Molecular Reconstruction of *Sleeping Beauty*, a *Tc1*-like Transposon from Fish, and Its Transposition in Human Cells

Zoltán Ivics,[*][‖] **Perry B. Hackett,**[*][†]
Ronald H. Plasterk,[‡] **and Zsuzsanna Izsvák**[*][§][‖][#]
[*]Department of Genetics and Cell Biology
[†]Institute of Human Genetics
University of Minnesota
St. Paul, Minnesota 55108-1095
[‡]Division of Molecular Biology
Netherlands Cancer Institute
Amsterdam, 1066CX
The Netherlands
[§]Institute of Biochemistry
Biological Research Center of Hungarian
    Academy of Sciences
Szeged 6701
Hungary

Summary

Members of the *Tc1/mariner* superfamily of transposons isolated from fish appear to be transpositionally inactive due to the accumulation of mutations. Molecular phylogenetic data were used to construct a synthetic transposon, *Sleeping Beauty*, which could be identical or equivalent to an ancient element that dispersed in fish genomes in part by horizontal transmission between species. A consensus sequence of a transposase gene of the salmonid subfamily of elements was engineered by eliminating the inactivating mutations. *Sleeping Beauty* transposase binds to the inverted repeats of salmonid transposons in a substrate-specific manner, and it mediates precise cut-and-paste transposition in fish as well as in mouse and human cells. *Sleeping Beauty* is an active DNA-transposon system from vertebrates for genetic transformation and insertional mutagenesis.

Introduction

Apart from their impact on genome organization and evolution, transposable elements may serve as valuable tools for genetic analyses. In vertebrates, the discovery of DNA-transposons, mobile elements that move via a DNA intermediate, is relatively recent (Radice et al., 1994). Since then, members of the *Tc1/mariner* as well as the hAT (*hobo/Ac/Tam*) superfamilies of eukaryotic transposons have been isolated from different fish species, Xenopus, and human genomes (Oosumi et al., 1995; Ivics et al., 1996; Koga et al., 1996; Lam et al., 1996b), and the enormous potential of these elements to investigate vertebrate genomes has been constantly under discussion.

These transposable elements transpose through a cut-and-paste mechanism; the element-encoded transposase catalyzes the excision of the transposon from its original location and promotes its reintegration elsewhere in the genome (Plasterk, 1996). Autonomous members of a transposon family can express an active transposase, the *trans*-acting factor for transposition, and thus are capable of transposing on their own. Nonautonomous elements have mutated transposase genes but may retain *cis*-acting DNA sequences necessary for transposition. These sequences, usually embedded in the terminal inverted repeats (IRs) of the elements, are required for mobilization in the presence of a complementary transposase. These features make DNA-transposons potential vectors for moving exogenous DNA sequences into chromosomes.

Unfortunately, not a single autonomous element has been isolated from vertebrates; they seem to be defective for having mutations as a result of a process called "vertical inactivation" (Lohe et al., 1995). According to a phylogenetic model (Hartl et al., 1997), the ratio of nonautonomous to autonomous elements in eukaryotic genomes increases due to the *trans*-complementary nature of transposition. This process leads to a state where the ultimate disappearance of active, transposase-producing copies in a genome is inevitable. Consequently, DNA-transposons can be viewed as transitory components of genomes, which, in order to avoid extinction, must find ways to establish themselves in a new host. Indeed, horizontal gene transmission between species is thought to be one of the important processes in the evolution of transposons (Kidwell, 1992; Lohe et al., 1995).

The natural process of horizontal gene transfer can be mimicked under laboratory conditions. In plants, transposable elements of the *Ac/Ds* and *Spm* families have been routinely introduced into heterologous species (Osborne and Baker, 1995). In animals, however, a major obstacle to the transfer of an active transposon system from one species to another has been that of species-specificity of transposition due to the requirement for factors produced by the natural host. For this reason, attempts to use the *P* element transposon of Drosophila melanogaster for genetic transformation of nondrosophilid insects, zebrafish, and mammalian cells have been unsuccessful (Rio et al., 1988; Handler et al., 1993; Gibbs et al., 1994). In contrast to *P* elements, there are indications that members of the *Tc1/mariner* superfamily of transposable elements may not be as demanding for species-specific factors for their transposition. First, these elements are extraordinarily widespread in nature, ranging from single-cellular organisms to humans (Plasterk, 1996). In addition, recombinant Tc1 and mariner transposases expressed in E. coli are sufficient to catalyze transposition in vitro (Lampe et al., 1996; Vos et al., 1996). Furthermore, gene vectors based on *Minos*, a Tc1-like element (TcE) endogenous to Drosophila hydei, were successfully used for germline transformation of the fly Ceratitis capitata (Loukeris et al., 1995), and the *mariner* element from Drosophila mauritiana was capable of undergoing transposition in the protozoan Leishmania (Gueiros-Filho and Beverley, 1997).

[#]To whom correspondence should be addressed.
[‖]Present address: Division of Molecular Biology, Netherlands Cancer Institute, Amsterdam, 1066CX, The Netherlands.
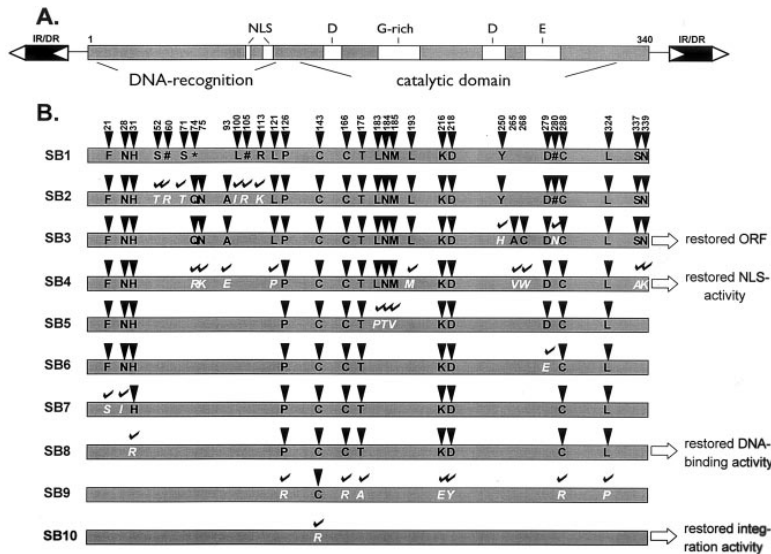
Figure 1. Molecular Reconstruction of a Salmonid *Tc1*-like Transposase Gene

(A) Schematic map of a salmonid TcE with the conserved domains in the transposase and IR/DR flanking sequences.

(B) The strategy of first constructing an open reading frame for a salmonid transposase and then systematically introducing amino acid replacements into this gene is illustrated. Amino acid residues are typed black when different from the consensus, and their positions within the transposase polypeptide are indicated with arrows. Translational termination codons appear as asterisks; frame shift mutations are shown as number signs. Amino acids changed to the consensus are checkmarked and typed white italic. In the right margin, the various functional tests that were done at certain stages of the reconstruction procedure are indicated.

For the above reasons, members of the *Tc1/mariner* superfamily are valuable candidates for being developed as wide(r) host-range transformation vectors.

