Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biotechnology 123 (2006) 397–412
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

Review

Fish transposons and their potential use in aquaculture

C. Tafalla[a], A. Estepa[b], J.M. Coll[c,*]

[a] *CISA, Sanidad Animal, INIA, Valdeolmos, 28130 Madrid, Spain*
[b] *IBMC, Universidad Miguel Hernandez, 03202 Elche Alicante, Spain*
[c] *SGIT, Biotecnología. INIA, Crt. La Coruña, Km. 7, 28040 Madrid, Spain*

Received 13 September 2005; received in revised form 2 November 2005; accepted 15 December 2005

Abstract

A large part of repetitive DNA of vertebrate genomes have been identified as transposon elements (TEs) or mobile sequences. Although TEs detected to date in most vertebrates are inactivated, active TEs have been found in fish and a salmonid TE has been successfully reactivated by molecular genetic manipulation from inactive genomic copies (*Sleeping Beauty*, SB). Progress in the understanding of the dynamics, control and evolution of fish TEs will allow the insertion of selected sequences into the fish genomes of germ cells to obtain transgenics or to identify genes important for growth and/or of somatic cells to improve DNA vaccination. Expectations are high for new possible applications to fish of this well developed technology for mammals. Here, we review the present state of knowledge of inactive and active fish TEs and briefly discuss how their possible future applications might be used to improve fish production in aquaculture.

Keywords: Fish transposons; Aquaculture; Repetitive DNA; Sleeping beauty; DNA vaccination

Contents

* Corresponding author. Tel.: +34 1 3476850; fax: +34 1 3572293.
E-mail address: juliocoll@inia.es (J.M. Coll).

0168-1656/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2005.12.019

1. Introduction

At least 40% of the vertebrate genomes contain highly and moderately repeated (Britten and Kohne, 1968) transposon elements (TE) or mobile sequences (Sherratt, 1995). TEs occur in families, copies of which are dispersed throughout the genome. The number of copies per genome can vary from 10 to several thousands, depending on the family of TEs and the species concerned.

On the basis of their structure and presumed mechanism of transposition, TEs fall into two major classes: (i) RNA-based, which transpose DNA sequences after reverse transcription of an RNA intermediate to DNA and (ii) DNA-based, which transpose directly from DNA to DNA.

There are three types of RNA-based TEs, short interdispersed elements (SINEs), long interdispersed elements (LINEs) and retroviruses. The SINEs follow an intermediate RNA to transpose from one locus to another in genomic DNA and are dependent on the presence of a reverse transcriptase provided by separate partner LINEs. The identified fish SINEs (retroposons) are of low size (200–300 bp) and high number of copies (10,000–100,000 copies per haploid genome) and have no terminal repeats (TR). The retroviruses have ∼600 bp long terminal repeats (LTRs) and encode reverse transcriptase and other polypeptides. A few retroviruses have been detected in fish (Poulet et al., 1994) belonging to both endogenous (Hronek et al., 2004) and exogenous (Shen and Steiner, 2004) virus. They will not be considered further in this review.

Most DNA-based TEs are 1000–5000 bp and contain 5′ and 3′ inverted terminal repeats (ITR) of 10–200 bp long. The DNA-based TEs depend on the presence of an active transposase to cut and paste the sequences flanked by the two ITR, to transpose from one locus to another in genomic DNA. The complete sequence of the transposase can be included (autonomous elements) or not (nonautonomous elements) into the TE sequence. The majority of the known DNA-based fish TEs are either Tc1-like of intermediate size (1200–2500 bp) and a moderate number of copies (1000–10,000 copies per haploid genome) or Tol2 of higher size (4700 bp) and lower number of copies (10–30 per haploid genome). Smaller, miniature inverted-repeat TEs (MITEs) have been also isolated from zebrafish (1000–10000 copies per haploid genome) (Izsvak et al., 1999).

In contrast to mammals, a few transpositions and complete transposase sequences have been detected in the TEs of some fish genomes. Here, we review the actual state of knowledge of active and inactive fish TEs and the possible future use of that knowledge to improve different aspects of aquaculture. Fish TEs could be used in commercial fish to generate a large number of transgenics (Davidson et al., 2003) to increase the probabilities to find among those the ones to improve production, resistance to diseases or to obtain biopharmaceuticals (Rocha et al., 2003), to identify commercially relevant genes by TE tagging to apply traditional genetic selection methods and/or to improve fish DNA vaccination vectors which would decrease the required dosages or increase the duration of immunisation (Coll, 2001; Fernandez-Alonso et al., 2001; Lorenzen and LaPatra, 2005; Lorenzen et al., 1998; Purcell et al., 2004; Takano et al., 2004; Traxler et al., 1999).

2. RNA-based fish retroposons SINEs/LINEs

RNA-based fish retroposons or mobile short interdispersed elements (SINEs) are highly repetitive (10,000–100,000 copies per haploid genome),

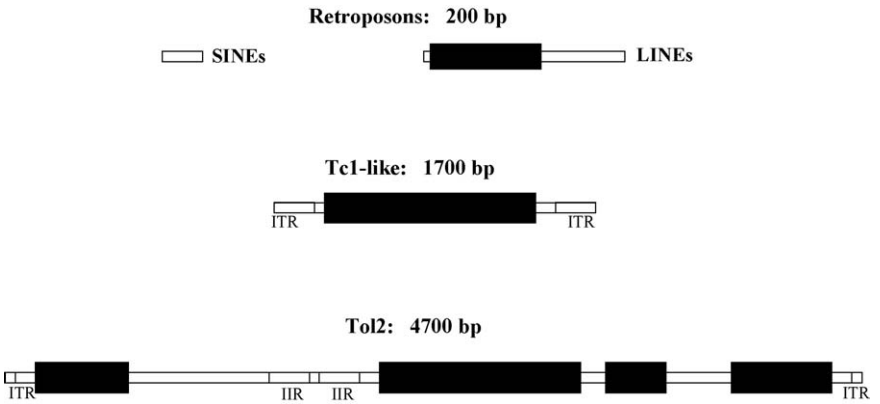

Fig. 1. Scheme of the main families of characterised fish TE. Most studied types of fish TEs belong to the groups of RNA-based TEs or retroposons (SYNEs and LINEs) and DNA-based TEs, Tc1-like or Tol2. Black rectangles illustrate the approximate location of the enzymes used for fish TEs to move. Those enzymes are reverse transcriptase in retroposons and transposase in DNA-based elements. ITR, inverted terminal repeats which flank the enzyme genes in the DNA TEs. IIR, internal inverted repeats that are found inside the Tol2. In the case of retroposons, an RNA from SINEs is copied to cDNA by the reverse transcriptase (black rectangle) coded in LINEs. In the case of DNA-based TEs cut-and-paste mechanisms are catalysed by a transposase (black rectangle) coded in an autonomous element which may or may not be in the same element.

200–300 bp sequences described before in humans and mice (Fig. 1) (Ohshima and Okada, 2005). They were first identified in fish in the chum salmon (*Oncorhynchus keta*) as a ∼200 bp RNA band transcribed in vitro from the genomic DNA by RNA polymerase III (Matsumoto et al., 1984). Sequences among different copies of the same retroposon varied between 1 to 15%. All the copies have a 5′ tRNA-like ∼80 bp sequence that includes two RNA polymerase III promoter binding sites of ∼10 bp each. They contain flanking 6–7 bp direct terminal repeats (DTR) and encode no protein gene (s) (Matsumoto et al., 1986).

SINEs are mobilised by the reverse transcriptase encoded in a longer partner member: the so-called long interdispersed elements (LINEs) (Fig. 1) (Terai et al., 1998). LINEs are about 1000 bp long, encode a reverse transcriptase gene of about 500 bp and contain a 3′ sequence similar to the corresponding SINEs partner. Pairs of LINEs/SINEs elements have been identified in the rainbow trout *O. mykiss* (Smit and Riggs, 1995; Winkfein et al., 1988).

Three retroposon families of 150–180 bp containing flanking short DTR were described in salmonids and named by the unique restriction enzyme sequence found inside each of them (Table 1). They were: the SmaI family restricted to chum and pink salmon in the genus *Oncorhynchus*, the FokI family confined to all known species of the genus *Salvelinus* and the HpaI family present in all species of the subfamily *Salmoniae* including the genera *Hucho*, *Salvelinus*, *Salmo* and *Oncorhynchus* (Kido et al., 1991, 1994). All of them could be identified by hybridisation of genomic libraries to the tRNA-Lys because of the presence of a 5′ tRNA homologous sequence in each of their copies. Confirmation of its presence in several species of salmonids and other fish could then be demonstrated by PCR using two primers located in the consensus sequences for each of the retroposon families.

