclear parasites, are transferred vertically in the process of sexual reproduction (Arkhipova and Meselson 2002).

LINE elements usually have two open reading frames (ORFs). One encodes a protein that binds nucleic acids, the second encodes a protein which has endonuclease and reverse transcriptase activities (Kazazian 2004).

During the process of transposition, proteins encoded by LINES form a ribonucleoprotein complex that moves to the target site in a host chromosome. Endonuclease cleaves a genomic DNA strand and then LINE-encoded reverse transcriptase translates genetic information from RNA to DNA, using as a starter the hydroxylic group formed as a consequence of cleavage. Next the newly synthesized DNA integrates in the genomic DNA of the host.

The mechanisms of integration are not yet well known (Sugano et al. 2006). The insertion of some non-LTR retrotransposons takes place in specific sites in the genome. For instance, integration of R1 and R2 elements in *Drosophila melanogaster* and *Bombyx mori* takes place in rRNA gene (Kazazian 2004). It is known, that heT-A and TART elements are specifically located on the ends of chromosome in *D. melanogaster*, maintaining the stability of telomeres (Pardue et al. 2005). In mammals however, LINE elements are scattered throughout many sites in the genome. This variety results from the tendency of endonuclease to cleave a short consensus sequence 5'-TTTT/A-3' between the last T and A (Kazazian 2004).

It is assumed, that LINE retrotransposons are transmitted vertically, although there is a hypothesis of the horizontal transfer of Rex1 element between *Anguilla* (eel) genus and evolutionally distant species *Battrachocottus baikalensis*. This hypothesis is based on the discovery that the percentage of synonymous substitutions between genes which code a reverse transcriptase of Rex1 element in *Anguilla* and in *Battrachocottus baikalensis* is lower than between most other genes found in these species. But that observation cannot be an incontrovertible proof for horizontal transfer in these species (Volff et al. 2000).

SINE elements show similarity to LINES, but are shorter (on average from 100 to 500 bp), simpler in structure and almost totally dependent on the function of reverse transcriptase and endonuclease encoded by LINES (Weiner 2002). They are numerously represented; in the human genome there are about 1,500,000 copies, which is approximately 13% of the whole genome (Kawagoe-Takaki et al. 2006). Because they make up a large part of the size of the genome, they significantly influence its complexity and evolution (Kazazian 2004).

SINE elements do not encode enzymes necessary in the process of transposition. Typical SINE elements consist of two parts: a region related, and a region unrelated, to tRNA. For instance, the Sma1 SINE element is found in the genome of salmon (Kawagoe-Takaki et al. 2006). Exceptions to this are Alu elements, which occur in human genome and B1 found in mice, both of which contain a region related to 7SL RNA instead of tRNA (Bowen and Jordan 2002; Okada 1991).

Often the SINE elements neighbouring the LINE elements, share a conserved poly(A) 3'-end (Kajikawa et al. 2005). This specific structure enables the mobilization of the Alu element by the LINE element. The reverse transcriptase recognizes poly(A) region and binds to the 3'-end. A similar mechanism of mobilization was described in the case of SINE (UnaSINE1) and LINE (UnaL1) elements in eel (*Anguilla*) (Kajikawa and Okada 2002).

Alu elements are the most active and dominant type of SINES in the human genome. Polymorphism of Alu elements, shown by their presence or absence, contributes to differences between human populations and therefore is a source of useful markers for studies on population and evolutionary genetics. It is estimated that about 5000 Alu elements are specific only for humans. About 25% of them were inserted so recently, that they are polymorphic for different populations, families or even individuals (Wang et al. 2006).

Alu elements can contribute to the emergence of minisatellites, which are repetitive, short segments of DNA (from 10 to 100 bp). In humans, for example, minisatellites correspond to 44 bp of consensus Alu sequence (Jurka and Gentles 2006).

1.3.3. Penelope elements

Classification of retrotransposons of the Penelope group is not easy. They have features characteristic for both LTR and non-LTR retrotransposons. A Penelope element found in *Drosophila virilis* was first described by Evgenev et al. (1997). It contains simple LTR sequences, which suggests that the element belongs to the LTR-retrotransposons class. However the Penelope elements lack several genes specific for that class: *gag*, protease and RNase H. Therefore Penelope elements were classified as non-LTR retrotransposons. Heterogeneity in the structure of Penelope elements may indicate the diversity of their mechanisms of retrotransposition. The evolutionary position of this group remains unsolved (Eickbush and Jamburuthugoda 2008). Elements related to Penelope described in *D. virilis* – Poseidon and Neptune – were found in many teleosts including *Oryzias latipes* and *Tetraodon nigroviridis*, some of them are found in numerous copies (Volff et al. 2001).

1.4. Class II of eukaryotic transposons

Transposons from class II conduct transposition via a DNA intermediate. The mechanism involves cutting out the element and reintegrating it in a new site in the genome (cut-and-paste) (Miskey et al. 2005). An exception is the family of Helitrons transposons. This family uses, in the process of transposition, the rolling circle model (RC), similar to the replication process in bacteriophages (Kidwell 2002). Most representatives of class II contain in their structure characteristic ITR elements (Figure 1).

The characteristic, conserved amino acid motif DDE or DDD is present in all known groups of transposing elements, including bacterial insertion sequences, this may suggest their common origin (Capy et al. 1997). The members of class II of DNA transposons are very diverse, and include many superfamilies such as: Tc1/mariner, hAT (hobo, Activator, Tam-3), P elements, MuDR elements (Mu, MULEs), PIF-Harbinger and others.

The length of class II elements varies widely. The longest is MuDR family (400 - 20,000 bp) and Helitrons (5,500 - 17,500 bp). Tc1-mariner elements are 1,000 - 3,500 bp long (Kidwell 2002).

A characteristic feature, common for class I and class II elements, is that most of the elements are inactive because of mutations, and some of them are only short fragments of full-size elements. Class II transposons usually occur in a smaller number of copies compared to retrotransposons, which may be explained by the replicative model of transposition of the latter (Kidwell 2002).

The Tc1/mariner superfamily of transposons is one of the most diverse and widespread of class II elements. Mariner-like Elements (MLEs) and Tc1-like Elements (TLEs) are distinct families, for which the main division criterion is the composition of the cation-binding domain: D,D,E/D. The MLE and TLE elements are divided into subfamilies, based on similarity of sequences. Both MLEs and TLEs have a very wide range of hosts (Hartl et al. 1997). They frequently occur in insect genomes and other invertebrates. Their presence was also noted in vertebrates, including human (Auge-Gouillou et al. 1995; Robertson and Zumpano 1997) and plants (Feschotte et al. 2003).

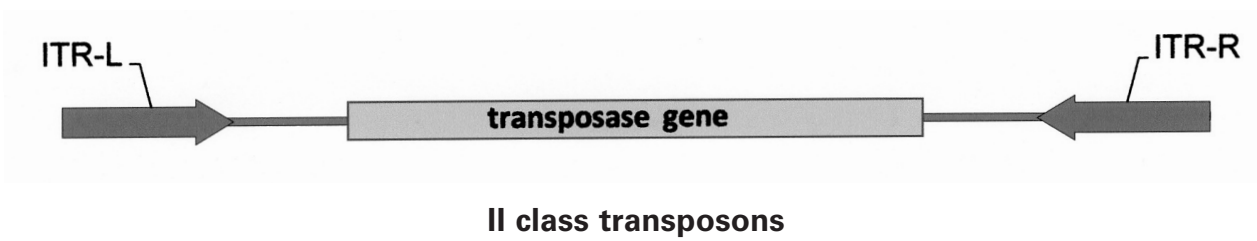

Figure 1. Scheme of structure of class II transposons. Grey rectangle represents transposase gene. On both ends there are inverted terminal repeats ITR, ITR-L – left one, ITR-R – right one.