There can be two major strategies to obtain an active transposon system for any organism: find one or make one. From all DNA-transposons found so far in vertebrates, TcEs from teleost fish are by far the best characterized (Goodier and Davidson, 1994; Radice et al., 1994; Izsvak et al., 1995; Ivics et al., 1996; Lam et al., 1996a). In the course of our search for an active fish transposon, we characterized *Tdr1* in zebrafish (Danio rerio) (Izsvak et al., 1995), as well as other closely related TcEs from nine additional fish species (Ivics et al., 1996). Similarly to all other TcEs, the fish elements can be typified by a single gene encoding a transposase enzyme flanked by IRs. Unfortunately, all of the fish elements isolated so far are probably inactive due to several mutations in their transposase genes. Molecular phylogenetic analyses have shown that the majority of the fish TcEs can be classified into three major types: zebrafish-, salmonid- and Xenopus *TXr*-type elements (Izsvak et al., 1997), of which the salmonid subfamily is probably the youngest and thus most recently active (Ivics et al., 1996). In addition, examination of the phylogeny of salmonid TcEs and that of their host species provided important clues about the ability of this particular subfamily of elements to invade and establish permanent residences in naive genomes through horizontal transfer, even over relatively large evolutionary distances (Ivics et al., 1996).

There are two fundamental components of any mobile cut-and-paste type transposon system: a source of an active transposase, and the DNA sequences that are recognized and mobilized by the transposase. Both the transposase coding regions and the IRs of salmonid-type TcEs accumulated several mutations, including point mutations, deletions, and insertions, and they show about 5% average pairwise divergence (Ivics et al., 1996). We used the accumulated phylogenetic data to reconstruct a transposase gene of the salmonid subfamily of fish elements. We expressed this transposase and show that it is capable of catalyzing transposition of engineered, nonautonomous salmonid elements in fish as well as in mammalian cells. This transposon system, which was awakened from a long evolutionary sleep and which we named *Sleeping Beauty* (*SB*), can be developed as a powerful tool for germline transformation and insertional mutagenesis in vertebrates, with the potential to be applicable in other organisms as well.

Results

The Transposase: Reconstruction of a Transposase Gene

Sequence alignment of 12 partial salmonid-type TcE sequences found in 8 fish species (available under DS30090 from FTP.EBI. AC.UK in directory pub/databases/embl/align) allowed us to derive a majority-rule consensus sequence and to identify conserved protein and DNA sequence motifs that likely have functional importance (Figure 1A).

Conceptual translation of the mutated transposase open reading frames revealed five regions that are highly conserved in all TcE transposases: a bipartite nuclear localization signal (NLS) in the N-terminal half of the transposase with a possible overlap with the DNA-binding domain (Ivics et al., 1996), a glycine-rich motif close to the center of the transposase without any known function at present, and three segments in the C-terminal half comprising the DDE domain (Doak et al., 1994) that catalyzes the transposition (Craig, 1995) (Figure 1A). Multiple sequence alignment also revealed a fairly random distribution of mutations in transposase coding sequences; 72% of the base pair changes had occurred at nonsynonymous positions of codons. The highest mutation frequencies were observed at mutable CpG dinucleotide sites (Yoder et al., 1997). Although amino acid substitutions were distributed throughout the transposases, fewer mutations were detected at the conserved motifs (0.07 nonsynonymous mutation per codon), as compared to protein regions between the conserved domains (0.1 nonsynonymous mutation per codon). This implies that some selection had maintained
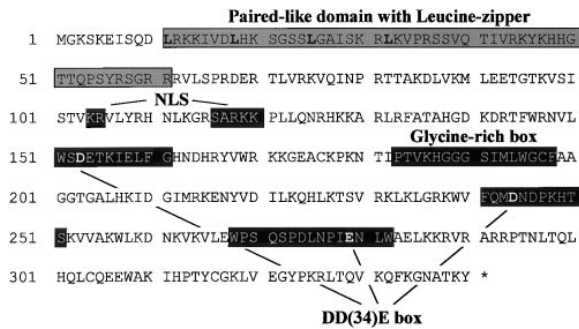
Paired-like domain with Leucine-zipper

```
  1  MGKSKEISQD LRKKIVDLHK SGSSLGAISK RLKVPRSSVQ TIVRKYKHHG

 51  TTQPSYRSGR RRVLSPRDER TLVRKVQINP RTTAKDLVKM LEETGTKVSI
               ━━ NLS ━━
101  STVKRVLYRH NLKGRSARKK PLLQNRHKKA RLRFATAHGD KDRTFWRNVL
                                            Glycine-rich box
151  WSDETKIELF GHNDHRYVWR KKGEACKPKN TIPTVKHGGG SIMLWGCFAA

201  GGTGALHKID GIMRKENYVD ILKQHLKTSV RKLKLGRKWV FQMDNDPKHT

251  SKVVAKWLKD NKVKVLEWPS QSPDLNPIEN LWAELKKRVR ARRPTNLTQL

301  HQLCQEEWAK IHPTYCGKLV EGYPKRLTQV KQFKGNATKY *
                        DD(34)E box
```

Figure 2. Amino Acid Sequence of the Consensus SB10 Transposase

The major functional domains are highlighted.

the functional domains before inactivation of transposons took place in host genomes. The identification of these putative functional domains was of key importance during the reactivation procedure.

The first step of reactivating the transposase gene was to restore an open reading frame (SB1 through SB3 in Figure 1B) from bits and pieces of two inactive TcEs from Atlantic salmon (Salmo salar) and a single element from rainbow trout (Oncorhynchus mykiss) (Radice et al., 1994), by removing the premature translational stop codons and frameshifts. SB3, a complete open reading frame, was tested in an excision assay similar to that described in Handler et al. (1993), but no detectable activity was observed. Due to nonsynonymous nucleotide substitutions, the SB3 polypeptide differs from the consensus transposase sequence in 24 positions (Figure 1B), from which nine are in the putative functional domains and therefore are probably essential for transposase activity. Consequently, we undertook a dual gene reconstruction strategy. First, the putative functional protein domains of the transposase were systematically rebuilt one at a time, and then their corresponding biochemical activities were tested independently. Second, in parallel with the first approach, a full-length gene was synthesized by extending the reconstruction procedure to all of the 24 mutant amino acids in the putative transposase.

Accordingly, a series of constructs was made to bring the coding sequence closer, step-by-step, to the consensus using PCR mutagenesis (SB4-SB10 in Figure 1B). As a general approach, the sequence information predicted by the majority-rule consensus was followed. However, at some codons deamination of ^{5m}C residues of CpG sites occurred, and C→T mutations had been fixed in many elements. In one position, where TpGs and CpGs were represented in equal numbers in the alignment, a CpG sequence was carefully chosen to encode R(288). The result of this extensive genetic engineering is a synthetic, putative transposase gene encoding 340 amino acids (SB10 in Figures 1B and 2).

The Transposable Substrate DNA

In contrast to the prototypic Tc1 transposon from Caenorhabditis elegans, which has short 54 bp IRs flanking its transposase gene, most TcEs from fish belong to the IR/DR subgroup of TcEs (Izsvak et al., 1995; Ivics

et al., 1996), which have 210–250 bp IRs at their termini and directly repeated DNA sequence motifs (DRs) at the ends of each IR (Figure 1A). However, the consensus IR sequences are not perfect repeats indicating that, in contrast to most TcEs, these fish elements naturally possess imperfect inverted repeats. The match is less than 80% at the center of the IRs, but it is perfect at the DRs, suggesting that this nonrandom distribution of dissimilarity could be the result of positive selections that have maintained functionally important sequence motifs in the IRs (Figure 1A). Therefore, we suspected that DNA sequences at and around the DRs might carry cis-acting information for transposition, and mutations within the IRs but outside the DRs would probably not impair the ability of the element to transpose. As a model substrate, we chose a single salmonid-type TcE from Tanichthys albonubes (hereafter referred to as T) whose sequence is only 3.8% divergent from the salmonid consensus (Ivics et al., 1996) and has intact DR motifs.