Using similar methods, a zebrafish specific retroposon called DANA "*mermaid*" comprising about 10% of its genome was also detected (Ivics et al., 1996a; Shimoda et al., 1996). Other SINEs without tRNA sequences were discovered in *O. mykiss* because of their insertion into protamine and histone genes (Winkfein et al., 1988).

Evolution of African cichlid (AFC) SINEs of 292–328 bp have been studied in more than 100 different groups of cichlid fish in lake Tanganyika (Takahashi et al., 1998). The number of AFC copies were estimated between 2000 and 10,000 copies per haploid genome (0.06–0.6% of the genome).

RNA-based transposable elements have not been exploited as experimental tools to the same extent

Table 1
Properties of the best-characterised fish TEs

Type	Name	Fish group	Size (bp)[a]	TR (bp)[a]	Copies/haploid genome	Active	GenBank number	Reference
Retro	SmaI	*Oncorhynchus* sp.	200	7 (DTR)	10000–100000	?	D90289–D90300	Kido et al. (1994)
	FokI	*Salvelinus* sp.	200	7 (DTR)	10000–100000	?	D90289–D90300	Kido et al. (1994)
	HpaI	*Salmonidae*	200	7 (DTR)	10000–100000	?	D32145–D32164	Kido et al. (1994)
	AFC	*Cichlids*	300	–	2000–10000	?	AB009699–AB009714	Takahashi et al. (1998)
	RomI	*Oncorhynchus mykiss*	600	20 (DTR)	4000	?	–	Winkfein et al. (1988)
Tc1	Tip1	*Ictalurus punctatus*	1135	85 (ITR)	?	?	X52617	Henikoff (1992)
	Tes1	*Eptatretis stouti*	1499	66 (ITR)	?	?	M93037–M93040	Heierhorst et al. (1992)
	Ssal1 (SALT)	*Salmo salar*	1535	35 (ITR)	15000	No	L22865	Goodier and Davidson (1994)
	Tel1	*Esox lucius*	1009	223 (ITR)	?	No	L41172	Ivics et al. (1996b)
	Tdr1	*Danio rerio*	1205	208 (ITR)	1000	No	L12210, L33472, L33469	Izsvak et al. (1995), Radice et al. (1994)
	Tdr2	*Danio rerio*	1100	100 (ITR)	1000	No	AJ242983 L48874	Gottgens et al. (1999) Ivics et al. (1996b)
	Tcc1	*Cyprinus carpio*	866	209 (ITR)	?	No	L48683	Ivics et al. (1996b)
	Tcch1	*Chionodraco hamatus*	802	80	?	No	AF305833, AY008266	Capriglione et al. (2002)
	Tom1	*Oncorhynchus mykiss*	1689	221 (ITR)	?	No	L12209	Radice et al. (1994)
	Tss1	*Salmo salar*	1619	226 (ITR)	?	No	L12206–L12208	Ivics et al. (1996b), Radice et al. (1994)
	Tss2						L22865	
	Tzf	***ZebraFish*, *Danio rerio***	**1600**	**200 (ITR)**	**700**	**Yes?**	**U51226–U51230**	**Lam et al. (1996)**
	Tpp1 (PPTN)	***Pleuronectes* platessa**	**1627**	**200 (ITR)**	**150**	**Yes?**	**AJ249083, AJ249085, etc.**	**Leaver (2001)**
	Tsn1	***Salvelinus* namaycush**	**1643**	**225 (ITR)**	**Many**	**Yes?**	**AF017232-34**	**Reed (1999)**
	SB10	***Salmo salar* + *Oncorhynchus mykiss***	**1639**	**225 (ITR)**	**Synthetic**	**Yes**		**Ivics et al. (1997)**
hAT	**Tol2**	***Oryzias* latipes**	**4700**	**12 (ITR)**	**10–30**	**Yes**	**D84375**	**Koga et al. (1996)**

Within most of the families there exist or have existed at a previous time point both autonomous elements that can stimulate their own transposition and nonautonomouss elements that can only transpose in the presence of transposase from an autonomous element of the same family. T, TE. TR, terminal repeats. ITR, inverted TR. DTR, direct TR. Underlined, the initials taken to describe the TE types. Bold, active TEs.

[a] Sizes are approximated.

as have some of the DNA-based elements. The relative complexity of their transposition mechanism and the high cost for facilities to prevent possible infections and spread, will probably preclude their widespread use as experimental tools.

3. Inactive Tc1-like fish TEs

Years ago, TE copies containing inverted terminal repeats (ITR) were only known in maize (Ac/Ds elements), *Drosophila* (P element) and *Caenorhabditis* (Tc1: T, transposon; c, initial of the genus/specie and 1, order number of the TE described for that specie). Tc1-like TEs (also known as Tc1/mariner) were 1200–2500 bp in length containing a transposase gene and two flanking ITR of 50–200 bp (Fig. 1).

Although transposition of Tc1-like TEs required endogenous proteins, they were not species-specific (Lampe et al., 1996) contributing to their wide spread of Tc1-like elements in nature. However, the Tc1-like copies isolated from most vertebrates are dead remnants of once active TEs that have become inactivated by mutations, deletions and/or stochastic loss of autonomous copies after successfully colonising genomes.

In fish, Tc1-like elements were first reported in 1992 as gene insertions. Those were in the immunoglobulin genes of catfish (*Ictalurus punctatus*), Tip1 (Henikoff, 1992) and in the vasotocin gene of the pacific hagfish (*Eptatretis stouti*), Tes1 (Heierhorst et al., 1992) (Table 1).

Tc1-like TEs of 1600–1700 bp were then identified by PCR by using a single primer derived from the ITR of Tc1 in zebrafish (*Danio rerio*, Tdr1), rainbow trout (*O. mykiss*, Tom1) and Atlantic salmon (*Salmo salar*, Tss1) (Radice et al., 1994). Copies of Tc1-like elements revealed considerable intra-individual as well as inter-specific sequence variations. Thus, sequence similarities throughout a canonical 320–343 amino acid transposase from several clones showed that Tss1, Tom1, Tip1 and Tes1 have 77, 69, 43 and 34% amino acid identities with Tdr1 (Izsvak et al., 1995). All the Tc1-like sequences containing ITR of about 200 bp were flanked by TA dinucleotides, which are target site duplications. None of the copies of these fish TEs were active because many contain incomplete or nonautonomous elements consisting in sequences with deletions, frameshift mutations or premature termination codons in their transposase genes (Ivics et al., 2004).

Further examination of sequences from many Tdr1 copies (Izsvak et al., 1995) revealed that a deletion derivative of 1250 bp was its dominant form (the deletion included the N-terminal DNA-binding region of the transposase). Estimations showed that Tdr1 was present in about 1000 copies per haploid zebrafish genome (0.07% of the genome). The Tdr1 copies contained flanking ITR of 208 bp which were each flanked by short DTR of 12–21 bp (the so-called IR/DR structure) (Fig. 1). Tdr1 probes did not hybridised with other zebrafish species or carp (*Cyprinus carpio*) but intensive hybridisation were observed with chinook salmon (*Oncorhynchus tschawitscha*) and northern pike (*Esox lucius*).

Tdr2 was later identified by hybridisation of a Tc1-like sequence found in a developmental zebrafish gene to the complete zebrafish genomic sequence. Fragments of Tdr2 were then detected in several expressed sequence tag sequences from the same genome (Table 1). Tdr2 expands 1100 bp contained flanking ITR of 100 bp and were present in about 1000 copies per haploid zebrafish genome. The 5′ ITR is itself composed of another ITR, an original structure only found in Tdr2. All the sequenced copies were inactivated by mutations (Gottgens et al., 1999; Ivics et al., 1996b).

Other Tc1-like TEs, the so-called Ssal1 (SALT) of 1535 bp were independently identified in Atlantic salmon by hybridisation of a genomic library with the 580 and 800 bp bands obtained by PstI digestion of its genomic DNA. By densitometry of ethidium bromide stained gels, the two bands accounted for 0.2% of the genomic DNA. By hybridisation with Ssal1 probes, 335 positive clones were detected among 3600 clones of a size of 15,000 bp each, which corresponded to about 15,000 copies per haploid genome (Goodier and Davidson, 1994). Ssal1 TEs contained an open reading frame corresponding to a transposase gene with a high number of stop codons and therefore are probably not active. Ssal1 copies were also found in other salmonids (*Salmo*, *Salvelinus* and *Oncorhynchus*) but not in any other fish species studied (Goodier and Davidson, 1994).

