

The *Sleeping Beauty* Transposable Element: Evolution, Regulation and Genetic Applications

Zoltán Ivics¹, Christopher D. Kaufman¹, Hatem Zayed¹, Csaba Miskey¹, Oliver Walisko¹ and Zsuzsanna Izsvák^{1,2}

¹Max Delbrück Center for Molecular Medicine, Robert Rössle Str. 10, D-13092, Berlin, Germany

²Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary

Abstract

Members of the Tc1/*mariner* superfamily of transposable elements isolated from vertebrate species are inactive due to the accumulation of mutations. A representative of a subfamily of fish elements estimated to be last active >10 million years ago has been reconstructed, and named *Sleeping Beauty* (SB). This element opened up new avenues for studies on DNA transposition in vertebrates, and for the development of transposon tools for genetic manipulation in important model species and in humans. Multiple transposase binding sites within the terminal inverted repeats, a transpositional enhancer sequence, unequal affinity of the transposase to the binding sites and the activity of the cellular HMGB1 protein all contribute to a highly regulated assembly of SB synaptic complexes, which is likely a requirement for the subsequent catalytic steps. Host proteins involved in double-strand DNA break repair are limiting factors of SB transposition in mammalian cells, underscoring evolutionary, structural and functional links between DNA transposition, retroviral integration and V(D)J recombination. SB catalyzes efficient cut-and-paste transposition in a wide range of vertebrate cells in tissue culture, and in somatic tissues as well as the germline of the mouse and zebrafish *in vivo*, indicating its usefulness as a vector for transgenesis and insertional mutagenesis.

The Evolutionary Life-Cycle of DNA Transposons

Large fractions of genomes can be composed of transposable element sequences. The Human Genome Project revealed that approximately 45% of the human genome is transposon-derived (I. H. G. S. C. 2001); nevertheless, most of these elements are inactive. Three evolutionary processes were proposed to describe the "life-cycle" of a DNA transposon in a genome (Lohe *et al.*, 1995; Hartl *et al.*, 1997). In the absence of selection pressure, "vertical inactivation" leads to accumulation of mutations in the transposon sequence. DNA transposons consist of

two components: a transposon DNA and a recombinase that can *trans*-activate the transposon. Since transposase-defective (nonautonomous) elements can be as good "jumpers" as the autonomous copies, the ratio of autonomous to nonautonomous elements decreases with time, eventually resulting in the complete disappearance of active transposons in a given genome: a process termed "stochastic loss". "Horizontal transfer" is able to rescue the active transposon by invasion of the germline of a naive genome, so that the cycle can start over again. 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. Due to these processes, not a single autonomous element has been isolated from vertebrates, hindering studies on the basic mechanisms of DNA transposition, and prohibiting the development of transposon-based tools for vertebrate genetics.

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

From all DNA-transposons found so far in vertebrates, members of the Tc1/*mariner* superfamily from teleost fish are by far the best characterized (Goodier and Davidson, 1994; Radice *et al.*, 1994; Izsvák *et al.*, 1995; Ivics *et al.*, 1996). In contrast to the P element transposon which is restricted to the *Drosophila* genus (Rio *et al.*, 1988), Tc1/*mariner* elements are extremely widespread in nature (Plasterk *et al.*, 1999). This indicates that host requirements for Tc1/*mariner* transposition are not that tight and that elements might be promiscuous in evolutionary terms. Molecular phylogenetic analyses have shown that the majority of the fish Tc1-like elements can be classified into three major types: zebrafish-, salmonid- and *Xenopus* T_{tr}-type elements (Ivics *et al.*, 1996), of which the salmonid subfamily is probably the youngest and thus most recently active.