The T element and the putative, synthetic transposase expressed by the reconstructed SB10 transposase gene together constitute the Sleeping Beauty transposon system. The efficacy of our phylogenetic approach for constructing this active transposon is demonstrated by the activities of SB, including nuclear localization, DNA binding, and integration.

**DNA-Binding Activities of Sleeping
Beauty Transposase**

The reconstituted functional transposase domains were tested for activity. First, a short segment of the SB4 transposase gene (Figure 1B) encoding an NLS-like protein motif was fused to the lacZ gene. The transposase NLS was able to transfer the cytoplasmic marker protein, β-galactosidase, into the nuclei of cultured mouse cells (Ivics et al., 1996), supporting our supposition that we could predict both amino acid sequences and functions of domains and resurrect a full-length, multifunctional enzyme.

Once in the nucleus, a transposase must bind specifically to its recognition sequences in transposon DNA. The specific DNA-binding domains of both Tc1 and Tc3 transposases have been mapped to the N-terminal regions (Colloms et al., 1994; Vos and Plasterk, 1994), suggesting that analogous sequences are likely to encode specific DNA-binding functions in virtually all TcE transposases. Therefore, a gene segment encoding the first 123 amino acids of SB (N123), which includes the NLS, was reconstructed (SB8 in Figure 1B) and expressed in E. coli. N123 was purified via a C-terminal histidine tag as a 16 kDa polypeptide (Figure 3A).

Upon incubation of N123 with a radiolabeled 300bp DNA fragment comprising the left IR of T, nucleoprotein complexes were observed in a mobility shift assay (Figure 3B, left panel, lane 3), as compared to samples containing extracts of bacteria transformed with the expression vector only (lane 2) or probe without any protein (lane 1). Unlabeled IR sequences of T added in excess to the reaction as competitor DNA inhibited binding of the probe (lane 4), whereas the analogous region of a cloned Tdr1 element from zebrafish did not appreciably compete with binding (lane 5). Thus, N123 contains the
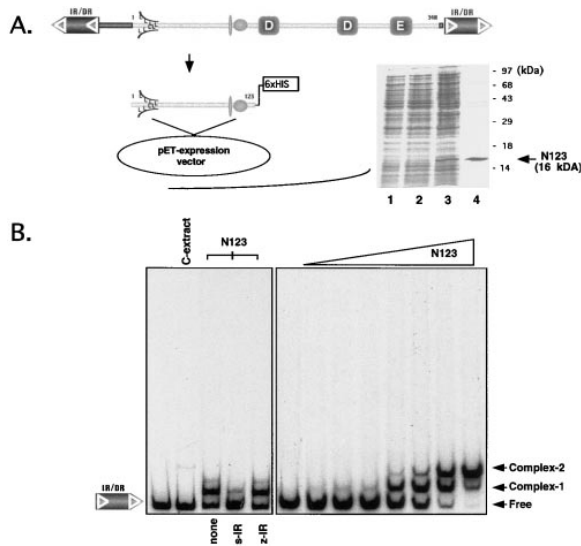
Figure 3. DNA-Binding Activities of an N-Terminal Derivative of the *Sleeping Beauty* Transposase

(A) SDS-PAGE analysis showing the steps in the expression and purification of N123. Lanes: (1) extract of cells containing expression vector pET21a; (2) extract of cells containing expression vector pET21a/N123 before induction with IPTG; (3) extract of cells containing expression vector pET21a/N123 after induction with IPTG; (4) partially purified N123 using Ni2+-NTA resin. Molecular weights in kDa are indicated on the right.

(B) Mobility shift analysis of the ability of N123 to bind to the inverted repeats of fish transposons. Lanes: (1) probe only; (2) extract of cells containing expression vector pET21a; (3) 10,000-fold dilution of the N123 preparation shown in lane 4 of Panel A; (4) same as lane 3 plus a 1000-fold molar excess of unlabeled probe as competitor DNA; (5) same as lane 3 plus a 1000-fold molar excess of an inverted repeat fragment of a zebrafish *Tdr1* element as competitor DNA; (6–13) 200,000-, 100,000-, 50,000-, 20,000-, 10,000-, 5,000-, 2,500-, and 1,000-fold dilutions of the N123 preparation shown in lane 4 of (A).

specific DNA-binding domain of SB transposase, which seems to be able to distinguish between salmonid-type and zebrafish-type TcE substrates.

The number of nucleoprotein complexes at increasingly higher N123 concentrations indicated two protein molecules bound per IR (Figure 3B, right panel), consistent with either two binding sites for the transposase within the IR or a transposase dimer bound to a single site. Transposase binding sites were further analyzed and mapped in a DNaseI footprinting experiment. Using the same fragment of *T* as above, two protected regions close to the ends of the IR probe were observed (Figure 4A). The two 30bp footprints cover the subterminal DR motifs within the IRs, thus the DRs are the core sequences for DNA-binding by N123. The DR motifs are almost identical between salmonid- and zebrafish-type TcEs (Izsvak et al., 1995). However, the 30bp transposase binding sites are longer than the DR motifs and contain 8 and 7 bp in the outer and internal binding sites, respectively, that are different between the zebrafish- and the salmonid-type IRs (Figure 4B).

Although there are two binding sites for transposase near the ends of each IR, apparently only the outer sites are utilized for DNA cleavage and thus for the excision

of the transposon. Sequence comparison shows that there is a 3bp difference in composition and a 2bp difference in length between the outer and internal transposase binding sites (Figure 4C). In summary, our synthetic transposase protein has DNA-binding activity, and this binding appears to be specific for salmonid-type IR/DR sequences.

Integration Activity of *Sleeping Beauty* in Heterologous Vertebrate Cells

In addition to the abilities to enter nuclei and specifically bind to its sites of action within the inverted repeats, a fully active transposase is expected to excise and integrate transposons. In the C-terminal half of the SB transposase, three protein motifs make up the DD(34)E catalytic domain, which contains two invariable aspartic acid residues, D(153) and D(244), and a glutamic acid residue, E(279), the latter two separated by 34 amino acids (Figure 2). An intact DD(34)E box is essential for catalytic functions of Tc1 and Tc3 transposases (van Luenen et al., 1994; Vos and Plasterk, 1994).