4. Active Tc1-like fish TEs

In contrast to other vertebrates, active transposition in zebrafish (Lam et al., 1996) or complete Tc1-like TE sequences suggesting active Tc1-like transposition in *Salvelinus namaycush* and *Pleuronectes platessa* (Leaver, 2001; Reed, 1999) have been described or detected.

Thus, by hybridisation with probes derived from the transposases Tc1 + Tes1, about 700 copies of a new zebrafish TE were first detected and called Tzf (Lam et al., 1996). The consensus of several copies defined a 1621 bp candidate to an active transposase similar to the Ssal1 and Tdr2 (Ivics et al., 1996b). When comparison of digested DNA mixed from father and mother zebrafish were made with each progeny DNA by two dimensional electrophoresis followed by hybridisation, new spots were detected, demonstrating Tzf transposition activity had taken place during gametogenesis in the germ line of the parent fish. On the average there were 3.9 new spots per each progeny individual zebrafish DNA. Since it is not clear that Tzf does move (the authors pointed out that they saw new spots but never saw a spot disappear, this TE have an ? in activity in Table 1).

Complete full length of Tc1-like TE copies were detected in the genome of the lake trout *Salvelinus namaycush*, Tsn1 (Reed, 1999) by using primers from the Tdr1 for PCR amplification (Izsvak et al., 1995). The consensus sequence of full length Tsn1 was 1643 bp (corresponding to a complete transposase of 340 amino acids) containing flanking ITR of 225 bp. Although even one amino acid change could destroy transposase activity, the amino acid consensus sequence of Tsn1 only differed by three amino acids from the *Sleeping Beauty* (SB) active transposase (see later), suggesting that some of the Tsn1 copies might still be active. However, since to detect activity there also needs to be an enhancer/promoter source as well and often there is none (the ITR seems to lack promoter activity) it is not known if those TEs have retained activity. Sequence comparisons indicated that Tsn1 is common to some salmonid and cyprinid fish.

Complete full length copies of Tpp1 (PPTN) were also found among the 150 copies per haploid genome of *P. platessa* (Leaver, 2001). The Tpp1 sequence was obtained from genomic DNA by PCR with a primer derived from the ITR of the TE sequence identified within the glutathione *S*-transferase gene from *Pleuronectes*. The existence of ∼60% of full length (1627 bp) copies of the transposase suggests that Tpp1 is still active. Similar sequences were also detected in *Pleuronectes flexus* and *Limanda limanda*. The corresponding but different and incomplete sequences were also found among 600 copies per haploid genome in *Salmo salar* (SSTN) and *Salmo trutta*.

5. Reactivating and improving Tc1-like TEs: the SB model

From defective copies of Tss1 and Tom1 and using additional site directed mutagenesis, a Tc1-like transposase with activity in mammal cells was obtained. The active TE consisted in a 1600 bp transposase gene containing flanking ITR of 225 bp. It was called *Sleeping Beauty* or SB (Ivics et al., 1997).

To control stable insertion in the genomes of desirable genes based on SB, dual transposition systems (SB/ITR) were then developed. A donor plasmid with an expression cassette between two ITR (5′ ITR-promoter-gene-terminator-3′ ITR) or ITR construct, was used together with a helper plasmid with the SB transposase sequence under the control of a promoter (most of the times the SB protein or its mRNA are used instead of the SB plasmid to avoid insertion of the SB gene into the genomes). Dual SB/ITR transposition systems were functional in many mammalian and fish cell lines (Dupuy et al., 2002; Izsvak et al., 2000) as well as in transgenic mice (Dupuy et al., 2002; Horie et al., 2003; Masuda et al., 2004; Yant et al., 2000; Yusa et al., 2004), medaka (Grabher et al., 2003) and zebrafish (Davidson et al., 2003).

The SB transposition required an active reconstructed transposase, two transposase binding sites of 30 bp in the 225 bp ITR (Izsvak et al., 2000, 2002) and cellular factor (s) (Yant and Kay, 2003; Zayed et al., 2003), to cut the sequence bracketed by the ITR in the ITR construct and paste it into the genome at target TA sites (Plasterk et al., 1999). For optimal use in HeLa human cells, mutations have been scanned to increase the efficiency of transposition and introduced both in the SB10 first active version (Geurts et al., 2003; Ivics et al., 1996b; Plasterk et al., 1999; Yant et al., 2004; Zayed et al., 2004) and in the first ITR1 sequences (Cui et al., 2002; Izsvak et al., 2000; Yant et al., 2004).

Mutant SB11, SB12 and HSB3 versions of the transposase improved two- to eight-fold the efficiency of transposition. Combination of some of those new transposase with improved ITR2 or ITR3 versions produced about a 14-fold increase in the efficiency of transposition in comparison to the original.

Promoter improvement can also be added to both the SB transposase and to the ITR construct. The use of different promoters would depend on the application involved as shown for gene therapy or insertional mutagenesis examples in mammals. Thus, the best gene expression in a gene therapy in vivo mice model were obtained with four-fold less active promoters in the SB transposase than in the gene of the ITR construct (Mikkelsen et al., 2003). On the contrary, when used for massive insertional mutagenesis in mice, transposition was too low and local (Fischer et al., 2001), so both the SB11/ITR2 improved versions and a smaller size/highly efficient promoter on the ITR2 construct were used to increase efficiency (Collier et al., 2005a; Dupuy et al., 2005).

Other additional improvements to optimise transposition efficiency for fish cells, might still be incorporated as our knowledge of the SB/ITR systems progresses and this, in turn, could boost their possible applications to aquaculture (Tafalla et al., submitted for publication. On the other hand, an active Tc1-like element has recently been constructed from frogs called the *Frog Prince* (Miskey et al., 2003). This new synthetically active TE and/or some future ones from other species might also be applied to fish cells.

6. Active Tol2 TEs of medaka

The Tol2 TEs were discovered as a TE-like insertion into the tyrosinase gene of medaka (*Oryzias latipes*) (Koga et al., 1996) causing a loss-of-function mutation in the tyrosinase gene resulting in a naturally occurring albino phenotype. The Tol2 insertion was a 4700 bp TE containing flanking ITR of 17 and 19 bp, two internal inverted repeats (IIR) of ∼300 bp and present in 10–30 copies per haploid genome of nearly identical sequences (Koga and Hori, 1999). The Tol2 contained four open reading frames of an specific transposase belonging to the hAT TE family (Fig. 1). Its insertion site showed an 8 bp target site sequence duplication (Koga and Hori, 2000).

The observation of Tol2 excision during medaka embryogenesis (Koga et al., 1996) and the presence of complete transposase open reading frames, suggested the existence of active transposases. Active transposition was demonstrated by inducing transposase-dependent excision in cells (Kawakami et al., 2004a) and by coinjecting fertilised medaka eggs (from a medaka specie without Tol2) with a Tol2 derivative containing an antibiotic resistance gene, an mRNA from the putative Tol2 transposase and a target plasmid (Koga and Hori, 2000).

Further work showed that a similar Tol2/ITR system successfully worked in zebrafish eggs, which do not contain Tol2 (Kawakami et al., 2000). Thus, a Tol2 ITR construct designed by deleting a fragment of 1500 bp of the transposase and by inserting antibiotic resistant genes in its place, coinjected in zebrafish eggs with mRNA for the Tol2 transposase was able to transpose. Long (four exons) or short (three exons) transposase mRNAs exerted positive and negative effects on excision, respectively, suggesting an internal pathway of regulation of the Tol2 activity (Tsutsumi et al., 2003). Furthermore, the Tol2 transposase showed an extranuclear localisation which might explain its low transposition activity in its original host species (Iida et al., 2004). Improvements in mRNA in vitro production and inclusion of a promoter-green fluorescent protein, GFP-poly(A) terminator cassette in the Tol2 ITR construct, increased in vivo transmission of the GFP gene to about 50% of the zebrafish progeny (Kawakami et al., 2004b).

Because medaka Tol2 transposes in cells from other species and, although there is no limit to capacity for genes in SB or probably in any other TE (Horie et al., 2003; Izsvak et al., 2000; Karsi et al., 2001), Tol2 TE might have a larger size capacity for inserting genes than the Tc1-like elements (Kawakami et al., 2004a; Kawakami and Noda, 2004). The expectations of their possible biotechnological applications are high (Koga, 2004).

7. Other fish TEs

New families of fish TEs are being identified by computer assisted searches of sequences related to known TE from other species, including prokarytes, in the banks from fish genome projects now available.