1.4.1. Mariner elements

The family of mariner transposons contains DNA fragments 1,200 - 3,500 bp long. They include a single gene, without introns, encoding a transposase composed of about 350 amino acids. The gene is flanked by two ITRs, each 19-40 bp long (Casse et al. 2006). Most of the mariner elements are inactive because of numerous mutations: presence of stop codons, deletions, insertions or substitutions (Lohe et al. 1997).

Currently the mariner family is divided into five subfamilies: cecropia, elegans/briggsae, irritans, mauritiana and mellifera/capitata (Bigot et al. 2005). Discovery of the Tvmar1 element, related to MLEs, in the genome of protozoon *Trichomonas vaginalis*, suggests that a sixth family exists (Silva et al. 2004).

The mariner element (Mos1) was first identified and described in *Drosophila mauritiana* (Jacobson and Hartl 1985; Jacobson et al. 1986) as an insertion in white eye gene. Later works confirmed the presence of related elements in insect genomes such as: *D. melanogaster* (Garza et al. 1991), *Chrysoperla plorabunda* (Robertson and Lampe 1995), *Hyalophora cecropia* (Lidholm et al. 1991) and *Solenopsis invicta* (Krieger and Ross 2003), as well as in flatworms (*Dugesia tigrina*) (Garcia-Fernandez et al. 1993) and roundworms (*C. elegans*) (Sedensky et al. 1994).

The presence of deletions in the mariner sequence from *H. cecropia* differentiates it from those found in *D. mauritiana*. Both have stop codons in the transposase gene, resulting in lack of activity. Sequence analysis of the elements found in *C. elegans* also indicates their lack of activity (Sedensky et al. 1994). Full-length mariner elements identified in silkworm *Antheraea mylitta* are inactive because of point mutations (Prasad and Nagaraju 2003). However, there are many examples of elements from the mariner family, whose sequences include intact reading frames and complete ITR sequences, hence it may be concluded that they remain active.

A typical phylogenetic tree of mariner transposons reveals a characteristic lack of compatibility with the phylogeny of their hosts. This shows the special predispositions of that group for horizontal transfer (HT) (Lampe et al. 2003). Probably for no other family of transposons, have so many examples of HT been proven. This phenomenon was described not only between species belonging to the same genus, but also between higher taxons (Garcia-Fernandez et al. 1995; Lampe et al. 2003).

A sequence from *D. tigrina* has an intact reading frame, which may suggest that the element has the ability to transpose (Garcia-Fernandez et al. 1993). A similar thesis was stated in relation to a mariner element from *T. vaginalis* (Silva et al. 2004).

Sequence analysis of Mboumar element located in satellite DNA (stDNA) in genomes of three species of ants from *Messor* genus (*M. bouvieri*, *M. barbarus*, *M. structor*) was performed. Results suggest, that it might be an active element. Insertion of this Mboumar element was also found inside MITEs elements. Studies on these elements show, that genomic mobility of stDNA repetitions mediated by transposing elements, might be an important molecular mechanism of DNA conservation (Palomeque et al. 2006).

Mariner elements from the irritans subfamily were identified in the genome of hydrothermal crab *Bythograea thermydron* (Bytmar1 element). Very few of their copies includes full-length sequences, which suggests the presence of HT (Case et al. 2006).

The distribution of Mos1 element inside endemic African species *Drosophila teissieri* is heterogeneous in separate populations. In the Brazzaville population, full length elements were discovered, suggesting their potential activity. Other studied populations had a deletion of 500 bp long segment in Mos1 sequence. Based on analysis of complete Mos1 sequences, it was stated that their phylogeny is compatible with the model of vertical transmission from a common ancestor. Therefore, in that case, the influence of HT on the distribution of mariner elements may be precluded (Brunet et al. 1996).

1.4.2. Tc1-like transposons

The family of Tc1 transposons includes elements homological to Tc1 transposon, first described in *C. elegans* (Rosenzweig et al. 1983a). The name is an abbreviation based on: T - transposon, c - ceanorhabditis, 1 - ordinal number of the transposon identified in that species, conferred according to discovery order.

ITRs are 50 to 200 bp long and have on their end conserved sequences TACAGTG. Most Tc1-like transposons have defects in their transposase genes because of numerous mutations resulting in inactivation of the transposon. An inactive transposon is a permanent part of genome, defined as a “genetic fossil” (Miskey et al. 2005; Sinzelle et al. 2006a). The presence of relatives to Tc1 transposon was observed in genomes of numerous species of fish and amphibians (Radice et al. 1994). In fish these elements were first discovered in channel catfish (*Ictalurus punctatus*, Tip1) as insertions in the immunoglobulin gene (Henikoff 1992) and in the genome of *Eptatretis stouti* (Tes1) as insertions in a vasotocin gene (Heierhorst et al. 1992). These transposons were then discovered in *D. rerio* (Tdr1), rainbow trout (*Oncorhynchus mykiss*, Tom1) and Atlantic salmon (*Salmo salar*, Tss1) (Radice et al. 1994).

The sequences of these transposons vary significantly, both intra-individual and interspecies. The amino acid sequence of transposase encoded by Tss1, Tom1, Tip1 and Tes1 elements when compared to the Tdr1 element, shows 77, 69, 43 and 34% of similarity, respectively (Izsvak et al. 1995). Dinucleotide TA is a target integration site of transposon and is located on the ends of both ITRs, as in *C. elegans* (Ketting et al. 1997; Rosenzweig et al. 1983b). All sequences of these elements contain ITRs about 200 bp long. None of the copies of these transposons is active, because of

numerous mutations, reading frame shifts and stop codons in the transposase gene (Ivics et al. 2004).

Research on the Tdr1 element shows, that the dominant form is deletion derivate 1,250 bp long. Deletion in the transposon included the N-terminal part of the DNA-binding region of transposase. It has been estimated, that Tdr1 is present in about 1,000 copies per haploid genome of *D. rerio*, which is about 0.07% of its whole genome.

Both ITRs of Tdr1 transposon contain in their sequence simple short repeats (Direct Terminal Repeats, DTR, 12-22 bp long), called IR/DR structures. Molecular probes complementary to Tdr1 did not hybridize with DNA of other species of *Danio* or with DNA of carp (*Cyprinus carpio*), whereas an intense hybridization signal occurred for *Oncorhynchus tshawytscha* and pike (*Esox lucius*) (Izsvak et al. 1995).

The discovery of a new Tc1-like transposons class called Tdr2 in *D. rerio*, showed the similarity of the transposons to the Tc3A element from *C. elegans*, suggesting the distant evolutionary origin of that family (Gottgens et al. 1999). Tdr2 is about 1000 bp long, flanked by ITRs 100 bp long and occurs in about 1,000 copies per haploid genome of *Danio*. All sequenced copies appeared to be inactive because of many mutations (Ivics et al. 1996).

Independently of research on transposons in *D. rerio*, other Tc1-like elements were identified in the Atlantic salmon genome Ssal1 (SALT). They are 1,535 bp long and occur in 15,000 copies per haploid genome (Goodier and Davidson 1994). Sequence analysis of the Ssal1 transposon revealed the presence of many stop codons in the transposase gene. Therefore, the group of Ssa1 elements is also included in non-autonomic elements (Goodier and Davidson 1994).

The genome of salmon (*Salmo salar*) was analysed for the presence of transposons (de Boer et al. 2007). Over 250 transposon sequences were found, that belong to 14 different families, grouped in two separate classes: Tc1-like and piggyback-like transposons. Several of these families showed similarity to sequences in other species of fish, lamprey, flatworms and other parasites (*Schistosoma japonicum*).

Ivics et al. (1996) classified the Tc1-like transposons, based on similarity of sequence and grouped them into three clusters: A, B and C. Despite the close relationship of cluster A transposons, their hosts represent two distant orders of teleost fish: *Cypriniformes* and *Salmoniformes*. Evolutionarily, lines of the fish diverged 130 million years ago. The Tsn1 transposon from *Salvelinus namaycush* was also classified to cluster A (Reed 1999).