To detect chromosomal integration events into the chromosomes of cultured cells, a two-component assay system was established. The assay is based on *trans*-complementation of two nonautonomous transposable elements, one containing a selectable marker gene (donor) and another that expresses the transposase (helper) (Figure 5A). The donor, pT/neo, is an engineered, *T*-based element that contains an SV40 promoter-driven *neo* gene flanked by the terminal IRs of the transposon, containing binding sites for the transposase. The helper construct expresses the full-length *SB10* transposase gene driven by a human cytomegalovirus (CMV) enhancer/promoter. In the assay, the donor plasmid is cotransfected with the helper or control constructs into cultured vertebrate cells, and the number of cell clones that are resistant to the neomycin-analog drug G-418 due to chromosomal integration and expression of the *neo* transgene serves as an indicator of the efficiency of gene transfer. Host-requirements of transposase activity were assessed using three different vertebrate cell lines: EPC from carp, LMTK from mouse, and HeLa from human. Carp EPC cells are expected to provide a permissive environment for transposition, because the carp genome contains endogenous, but probably inactive, salmonid-type TcEs (Ivics et al., 1996). Moreover, if *SB* is not strictly host-specific, transposition could also occur in phylogenetically more distant vertebrate species. Using the assay system shown in Figure 5A, enhanced levels of transgene integration were observed in the presence of the helper plasmid: more than 2-fold in carp EPC cells (not shown), more than 5-fold in mouse LMTK cells (not shown), and more than 20-fold in human HeLa cells (Figures 5B and 6). Figure 5B shows five plates of transfected HeLa cells that were placed under G-418 selection and were stained with methylene blue two weeks posttransfection. The stainings clearly demonstrate a significant increase in integration of *neo*-marked transposons into the chromosomes of HeLa cells when the SB transposase-expressing helper construct was cotransfected (plate 2), as compared to a control cotransfection of the donor plasmid plus a construct in
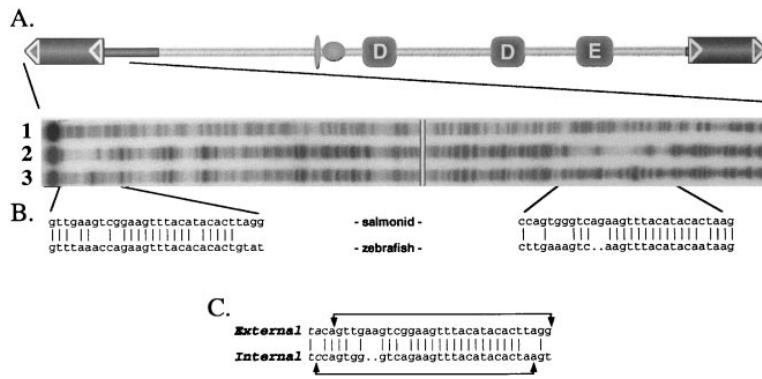
A.

B.
```
gttgaagtcggaagtttacatacacttagg        - salmonid -        ccagtgggtcagaagtttacatacactaag
|||  |  |||   |||  |||||||||||                           |||  ||| |||||||||||||||||
gtttaaaccagaagtttacacacactgtat        - zebrafish -       cttgaaagtc..aagtttacatacaataag
```

C.
```
External  tacagttgaagtcggaagtttacatacacttagg
          | ||||  ||| ||||||||||||||||||
Internal  cccagtgg..gtcagaagtttacatacactaagt
```

Figure 4. *Sleeping Beauty* Transposase Has Two Binding Sites in Each of the Inverted Repeats of Salmonid Transposons

(A) DNase I footprinting of nucleoprotein complexes formed by a 500-fold dilution of the N123 preparation shown in Figure 3A, lane 4, and the same probe as in Figure 3B. Maxam-Gilbert sequencing of purine bases in the same DNA was used as a marker (lane 1). Reactions were run in the presence (lane 2) or absence (lane 3) of N123. Only one strand of DNA is shown.

(B) Sequence comparison of the salmonid transposase binding sites shown in (A) with the corresponding sequences in zebrafish *Tdr1* elements.

(C) Sequence comparison between the outer and internal transposase binding sites in salmonid transposons.

which the transposase gene was cloned in an antisense orientation (pSB10-AS; plate 1). Consequently, SB transposase appears to be able to increase the efficiency of transgene integration, and this activity is not restricted to fish cells.

To map transposase domains necessary for chromosomal integration, a frameshift mutation was introduced into the *SB* transposase gene that brought a translational stop codon into frame nine codons following G(161). This construct, pSB10-ΔDDE, expresses a truncated transposase polypeptide that contains specific DNA-binding and NLS domains, but it lacks the catalytic domain. The transformation rates obtained using this construct (plate 3 in Figure 5B) were comparable to those obtained with the antisense control (Figure 6). This result confirms that the presence of a full-length transposase protein is necessary and that DNA-binding and nuclear transport activities themselves are not sufficient for the observed enhancement of transgene integration.

As a further control of transposase requirement, we tested the integration activity of an earlier version of the SB transposase, SB6, which differs from SB10 at 11 residues (Figure 1B). Again, the number of transformants observed using this construct (plate 4 in Figure 5B) was about the same as in the antisense control experiment (Figure 6), indicating that the amino acid replacements that we introduced into the transposase gene were critical for transposase function. In summary, the three controls shown in plates 1, 3, and 4 of Figure 5B establish the *trans*-requirements of enhanced, SB-mediated transgene integration.

But is this phenomenon also dependent on the DNA substrate used in our assay? One of the IRs of the *neo*-marked transposon substrate was removed, and the performance of this construct, pT/neo-ΔIR in Figure 5, was tested for integration. The transformation rates observed with this plasmid (plate 5 in Figure 5B) were more than 7-fold lower than those with the full-length donor (Figure 6). These results indicate that both IRs flanking the transposon are required for efficient transposition and thereby establish some of the *cis*-requirements of the two-component *SB* transposon system.

Taken together, the dependence on a full-length transposase enzyme and two inverted repeats at the ends of the transposon suggests that enhanced transgene
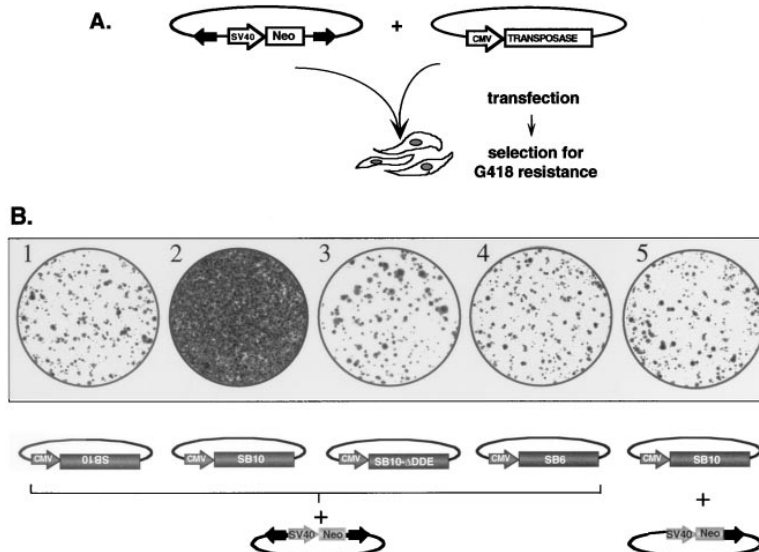

Figure 5. Integration Activity of *Sleeping Beauty* in Human HeLa Cells

(A) Genetic assay for *Sleeping Beauty*-mediated transgene integration in cultured cells.