All of the fish elements isolated so far appear to have undergone vertical inactivation, and accumulated several mutations in their transposase genes. In an attempt to derive an active Tc1-like transposon from vertebrates, we have analyzed the salmonid subfamily of elements that was presumed to be active more than 10-15 million years ago, and appeared to have been able to invade different fish genomes through horizontal transmission (Ivics *et al.*, 1996). We reasoned that a consensus sequence generated from a sequence alignment of defective copies isolated from different fish genomes would likely represent an active archetypal sequence. We engineered this sequence to reconstruct an active ancestral element, which was named *Sleeping Beauty* (SB) (Ivics *et al.*, 1997). The SB transposon system consists of two main functional components: the transposase encoded by a synthetic gene, and a cloned, nonautonomous, salmonid-type element that

*For correspondence. Email zivics@mdc-berlin.de.

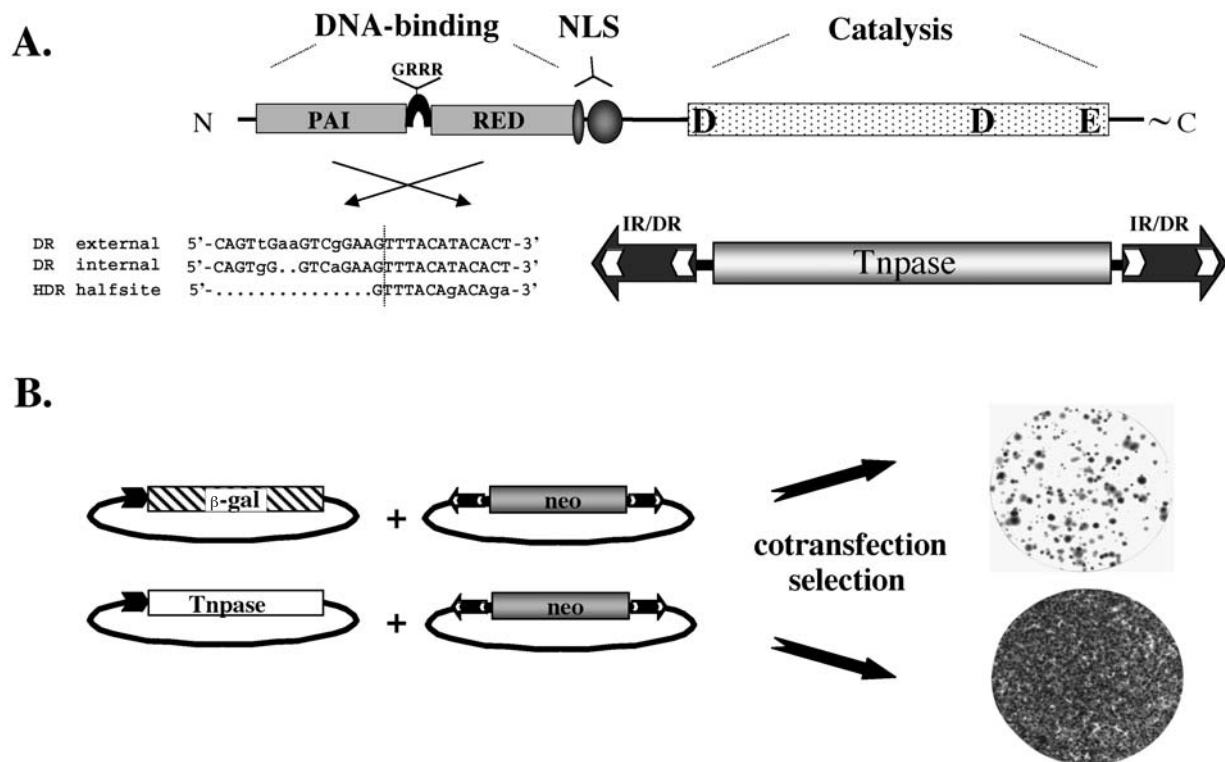


Figure 1. *Sleeping Beauty* transposition in cultured vertebrate cells. (A) Schematic representation of the two major components of the *Sleeping Beauty* transposable element system. In nature, the terminal inverted repeats (black arrows) flank a gene encoding the transposase (Tnpase). The inverted repeats of SB elements have a characteristic structure (IR/DR), and contain two binding sites for the transposase (white arrows) per repeat. The transposase has an N-terminal, paired-like DNA-binding domain consisting of two helix-turn-helix motifs, PAI and RED, with a GRRR-motif (AT-hook) between them. The RED subdomain overlaps with a nuclear localization signal (NLS), which is followed by the catalytic domain responsible for the DNA cleavage and joining reactions and characterized by the conserved DDE signature. The PAI subdomain recognizes the 3', whereas the RED subdomain recognizes the 5'-half of the bipartite recognition sequence. A sequence comparison between the outer and inner binding sites and the HDR enhancer is shown. (B) Assay for transposition in cultured cells. Under experimental conditions the two components of the transposon system are separated. A selectable marker such as an antibiotic resistance gene (*neo*) is cloned between the inverted repeats of the SB transposon. Transposon donor plasmids are introduced into cells together with transposase-expressing helper plasmids by transfection. In control transfections, a plasmid expressing a nonrelevant protein (β -gal) substitutes for the transposase. Cells are placed under antibiotic selection; only cells that express the antibiotic resistance gene due to chromosomal integration survive. Resistant cells give rise to colonies that can be harvested for DNA analysis, picked and expanded into larger cultures or stained for documentation. Shown are two petri dishes with stained human HeLa cell colonies obtained in the absence (upper dish) or in the presence (lower dish) of transposase. The marked difference in the numbers of resistant clones is due to transposition of the marked transposable elements into chromosomes.