In the B cluster there are representatives of *Cypriniformes* order, including the Tdr1 transposon. Elements in this cluster have a higher level of divergence than elements in the A cluster, and they contain in their sequences multiple insertions and deletions. Cluster C contains Tss2 from *S. salar* and Tdr2 from *D. rerio*. It is worth emphasizing, that in genomes of *S. salar* and *D. rerio* several different evolutionary lineages of Tc1-like transposons coexist (Izsvak et al. 1995).

In the fish *Ictalurus punctatus*, non-autonomous Tipnon sequences related to Tc1-like elements have been described. The Tipnon sequences are only 500 bp long and their ITR sequences are 29 bp long and identical to Tdr1 ITRs. However comparison of inner sequences of Tipnon and Tdr1 shows little similarity between these elements. Tipnon sequences are classified as non-autonomous elements, but they can be mobilized by transposons encoded by autonomous elements (Liu et al. 1999).

Tcch1 transposons, about 800 bp long, have been found in the genome of Antarctic fish *Chiondraco hamatus*. They are similar in general structure to Tc1 elements, but they have several deletions in the DNA sequence, located mainly in 5' region (Capriglione et al. 2002).

There are examples of autonomous Tc1-like transposons, which can be considered to be currently active. Tzf element, identified in *D. rerio* genome shows some relationship with Tdr1 found in the same species. Results of two-dimensional (2D) electrophoresis, combined with hybridization of parent and offspring DNA, suggest a potential activity of Tzf transposon (Lam et al. 1996a). It has been suggested that full length Tc1-like sequences (Tsn1), found in *S. namaycush*, have the ability of active transposition (Reed 1999). The consensus amino acid sequence of Tsn1 element differs in only three amino acids from an active transposase in the reconstructed *Sleeping Beauty* transposon (Ivics et al. 1997), perhaps suggesting that some copies of Tsn1 remain still active.

PPTN transposons from Tc1-like family were found in gene clusters which encode glutathione S-transferase in flatfish *Pleuronectes platessa* (Leaver 2001). Some copies of PPTN transposon were full-length and contained an intact transposase gene, which may suggest the possibility of active transposition (Leaver 2001).

Tc1-like transposons are also widespread in amphibian genomes. They were identified in frogs *Rana catesbeiana* and *Xenopus laevis*, in the latter the coexistence of several evolutionary lineages of Tc1-like transposons was observed (Lam et al. 1996b).

In the genome of *Rana esculenta* a highly repetitive R.e./Tc1 subfamily was identified. A single repetition contains a fragment of Tc1 transposon, which is very similar to a transposon found in the fly *Haematobia* irritans, this is flanked by two short simple repetitions DR (Pontecorvo et al. 2000).

In the genome *Xenopus tropicalis*, the coexistence of seven different and distinct evolutionary lineages of Tc1-like transposons was described (Sinzelle et al. 2005). Three of them are homological to the transposons identified in fish *D. rerio* (Tdr1, Tdr2, Tzf) and *S. salar* (Tss, SALT1).

2. HORIZONTAL TRANSFER (HT) AS A PHENOMENON WHICH CONDITIONS THE FUNCTIONING OF TE

Most transposable elements are defective in the transposase gene. As noted above, transposons that contain a transposase pseudogene become permanent components of a genome

(genetic fossils accumulating mutations). If complete ITR sequences are preserved, this in some cases gives them the chance of being mobilized by an active transposase (*trans* mobilization).

Loss of activity, distribution and sequence polymorphism of transposing elements may be explained by an appropriate evolutionary model. The model assumes, that horizontal transfer (HT) is a process enabling transposons to begin their life cycle as a result of being transmitted from one host species to another (Lohe et al. 1997; Miskey et al. 2005; Sinzelle et al. 2006a).

The mechanism of transposon invasion of cells, which produce gametes of an organism, is not yet known. Transposons themselves are not infectious factors, therefore invasion probably takes place by means of a vector, for instance, either a virus (Ivics et al. 1996) or intracellular and extracellular parasites (Kidwell 1992).

Research by Jehle et al. (1998) demonstrated the presence of the TCp3.2 transposon, which belongs to Tc1-like group, in genome of a baculovirus that infects a species of butterfly – *Cydia pomonella*. In genomes of that species 10 copies of TCp3.2 have been identified. On the basis of phylogenetic analysis a hypothesis was put forward, that the TCp3.2 transposon was inserted into the virus genome during infection of the hosts larvae. This example of HT of transposon from host genome to virus genome suggested a potential role of baculoviruses as interspecies vectors to mediate horizontal transmission of transposons among insects (Arends et al. 2005; Jehle et al. 1998). Newly acquired transposon colonizes host genome by amplification of its copies in the process of transposition. Then, as a result of sexual reproduction of the host, the transposon spreads in the population. This process of transferring transposon copies to host offspring is called vertical transmission.

However, transposons do not come under positive selection, therefore in proportion to elapsed time, random mutations are accumulated in their sequences. As a result, the proportion of inactive copies of a transposon rises. This process is called vertical inactivation (Lohe et al. 1997). Simultaneously, the mutated copies may turn into negative regulators of transposition. Transposase produced by active copies of transposons, may accidentally influence the inactive copies which it can also recognize. Finally, often mutations will eventually inactivate all copies. That way, with time, the rate of propagation falls, and eventually as a result of genetic drift, transposons may be removed from genome in the process of stochastic loss (Miskey et al. 2005). Therefore, in order to survive, transposons must undergo the process of HT to new organisms and start their life cycle over again. It is thought, that HT is a selection tool, which maintains "live" transposons, because only the active elements are able to start a new evolutionary lineages in new populations and species (Sinzelle et al. 2006a).

The HT phenomenon is well documented for the P transposon, first discovered in *D. melanogaster*. It is suggested that P elements have spread fast in natural populations of *D. melanogaster* in the past few decades and seized genome of that species by horizontal transmission from *Drosophila willistoni* (Daniels et al. 1990). HT is also considered the main mechanism for spreading mariner elements, especially in insects, as seen in the presence, in *Drosophila erecta*, of an MLE copy from the *melifera* subfamily, which are 97% similar to MLE found in *Ctenocephalis felis*. Such a close relationship may prove that HT occurred between ancestors of *C. felis* and *D. erecta* (Lohe et al. 1997).

MLE elements from the same subfamily were found in genomes of insects from four orders: earwig *Forficula auricularia*, bee *Apis mellifera*, Mediterranean fruit fly *Ceratitis capitata* and beetle *Epicauta funebris*. Phylogenetic analysis showed the close relationship of MLE from these species and HT between these orders might be the explanation (Lampe et al. 2003).

Relatively recently, HT occurred between evolutionarily distant insects including fruit fly *Drosophila ananassae*, fly *Haematobia irritans*, mosquito *Anopheles gambiae* and green lacewing *Chrysoperla plorabunda* (Robertson and Lampe 1995). Moreover, HT has been documented not only in insects. MLE elements were identified in two evolutionarily distant crustaceans from the amphipods and decapods orders. These MLE are 99.5% similar between the two species (Casse et al. 2006). It is also suggested that HT occurred between more distant taxons: a 75% similarity of MLE elements in flatworms and arthropods may prove that HT occurred between these two taxons (Garcia-Fernandez et al. 1995). MLE elements from the irritans subfamily are present in genomes of chordates (Sinzelle et al. 2006a). It is suggested that these transposons are maintained in the genome by vertical transmission. Phylogenetic incompatibility of MLE elements and host DNA may be explained by a different rate of evolution of these sequences, depending on host species, by stochastic loss and by polymorphism of the ancestors' genome.