Table 2
Properties of some new fish TEs discovered by searching in recently published sequences

Type	No	Fish	Size (bp)	TR (bp)[a]	Copies/haploid genome	A[b]	GenBank number	Reference
Merlin	1	*Danio rerio*	10173	177 (ITR)	50	No	AL845359	Feschotte (2004)
	2	*Danio rerio*	1371	462 (ITR)	500	No	AL845359	Feschotte (2004)
	3	*Danio rerio*	239	29 (ITR)	8000	No	AL845359	Feschotte (2004)
Harbinger	1	*Danio rerio*	4016	16 (ITR)	<10	Yes	AL591210	Kapitonov and Jurka (2004)
					>2000	No		
	2	*Danio rerio*	3727	13 (ITR)	100	Yes	ctg25784.2	Kapitonov and Jurka (2004)
	3	*Danio rerio*	3599	12 (ITR)	<10	Yes	ctg25396.1	Kapitonov and Jurka (2004)
					>1000	No	~HARBII protein	
PiggyBAC	1	*Fugu rubripes*	2472	13 (ITR)	200	No	AC1562662[c]	Sarkar et al. (2003)
	1–5	*Danio rerio*	?	?	?	No	XM677997[d]	Sarkar et al. (2003)
P-like	1	*Danio rerio*	11149	13 (ITR)	1	Yes	BX511023	Hammer et al. (2005)
					3	No		

No, number.
[a] Sizes are approximated.
[b] Etimations of TE activity are based only in the autonomous/nonautonomous characteristics of the elements detected in the sequences of the GenBank. However, the nonautonomous elements could be mobilised by transposases encoded in autonomous non-yet detected elements.
[c] Longest sequence of fugus in the GenBank.
[d] Five piggyBAC-like defective elements were found in zebrafish. Only one sequence source is given.

Such is the case of the DNA-based Merlin/IS1016 group of bacterial insertion sequences (IS) present in the parasitic blood fluke *Schistosoma mansoni* (Feschotte, 2004). Thus, three families of nonautonomous Merlin-like elements of different sizes and copy numbers have been found in the genome sequences of zebrafish (Table 2) as well as in other eukaryotic organisms including frogs and man.

Harbinger transposases are a family of widely distributed (from bacteria to animals and plants, including vertebrates) autonomous DNA-based TEs requiring two separated proteins to transpose: one with DNA-binding activity, the other with the transposase activity. Harbinger-like TEs have been found among the sequences in gene banks in many species including the fish fugu, medaka, trout and zebrafish (Kapitonov and Jurka, 2004). Three families of Harbinger-like sequences divergent but related, have been described in zebrafish (Table 2). Only a few copies of those Harbinger-like elements were autonomous while several thousand copies were nonautonomous. Harbinger3-like elements from zebrafish were characterised by a striking preference for a 17 bp target site never seen previously in any other TE.

About 200 copies of inactivated piggyBac-like sequences were identified in the pufferfish (*Fugu rubripes*), in zebrafish and in many mammalian species (Sarkar et al., 2003). PiggyBac-like DNA-based TE were of 2472 bp encoding an 594 amino acid transposase and containing flanking ITR of 13 bp inserted into TTAA target sites (originally isolated from the moth *Trichoplusia*). Unique piggyBac-like sequences have been found presumably dedicated to cellular functions (known as "domesticated" TE) in the *Fugu* genome (Sarkar et al., 2003).

P-like elements, related to the first DNA-based TEs characterised in *Drosophila*, have been recently identified in the zebrafish genome by computer searches (Hammer et al., 2005). Truncated P-like elements have been detected before in other insect species but "domesticated" forms beneficial to the host genome, truncated and complete copies were not. The Pdre2 P-like element in zebrafish was the only complete copy among the four copies of P-like elements detected in the zebrafish genome (Table 2). It is 11,149 bp long, containing an open reading frame of 2631 bp in three exons (corresponding to a transposase of 877 amino acids), 13 bp flanking ITR and 8 bp target site duplications. A maximum DNA sequence divergence of 3.5% was obtained when comparing the internal regions of all four P-like elements.

TEs of different types are potentially capable of inserting into each other to form complex TEs. Evidence for some of those have been found in zebrafish (Ivics et al., 1996a; Izsvak et al., 1999). Other, yet unidentified, complex TEs surely exist not only in zebrafish but in many other fish genomes, but these remain to be investigated.

8. Control of efficiency of transposition and number of copies

The control of exogenously induced transposition, efficiency is different whether they are made in the absence (mammals) or in the presence (fish) of endogenous TEs. Thus, in the absence of endogenous TEs related sequences, once the ITR constructs are inserted into the genome, the rate of TEs excision depends on both the number of mobilizable copies inserted and the level of exogenous specific transposase supplied. However, in fish, that already have endogenous TEs related sequences in their genomes, both the activity of exogenous transposase and the ITR copies added might be tightly regulated to avoid excessive transposition activities which would cause lethal effects. Although, diploid genomes which provide two copies of each gene, represent a first line of control of the possible negative effects of transposition, excessive transposition would cause generation of numerous random insertions in genes and/or promoters (thereby causing numerous loss-of-functions). However, adverse effects of TE insertional mutagenesis are hardly seen (Essner et al., 2005b; Kazazian, 2004; Keng et al., 2005; Yusa et al., 2004) unless: (i) the offspring are inter-mated to detect homozygous mutations or (ii) the inserted TE is a construct with powerful enhancer/promoter combinations (Collier et al., 2005a,b).

A way to control transposition is by inactivating most of the transposase copies as it happens with Tc1-like transposases in vertebrates. In vertebrates most of the copies of the existing transposase-like sequences detected at the moment are truncated by deletions, internal stop codons and/or frameship mutations, resulting in non-detectable transpositions.

On the other hand, in addition to complete sequences of both transposase and ITR, the presence of a promoter at the 5′ end and a poly(A) signal at the 3′ end are both required for a high level of endogenous transposase expression. Few studies have been made about the existence of such putative promoters in fish TE. The possibility that ITR sequences and their adjacent sequences might contain some of the required elements with enhancer/promoter activities, as it happens in fish retrovirus with the LTR flanking sequences (Hronek et al., 2004) or in Tol2 in which a promoter activity was detected in their IIR (Tsutsumi et al., 2003), has not been reported.

Several host endogenous enzymes might play a role in the repair of both TE excision and/or integration in cells as it occurs with the SB (Izsvak et al., 2000; Yant and Kay, 2003; Zayed et al., 2003). Excision of the sequences at the ends of the ITRs generates double-strand breaks with 3 bp of single stranded sequences at the 3′ ends of each strand that invade the DNA target to generate a 6 bp single stranded gap at each flank of the TE at the insertion site (Plasterk et al., 1999). Failure to repair any of these gaps might cause cell death or reduce the efficiency of transposition (Izsvak et al., 2004). On the other hand, required DNA bending proteins for transposition could be also used to control the efficiency of transposition (Zayed et al., 2003).

The effect of the ratio of expression of exogenously added transposase and ITR construct in transposition efficiency has been studied for SB in human HeLa cells. Optimal transpositions were obtained in a narrow window and excess of transposase caused overexpression inhibition (Geurts et al., 2003). In vivo, the inhibitory effect of SB transposase overproduction was also detected by comparing 10 different promoters. Best gene expression was obtained with SB promoters 30–40-fold less active than CMV or four-fold less active than the RSV used for the gene in the ITR construct (Mikkelsen et al., 2003).

The size of the TE and the endogenous or exogenous number of existing or induced copies per haploid genome seem to be somehow controlled. Thus, an inverse correlation between the size of existing fish TE and their number of copies present in the corresponding genomes have long been recognised (Tables 1 and 2 and Fig. 2). The longer the sequence in between the ITR, the lower the number of endogenous copies in the genomes (for instance, there are 10,000–100,000 copies of retroposons of ∼200 bp, 1000–1500 copies of Tc1-like TEs of ∼1500 bp and 10 copies of Tol2 of 4700 bp).

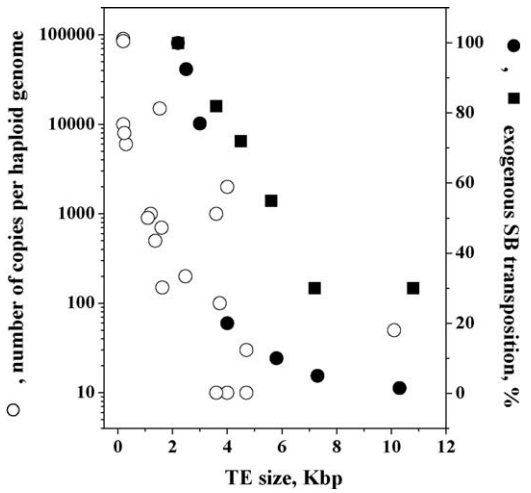

Fig. 2. Representations of the size of endogenous fish TE with their number of copies per haploid genome (open circles) and of the size of exogenous SB and their relative transposition efficiencies (black circles and squares). Open circles, number of copies per haploid genome of endogenous fish TE (approximated data form Tables 1 and 2). Black circles, approximated transposition efficiencies obtained with different size constructs between the ITR of SB expressed in percentage with respect to the 2.2 kbp pT1/SVneo plasmid (Izsvak et al., 2000). Black rectangles, approximated transposition efficiencies obtained with different size constructs between the ITR of SB expressed in percentage with respect to the 2.2 kbp pT1/SVneo plasmid corrected by variation in transfection efficiencies with size of the constructs (Geurts et al., 2003).