(B) Petri dishes with stained colonies of G-418-resistant HeLa cells that have been transfected with different combinations of donor and helper plasmids. Plates: (1) pT/neo plus pSB10-AS; (2) pT/neo plus pSB10; (3) pT/neo plus pSB10-ΔDDE; (4) pT/neo plus pSB6; (5) pT/neo-ΔIR plus pSB10.
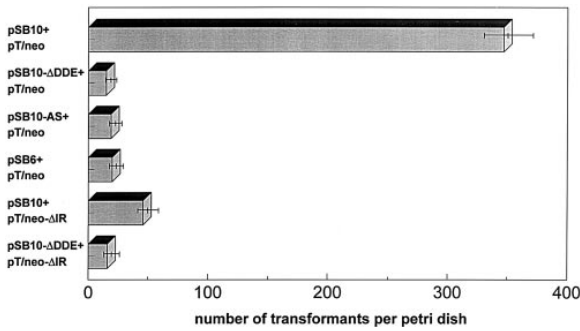
Figure 6. Enhanced Transgene Integration in Human HeLa Cells Is Dependent on Full-Length *Sleeping Beauty* Transposase and a Transgene Flanked by Transposon-Inverted Repeats

Quantitative representation of the *cis-* and *trans-*requirements of SB transposition in HeLa cells. Different combinations of the indicated donor and helper plasmids were cotransfected into cultured HeLa cells, and one-tenth of the cells as compared to the experiments shown in Figure 5 were plated under selection in order to be able to count transformants. The efficiency of transgene integration is scored as the number of transformants that survive antibiotic selection. Numbers of transformants at right represent the numbers of G-418-resistant cell colonies per dish. Each number is the average obtained from three transfection experiments.

integration in our assay is the result of active *Sleeping Beauty* transposition from extrachromosomal plasmids into the chromosomes of vertebrate cells.

Cut-and-Paste Transposition of *Sleeping Beauty* to Human Chromosomes

To examine the structures of integrated transgenes, colonies of transformed HeLa cells growing under G-418 selection from an experiment similar to that shown in plate 2 in Figure 5B were picked and their DNAs analyzed using Southern hybridization. Genomic DNA samples were digested with a combination of five restriction enzymes that do not cut within the 2233 bp *T/neo* marker transposon and were hybridized with a *neo*-specific probe (Figure 7A). The hybridization patterns indicate that all of the analyzed clones contained integrated transgenes in the range of 1 (lane 4) to 11 (lane 2) copies per transformant. Moreover, these multicopy insertions appear to have occurred in different locations in the

human genome, consistent with a transpositional mechanism of transgene integration.

Members of the *Tc1/mariner* superfamily always integrate into a TA target dinucleotide which is duplicated upon insertion (Plasterk, 1996). Thus, the presence of duplicated TA sequences flanking an integrated transposon is a hallmark of TcE transposition. To reveal such sequences, junction fragments of integrated transposons and human genomic DNA were isolated using a ligation-mediated PCR assay (Devon et al., 1995). We have cloned and sequenced junction fragments of five integrated transposons, all of them showing the predicted sequences of the IRs, which continue with TA dinucleotides, and sequences that are different in all of the junctions and different from the plasmid vector sequences originally flanking the transposon in pT/neo (Figure 7B). The same results were obtained from nine additional junctions containing either the left or the right IR of the transposon (data not shown). These results indicate that the marker transposons had been precisely excised from the donor plasmids and subsequently spliced into various locations in human chromosomes. Next, the junction sequences were compared to the corresponding "empty" chromosomal regions cloned from wild-type HeLa DNA. As shown in Figure 7B, all of these insertions had occurred into TA target sites, one of them within an L1 retrotransposon, which were subsequently duplicated to result in TAs flanking the integrated transposons. These data demonstrate that *Sleeping Beauty* utilizes the same cut-and-paste-type mechanism of transposition as other members of the *Tc1/mariner* superfamily and that fidelity of the reaction is maintained in heterologous cells.

Discussion

Revival of *Sleeping Beauty*, an Ancient *Tc1*-like Transposon from Teleost Fish

DNA-transposons, including members of the *Tc1/mariner* superfamily, are ancient residents of vertebrate genomes (Radice et al., 1994; Smit and Riggs, 1996). However, neither autonomous copies of this class of transposon, nor a single case of a spontaneous mutation caused by a transposon insertion has been proven in
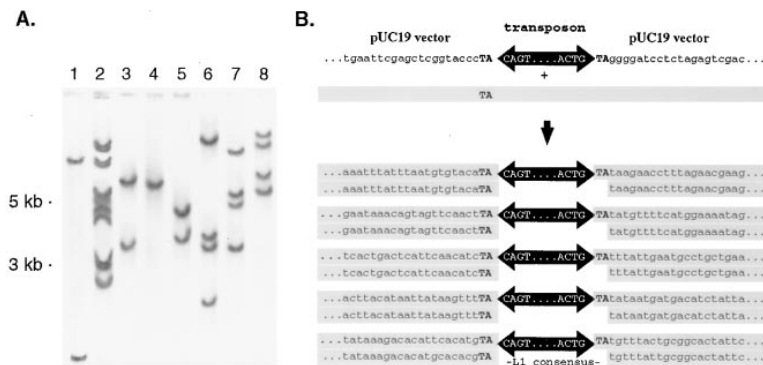

Figure 7. Transposition of Neomycin Resistance-Marked Transposons into the Chromosomes of Human HeLa Cells

(A) Southern hybridization with a *neo*-specific radiolabeled probe of genomic DNA samples prepared from individual HeLa cell clones that had been cotransfected with pT/neo and pSB10 and survived G-418 selection. Genomic DNAs were prepared as described in Ivics et al. (1996) and digested with restriction enzymes NheI, XhoI, BglII, SpeI and XbaI, of which none cuts within the *neo*-marked transposon, prior to agarose gel electrophoresis and blotting.

(B) Junction sequences of *T/neo* transposons integrated into human genomic DNA. On top, the donor site is shown with plasmid vector sequences originally flanking the transposon (black arrow) in pT/neo. Human genomic DNA serving as target for transposon insertion is illustrated as a gray box. Cut-and-paste transposition into TA target dinucleotides (typed bold) results in integrated transposons flanked by TA duplications. IR sequences are uppercase, whereas flanking sequences are lowercase.

vertebrate animals, in contrast to retroposons whose phylogenetic histories of mutating genes in vertebrates are documented (Izsvak et al., 1997). This suggests that DNA-transposons either have been inactivated in most vertebrate genomes or their activity is tightly regulated. Failure to isolate active elements from vertebrates has greatly hindered ambitions to develop transposon-based vectors for germline transformation and insertional mutagenesis in this group of animals. However, the apparent capability of salmonid TcEs for horizontal transmission between two teleost orders (Ivics et al., 1996) suggested that this particular subfamily of fish transposons might be transferred over even larger evolutionary distances.

Therefore, based on phylogenetic data collected from TcE transposons in different teleost fish species and comparative analysis of functional transposase domains, we reconstructed an active salmonid-type *Tc1*-like transposon, *Sleeping Beauty*. This transposon system consists of two components: a synthetic gene encoding a transposase, and a cloned, nonautonomous salmonid-type element that carries the inverted repeats of the transposon substrate DNA. When put together, these two components represent, or are very similar to, a sequence of an active transposon, which was in its prime approximately 10 million years ago (Ivics et al., 1996) and was able to infest teleost genomes through horizontal transmission.