The similarity of sequences of transposons that were indentified in Cypriniformes and Salmoniformes (grouped in cluster A) is 90% (Ivics et al. 1996), perhaps suggesting that HT occurred. Transposons, which are degenerated Tc1-like elements related to PPTN from plaice (*Pleuronectes platessa*), were identified in genomes of evolutionally distant species like salmon (*S. salar*) and frog (*Rana temporaria*). This may also suggest, that horizontal transfer occurred.

The Minos element from the Tc1/mariner family (Arca and Savakis 2000), first discovered in *Drosophila hydei* (Franz and Savakis 1991), was found in numerous species of *Drosophila*. Based on several kinds of sequence analysis, it is suggested that five HT events combined with vertical inactivation and stochastic loss, can explain the distribution of Minos in the examined species of *Drosophila* (de Almeida and Carareto 2005).

3. THE INFLUENCE OF TRANSPOSING ELEMENTS ON GENOME EVOLUTION

Since the discovery of transposons by McClintock (1950, 1953) and their role in generating genetic changes, the nature of these elements and their impact on genome evolution have

been discussed. Genomes, particularly eukaryotic, are complex and dynamic units, where the coding sequences constitute only a small part. The majority are non-coding sequences, composed of repetitive elements. In mammals only 2% of genome appears to be functional exons whereas about 50% of the genomes is comprised of repetitive elements originating probably from mobile elements (Makalowski 2000). Transposable elements constitute a large fraction of genome, i.e. in maize it is 50%, in some other plants it is as much as 95%. Insertion sequences in some species of bacteria constitute 10% of genome, whereas in mammals 20% of genome are retrotransposons, especially LINE and SINE elements (Lerat et al. 1999; Shapiro 1999).

While studying the nature of TE and their influence on genome evolution, two different ideas develop. The first originates in the conception of the paradigm of phenotype which is derived from neo-Darwinian theory. This theory states that genes ensure their survival and representation in following generations by providing a selective advantage to the host (Bowen and Jordan 2002). A corollary of this proposes that the presence of TEs must provide a selective benefit to the host genome.

Two important papers, published simultaneously (Doolittle and Sapienza 1980; Orgel and Crick 1980), opposed this neo-Darwinism. They concluded that the appearance and spread of transposing elements may be explained solely by their ability to self-replicate in genome. Therefore, the evolutionary success of these elements is irrelevant to gaining selective advantage for the host genome. A theoretical model was developed, showing that TEs can spread and remain in natural populations even if they generate changes disadvantageous for the host (Hickey 1982). These discoveries are the basis of the theory of selfish DNA with reference to TE. This theory emphasizes the parasitic nature of transposons. Their parasitic status is determined by their inability to reproduce beyond the host genome which provides the mechanism essential for the transposon to perform replication and transcription. TEs are transferred to offspring together with the host genes, and their evolution depends on interactions with other genome components – genes and other TE (Le Rouzic et al. 2007).

However, uncritical approval of this latter concept leads to a radical opinion on the evolutionary importance of transposons - that TEs are only genomic parasites and that their selfish nature excludes a significant influence on genomic evolution. Such a conclusion would discourage research on an evolutionary role for transposons. This latter concept also fails to account for examples of transposons that, despite their undeniable parasitic nature, were adapted in genome and in many ways can well serve their hosts (Bowen and Jordan 2002). Moreover, horizontal transfer, which is characteristic of TEs, is considered to be one of main driving forces of evolution (Syvanen 1994). On the basis of this information it can be concluded, that TEs influence genomic evolution. With their abundance and dispersion in contemporary genomes, they fulfill functions advantageous to the host and as demonstrated in laboratory tests, they provide an ability to generate genomic changes. This leads to the conclusion, that TEs used to, and still do, play an important role in genome reorganization (Shapiro 1999).

Interactions between TEs and their host genome are poorly understood, but examples of their influence on the functioning, structure and evolution of genome have been observed. TE interfere with the genomic environment influencing its evolution and they increase the genome plasticity (Makalowski 2000). It is thought, that non-coding sequences, including non-autonomous TE, determine the overall architecture of the genome, integrate coding sequences into functional groups and also participate in the organization of chromatin (Shapiro 1999). Many interactions of this kind are considered to be the domestication of TE by host genome or the co-evolution of these two units (Capy et al. 2000).

Prokaryotic transposons can carry genes encoding resistance to antibiotics (Recchia and Hall 1997) indicating the role of TEs in generating evolutionary changes by delivering new functional systems. Retrotransposons LINE (heT-A and TART) in *D. melanogaster*, located on chromosome ends, protect them from degradation (Pardue et al. 1997, 2005).

Another example of the influence of TEs on genome evolution are LTR-retrotransposons, which can partly or completely regulate the activity of the host gene (McDonald et al. 1997). Ty sequences in genome of *Saccharomyces cerevisiae*, are often inserted in tRNA genes influencing the profile of their expression (Hani and Feldmann 1998).

TEs are a source of motifs, which can take a part in a controlling RNA transcription. Retrotransposons may be a signal for polyadenylation – it was shown that in 20 vertebrate genes, the poly(A) sequences derive from retrotransposons (Makalowski 2000). Transposon sequences may also play an important role in homologous recombination. Thanks to the similarity of their sequences, they can cause an exchange of chromatin fragments, leading to deletions or duplications of genome fragments. Although most observed recombinations cause pathological effects (Deiniger and Batzer 1999; Kazazian 1998), there can be positive effects. For example, in humans, the family of glycophorin (transmembrane protein of erythrocytes) developed as a result of several duplications generated by recombination between Alu elements (Makalowski et al. 1994). Insertion of transposons from Tc1/mariner superfamily could probably have been adapted during evolution of the vertebrate immune system in order to provide mechanisms enabling proliferation of immunoglobins (Agrawal et al. 1998; Hurst and Werren 2001; Van Gent et al. 1996).

Retrotransposons, thanks to the activity of reverse transcriptase, can take part in the mechanism of gene shuffling by incorporating additional genomic sequences to a new locus (Pickeral et al. 2000). Induced rearrangement of chromosome segments includes genomic changes like inversions, translocations, duplications and the generation of tandem structures (Moran 1999). There is also a hypothesis assuming that accumulation of TEs plays a key role in maintaining the critical volume of the nucleus (Cavalier-Smith and Beaton 1999).

TEs deliver genetic material to their hosts, which can be tested by natural selection for functions that are advantageous to the host (Hurst and Werren 2001). This does not contradict the fact, that TEs can remain in genomes, thanks to their ability to self-replicate. Benefits for hosts may be a consequence of the presence of TEs in genome rather then a reason for it (Charlesworth et al. 1994; Edgell et al. 1996; Zeyl and Bell 1996).

4. EVOLUTIONARY DYNAMICS OF TRANSPOSABLE ELEMENTS

The evolutionary dynamics of TEs is a complex phenomenon because it results from interactions between the following processes: their natural ability of amplification, selection at the level of host, regulation of transposition, and genetic drift (Le Rouzic and Capy 2005). The evolutionary success of TEs in generating an abundance of copies represented in genomes when set against the harmful effects of their activity seems paradoxical.

The parasitic nature of TEs cannot be denied because their mutagenic activity decreases host fitness. They are responsible for 50% of harmful mutations in *Drosophila* (Finnegan 1992b) and 10% in mice (Kazazian 1998). Their success is related to their ability to amplify, thanks to which they can effectively invade the genomes. Without any evolutionary force limiting the amplification of TE, increasing number of copies would lead to the destruction of the host genome. Experimental data suggest, that selection is the main factor limiting excessive multiplication of copies (Charlesworth et al. 1992; Hoogland and Biemont 1996). This functions together with self regulation of TEs. Self regulation is thought to be a successful parasitic strategy, because it controls, if necessary, the rate of transposition. Transposons with a constant rate of transposition would be removed from the genome. Too slow a rate of transposition would lead to removal by genetic drift. Too high a rate would effect in excessive amplification, leading to destruction of the host genome. Transposons, whose rates of transposition are regulated are able to successfully colonize populations, thanks to an initial rapid invasion followed by a reduction in their activity (Le Rouzic and Capy 2005).