carries the inverted repeats of the transposon that are recognized by the transposase (Figure 1A).

The two components of the SB system have to at least temporarily coexist in a cell for transposition to occur. We established an *in vivo* transposition assay to detect SB transposition events from plasmids to the chromosomes of vertebrate cells (Ivics *et al.*, 1997). The assay, shown in Figure 1B, is based on cotransfection of a transposon donor plasmid and a transposase-expressing helper plasmid into cultured cells. The transposable element carries an antibiotic resistance gene such as *neo*, so that inserted transposons that express their genes confer an antibiotic-resistant phenotype to cells. Cells are then placed under G-418 selection, and resistant colonies counted. The ratio between numbers obtained in the presence and absence of transposase is the readout of the assay, and is a measure of the efficiency of transposition. The result of a typical experiment in human HeLa cells is shown in Figure 1B.

Horizontal Gene Transfer in the Lab: Transposition of *Sleeping Beauty* in Diverse Vertebrate Cells

In plants, transposable elements of the *Ac/Ds* and *Spm* families have been routinely introduced into heterologous species (Haring *et al.*, 1991). No obvious barriers existed *a priori* that would restrict the activity of SB in heterologous species: the expression of transposase in any host should be sufficient to trigger transposition. The observation that both Tc1 (Vos *et al.*, 1996) and *mariner* transposases (Lampe *et al.*, 1996) produced in *E. coli* are able to catalyze transposition *in vitro*, and the several successful gene transfer experiments with various Tc1/*mariner* elements into heterologous hosts (reviewed in Plasterk *et al.*, 1999) seemed to support this assumption. The ancient salmonid transposon was clearly able to invade certain teleost fish species; nevertheless, it is absent from other vertebrate species, suggesting some limitations to its spread. A “zoo”

Table 1. Horizontal gene transfer in the laboratory. Summary of tissue culture transformation experiments using *Sleeping Beauty* transposons as vectors. Transpositional efficiency is expressed as the ratio between the number of G-418-resistant cell clones obtained in the presence versus in the absence of SB transposase. + 1-3-fold, ++ 3-5-fold, +++ 5-10-fold, ++++ 10-20-fold, +++++ >20-fold.