Reconstructions of ancestral genes by predicting an archetypal sequence using parsimony analysis have been reported (Jermann et al., 1995). However, such a strategy requires vertical transmission of a gene through evolution for phylogenetic backtracking to the root sequence. Because parsimony analysis could not resolve the phylogenetic relationships between salmonid TcEs, we took a different approach by reconstructing a consensus sequence from inactive elements belonging to the same subfamily of transposons. A similar approach has been used for the resurrection of a functional promoter of L1 retrotransposons in mouse (Adey et al., 1994). However, such a strategy for obtaining an active gene is not without risks. The consensus sequence of transposase pseudogenes from a single organism might simply reflect the mutations that had occurred during vertical inactivation and had subsequently been fixed in the genome as a result of amplification of the mutated element. For instance, most *Tdr1* elements isolated from zebrafish contain a conserved 350bp deletion in the transposase gene (Izsvak et al., 1995); therefore, their consensus is expected to encode an inactive element. In contrast, because independent fixation of the same mutation in different species is unlikely, we derived a consensus from inactive elements of the same subfamily of transposons from several organisms to provide the sequence of an active transposon.

To examine the potential usefulness of *Sleeping Beauty* as a molecular tool for vertebrate genetics, we addressed the following three questions: (1) Is it active? (2) Is it transferable to different vertebrate genomes? (3) Is it sufficiently specific for its substrate so that it does not interfere with endogenous elements of other TcE subfamilies already present in genomes?

Transposition of *Sleeping Beauty* in Vertebrate Cells

We have shown that *Sleeping Beauty* is a fully functional transposon system that can perform all the complex steps of cut-and-paste DNA transposition; that is, the nuclear-localized transposase is able to recognize and to excise its specific DNA substrate from an ectopic plasmid and to insert it into chromosomes.

Upon cotransfection of the two-component *SB* transposon system into cultured vertebrate cells, transpositional activity is manifested as enhanced integration of the transgene serving as the DNA substrate for transposase. DNA-binding and nuclear targeting activities alone did not increase transformation frequency. Although not sufficient, these functions are probably necessary for transposase activity. Indeed, a single amino acid replacement in the NLS of mariner is detrimental to overall transposase function (Lohe et al., 1997). The inability of SB6, a mutated version of the transposase, to catalyze transposition justified our efforts to reconstruct a consensus sequence. Notably, 3 of the 11 amino acid substitutions that SB6 contains, F(21), N(28), and H(31), are within the specific DNA-binding domain (Figures 1 and 2). Sequence analysis of the *paired*-like DNA-binding domain of fish TcE transposases indicates that an isoleucine residue in position 28 is conserved between the transposases and the corresponding positions in the Pax proteins (Ivics et al., 1996). Thus, we predict that this motif is crucial for DNA-binding activity. Moreover, SB transposase exhibits strong substrate-preference for those engineered transposons that have both of the terminal inverted repeats. In contrast to P element transposase, which requires both transposon ends for cleavage (Beall and Rio, 1997), Tc1 transposase is able to cleave substrates that have only one of the transposon IRs (Vos et al., 1996). Thus, we expect that SB transposase could introduce a single double-strand DNA break at the solitary IR of pT/neo-ΔIR. Such linearized donor constructs would have elevated competence in genomic insertion compared to circular molecules. Indeed, we observed an approximately 3-fold stimulation in integration of pT/neo-ΔIR by SB transposase (Figure 6). Transposition in our assay can only occur if both components of the *SB* system are present in the same cell. Once that happens, multiple integrations might take place, as evidenced by our finding of up to 11 transposon insertions in neomycin-resistant cell clones.

We have observed transposition of synthetic salmonid transposons in fish, mouse, and human cells. In addition, transposition of genetically marked transposons in a plasmid-to-plasmid transposition assay was significantly enhanced in microinjected zebrafish embryos in the presence of transposase (data not shown). Moreover, precision of the cut-and-paste reaction was remarkably maintained in heterologous cells; we have not encountered a single junction sequence that did not contain the predicted ends of the element and TA target site duplications flanking the insertion. Consequently, *SB* apparently does not need any obvious species-specific factor that would restrict its activity (i.e., efficiency and fidelity) to its original host. This observation is supported by the wide species-distribution and apparent success of the *Tc1/mariner* superfamily in the animal

kingdom. Interestingly, the most significant enhancement, about 20-fold, of transposition was observed in human cells. Possible explanations of this phenomenon might be the absence of repressor-like factors and/or presence of stimulatory factors in human cells.

Specific DNA-Binding Activity of Sleeping Beauty Transposase

There are at least two distinct subfamilies of TcEs in the genomes of Atlantic salmon and zebrafish, *Tss1* and *Tss2* (Goodier and Davidson, 1994; Radice et al., 1994), and *Tdr1* and *Tdr2* (Radice et al., 1994; Izsvak et al., 1995; Ivics et al., 1996; Lam et al., 1996a), respectively. Elements of different subfamilies from different species may be more alike (e.g., *Tss1* and *Tdr1* have about 70% nucleic acid identity) than those in the same species. For example, *Tdr1* and *Tdr2* are characteristically different in their encoded transposases and their inverted repeat sequences and share only about 30% nucleic acid identity. It may be that subfamilies of transposons within the same genome should be significantly different from each other in order to avoid cross-mobilization. A major question is whether substrate recognition of transposases is sufficiently specific to prevent activation of transposons of closely related subfamilies.

We have shown that the 12 bp DRs of salmonid-type elements, identical to the DRs of zebrafish-type TcEs, are part of the binding sites for SB transposase. However, these binding sites are 30 bp long. Thus, specific DNA binding also involves DNA sequences around the DRs, which are variable between TcE subfamilies in fish. Such a difference in the sequences of transposase binding sites might explain the inability of N123 to bind zebrafish *Tdr1* IRs efficiently, and it may enable the transposase to distinguish even between closely related TcE subfamilies. Indeed, mutations of four base pairs of the 20bp Tc1 binding site can abolish binding of transposase (Vos and Plasterk, 1994). We propose that the DR core motifs are primarily involved in transposase binding, while sequences around the DR motifs may provide the specificity for this binding.

SB transposase has four binding sites in its transposon substrate DNA that are located at the ends of the IRs. However, a zebrafish *Tdr1* element lacking an internal transposase binding site was apparently able to transpose (data not shown). This observation is in agreement with the finding that removal of internal transposase binding sites from engineered *Tc3* elements did not reduce their ability to transpose (Colloms et al., 1994), suggesting that the presence of internal transposase binding sites is not essential for transposition. Multiple binding sites for proteins, including transposases, are frequently associated with regulatory functions (Gierl et al., 1988). Consequently, it is probable that the multiple binding sites for transposases in the IR/DR group of TcEs serve a yet unknown regulatory purpose affecting transposition, which might represent an advanced level of control in the delicate balance of transposon–host relationship.

Potential Use of *Sleeping Beauty* as a Vector System for Insertional Mutagenesis or Transgenesis in Vertebrates

Due to their inherent ability to move from one chromosomal location to another within and between genomes, transposable elements revolutionized genetic manipulation of certain organisms including bacteria (Gonzales et al., 1996; Lee and Henk, 1996), Drosophila (Bellen et al., 1989; Spradling et al., 1995), C. elegans (Plasterk, 1995) and a variety of plant species (Osborne and Baker, 1995). However, the majority of animal model organisms as well as species of economic importance lack such a tool. For its simplicity and apparent ability to function in diverse organisms, *Sleeping Beauty* should prove useful as an efficient vector for transposon tagging, enhancer trapping, and transgenesis in species in which DNA transposon technology is currently not available.