The influence of genetic drift on the dynamics of transposons is not understood very well. Theoretical studies show that in small populations, TE families with several copies per individual, would undergo random loss (Brookfield and Badge 1997). Estimation of the frequency of TE distribution in natural, large populations shows, that the influence of genetic drift is unimportant (Biemont et al. 1994). Hence the evolutionary dynamics of transposons is based on balancing different evolutionary forces. Too aggressive behaviour by transposons would lead to the destruction of host genome and then to degradation of the transposons. On the other hand, species that react too repressively to the presence of TEs in their genomes, would lose them and deprive themselves of an important source of genetic diversity (Kazazian 2004).

5. TRANSPOSONS AS USEFUL GENETIC TOOLS

Apart from the undeniable profits for the genome resulting from the genetic changes generated by transposons, their great potential as useful genetic tools should be emphasized. Transposons can be used in biotechnology (Clark et al. 2004; Sasakura et al. 2003), medicine (Aronovich et al. 2007, 2008) and genetics (Kawakami 2005).

A transposase gene was reactivated from defective copies of Tss1 transposon from *S. salar* and Tom1 from *Oncorhynchus mykiss*, by site-specific mutagenesis. The active transposon 1600 bp long was called Sleeping Beauty, SB (Ivics et al. 1997). In order to control stable insertions of desired genes in genomes, a transposing system, based on SB was constructed. It consists of donor plasmid, containing an expression cassette, located between two ITRs, and of an helper plasmid with a sequence of SB transposase gene. This system functions in many fish (Izsvak et al. 2000) and mammalian cell cultures (Dupuy et al. 2002) and in transgenic organisms: mice (Masuda et al. 2004; Yusa et al. 2004), medakafish (Grabher et al. 2003) and *Danio* (Davidson et al. 2003).

The introduction of suitable mutations in the SB transposase gene and in the ITR sequences, increased the effectiveness of transposition 2 to 8 times. Thanks to successful combinations of modified transposases and ITR sequences, an increase of transposition was achieved, 14-fold over the original (Cui et al. 2002; Geurts et al. 2003; Yant et al. 2004) and more recently up to 50-fold more active (Mates et al. 2009).

Depending on the application, different sets of promoters are used. In gene therapy in vivo using a single plasmid for both transposon and transposase in mice, the best expression of gene occurs with the use of a moderate promoter for the transposase gene (Mikkelsen et al. 2003). However, when the system was used to perform insertional mutagenesis in mice, it was optimal to use a modified promoter in the ITR construct in order to increase the efficiency of transposition (Collier et al. 2005; Dupuy et al. 2005).

Active transposons, based on a sequence of Tc1-like transposon from a frog *Rana pipiens* were constructed and called *Frog Prince*, FP (Miskey et al. 2003). They showed 50% similarity in sequence to SB. Their functionality in cells of fishes, amphibians and mammals was proved. FP are regarded as the most effective transposons in vertebrates, and show twice the activity of SB transposon in cells of *Danio*. That increases possible applications in genetic analysis of vertebrates (Miskey et al. 2003). Both SB and FP systems are used in generating lines of transgenic fish (Grabher et al. 2003) and amphibians (Sinzelle et al. 2006b), in which transgene integrated with genome is transferred to offspring. In this way transposons inactive for 10 million years, whose activities have been reconstructed by molecular methods, are able to perform an effective transfer of DNA sequences to chromosomes in order to incorporate desired, new phenotype features (Wadman et al. 2005).

REFERENCES

Agrawal, A., Q.M. Eastman, D.G. Schatz. 1998. Transposition mediated RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394: 744-751.

Alix, K., C.D. Ryder, J. Moore, G.J. King, J.S. Heslop-Harrison. 2005. The genomic organization of retrotransposons in *Brassica olaracea*. Plant Molecular Biology 59: 839-851.

Arca, B., C. Savakis. 2000. Distribution of the transposable element Minos in the genus *Drosophila*. Genetica 108: 263-267.

Arends, H.M., D. Winstanley, J.A. Jehle. 2005. Virulence and competitiveness of *Cydia pomonella granulovirus* mutants: parameters that do not match. Journal of General Virology 86: 2731-2738.

Arkhipova, I.R., M. Meselson. 2000. Transposable elements in sexual and ancient asexual taxa. Proceedings of the National Academy of Sciences of the United States of America 97: 14473-14477.

Aronovich, E., J.B. Bell, L.R. Belur. 2007. Sleeping Beauty transposon-mediated correction of mucopolysaccharidosis type I. Molecular Genetics and Metabolism 92: 27.

Aronovich, E., J.B. Bell, L.R. Belur, R. Gunther, K. Brenda, S. Kahn, P. Schachern, R. Ehlenfeldt, P. Gupta, C. Carlson, C. Whitley, R.S. McIvor, P. Hackett. 2008. Sleeping Beauty transposon-mediated long-term-L-irudonidase expression and correction of lysosomal pathology in NOD/SCID mice with MPS I. Molecular Genetics and Metabolism 93: 15.

Auge-Gouillou, C., Y. Bigot, N. Pollet, M.-H. Hamelin, M. Meunier-Rotival, G. Periquet. 1995. Human and other mammalian genomes contain transposons of the mariner family. FEBS Letters 368: 541-546.

Biemont, C., F. Lemeunier, M.P. Garcia Guerreiro, J.F. Brookfield, C. Gautier. 1994. Population dynamics of the *copia*, *mdg1*, *mdg3*, *gypsy* and *P* transposable elements in a natural population of *Drosophila melanogaster*. Genetical Research 63: 197-212.

Bigot, Y., B. Brillet, C. Auge-Gouillou. 2005. Conservation of palindromic and mirror motifs within inverted terminal repeats of mariner-like elements. Journal of Molecular Biology 351: 108-116.

Bowen, N.J., I.K. Jordan. 2002. Transposable elements and the evolution of eukaryotic complexity. Current Issues in Molecular Biology 4: 65-76.

Brookfield, J.F., R.M. Badge. 1997. Population genetics models of transposable elements. Genetica 52: 281-294.

Brown, N.L., L.R. Evans. 1991. Transposition in prokaryotes: transposon Tn501. Research in Microbiology 142: 689-700.

Brunet, F., F. Godin, C. Bazin, J.R. David, P. Capy. 1996. The mariner transposable element in natural populations of *Drosophila teissieri*. Journal of Molecular Evolution 42: 669-675.

Butler, M., T. Goodwin, R. Poulter. 2001. An unusual vertebrate LTR retrotransposon from the cod *Gadus morhua*. Molecular Biology and Evolution 18: 443-447.

Capriglione, T., G. Odierna, V. Caputo, A. Canapa, E. Olmo. 2002. Characterization of a Tc1-like transposon in the Antarctic ice-fish, *Chiondraco hamatus*. Gene 295: 193-198.

Capy, P., T. Langin, D. Higuet, P. Maurer, C. Bazin. 1997. Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? Genetica 100: 63-72.

Capy, P., G. Gasper, C. Biemont, C. Bazin. 2000. Stress and transposable elements: co-evolution or useful parasites? Heredity 85: 101-106.

Casse, N., Q.T. Bui, V. Nicolas, S. Renault, Y. Bigot, M. Laulier. 2006. Species sympatry and horizontal transfer of mariner transposons in marine crustacean genomes. Molecular Phylogenetics and Evolution 40: 609-619.

Cavalier-Smith, T., M.J. Beaton. 1999. The skeletal function of non-genic nuclear DNA: new evidence from ancient cell chimaeras. Genetica 106: 3-13.

Charlesworth, B., A. Lapid, D. Canada. 1992. The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*: interferences in the nature of selection against elements. Genetical Research 60: 115-130.

Charlesworth, B., P. Sniegowski, W. Stephan. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371: 215-220.