Although there still would be a 30% transposition with exogenous TE carrying >10 kbp (Geurts et al., 2003), the efficiency of exogenous SB-mediated transposition has been also shown to be inversely proportional to ITR construct size (Geurts et al., 2003; Izsvak et al., 2000). Fig. 2 shows a composition of endogenous and exogenous data suggesting that those two different observations might correlate due to the existence of a common mechanism of regulation. Since other factors are involved when delivering exogenous TE as a plasmid, for instance the outer distance of the ends of the TE (Izsvak et al., 2000) or the efficiency of transfection due to the use of different expression cassettes (Karsi et al., 2001), the correlation might be not significant.

The Tol2 TE has shown two other possibilities to control fish transposition. Thus, by coinjecting Tol2 vectors together with long (four exons) or short (three exons) Tol2 mRNA into medaka eggs, it was shown that the long and short Tol2 mRNAs exerted positive and negative effects on excision, respectively. These results suggested a new internal pathway of regulation of the Tol2 activity (Tsutsumi et al., 2003). In addition, the extranuclear localisation of Tol2 transposase (Iida et al., 2004) might provide another control, in contrast to the intranuclear localisation of the Tc1-like transposases (Ivics et al., 1997; Yant et al., 2004), confirmed by following localisation with SB10-GFP fusion constructs in cells (unpublished).

9. Possible influences of endogenous fish TE in exogenous TE-induced transposition

A TE taken from one specie (salmon/trout) and introduced into another specie (mammal), will usually be placed into a genome that has no related TE sequences. This situation simplifies subsequent manipulation and reduces the chances of transposition being affected by possible endogenous controls.

However, when inactivated copies of an specific TE are still present in the host genome such as it seems to occur in mammalian cells with retroposons (estimates of about 100 active LINEs hoping all over the genome) (Kazazian, 2004) and it could happen in many commercially important fish, what would be the effects of introducing active TEs? Would the introduction of exogenous specific transposase in fish genomes induce endogenous transposition events of the present specific TEs? Would it cause excessive fish genome damages? What would happen if there is a low specific endogenous TE transposase activity present in fish cells? These and many more questions need to be answered before one TE system could be applied to the germ line (transgenic fish, gene identification) or to somatic (DNA vaccination) fish cells.

To have a preliminary estimation of some of the possibilities mentioned above, we have compared the published Tc1-like endogenous sequences from carp (L48683) or trout (L12209) to the existing active SB10/ITR1 sequences) (Ivics et al., 1997). Maximal identities of 78 or 88% were obtained between the SB10 active transposase sequence (Ivics et al., 1997) and the carp (nucleotides 336–1225) or the trout (nucleotides 370–1397) corresponding sequences. Similarly, the ITR sequences from carp or trout were 84 or 87% (5′ ITR) and 83 or 82% (3′ ITR) identical to the corresponding ITR1 sequences from the ITR construct (plasmid pT H/B) (Ivics et al., 1997) (unpublished anal-

ysis). Because there are ∼1000 copies of those similar sequences per haploid genome in both carp and trout, there might exist among those some identical copies to SB/ITR. Also, as has been mentioned above, since naturally active transpositions have been detected in some fish (Kawakami et al., 2000; Koga and Hori, 2000; Lam et al., 1996; Tsutsumi et al., 2003) and complete, potentially active transposase gene(s) have also been detected (Leaver, 2001; Reed, 1999), it seems likely that active TEs related or not to the SB/ITR TE might still be active in particular fish species. Therefore, the use of the SB/ITR, at least in carp (*Cyprinid*) or rainbow trout (*Salmonid*), will be difficult to predict since we do not know yet the possible effect(s) of those endogenous sequences on exogenous SB/ITR-induced transposition. The above studied situation might be similar for other still unexplored fish TEs.

10. Transgenic fish by TEs

The SB/ITR system, the first example of a reactivated vertebrate transposon, has been mainly developed and optimised to be used for human gene therapy vectors (transformation of somatic cells) (Essner et al., 2005a; Izsvak and Ivics, 2003; Liu et al., 2004; Miskey et al., 2005), for transgenic mice generation (Dupuy et al., 2005, 2002; Ivics et al., 2004; Miskey et al., 2005) or for insertional mutagenesis in mice (Clark et al., 2004) (Table 3). For stable transgenesis, massive transpositions need to be avoided (Davidson et al., 2003; Grabher et al., 2003), whereas for insertional mutagenesis the efficiencies of transposition and the number of copies need to be increased (Collier et al., 2005a; Dupuy et al., 2005).

Methods based and/or related to the SB/ITR system and/or other fish TEs, could also be applied to model fish and to commercially important fish species to improve immunisation by DNA or to obtain transgenic fish and to identify important fish genes. Insertional mutagenesis for basic research has been widely applied to zebrafish (Davidson et al., 2003; Kawakami et al., 2000; Wadman et al., 2005).

However, in vivo use of TEs in mammals is favoured because of the absence of active copies of the corresponding transposases prevents excision and increases stability of the inserted genes. The possible existence of active copies of some TEs in fish, changes that situation but it might also offer some other possibilities for manipulation. In other words, the presence of active transposases in fish might be an advantage (only the donor plasmid will be required or the efficiency of transposition might be higher) or a disadvantage (low stability of the inserts). The low stability of the inserts might not be an important disadvantage because there is no requirement to be stable enough to pass onto the next generation (to increase the duration of an immunisation by DNA, for instance, will only require some months of expression).

Although the SB/ITR system have already been used to express the green fluorescent protein (GFP) in zebrafish (Davidson et al., 2003) and medaka (Grabher et al., 2003), arguing about the absence or, at least, the existence of a low endogenous activity of the SB transposase in both fish species, molecular characterisation of those fish lines are not so clear to interpret. Thus,

Table 3
Some of the possible applications of TEs in fish

	Purpose	Results	Reference
Loss-of-function	Gene identification	Gene-tagged deficient mutants	Clark et al. (2004), Gollin et al. (2002)
Gain-of-function	Transgenic vector	Transgenics	Davidson et al. (2003)
	Enhancer traps	Enhancer identification	Balciunas et al. (2004), Grabher et al. (2003)
	Promoter traps	Promoter identification	Westerfield et al. (1992)
	Poly(A) traps	Poly(A) identification	Yoshida et al. (1995), Zambrowicz et al. (1998)
	Immunisation[a]	DNA vaccination	Coll (2001)
	Oncogene induction	Oncogene-tagged identification	Collier et al. (2005a), Dupuy et al. (2005)

Tabulated and modified from Wadman (Wadman et al., 2005), Hackett (Hackett and Alvarez, 2000) and Coll (Coll, 2001). The ideal transposable system for use in loss- or gain-of-functions should have, at least, the following properties: (i) the possibility to develop a dual system, separating the transposase from the donor sequence, (ii) a genetic background free of the TE to avoid control mechanism possible interferences and (iii) a rate of transposition adequated to the goals pursued by the transposition.

[a] Preliminary evidence.

sequencing of the junctions of the ITR insertions or of the SB-induced excision of previously inserted ITR constructs (Liu et al., 2004) showed that in contrast to 10 in 10 zebrafish lines being the result of SB-induced transposition events, only 1 in 10 of the medaka lines was the result of SB transposition, the other nine being the result of unspecific recombination (Davidson et al., 2003; Grabher et al., 2003).

The effective concentration of the ITR constructs, the molecular ratio between exogenous transposases and ITR, the possible influence of fish species–specific host factors, the interference of possible endogenous TEs-related sequences and/or other still unknown control mechanisms, might be some of the factors to be taken into account when TE technology would be applied to transgenesis of commercially important fish species such as salmonids.