Experimental Procedures

Recombinant DNA
Gene Reconstruction–Phase 1: Reconstruction of a Transposase Open Reading Frame
The *Tss1.1* element from Atlantic salmon (GenBank accession number L12206) was PCR-amplified using primers flanking the defective transposase gene, FTC-Start and FTC-Stop, to yield *SB1*. Next, a segment of the *Tss1.2* element (L12207) was amplified using primers FTC-3 and FTC-4 and then with FTC-3 and FTC-5. The PCR product was digested with restriction enzymes NcoI and BlpI, underlined in the primer sequences, and cloned to replace the corresponding fragment in *SB1* to yield *SB2*. Then, an approximately 250 bp HindIII fragment of the *Tsg1* element from rainbow trout (L12209) was isolated and cloned into the respective sites in *SB2* to result in *SB3*. The *Tss1* and *Tsg1* elements were described in Radice et al. (1994) and were kind gifts from S. Emmons.

Primers for *SB1-SB3*: FTC-Start: 5′-CCTCTAGGATCCGACATC ATG; FTC-Stop: 5′-TCTAGAATTCTAGTATTTGGTAGCATTG; FTC-3: 5′-AACACCATGGGACCACGCAGCCGTCA; FTC-4: 5′-CAGGTTATG TCGATATAGGACTCGTTTTAC; FTC-5: 5′-CCTTGCTGAGCGGCCTT TCAGGTTATGTCG.

Gene Reconstruction-Phase 2: Site-Specific PCR Mutagenesis of the SB3 ORF to Introduce Consensus Amino Acids
For PCR mutagenesis, two methods were used: megaprimer PCR (Sarkar and Sommer, 1990) from *SB4* through *SB6*, and ligase chain reaction (Michael, 1994) for steps *SB7* to *SB10*.

Primers for *SB4*: FTC-7: 5′-TTGCACTTTTCGCACCAA for Gln→ Arg(74) and Asn→Lys(75); FTC-13: 5′-GTACCTGTTTCCTCCA GCATC for Ala→Glu(93); FTC-8: 5′-GAGCAGTGGCTTCTTCCT for Leu→Pro(121); FTC-9: 5′-CCACAACATGATGCTGCC for Leu→ Met(193); FTC-10: 5′-TGGCCACTCCAATACCTTGAC for Ala→Val(265) and Cys→Trp(268); FTC-11: 5′-ACACTCTAGACTAGTATTTGGTAG CATTGCC for Ser→Ala(337) and Asn→Lys(339).

Primer for *SB5*: B5-PTV: 5′-GTGCTTCACGGTTGGGATGGTG for Leu→Pro(183), Asn→Thr(184), and Met→Val(185).

Primer for *SB6*: FTC-DDE: 5′-ATTTTCTATAGGATTGAGGT CAGGGC for Asp→ Glu(279).

Primers for *SB7* and *SB8*, in two steps: PR-GAIS: 5′-GTCTGGTTCATCCTTGGGAGCAATTTCCAAACGCC for Asn→Ile(28), His→Arg(31), and Phe→Ser(21).

Primers for *SB9*: KARL: 5′-CAAAACCGACATAAGAAAGCCAGAC TACGG for Pro→Arg(126); RA: 5′-GACCATCGTTATGTTTGGAGGAA GAAGGGGGAGGCTTGCAAGC CG for Cys→Arg(166) and Thr→ Ala(175); EY: 5′-GGCATCATGAGGAAGGAAAATTATGTGGATATA TTG for Lys→Glu(216) and Asp→Tyr(218); KRV: 5′-CTGAAAAAG CGTGTGCGAGCAAGGAGGCC for Cys→Arg(288); VEGYP: 5′-GTGGAAGGCTACCCGAAACGTTTGACC for Leu→Pro(324).

Primer for *SB10*: FATAH: 5′-GACAAAGATCGTACTTTTTGGA GAAATGTC for Cys→Arg(143).

Plasmids
For pSB10, the *SB10* transposase gene was cut with BamHI and EcoRI, whose recognition sequences are incorporated (and underlined above) in the primers FTC-Start and FTC-Stop; it was then filled in with Klenow and cloned into the Klenow-filled NotI sites of CMV-bgal (Clontech), replacing the *lacZ* gene originally present in this plasmid. Because of the blunt end cloning, both orientations of the gene insert were possible to obtain, and the antisense directions

were used as a control for transposase (pSB10-AS). For pSB10-ΔDDE, pSB10 was cut with MscI, which removes 322 bp of the transposase coding region, and recircularized. Removal of the MscI fragment from the transposase gene deleted much of the catalytic DDE domain and disrupted the reading frame by introducing a premature translational termination codon.

A TcE from Tanichthys albonubes (L48685) was cloned into the SmaI site of pUC19 to result in pT. The donor construct for the integration assays, pT/neo, was made by cloning, after Klenow fill-in, an EcoRI/BamHI fragment of the plasmid pRc-CMV (Invitrogen) containing the SV40 promoter/enhancer, the neomycin resistance gene and an SV40 poly(A) signal into the StuI/MscI sites of pT. The StuI/MscI double digest of T leaves 352 bp on the left side and 372 bp on the right side of the transposon, containing the terminal inverted repeats. An EcoRI digest of pT/neo removed a 350 bp fragment including the the left inverted repeat of the transposon, and this plasmid, designated pT/neo-ΔIR, was used as a control for the substratedependence of transposition.

For the expression of an N-terminal derivative of SB transposase, a gene segment of $SB8$ was PCR-amplified using primers FTC-Start and FTC-8, 5′-phosphorylated with T4 polynucleotide kinase, digested with BamHI, filled in with Klenow, and cloned into NdeI/EcoRI digested expression vector pET21a (Novagen) after Klenow fill-in. This plasmid, pET21a/N123, expresses the first 123 amino acids of the transposase (N123) with a C-terminal histidine tag.

Analysis of Junction Fragments

Splinkerette-PCR was done as described (Devon et al., 1995) on Sau3AI-digested genomic DNA samples prepared from individual, G-418-resistant HeLa cell clones. For the "empty" genomic regions corresponding to the insertion sites wild-type HeLa DNA was used. Primers for nested PCR amplifying the left junctions were T/DR 5′ AGTGTATGTAAACTTCTGACCCACTGG and T/BAL 5′ CTTGTGT CATGCACAAAGTAGATGTCC, whereas for the right junctions were T/JOBB1 5′ TTTACTCGGATTAAATGTCAGGAATTG and T/JOBB2 5′ TGAGTTTAAATGTA TTTGGCTAAGGTG. PCR products were either directly sequenced or cloned into pUC19 vector prior to sequencing.

Recombinant Protein Expression and Purification

Induction of N123 was in E. coli strain BL21(DE3) (Novagen) by the addition of 0.4 mM IPTG at 0.5 OD_{600} and continued for 2.5 hr at 30°C. Cells were sonicated in 25 mM HEPES (pH 7.5), 1 M NaCl, 15% glycerol, 0.25% Tween 20, 2 mM β-mercaptoethanol, 1 mM PMSF, and 10 mM imidazole (pH 8.0) was added to the soluble fraction before it was mixed with Ni^{2+}-NTA resin (Qiagen) according to the recommendations of the manufacturer. The resin was washed with sonication buffer containing 30% glycerol and 50 mM imidazole; bound proteins were eluted with sonication buffer containing 300 mM imidazole and dialyzed overnight at 4°C against sonication buffer without imidazole.