Clark, K.J., A.M. Geurts, J.B. Bell, P. Hackett. 2004. Transposon vectors for gene-trap insertional mutagenesis in vertebrates. Genesis 39: 225-233.

Collier, L.S., C.M. Carlson, S. Ravimohan, A.J. Dupuy, D.A. Largaespada. 2005. Cancer gene discovery in solid tumors using transposon-based somatic mutagenesis in the mouse. Nature 436: 272-276.

Cui, Z., A.M. Geurts, G. Liu., D. Kaufman, P.B. Hackett. 2002. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. Journal of Molecular Biology 318: 1221-1235.

Dalle Nogare, D.E., M.S. Clark, G. Elgar, I.G. Frame, R.T.M. Poulter. 2002. Xena, a full-length basal retroelement from tetraodontid fish. Molecular Biology and Evolution 19: 247-255.

Daniels, S.B., K.R. Peterson, L.D. Strausbaugh, M.G. Kidwell, A. Chovnick. 1990. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. Genetics 124: 339-355.

Davidson, A.E., D. Balciunas, D. Mohn, J. Shaffer, S. Hermanson, S. Sivasubbu, M.P. Cliff, P.B. Hackett, C. Ekker. 2003. Efficient gene delivery and gene expression in zebrafish using the sleeping beauty transposon. Developmental Biology 263: 191-202.

de Almeida, L.M., C.M.A. Carareto. 2005. Multiple events of horizontal transfer of Minos transposable element between *Drosophila* species. Molecular Phylogenetics and Evolution 35: 583-594.

de Boer, J.G., R. Yazawa, W.S. Davidson, B.F. Koop. 2007. Bursts and horizontal evolution of DANN transposons in the speciation of pseudotetraploid salmonids. BMC Genomics 8: 422.

Deininger, P.L., M.A. Batzer. 1999. Alu repeats and human disease. Molecular Genetics and Metabolism 67: 183-193.

Doolittle, W.F., C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284: 601-603.

Dupuy, A.J., K. Clark, C.M. Carlson, S. Fritz, A.E. Davidson, K.M. Markley, K. Finley, C.F. Fletcher, S.C. Ekker, P.B. Hackett, S. Hom, D.A. Largaespeda. 2002. Mammalian germ-line transgenesis by transposition. Proceedings of the National Academy of Sciences of the United States of America 99: 4495-4499.

Dupuy, A.J., K. Akagi, D.A. Largaespeda, N.G. Copeland, N.A. Jenkins. 2005. Cancer biology: Sleeping Beauty awakens. Nature 436: 184-186.

Edgell, D.R., N.M. Fast, W.F. Doolittle. 1996. Selfish DNA: the best defense is a good offence. Current Biology 6: 385-388.

Eickbush, T.H., A.V. Furano. 2002. Fruit flies and human respond differently to retrotransposons. Current Opinion in Genetics & Development 12: 669-674.

Eickbush, T.H., V.K. Jamburuthugoda. 2008. The diversity of retro-transposons and the properties of their reverse transcriptases. Virus Research 134: 221-234.

Evgenev, M.B., H. Zelentsova, N. Shostak, M. Kozitsina, V. Barskyi, D.H. Lankenau, V.G. Corces. 1997. Penelope, a new family of transposable elements and its possible role in hybrid disgenesis in *Drosophila virilis*. Proceedings of the National Academy of Sciences of the United States of America 94: 196-201.

Feschotte, C., L. Swamy, S.R. Wessler. 2003. Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with Stowaway miniature inverted repeat transposable elements (MITEs). Genetics 163: 747-758.

Filee, J., P. Siguier, M. Chandler. 2007. Insertion sequence diversity in Archaea. Microbiology and Molecular Biology Reviews 71: 121-157.

Finnegan, D.J. 1989. Eukaryotic transposable elements and genome evolution. Trends in Genetics 5: 103-107.

Finnegan, D.J. 1992a. Transposable elements. Current Opinion in Genetics and Development 2: 861-867.

Finnegan, D.J. 1992b. The genome of *Drosophila melanogaster*. Academic Press, New York: 1096-1107.

Fisher, C., L. Bouneau, J.-P. Coutanceau, J. Weissenbach, J.-N. Volff, C. Ozouf-Costaz. 2004. Global heterochromatic colocalization of transposable elements with minisatellites in the compact genome of the pufferfish *Tetraodon nigroviridis*. Gene 336: 175-183.

Franz, G., C. Savakis. 1991. Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposon. Nucleic Acids Research 19: 6646.

Furano, A.V., D.D. Duvernell, S. Boissinot. 2004. L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. Trends in Genetics 20: 9-14.

Garcia-Fernandez, J., G. Marfany, J. Baguna, E. Salo. 1993. Infiltration of mariner elements. Nature 364: 109-110.

Garcia-Fernandez, J., J.R. Bayascas-Ramirez, G. Marfany, A.M. Munoz-Marmol, A. Casali, J. Baguna, E. Salo. 1995. High copy number of highly similar mariner-like transposons in planarian (Platyhelminthe): evidence for a trans-phyla horizontal transfer. Molecular Biology and Evolution 12: 421-431.

Gardner, M.J., N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M.W. Mather, A.B. Vaidya, D.M.A. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Fraser, B. Barrell. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419: 498-511.

Garza, D., M. Medhora, A. Koga, D.L. Hartl. 1991. Introduction of the transposable element mariner into the germline of *Drosophila melanogaster*. Genetics 128: 303-310.

Geurts, A.M., Y. Yang, K.J. Clark, G. Liu, Z. Cui, A.J. Dupuy, J.B. Bell, D.A. Largaespada, P.B. Hackett. 2003. Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. Molecular Therapy 8: 108-117.

Goodier, J.L., S.D. Davidson. 1994. Tc1 transposon-like sequences are widely distributed in salmonids. Journal of Molecular Biology 241: 26-34.

Gottgens, B., L.M. Barton, D. Grafham, M. Vaudin, A.R. Green. 1999. Tdr 2, a new zebrafish transposon of the Tc1 family. Gene 239: 373-379.

Grabher, C., T. Henrich, T. Sasado, A. Arenz, J. Wittbrodt, M. Furutani-Seiki. 2003. Transposon-mediated enhancer trapping in medaka. Gene 322: 57-66.

Hani, J., H. Feldmann. 1998. tRNA genes and retroelements in the yeast genome. Nucleic Acids Research 26: 689-696.

Hartl, D.L., A.R. Lohe, E.R. Lazovskaya. 1997. Modern thoughts on an ancient mariner: function, evolution, regulation. Annual Review of Genetics 31: 337-358.

Heierhorst, J., K. Lederis, D. Richter. 1992. Presence of a member of the Tc1-like transposon family from nematodes and *Drosophila* within the vasoticin gene of a primitive vertebrate, the Pacific hagfish *Eptatretis stouti*. Proceedings of the National Academy of Sciences of the United States of America 89: 6798-6802.

Henikoff, J. 1992. Detection of *Caenorhabditis* transposons homologs in diverse organisms. New Biology 4: 382-388.

Hickey, D.A. 1982. Selfish DNA: a sexually transmitted nuclear parasite. Genetics 101: 519-531.

Hirsch, H., H. Saedler, P. Starlinger. 1972. Insertion mutations in the control region of the galactose operon of *E. coli*. Molecular and General Genetics 115: 266-276.

Hoogland, C., C. Biemont. 1996. Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintance of insertion site number. Genetics 144: 197-204.

Hurst, G.D.D., J.H. Werren. 2001. The role of selfish genetic elements in Eukaryotic evolution. Nature Reviews Genetics 2: 597-606.

Ivics, Z., Z. Izsvak, A. Minter, P.B. Hackett. 1996. Identification of functional domains and evolution of Tc1-like transposable elements. Proceedings of the National Academy of Sciences of the United States of America 93: 5008-5013.