Organism	Cell line	Activity	Organism	Cell line	Activity
Human	Hela	+++++	Chicken	DT40	+++++
	Jurkat	+	Frog	A6	+++++
	HuH7	+++	Zebrafish	ZF4	+
	P19	+		PAC2	+
Monkey	Cos7	+	Carp	EPC	+
Mouse	LMTK	++	Sea bream	SAF1	+
	3T3	++	Trout	RTG	++
	ES	++	Medaka	OLF136	+
Hamster	CHO-K1	++++	Swordtail	A2	+
Rabbit	SIRC	+	Fathead minnow	FHM	+++++
Dog	MDCK	+			
Cow	MDBK	+++			
Sheep	MDOK	+			

experiment, in which SB's activity was tested in cells of different vertebrate classes, raised some interesting points (Izsvák *et al.*, 2000). Cell lines from different fish species, from mouse, human, frog, chick, sheep, cow, dog, rabbit, hamster and monkey were tested, using the *in vivo* transposition assay described above. As summarized in Table 1, SB was able to increase the frequency of transgene integration in all of these cell lines, indicating that SB is active in most (if not all) vertebrate species. We found extensive variation in the extent to which transposase stimulates integration between different species and even between different cell lines of the same species (Table 1). Unexpectedly, evolutionary distance of the recipient species (as compared to bony fishes from which SB originates) was not the main factor affecting transpositional efficiency, since the highest activity was observed in human cells. One factor that can clearly affect the activity of SB in these experiments is the varying levels of transfectability of the different cell lines, which can influence both the amount of transposase and the number of available transposon substrate molecules per cell. This might explain the difference in transpositional efficiency in cells derived from the same species. Furthermore, despite the widespread nature of Tc1/*mariner* elements, the involvement of certain host factors in transposition cannot be ruled out. Interaction of such factors with the transpositional machinery might lead to different efficiencies of transposition in different species and/or cells. Indeed, we have found that the relative inefficiency of transposition correlated with a decreased level of precision of transposon integration in some cell lines (Izsvák *et al.*, 2000).

Another line of experiments approached the question from the other direction. When tested under identical experimental conditions in the same cell line, *Sleeping Beauty* outperformed other members of the Tc1/*mariner* superfamily (Fischer *et al.*, 2001). SB activity was about 10-25-fold higher in human cells compared to Tc1, Tc3, the *Mos1* and *Himar1* *mariner* elements and *Minos* (Fischer *et al.*, 1991; and our own unpublished results). One possible

explanation for the differences is that the activity of a transposon might depend on the actual phase of its evolutionary life cycle. The three main periods in the cycle are: 1) the invasion of a naïve genome, 2) establishment of the transposon in the new host and 3) propagation, accumulation of mutations and extinction. It is possible, that a transposable element has the highest activity in the period of invasion and, subsequently, certain regulatory mechanisms are established between the host and the transposon resulting in low(er) transpositional activity. The available data suggest that transposons begin to accumulate mutations after their arrival into a genome. Transposons are not always successful, they can fail to establish themselves in a given genome, and might be unable to achieve high copy numbers, or can even "die out" without successfully invading a new host. Accordingly, the *Sleeping Beauty* transposon, which is presumed to represent an ancient element capable of invading new genomes through horizontal transmission, might have a higher intrinsic transpositional activity than the other elements that were tested. Nevertheless, hyperactive mutants of the SB transposase can be selected (unpublished results), indicating that a plateau of possible intrinsic activity has not yet been reached.

Another explanation for the observed difference in the respective transpositional activities of different elements can be their different abilities to interact with host factors. Since SB is a vertebrate transposon, it might be better "tuned" to a vertebrate cellular environment than elements originating from invertebrates.

The Structure of the *Sleeping Beauty* Transposon

Conserved Protein Domains in the Transposase

The overall domain structure of the transposase is conserved in the entire Tc1/*mariner* superfamily (Plasterk *et al.*, 1999). Specific substrate recognition is mediated by an N-terminal, bipartite DNA-binding domain of the transposase (Figure 1A) (Vos and Plasterk, 1994; Pietrovski and Henikoff, 1997; Izsvák *et al.*, 2002). This DNA-binding domain has been proposed to consist of two helix-turn-helix (HTH) motifs, similar to the paired domain of some transcription factors in both amino acid sequence and structure (Franz *et al.*, 1994; Vos and Plasterk, 1994; Ivics *et al.*, 1996). The modular paired domain has evolved versatility in binding to a range of different DNA sequences through various combinations of its subdomains (PAI+RED) (Czerny *et al.*, 1993). The nucleotide sequences recognized by the composite paired domain are degenerate, the DNA-binding specificity is relaxed (Pellizzari *et al.*, 1999). The origin of the paired domain is not clear, but phylogenetic analyses indicate that it might have been derived from an ancestral transposase (Breitling and Gerber, 2000). Partially overlapping with the RED subdomain in the transposase is a nuclear localization signal (NLS in Figure 1A), flanked by phosphorylation target sites of casein kinase II (Ivics *et al.*, 1996). Phosphorylation of these sites is a potential checkpoint in the regulation of transposition. The NLS indicates that these transposons, unlike murine retroviruses, can take advantage of the receptor-mediated transport machinery of host cells for nuclear uptake of their