The capacity to select a desirable trait among new transgenic fish would depend on the number of transgenics that we are able to produce (Kawakami et al., 2004b). Therefore, a method that produces as many transgenics as possible (highest efficiency) should be applied to widen the possibilities to find the best one. Thus, injection of naked plasmids obtained a 5% efficiency (Stuart et al., 1990), of Tc obtained 37.5% efficiency (Raz et al., 1997), of pseudotyped retroviruses obtained 10% efficiency (Linney et al., 1999) and of I-SceI meganuclease obtained 30.5% efficiency (Thermes et al., 2002). On the other hand, fish transposons like SB obtained 5–31% efficiency (Davidson et al., 2003) and Tol2 obtained about 50% of the zebrafish progeny (Kawakami et al., 2004b).

The introduction of tissue-specific and/or inducible promoters could be used to obtain transgenic fish with some useful characteristics (increase growth rates, disease resistances, etc.), production of biopharmaceuticals or detection of toxic substances (Rocha et al., 2003–2005).

11. Possible use of TEs to identify genes important for aquaculture

The identification of genes important for aquaculture (for instance, those implicated in growth rates, feed conversion, resistance to diseases, etc.) could be approached by ITR tagging the germ line of the target species. A similar method to the natural onset of albino phenotypes in medaka by Tol2 insertional mutagenesis (Koga et al., 1996) and identical to the strategy used in mice to detect oncogenes (Dupuy et al., 2005) and/or to induce saturation mutagenesis (Keng et al., 2005).

To increase the number of insertional inactivations, the ITR constructs first inserted into the genome of new progeny can be further mobilised in the next generation by the addition of exogenous specific transposases to reinsert again into other genes. The inactivation of gene(s) by ITR insertional mutagenesis would result in an alteration of the selected phenotype. In this process, the inactivated gene(s) are tagged by the ITR construct and their sequence(s) can then be recovered by PCR and cloning (Collier et al., 2005a; Dupuy et al., 2005). Once the implicated gene(s) have been identified, traditional genetic selection methods could be applied to obtain the desirable traits.

12. Possible use of TEs to improve DNA vaccination

The existence of copies of TEs-related sequences in fish, might also influence transposition in DNA immunisation attempts using the TEs to induce somatic transposition.

For example, in the presence of endogenous active transposases, only the ITR construct plasmid would be required to obtain transposition of an antigenic gene, although that would be at the cost of a lower stability of the gene inserted. However, relatively lower stability might not be an important disadvantage in this case since there is no need to carry the gene to the next generation and the immunisation would probably last longer than if the gene had not been inserted.

On the other hand, in the presence of endogenous intact ITR sequences, the addition of exogenous TEs could induce their excessive mobilisation causing some gene insertion(s) but also massive cellular deaths. Cellular death could stimulate local inflammation of the tissues affected which in turn would increase cellular defenses and immunisation (Garver et al., 2005).

All these possible effects remain to be investigated but they might be viable methods to improve DNA vaccination in fish by use of TEs. Furthermore, if required to avoid excessive activity, inducible expression systems allowing precise control of temporal expression of the exogenous transposases could also

be incorporated to the helper plasmids (Rocha et al., 2005, 2003).

13. Concluding remarks

To find complete transposase and ITR sequences in any of the copies existing in vertebrates, has been a difficult task because of the extremely high background of incomplete or defective copies prevalent in the actual genomes. Nevertheless complete, hypothetically active copies of Tc1-like TEs, have been detected in flat fish (Tpp1, *Pleuronectes platessa*) (Leaver, 2001) and in *Salvelinus* (Tsn1, *Salvelinus namaycush*) (Reed, 1999). Also, active transposition has been demonstrated to occur during fish gametogenesis although only in zebrafish (Tc1-like TE: Tzf, zebrafish) (Lam et al., 1996) and in medaka (hAT-like TE: Tol2, *Oryzias latipes*) (Koga et al., 1996). Therefore, autonomous TEs do exist in fish and it seems likely that other active TEs might be discovered in the future in the genomes of commercial fish (Ivics and Izsvak, 2002).

As an example of another strategy, active TEs are being recovered by synthetic reconstruction throughout sequence comparison and by site-directed mutagenesis of incomplete copies. The best studied example is the reconstructed SB. SB has been mainly studied to be used for human gene therapy vectors (transformation of somatic cells) or for gene-promoter identification (transformation of one cell embryos in model animals) but it could also be applied to fish (Table 3).

Fish TEs offer new possibilities, many of them yet to be investigated, to increase aquaculture yields by manipulating the germ line of commercially important fish such as for the generation of a large number of transgenics (Davidson et al., 2003) to select among those the ones to increase production, resistance to diseases or to obtain biopharmaceuticals (Rocha et al., 2003) or to identify important genes by TE tagging to then apply traditional genetic selection methods.

The TEs might also be used for fish somatic transposition such as it is required to improve fish immunisation to develop new DNA vaccination vectors (Fernandez-Alonso et al., 2001; Lorenzen et al., 1998; Purcell et al., 2004; Takano et al., 2004; Traxler et al., 1999). Thus, the TEs could be used to decrease the dosage of the vectors now required for fish immunisation to economically viable levels or to increase the duration of the immune response to the required time in culture of the fish to be commercialised (Coll, 2001; Lorenzen and LaPatra, 2005) (Table 3).

Understanding the dynamics and evolution of fish TEs will allow to use them, to develop derivatives for further genetic manipulation or to design new biotechnological applications (Ivics et al., 2004), including those to aquaculture (Coll, 2001).

Acknowledgements

Thanks are given to Dr. R. Blasco by the experiments on nuclear localisation of SB10 by using SB10-GFP fusion constructs. We would like also to recognise here the helpful commentaries of all the reviewers of Journal of Biotechnology that have seen the submitted paper and have contributed to its improvement. This work was supported by the CICYT projects AGL2004-07404-02-01/ACU and AGL2005-339-01/ACU and the INIA projects RTA03217 and CPE03-016-C3, Spain.

References

Balciunas, D., Davidson, A.E., Sivasubbu, S., Hermanson, S.B., Welle, Z., Ekker, S.C., 2004. Enhancer trapping in zebrafish using the sleeping beauty transposon. BMC Genomics 5, 62–63.

Britten, R.J., Kohne, D.E., 1968. Repeated sequences in DNA. Science 161, 529–540.

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

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

Coll, J.M., 2001. El transposon SB de salmonidos como vector para transferencia de genes en vertebrados. Investigaciones Agrarias 16, 237–244.

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

Collier, L.S., Carlson, C.M., Ravimohan, S., Dupuy, A.J., Largaespada, D.A., 2005b. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 436, 221–226.

Cui, Z., Geurts, A.M., Liu, G., Kaufman, D., Hackett, P.B., 2002. Structure–function analysis of the inverted terminal repeats of the sleeping beauty transposon. J. Mol. Biol. 318, 1221–1235.

Davidson, A.E., Balciunas, D., Mohn, D., Shaffer, J., Hermanson, S., Sivasubbu, S., Cliff, M.P., Hackett, P.B., Ekker, C.,

2003. Efficient gene delivery and gene expression in zebrafish using the sleeping beauty transposon. Dev. Biol. 263, 191–202.

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

Dupuy, A.J., Clark, K., Carlson, C.M., Fritz, S., Davidson, A.E., Markley, K.M., Finley, K., Fletcher, C.F., Ekker, S.C., Hackett, P.B., Hom, S., Largaespeda, D.A., 2002. Mammalian germ-line transgenesis by transposition. Proc. Natl. Acad. Sci. U.S.A. 99, 4495–4499.

Essner, J.J., McIvor, R.S., Hackett, P.B., 2005a. Awakening gene therapy with sleeping beauty transposons. Curr. Opin. Pharmacol. 5, 513–519.

Essner, J.J., McIvor, R.S., Hackett, P.B., 2005b. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. Curr. Opin. Pharmacol. 436, 272–276.

Fernandez-Alonso, M., Rocha, A., Coll, J.M., 2001. DNA vaccination by immersion and ultrasound to trout viral haemorrhagic septicaemia virus. Vaccine 19, 3067–3075.

Feschotte, C., 2004. Merlin, a new superfamily of DNA transposons identified in deiverse animal genomes and related to bacterial IS1016 insertion sequences. Mol. Biol. Evol. 21, 1769–1780.

Fischer, S.E., Wienholds, E., Plasterk, R.H., 2001. Regulated transposition of a fish transposon in the mouse germ line. Proc. Natl. Acad. Sci. U.S.A. 98, 6759–6764.

Garver, K.A., Conway, C.M., Elliot, D.G., Kurath, G., 2005. Analysis of DNA-vaccinated fish reveals viral antigen in muscle, kidney and thymus and transient histopathological changes, Mar. Biotechnol.