Ivics, Z., P.B. Hackett, R.H. Plasterk, Z. Izsvak. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91: 501-510.

Ivics, Z., C.D. Kaufman, H. Zayed, C. Miskey, O. Walisko, Z. Izsvak. 2004. The sleeping beauty transposable element: evolution, regulation and genetic applications. Current Issues in Molecular Biology 6: 43-56.

Izsvak, Z., Z. Ivics, P.B. Hackett. 1995. Characterization of a Tc1-like transposable element in zebrafish (*Danio rerio*). Molecular and General Genetics 247: 312-322.

Izsvak, Z., Z. Ivics, R.H. Plasterk. 2000. Sleeping beauty, a wide host-range transposon vector for genetic transformation in vertebrates. Journal of Molecular Biology 302: 93-102.

Jacobson, J.W., D.L. Hartl. 1985. Coupled instability of two X-linked genes in *Drosophila mauritiana*: and somati mutability. Genetics 111: 57-65.

Jacobson, J.W., M.M. Medhora, D.L. Hartl. 1986. Molecular structure of a somatically unstable transposable germinal element in *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America 83: 8684-8688.

Jehle, J.A., A. Nickel, J.M. Vlak, H. Backhaus. 1998. Horizontal escape of the novel Tc1-like Lepidopteran transposon TCp3.2 into *Cydia pomonella* Granulovirus. Journal of Molecular Evolution 46: 215-224.

Jurka, J., A.J. Gentles. 2006. Origin and diversification of minisatellites derived from human Alu sequences. Gene 365: 21-26.

Kajikawa, M., N. Okada. 2002. LINEs mobilize SINEs in the eel through a shared 3' sequence. Cell 111: 433-444.

Kajikawa, M., K. Ichiyanagi, N. Tanaka, N. Okada. 2005. Isolation and characterization of active LINE and SINEs from the eel. Molecular Biology and Evolution 22: 673-682.

Kawagoe-Takaki, H., N. Nameki, M. Kajikawa, N. Okada. 2006. Probing the secondary structure of salmon SmaI SINE RNA. Gene 365: 67-73.

Kawakami, K. 2005. Transposon tools and methods in zebrafish. Developmental Dynamics 234: 244-254.

Kazazian, Jr., H.H. 1998. Mobile elements and disease. Current Opinion in Genetics & Development 8: 343-350.

Kazazian, H. 2004. Mobile elements: drivers of genome evolution. Science 303: 1626-1632.

Ketting, R.F., S.E.J. Fischer, R.H.A. Plasterk. 1997. Target choice determinants of the Tc1 transposon of *Caenorhabditis elegans*. Nucleic Acids Research 25: 4041-4047.

Kidwell, M.G. 1992. Horizontal transfer. Current Opinion in Genetics and Development 2: 868-873.

Kidwell, M.G. 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica 115: 49-63.

Krieger, M.J.B., K.G. Ross. 2003. Molecular evolutionary analyses of *mariners* and other transposable elements in fire ants (Hymenoptera: Formicidae). Insect Molecular Biology 12: 155-165.

L'Abee-Lund, T.M., H. Sorum. 2000. Functional Tn5393-like tranposon in the R plasmid pBRAS2 from the fish pathogen *Aeromonas salmonicida* subspecies *salmonicida* isolated in Norway. Applied and Environmental Microbiology 66: 5533-5535.

Lam, W.L., T.-S. Lee, W. Gilbert. 1996a. Active transposition in zebrafish. Proceedings of the National Academy of Sciences of the United States of America 93: 10870-10875.

Lam, W.L., P. Seo, K. Robison, S. Virk, W. Gilbert. 1996b. Discovery of Amphibian Tc1-like transposon families. Journal of Molecular Biology 257: 359-366.

Lampe, D.J., D.J. Witherspoon, F.N. Soto-Adames, H.M. Robertson. 2003. Recent horizontal transfer of mellifera subfamily mariner transposons into insect lineages representing four different