transposases. A characteristic GRPR-like motif (GRRR) between the two HTH motifs (Figure 1A) is similar to an AT-hook (Izsvák *et al.*, 2002), responsible for minor groove interactions in the Hin invertase of *Salmonella* (Feng *et al.*, 1994) and in the RAG1 recombinase of V(D)J recombination (Spanopoulou *et al.*, 1996).

The catalytic domain of the transposase, responsible for the DNA cleavage and joining reactions, is characterized by a conserved amino acid triad, the DDE motif (Figure 1A), which is found in a large group of recombinases, including retrotransposon and retrovirus integrases, bacterial IS element transposases (Doak *et al.*, 1994) and RAG1 (Kim *et al.*, 1999; Landree *et al.*, 1999). Within the catalytic domains of Tc1-like transposases, a conserved glycine-rich subdomain can be found (Ivics *et al.*, 1997). The function of this subdomain is unknown. In addition to the DDE-containing transposases and integrases (Dyda *et al.*, 1994; Davies *et al.*, 2000), crystallographic analyses of the catalytic domains of proteins whose functions are not obviously related to transposition, such as RNAase H (Katayanagi *et al.*, 1990) or RuvC (Ariyoshi *et al.*, 1994) have revealed a remarkably similar overall fold. The existence of a common structural motif that catalyses polynucleotidyl transfer reactions in diverse biological contexts suggests that the different specificities in binding to DNA might have evolved by the apparent acquisition of different DNA-binding domains in the evolution of DDE recombinases (Capy *et al.*, 1996).

Structure of the Transposon

Transposons are bracketed by terminal inverted repeats that contain binding sites for the transposase. Tc1/*mariner* elements have a roughly uniform size of approximately 1.6–1.7 kb, indicating a natural selection in genomes for this particular size. *Sleeping Beauty* has a pair of transposase-binding sites at the ends of the 200–250 bp long inverted repeats (IRs). Within each IR of SB, there are two transposase binding sites that contain short, 15–20 bp direct repeats (DRs). This special organization of inverted repeat, termed IR/DR (Figure 1A), is an evolutionarily conserved feature of a group of Tc1-like elements, but not that of the Tc1 element itself (Izsvák *et al.*, 1995; Plasterk *et al.*, 1999). The IR/DR subgroup is represented by the *Minos*, *Bar1*, *Selements* in flies (Franz and Savakis, 1991; Merriman *et al.*, 1995; Moschetti *et al.*, 1998), *Quetzal* elements in mosquitos (Ke *et al.*, 1996), *Txr* elements in frogs (Lam *et al.*, 1996) and at least three Tc1-like transposon subfamilies in fish (Ivics *et al.*, 1996). The spacing of about 200 bp between the outer and inner binding sites is conserved in all elements within the IR/DR group, but the actual DNA sequences are not similar, suggesting convergent evolution of the IR/DR-type repeats. The IR/DR group significantly differs from Tc1 or the *mariner* elements that are more simple and have repeats of less than 100 bp and a single transposase binding site per repeat. All four binding sites within the IR/DR structure are required for SB transposition (Izsvák *et al.*, 2000). The four binding sites are not identical, the outer ones are longer by two base pairs (Figure 1A). The IRs are not identical either; the left IR contains a sequence motif called the HDR, which resembles the 3'-half of the transposase binding sites (Figure 1A) (Izsvák *et*

al., 2002). A construct containing two left IRs transposes better than the wild-type transposon, but another version that has two right IRs has very poor mobility, indicating that the left and right IRs are functionally distinct (Izsvák *et al.*, 2002).