Gel Retardation and DNaseI Footprinting

A 300 bp EcoRI/HindIII fragment of pT comprising the left inverted repeat of the element was end-labeled using [α^{32}P]dCTP and Klenow. Nucleoprotein complexes were formed in 20 mM HEPES (pH 7.5), 0.1 mM EDTA, 0.1 mg/ml BSA, 150 mM NaCl, 1 mM DTT in a total volume of 10 μl. Reactions contained 100 pg labeled probe, 2 μg poly[dI][dC], and 1.5 μl N123. After 15 min incubation on ice, 5 μl of loading dye containing 50% glycerol and bromophenol blue was added and the samples loaded onto a 5% polyacrylamide gel. DNaseI footprinting was done using a kit from BRL according to the recommendations of the manufacturer.

Cell Transfections

Cells were cultured in DMEM supplemented with 10% fetal bovine serum, seeded onto 6 cm plates one day prior to transfection and transfected with 5 μg Elutip (Scleicher and Schuell)-purified plasmid DNA using Lipofectin from BRL. After 5 hr of incubation with the DNA-lipid complexes, the cells were "glycerol-shocked" for 30 s with 15% glycerol in phosphate-buffered saline (PBS), washed once with PBS, and then re-fed with serum-containing medium. Two days posttransfection, the transfected cells were trypsinized and resuspended in 2 ml of serum-containing DMEM, and either 1 ml or 0.1 ml aliquots of the cell suspensions were seeded onto several 10 cm plates in medium containing 600 μg/ml G-418 (BRL). After two weeks

of selection, cell clones were either picked and expanded into individual cultures, or fixed with 10% formaldehyde in PBS for 15 min, stained with methylene blue in PBS for 30 min, washed extensively with deionized water, air dried, and photographed.

References

Adey, N.B., Tollefsbol, T.O., Sparks, A.B., Edgell, M.H., and Hutchison, C.A., III, (1994). Molecular resurrection of an extinct ancestral promoter for mouse L1. Proc. Natl. Acad. Sci. USA 91, 1569–1573.

Beall, E.L., and Rio, D.C. (1997). Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev. 11, 2137–2151.

Bellen, H.J., O'Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K., and Gehring, W.J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev. 3, 1288–1300.

Colloms, S.D., van Luenen, H.G., and Plasterk, R.H. (1994). DNA binding activities of the Caenorhabditis elegans Tc3 transposase. Nucleic Acids Res. 22, 5548–5554.

Craig, N.L. (1995). Unity in transposition reactions. Science 270, 253–254.

Devon, R.S., Porteous, D.J., and Brookes, A.J. (1995). Splinkerettes—improved vectorettes for greater efficiency in PCR walking. Nucleic Acids Res. 23, 1644–1645.

Doak, T.G., Doerder, F.P., Jahn, C.L., and Herrick, G. (1994). A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc. Natl. Acad. Sci. USA 91, 942–946.

Gibbs, P.D., Gray, A., and Thorgaard, G. (1994). Inheritance of P element and reporter gene sequences in zebrafish. Mol. Mar. Biol. Biotech. 3, 317–326.

Gierl, A., Lutticke, S., and Saedler, H. (1988). TnpA product encoded by the transposable element En-1 of Zea mays is a DNA binding protein. EMBO J. 7, 4045–4053.

Gonzales, A.E., Glisson, J.R., and Jackwood, M.W. (1996). Transposon mutagenesis of Haemophilus paragallinarum with Tn916. Vet. Microbiol. 48, 283–291.

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

Gueiros-Filho, F.J., and Beverley, S.M. (1997). Trans-kingdom transposition of the Drosophila element mariner within the protozoan Leishmania. Science 276, 1716–1719.

Handler, A.M., Gomez, S.P., and O'Brochta, D.A. (1993). A functional analysis of the P-element gene-transfer vector in insects. Arch. Insect Biochem. Physiol. 22, 373–384.

Hartl, D.L., Lozovskaya, E.R., Nurminsky, D.I., and Lohe, A.R. (1997). What restricts the activity of mariner-like transposable elements? Trends Genet. 13, 197–201.

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

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

Izsvak, Zs., Ivics, Z., and Hackett, P.B. (1997). Repetitive elements and their genetic applications in zebrafish. J. Biochem. Cell Biol., in press.

Jermann, T.M., Opitz, J.G., Stackhouse, J., and Benner, S.A. (1995). Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. Nature *374*, 57–59.

Kidwell, M.G. (1992). Horizontal transfer. Curr. Opin. Genet. Dev. *2*, 868–873.

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

Lam, W.L., Lee, T.S., and Gilbert, W. (1996a). Active transposition in zebrafish. Proc. Natl. Acad. Sci. USA *93*, 10870–10875.

Lam, W.L., Seo, P., Robison, K., Virk, S., and Gilbert, W. (1996b). Discovery of amphibian Tc1-like transposon families. J. Mol. Biol. *257*, 359–366.

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

Lee, M.D., and Henk, A.D. (1996). Tn10 insertional mutagenesis in *Pasteurella multocida*. Vet. Microbiol. *50*, 143–148.

Lohe, A.R., Moriyama, E.N., Lidholm, D.A., and Hartl, D.L. (1995). Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol. Biol. Evol. *12*, 62–72.

Lohe, A.R., De Aguiar, D., and Hartl, D.L. (1997). Mutations in the *mariner* transposase: the D,D(35)E consensus sequence is nonfunctional. Proc. Natl. Acad. Sci. USA *94*, 1293–1297.

Loukeris, T.G., Livadaras, I., Arca, B., Zabalou, S., and Savakis, C. (1995). Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. Science *270*, 2002–2005.

Michael, S.F. (1994). Mutagenesis by incorporation of a phosphorylated oligo during PCR amplification. BioTechniques *16*, 410–412.

Oosumi, T., Belknap, W.R., and Garlick, B. (1995). *Mariner* transposons in humans. Nature *378*, 873.

Osborne, B.I., and Baker, B. (1995). Movers and shakers: maize transposons as tools for analyzing other plant genomes. Curr. Opin. Cell Biol. *7*, 406–413.

Plasterk, R.H. (1995). Reverse genetics: from gene sequence to mutant worm. In Methods in Cell Biology (New York: Academic Press), pp. 59–80.

Plasterk, R.H. (1996). The Tc1/*mariner* transposon family. Curr. Top. Microbiol. Immunol. *204*, 125–143.

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

Rio, D.C., Barnes, G., Laski, F.A., Rine, J., and Rubin, G.M. (1988). Evidence for Drosophila P element transposase activity in mammalian cells and yeast. J. Mol. Biol. *200*, 411–415.

Sarkar, G., and Sommer, S.S. (1990). The megaprimer method of site-directed mutagenesis. BioTechniques *8*, 404–407.

Smit, A.F., and Riggs, A.D. (1996). *Tiggers* and DNA transposon fossils in the human genome. Proc. Natl. Acad. Sci. USA *93*, 1443–1448.

Spradling, A.C., Stern, D.M., Kiss, I., Roote, J., Laverty, T., and Rubin, G.M. (1995). Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. Proc. Natl. Acad. Sci. USA *92*, 10824–10830.

van Luenen, H.G., Colloms, S.D., and Plasterk, R.H. (1994). The mechanism of transposition of Tc3 in *C. elegans*. Cell *79*, 293–301.

Vos, J.C., and Plasterk, R.H. (1994). Tc1 transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. EMBO J. *13*, 6125–6132.

Vos, J.C., De Baere, I., and Plasterk, R.H. (1994). Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev. *10*, 755–761.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. *13*, 335–340.

EMBL Reference Number