orders shows that selection acts only during horizontal transfer. Molecular Biology and Evolution 20: 554-562.
Leaver, M.J. 2001. A family of Tc1 – like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. Gene 271: 203-214.
Lerat, E., F. Brunet, C. Bazin, P. Capy. 1999. Is the evolution of transposable elements modular? Genetica 107: 15-25.
Le Rouzic, A., P. Capy. 2005. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. Genetics 169: 1033-1043.
Le Rouzic, A., S. Dupas, P. Capy. 2007. Genome ecosystem and transposable elements species. Gene 390: 214-220.
Lidholm, D.A., G.H. Gudmundsson, H.G. Boman. 1991. A highly repetitive, mariner-like element in the genome of *Hyalophora cecropia*. Journal of Biological Chemistry 266: 11518-11521.
Liu, Z., P. Li, H. Kucuktas, R. Dunham. 1999. Characterization of nonautonomous Tc1-like transposable elements of channel catfish (*Ictalurus punctatus*). Fish Physiology and Biochemistry 21: 65-72.
Lohe, A.R., D. De Auguiar, D.L. Hartl. 1997. Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional. Proceedings of the National Academy of Sciences of the United States of America 94: 1293-1297.
Makalowski, W., G.A. Mitchell, D. Labuda. 1994. Alu sequences in the coding regions of mRNA: a source of protein variability. Trends in Genetics 10: 188-193.
Makalowski, W. 2000. Genomic scrap yard: how genomes utilize all that junk. Gene 259: 61-67.
Masuda, K., S. Yamamoto, M. Endoh, Y. Kaneda. 2004. Transposon-independent increase of trancription by the sleeping beauty transposase. Biochemical and Biophysical Research Communications 317: 796-800.
Mátés, L., M.K.L. Chuah, E. Belay, B. Jerchow, N. Manoj, A. Acosta-Sanchez, D.P. Grzela, A. Schmitt, K. Becker, J. Matrai, L. Ma, E. Samara-Kuko, C. Gysemans, D. Pryputniewicz, C. Miskey, B. Fletcher, T. Vandendriessche, Z. Ivics, Z. Izsvák. 2009. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nature Genetics 41: 753-761.
McClintock, B. 1950. The origin and behavior of mutable loci in maize. Proceedings of the National Academy of Sciences of the United States of America 36: 344-355.
McClintock, B. 1953. Induction of instability at selected loci in maize. Genetics 38: 579.
McClintock, B. 1956. Controlling elements and the gene. Cold Spring Harbor Symposia on Quantitative Biology 21: 197.
McDonald, J.F., L.V. Matyunina, S. Wilson, I.K. Jordan, N.J. Bowen, W.J. Miller. 1997. LTR retrotransposons and the evolution of eukaryotic enhancers. Genetica 100: 3-13.
Mikkelsen, J.G., S.R. Yant, L. Meuse, Z. Huang, H. Xu, M.A. Kay. 2003. Helpar-independent sleeping beauty transposon - transposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo. Molecular Therapy 8: 654-665.
Miskey, C., Z. Izsvak, R.H. Plasterk, Z. Ivics. 2003. The Frog Prince: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells. Nucleic Acids Research 31: 6873-6881.
Miskey, C., Z. Izsvak, K. Kawakami, Z. Ivics. 2005. DNA transposons in vertebrate functional genomics. Cell and Molecular Life Sciences 62: 629-641.
Moran, J.V. 1999. Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay. Genetica 107: 39-51.
Okada, N. 1991. SINES: short interspersed repeated elements of the eukaryotic genome. Trends in Ecology and Evolution 6: 358-361.
Orgel, L.E., F.H. Crick. 1980. Selfish DNA: the ultimate parasite. Nature 284: 604-607.
Palomeque, T., J.A. Carillo, M. Munoz-Lopez, P. Lorite. 2006. Detection of a *mariner*-like element and a miniature inverted-repeat transposable element (MITE) asociated with the heterochromatin from ants of the genus *Messor* and their possible involvment for satellite DNA evolution. Gene 371: 194-205.
Pardue, M.L., O.N. Danilevskaya, K.L. Traverse, K. Lowenhaupt. 1997. Evolutionary links between telomers and transposable elements. Genetica 100: 73-84.
Pardue, M.L., S. Rashkova, E. Casacuberta, P.G. DeBaryshe, J.A. George, K.L. Traverse. 2005. Two retrotransposons maintain telomeres in *Drosophila*. Chromosome Research 13: 443-453.
Park, J.-M., G.M. Schneeweiss, H. Weiss-Schneeweiss. 2007. Diversity and evolution of Ty1-copia and Ty3-gypsy retroelements in the non-photosynthetic flowering plants *Orobanche* and *Phelipanche* (Orobanchaceae). Gene 387: 75-86.
Pickeral, O.K., W. Makalowski, M.S. Boguski, J.D. Boeke. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome Research 10: 411-415.
Pontecorvo, G., B. Felice, M. Carfagna. 2000. A novel repeated sequence DNA originated from a Tc1-like transposon in water green frog *Rana esculenta*. Gene 261: 205-210.
Prasad, M.D., J. Nagaraju. 2003. A comparative phylogenetic analysis of full-length mariner elements isolated from the Indian tasar silkmouth, *Antheraea mylitta* (Lepidoptera: saturniidae). Journal of Biosciences 28: 443-453.
Radice, A.D., B. Bugaj, D.H.A. Fitch, S.W. Emmons. 1994. Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. Molecular and General Genetics 244: 606-612.
Recchia, G.D., R.M. Hall. 1997. Origins of the mobile gene cassettes found in integrons. Trends in Microbiology 5: 389-394.
Reed, K.M. 1999. Tc1-like transposable elements in the genome of lake trout (*Salvelinus namaycush*). Marine Biotechnology 1: 60-67.
Robertson, H.M., D.J. Lampe. 1995. Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. Molecular Biology and Evolution 12: 850-862.
Robertson, H.M., K.L. Zumpano. 1997. Molecular evolution of an ancient mariner transposon, *Hsmar1*, in the human genome. Gene 205: 203-217.
Rosenzweig, B., L.W. Liao, D. Hirsh. 1983a. Sequence of the *C. elegans* transposable element Tc1. Nucleic Acids Research 11: 4201-4209.
Rosenzweig, B., L.W. Liao, D. Hirsh. 1983b. Target sequence for the *C. elegans* transposable element. Nucleic Acids Research 11: 7137-7140.
Sanmiguel, P., J.L. Bennetzen. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Annals of Botany 82: 37-44.
Sasakura, Y., S. Awazu, S. Chiba, S. Kano, N. Satoh. 2003. Application of Minos, one of the Tc1/mariner superfamily transposable elements, to ascidian embroys as a tool for insertional mutagenesis. Gene 308: 11-20.
Sedensky, M.M., S.J. Hudson, B. Everson, P.G. Morgan. 1994. Identification of a mariner-like repetitive sequence in *C. elegans*. Nucleic Acids Research 22: 1719-1723.
Shapiro, J.A. 1999. Transposable elements as the key to a 21st century view of evolution. Genetica 107: 171-179.
Sharov, A.A. 2006. Genome increase as a clock for the origin and evolution of life. Biology Direct 1: 17.
Siguier, P., J. Filee, M. Chandler. 2006. Insertion sequences in prokaryotic genomes. Current Opinion in Microbiology 9: 526-531.
Silva, J.C., F. Bastida, S.L. Bidwell, P.J. Johnson, J.M. Carlton. 2004. A potentially functional mariner transposable element in the protist *Trichomonas vaginalis*. Molecular Biology and Evolution 22: 126-134.
Simmen, M.W., A. Bird. 2000. Sequence analysis of transposable elements in the sea squirt, *Ciona intestinalis*. Molecular Biology and Evolution 17: 1685-1694.
Sinzelle, L., N. Pollet, Y. Bigot, A. Mazabraud. 2005. Characterization of multiple lineages of Tc1-like elements within the genome of the amphibian *Xenopus tropicalis*. Gene 329: 187-196.

Sinzelle, L., A. Chesneau, Y. Bigot, A. Mazabraud, N. Pollet. 2006a. The mariner transposons belonging to the irritans subfamily were maintained in Chordate genomes by vertical transmission. Journal of Molecular Evolution 62: 53-65.

Sinzelle, L., J. Vallin, L. Coen, A. Chesneau, D. Du Pasquier, N. Pollet, B. Demeneix, A. Mazabraud. 2006b. Generation of transgenic *Xenopus laevis* using the Sleeping Beauty transposon system. Transgenic Research 15: 751-760.

Sugano, T., M. Kajikawa, N. Okada. 2006. Isolation and characterization of retrotransposition-competent LINEs from zebrafish. Gene 365: 74-82.

Syvanen, M. 1994. Horizontal gene transfer: evidence and possible consequences. Annual Review of Genetics 28: 237-261.

Tan, H.M. 1999. Bacterial catabolic transposons. Applied Microbiology and Biotechnology 51: 1-12.

Terol, J., M.C. Castillo, M. Bargues, M. Perez-Alonso, R. de Frutos. 2001. Structural and evolutionary analysis of the *copia*-like elements in the *Arabidopsis thaliana* genome. Molecular Biology and Evolution 18: 882-892.

Tobes, R., E. Pareja. 2006. Bacterial repetitive extragenic palindromic sequences are DNA targets fo insertion sequence elements. BMC Genomics 7: 62.

Top, E.M., D. Springael. 2003. The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. Current Opinion in Biotechnology 14: 262-269.

Touchon, M., E.P.C. Rocha. 2007. Causes of insertion sequences abundance in prokaryotic genomes. Molecular Biology and Evolution 24: 969-981.

Van Gent, D.C., K. Mizuuchi, M. Gellert. 1996. Similarities between initiation of V(D)J recombination and retroviral integration. Science 271: 1592-1594.

Vicient, C.M., A.H. Schulman. 2002. Copia-like retrotransposons in the rice genome: few and assorted. Genome Letters 1: 35-47.

Volff, J.N., C. Korting, M. Schartl. 2000. Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. Molecular Biology and Evolution 17: 1673-1684.

Volff, J.N., U. Hornung, M. Schartl. 2001. Fish retroposons related to the Penelope element of *Drosophila virilis* define a new group of retrotransposable elements. Molecular Genetics and Genomics 265: 711-720.

Volff, J.N., L. Bouneau, C. Ozouf-Costaz, C. Fischer. 2003. Diversity of retrotransposable elements in compact pufferfish genomes. Trends in Genetics 19: 674-678.

Wadman, S.A., K.J. Clark, P.B. Hackett. 2005. Fishing for answers with transposons. Marine Biotechnology 7: 135-141.

Wang, J., L. Song, M.K. Gonder, S. Azrak, D.A. Ray, M.A. Batzer, S.A. Tishkoff, P. Liang. 2006. Whole genome computational comparative genomics: a fruitful approach for ascertaining Alu insertion polymorphisms. Gene 365: 11-20.

Weiner, A.M. 2002. SINEs and LINEs: the art of biting the hand that feeds you. Current Opinion in Cell Biology 14: 343-350.

Yant, S.R., J. Park, Y. Huang, J.G. Mikkelsen, M.A. Kay. 2004. Mutational analysis of the N-terminal DNA-binding domains of sleeping beauty transposase: critical residues for DNA binding and hyperactivity in mammalian cells. Molecular and Cellular Biology 24: 9239-9247.

Yusa, K., J. Takeda, K. Horie. 2004. Enhancement of sleeping beauty transposition by CpG methylation: possible role of heterochromatin formation. Molecular and Cellular Biology 24: 4004-4018.

Zeyl, C., G. Bell. 1996. Symbiotic DNA in eukaryotic genomes. Trends in Ecology and Evolution 11: 10-15.