Mechanism of Transposition

The transposase protein and the inverted repeats together engage in a series of molecular events that lead to the excision of the element from its DNA context and reintegration into a different locus, a process termed cut-and-paste transposition. The transposition process can arbitrarily be divided into at least four major steps: 1) binding of the transposase to its sites within the transposon IRs; 2) formation of a synaptic complex in which the two ends of the elements are paired and held together by transposase subunits; 3) excision from the donor site; 4) reintegration at a target site. On the molecular level, mobility of DNA-based transposable elements can be regulated by imposing constraints on transposition. One important form of transpositional control is represented by regulatory “checkpoints”, at which certain molecular requirements have to be fulfilled for the transpositional reaction to proceed. These requirements can operate at any of the four different stages of transposition listed above, and can be brought about by both element-encoded and host-encoded factors.

Specific DNA-binding by the Sleeping Beauty Transposase

Similar to the DNA-binding domain of the transposase, the binding sites also have a bipartite structure in which the 3'-part of the binding site is recognized by the PAI subdomain, whereas the 5'-sequences interact with the RED subdomain of the transposase (Figure 1A) (Izsvák *et al.*, 2002). Specificity of DNA-binding is predominantly determined by base-specific interactions mediated by the PAI subdomain (Izsvák *et al.*, 2002). The PAI subdomain also binds to the HDR motif within the left inverted repeat of SB, and mediates protein-protein interactions with other transposase subunits. Thus, the PAI subdomain is proposed to have at least three distinct functions: interaction with both the DRs and the HDR motif, and transposase oligomerization. In cooperation with the main DNA-binding domain, the GRRR motif was shown to function as an AT-hook, contributing to specific substrate recognition (Izsvák *et al.*, 2002). Although part of the NLS is included in the RED subdomain, it does not appear to contribute to DNA recognition. Domain swapping experiments have shown that primary DNA-binding is not sufficient to determine specificity of the transposition reaction. Zebrafish Tdr1 elements are closely related to SB, but are not mobilized by SB transposase. Comparison of the transposase binding site sequences of SB and Tdr1 elements revealed main differences in the 5'-half of the DRs. This sequence is contacted by the RED subdomain, indicating that the function of the RED is to enforce specificity at a later step in transposition. Substrate recognition of SB transposase is therefore sufficiently specific to prevent activation of transposons of closely related subfamilies.

The spacing between the DRs is conserved in the IR/DR group, and decreasing the distance between the DRs has a negative effect on transposition (Izsvák *et al.*, 2000). The transposase does not bind the DRs with equal affinity, it preferentially binds the internal recognition sequences (Cui *et al.*, 2002; Zayed *et al.*, 2003). Perhaps due to the two-base-pair difference in length, the helical phasing of the outer binding sites make transposase binding unfavored at these sites. The significance of this unequal affinity in binding is discussed in the next section.

Synaptic Complex Assembly, and the Role of the Multiple Binding Sites for the Transposase

A uniform requirement among transposition reactions is the formation of a nucleoprotein complex, before the catalytic steps can take place. This very early step, synaptic complex assembly, is the process by which the two ends of the elements are paired and held together by transposase subunits. *Sleeping Beauty* transposition is controlled at the level of complex assembly (Izsvák *et al.*, 2002). The paired-like DNA-binding domain forms tetramers in complex with transposase binding sites (Izsvák *et al.*, 2002). The necessary factors that are required for synaptic complex assembly of SB include the complete inverted repeats with four transposase binding sites, the HDR motif and tetramerization-competent transposase. These tetrameric complexes form only if all the four binding sites are present and they are in the in proper context. The HDR motif is important but not essential in transposition, and therefore can be viewed as a transpositional enhancer. Our findings suggest that the transpositional enhancer and the PAI subdomain of the transposase are stabilizing complexes formed by a transposase tetramer bound at the IR/DR. In contrast to Mu transposase, where the two

specificities of binding to the enhancer and to the recombination sites are encoded in two distinct domains (Leung *et al.*, 1989), the paired-like region of SB transposase combines these two functions in a single protein domain.