Geurts, A.M., Yang, Y., Clark, K.J., Liu, G., Cui, Z., Dupuy, A.J., Bell, J.B., Largaespada, D.A., Hackett, P.B., 2003. Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol. Ther. 8, 108–117.

Gollin, G., Amsteam, A., Sun, Z., Antonelly, A., Nissen, R.M., Hopkins, N., 2002. Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nat. Genet. 31, 135–140.

Goodier, J.L., Davidson, S.D., 1994. Tc1 transposon-like sequences are widely distributed in salmonids. J. Mol. Biol. 241, 26–34.

Gottgens, B., Barton, L.M., Grafham, D., Vaudin, M., Green, A.R., 1999. Tdr2, anew zebrafish transposon of the Tc1 family. Gene 239, 373–379.

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

Hackett, P.B., Alvarez, M.C., 2000. The molecular genetics of transgenic fish. Recent advances in Marine Biotechnology. In: Fingerman, M., Nagabhushanam, R. (Eds.), Aquaculture. Part B Fishes 4, pp. 77–145.

Hammer, S.E., Strehl, S., Hageman, S., 2005. Homologs of Drosophila P transposons were mobile in zebrafish but have been domesticated in a common ancestor of chicken and human. Mol. Biol. Evol. 22, 833–844.

Heierhorst, J., Lederis, K., Richter, D., 1992. Presence of a member of the Tc1-like transposon family from nematodes and Drosophila within the vasotocin gene of a primitive vertebrate, the Pacific hagfish *Eptatretis stouti*. Proc. Natl. Acad. Sci. U.S.A. 89, 6798–6802.

Henikoff, J., 1992. Detection of Caenorhabditis transposon homologs in diverse organisms. N. Biol. 4, 382–388.

Horie, K., Yusa, K., Yae, K., Odajima, J., Fisher, S.E., Keng, V.W., Hayakawa, T., Mizuno, A., Kondoh, G., Ijiri, T., Matsuda, Y., Plasterk, R.H., Takeda, J., 2003. Characterization of sleeping beauty transposition and its application to genetic screening in mice. Mol. Cell Biol. 23, 9189–9207.

Hronek, B.W., Meagher, A., Rovnak, J., Quackenbush, S.L., 2004. Identification and characterization of cis-acting elements residing in the Walleye dermal sarcoma virus promoter. J. Virol. 78, 7590–7601.

Iida, A., Tachibana, A., Hamada, S., Hori, H., Koga, A., 2004. Low transpsotion frequency of the medaka fish Tol2 element may be due to extranuclear localization of its transposase. Genes Genet. Syst. 79, 119–124.

Ivics, Z., Izsvak, Z., 2002. Transposable elements for transgenesis and insertional mutagenesis in vertebrates. A contemporary review of experimental strategies. Meth. Mol. Biol. 260, 255–276.

Ivics, Z., Izsvak, Z., Hackett, P.B., 1996a. Repeated sequence elements in zebrafish and their use in molecular genetic studies. Zebrafish Sci. Monitor 3, http:www//zfin.org/zfinfo/monitor/vol3.6/3.html.

Ivics, Z., Izsvak, Z., Hackett, P.B., 1997. Molecular reconstruction of sleeping beauty, a Tc1-like transposon system from fish and its transposition in human cells. Cell 91, 501–510.

Ivics, Z., Izsvak, Z., Minter, A., Hackett, P.B., 1996b. Identification of functional domains and evolution of Tc1-like transposable elements. Proc. Natl. Acad. Sci. U.S.A. 93, 5008–5013.

Ivics, Z., Kaufman, C.D., Zayed, H., Miskey, C., Wallsko, O., Izsvak, Z., 2004. The sleeping beauty transposable element: evolution, regulation and genetic applications. Curr. Issues Mol. Biol. 6, 43–56.

Izsvak, Z., Ivics, Z., 2003. Sleeping beauty transposition: biology and applications for molecular therapy. Mol. Ther. 9, 147–156.

Izsvak, Z., Ivics, Z., Hackett, P.B., 1995. Characterization of a Tc1-like transposable element in zebrafish (*Danio rerio*). Mol. Gen. Genet. 247, 312–322.

Izsvak, Z., Ivics, Z., Plasterk, R.H., 2000. Sleeping beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J. Mol. Biol. 302, 93–102.

Izsvak, Z., Ivics, Z., Shimoda, N., Mohn, D., Okamoto, H., Hackett, P.B., 1999. Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. J. Mol. Evol. 48, 13–21.

Izsvak, Z., Khare, D., Behlke, J., Heinemann, U., Plasterk, R.H., Ivics, Z., 2002. Involvement of a bifunctional, paired like DNA-binding domain and a transpositional enhancer in sleeping beauty transposition. J. Biol. Chem. 277, 34581–34588.

Izsvak, Z., Stuwe, E.E., Fuiedler, D., Katzer, A., Jeggo, P.A., Ivics, Z., 2004. Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. Mol. Cell 13, 279–290.

Kapitonov, V.V., Jurka, J., 2004. Herbinger transposons and an ancient HARB11 gene derived from a transposase. DNA Cell Biol. 23, 311–324.

Karsi, A., Moav, B., Hackett, P., Liu, Z., 2001. Effects of insert size on transposition efficiency of the sleeping beauty transposon in mouse cells. Mar. Biotechnol. (NY) 3, 241–245.

Kawakami, K., Imanaka, K., Itoh, M., Taira, M., 2004a. Excision of the Tol2 transposable element of the medaka fish *Oryzias latipes* in *Xenopus laevis* and *Xenopus tropiclis*. Gene 338, 93–98.

Kawakami, K., Noda, T., 2004. Transposition of the Tol2 elelment, an Ac-like element from the Japanese medaka fish *Oryzias latipes*, in mouse embryonic stem cells. Genetics 166, 895–899.

Kawakami, K., Shima, A., Kawakami, N., 2000. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. Proc. Natl. Acad. Sci. U.S.A. 97, 11403–11408.

Kawakami, K., Takeda, H., Kawakami, N., Kobayasi, M., Matsuda, N., Mishina, M., 2004b. A transposon-mediated gene trap approach identifies developmental regulated genes in zebrafish. Dev. Cell 7, 133–144.

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

Keng, V.W., Yae, K., Hayakawa, T., Mizuno, S., Uno, Y., Yusa, K., Kokubu, C., Kinoshita, T., Akagi, K., Jenkins, N.A., Copeland, N.G., Horie, K., Takeda, J., 2005. Region-specific saturation germline mutagenesis in mice using the Sleeping Beauty transposon system. Nat. Meth. 2, 763–769.

Kido, Y., Aono, M., Yamaki, T., Matsumoto, K., Murata, S., Saneyoshi, M., Okada, N., 1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. Proc. Natl. Acad. Sci. U.S.A. 88, 2326–2330.

Kido, Y., Himberg, M., Takasaki, N., Okada, N., 1994. Amplification of distinct subfamilies of short interdispersed elements during evolution of the salmonidae. J. Mol. Biol. 241, 633–644.

Koga, A., 2004. Transposition mechanisms and biotechnology applications of the medaka fish Tol2 transposable element. Adv. Biophys. 38, 161–180.

Koga, A., Hori, H., 1999. Homogeneity in the structure of the medaka fish transposable element Tol2. Genet. Res. 73, 7–14.

Koga, A., Hori, H., 2000. Detection of de novo insertion of the medaka fish transposable element Tol2. Genetics 156, 1243–1247.

Koga, A., Suzuki, M., Inagaki, H., Bessho, Y., Hori, H., 1996. Transposable element in fish. Nature 383, 30.

Lam, W.L., Lee, T.S., Gilbert, W., 1996. Active transposition in zebrafish. Proc. Natl. Acad. Sci. U.S.A. 93, 10870–10875.

Lampe, D.J., Churchill, M.E.A., Robertson, H.M., 1996. A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO J. 15, 5470–5479.

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.

Linney, E., Hardison, N.L., Lonze, B.E., Lyons, S., DiNapoli, L., 1999. Transgene expression in zebrafish: a comparison of retroviral vector and DNA injection approaches. Dev. Biol. 213, 207–216.

Liu, G., Aronovich, E.L., Cui, Z., Whitley, C.B., Hackett, P.B., 2004. Excision of sleeping beauty transposons: parameters and applications to gene therapy. J. Gene Med. 6, 574–583.

Lorenzen, N., LaPatra, S.E., 2005. DNA vaccines for aquacultured fish. Rev. Sci. Technol. Office Int. Epizooties 24, 201–213.

Lorenzen, N., Lorenzen, E., Eoner-jensen, K., Heppell, J., Wu, T., Davis, H., 1998. Protective immunity to VHS in rainbow trout (*Oncorhynchus mykiss*, Walbaum) following DNA vaccination. Fish Shellfish Immunol. 8, 261–270.

Masuda, K., Yamamoto, S., Endoh, M., Kaneda, Y., 2004. Transposon-independent increase of transcription by the sleeping beauty transposase. Biochem. Biophys. Res. Commun. 317, 796–800.

Matsumoto, K., Murakami, K., Okada, N., 1984. Pseudouridylic modification of a 6S RNA transcribed in vitro from highly repetitive and transcribable (Hirt) sequences of salmon total DNA. Biochem. Biophys. Res. Commun. 124, 514–522.

Matsumoto, K., Murakami, K., Okada, N., 1986. Gene for lysine tRNA1 may be a progenitor of the highly repetitive and transcribable sequences present in the salmon genome. Proc. Natl. Acad. Sci. U.S.A. 83, 3156–3160.

Mikkelsen, J.G., Yant, S.R., Meuse, L., Huang, Z., Xu, H., Kay, M.A., 2003. Helpar-independent sleeping beauty transposon-transposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo. Mol. Ther. 8, 654–665.

Miskey, C., Izsvak, Z., Kawakami, K., Ivics, Z., 2005. DNA transposons invertebrate functional genomics. Cell Mol. Life Sci. 62, 629–641.

Miskey, C., Izsvak, Z., Plasterk, R.H., Ivics, Z., 2003. The frog prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. Nucl. Acids Res. 31, 6873.

Ohshima, K., Okada, N., 2005. SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. Cytogenet. Genome Res. 110, 475–490.

Plasterk, R.H.A., Izsvak, Z., Ivics, Z., 1999. Resident Alliens. The Tc1/mariner superfamily of transposable elements. TIG 15, 326–332.

Poulet, F.M., Bowser, P.R., Casey, J.W., 1994. Retroviruses of fish, reptiles and molluscs. In: Levy, J.A. (Ed.), The Retroviridae, vol. 3. Plenum Press, NY, pp. 1–38.

Purcell, M.K., Kurath, G., Garver, K.A., Herwig, R.P., Winton, J.R., 2004. Quantitative expression profiling of imune response genes in rainbow trout following infectious haematopoietic necrosis virus (IHNV) infection or DNA vaccination. Fish Shellfish Immunol. 17, 447–462.

Radice, A.D., Bugaj, B., Fitch, D.H., Emmons, S.W., 1994. Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. Mol. Gen. Genet. 244, 606–612.

Raz, E., vanLuenen, H.G.A.M., Schaerringer, B., Plasterk, R.H.A., Driever, W., 1997. Transposition of the nematode Caenorhabditis elegans Tc3 element in the zebrafish *Danio rerio*. Curr. Biol. 8, 82–88.

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

Rocha, A., Ruiz, S., Coll, J.M., 2005. Improvement of transfection efficiency of epithelioma papulosum cyprini carp cells by

modification of their cell cycle and using an optimal promoter. Mar. Biotechnol. 6, 401–410.

Rocha, A., Ruiz, S., Estepa, A., Coll, J.M., 2003. Fish as biofactories: inducible genetic systems and gene targeting. Spanish J. Agric. Res. 1, 3–11.

Rocha, A., Ruiz, S., Estepa, A., Coll, J.M., 2004. Application of inducible and targeted gene-strategies to produce transgenic fish: a review. Mar. Biotechnol. 6, 118–127.

Sarkar, A., Sim, C., Hong, Y.S., Hogan, J.R., Fraser, M.J., Robertson, H.M., Collins, F.H., 2003. Molecular evolutionary analysis of the widespread piggyBac transposon family and related "domesticated" sequences. Mol. Genet. Genomics 270, 173–180.

Shen, C.H., Steiner, L.A., 2004. Genome structure and thymic expression of an endogenous retrovirus in zebrafish. J. Virol. 78, 899–911.

Sherratt, D.J., 1995. Mobile Genetic Elements, vol. 1. IRL Press, Oxford University Press, Oxford, United Kingdom, pp. 1–179.

Shimoda, N., Chevrette, M., Ekker, M., Kikuchi, Y., Hotta, Y., Okamoto, H., 1996. Mermaid: a family of short interdispersed repetitive elements widespread in vertebrates. Biochem. Biophys. Res. Commun. 220, 226–232.

Smit, A.F.A., Riggs, A.D., 1995. MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. Nucl. Acids Res. 23, 98–102.

Stuart, G.W., Vielkind, J.R., McMurray, J.V., Westerfield, M., 1990. Stable lines of transgenic zebrafish exhibit reproducible patterns of transgene expression. Development 109, 577–584.

Tafalla, C., Estepa, A., Coll, J.M. Development and possible applications to aquaculture of Sleeping Beauty, the first active vertebrate transposon obtained from fish. Aquaculture, submitted for publication.

Takahashi, K., Terai, Y., Nishida, M., Okada, N., 1998. A novel family of short interspersed repetitive elements (SINEs) from cichlids: the patterns of inserttion of SINEs at orthologous loci support the proposed monophyly of four major groups of cichlid fishes in lake Tanganyika. Mol. Biol. Evol. 15, 391–407.

Takano, T., Iwahori, A., Hirono, I., Aoki, T., 2004. Development of a DNA vaccine against hirame rhabdovirus and analysis of the expression of immune-related genes after vaccination. Fish Shellfish Immunol. 17, 367–374.

Terai, Y., Takahashi, K., Okada, N., 1998. SINE cousins: the 3′-end tails of the two oldest and distantly related families of SINEs are descendent from the 3′ ends of LINEs with the same genealogical origin. Mol. Biol. Evol. 15, 1460–1471.

Thermes, V., Grabher, C., Ristoratore, F., Bourrat, F., Choulika, A., Wittbrodt, J., Joly, J.S., 2002. I-SceI meganuclease mediates highly efficient transgenesis in fish. Mech. Dev. 118, 91–98.

Traxler, G.S., Anderson, E., LaPatra, S.E., Richard, J., Shewmaker, B., Kurath, G., 1999. Naked DNA vaccination of Atlantic salmon Salmosalar against IHNV. Dis. Aquatic Organ. 38, 183–190.

Tsutsumi, M., Koga, A., Hori, H., 2003. Long and short mRNAs transcribed from the medaka fish transposon Tol2 respectively exert positive and negative effects on excision. Genet. Res. Camb. 82, 33–40.

Wadman, S.A., Clark, K.J., Hackett, P.B., 2005. Fishing for answers with transposons. Mar. Biotechnol. 7, 135–141.

Westerfield, M., Wegner, J., Jegalian, B.G., DeRobertis, E.M., Puschel, A.W., 1992. Specific activation of mammalian Hox promoters in mosaic transgenic zebrafish. Genes Dev. 6, 591–598.

Winkfein, R.J., Moir, R.D., Krawetz, S.A., Blanco, J., States, J.C., Dixon, G.H., 1988. A new family of repetitive retroposon-like sequences in the genomes of the rainbow trout. Eur. J. Biochem. 176, 255–264.

Yant, S.R., Kay, M.A., 2003. Nonhomologous-end-joining factors regulate DNA repair fidelity during sleeping beauty element transposition in mammalian cells. Mol. Cell. Biol. 23, 8505–8518.

Yant, S.R., Meuse, L., Chiu, W., Ivics, Z., Izsvak, Z., Kay, M.A., 2000. Somatic integration and longterm transgene expression in normal and haemophilic mice using a DNA transposon system. Nat. Genet. 25, 35–41.

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

Yoshida, M., Yagi, T., Furuta, Y., Takayanagi, K., Kominami, R., Takeda, N., Tokunaga, T., Chiba, J., Ikawa, Y., Aizawa, S., 1995. A new strategy of gene trapping in ES cells using RACE. Transgenic Res. 4, 277–287.

Yusa, K., Takeda, J., Horie, K., 2004. Enhancement of sleeping beauty transposition by CpG methylation: possible role of heterochromatin formation. Mol. Cell Biol. 24, 4004–4018.

Zambrowicz, B.P., Friedrich, G.A., Buxton, E.C., Lilleberg, S.L., Person, C., Sands, A.T., 1998. Disruption and sequence identification of 2000 genes in mouse embryonic cells. Nature 392, 608–611.

Zayed, H., Izsvak, Z., Khare, D., Heinemann, U., Ivics, Z., 2003. The DNA-bending protein HMGB1 is a cellular cofactor of sleeping beauty transposition. Nucl. Acids Res. 31, 2313–2322.

Zayed, H., Izsvak, Z., Walisko, O., Ivics, Z., 2004. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Mol. Ther. 9, 292–304.