The role of HMGB1 in Sleeping Beauty Transposition: Ordered Assembly of Synaptic Complexes

Differential interactions between the transposon and host-encoded factors may result in limitation of host range. We have found that the high mobility group protein HMGB1 is required for efficient *Sleeping Beauty* transposition in mammalian cells (Zayed *et al.*, 2003). HMGB1 is an abundant, non-histone, nuclear protein associated with eukaryotic chromatin, and has the ability to bend DNA (Bustin, 1999). SB transposition was significantly reduced in HMGB1-deficient mouse cells. This effect was complemented by expressing HMGB1 and HMGB2, but not with the more distantly related HMGA1 protein. Overexpression of HMGB1 in wild-type cells enhanced transposition, indicating that HMGB1 is a limiting factor of transposition. HMGs have low affinity to standard, B-form DNA, and interactor proteins need to guide them to certain sites (Bustin, 1999). SB transposase was found to interact with HMGB1 *in vivo*, and to form a ternary complex with the transposase and transposon DNA, suggesting that the transposase may actively recruit HMGB1 to transposon DNA via protein-protein interactions.

Considering the significant drop of transposition activity in HMGB1-deficient cells, the role of HMGB1 in transposition is a critical one. HMGB1 was proposed to promote communication between DNA motifs within the transposon that are otherwise distant to each other, including the DRs, the transpositional enhancer and the

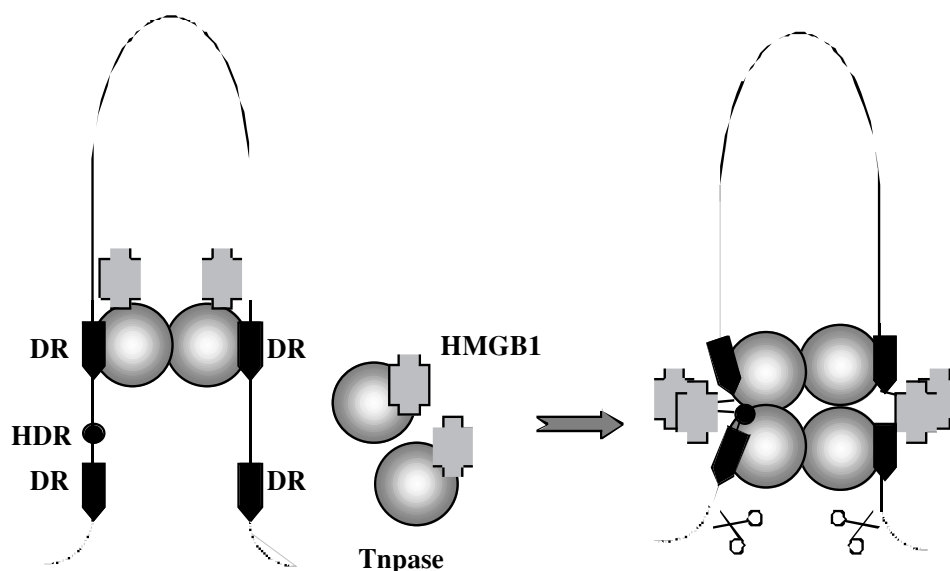


Figure 2. A proposed model for the role of HMGB1 in *Sleeping Beauty* synaptic complex formation. *Sleeping Beauty* transposase (gray spheres) recruits HMGB1 (striped objects) to the transposon inverted repeats. First, HMGB1 stimulates specific binding of the transposase to the inner binding sites. Once in contact with DNA, HMGB1 bends the spacer regions between the DRs, thereby assuring correct positioning of the outer sites for binding by the transposase. Cleavage (scissors) proceeds only if complex formation is complete. The complex includes the four binding sites (black boxes), the HDR enhancer sequence (black circle) and a tetramer of the transposase.

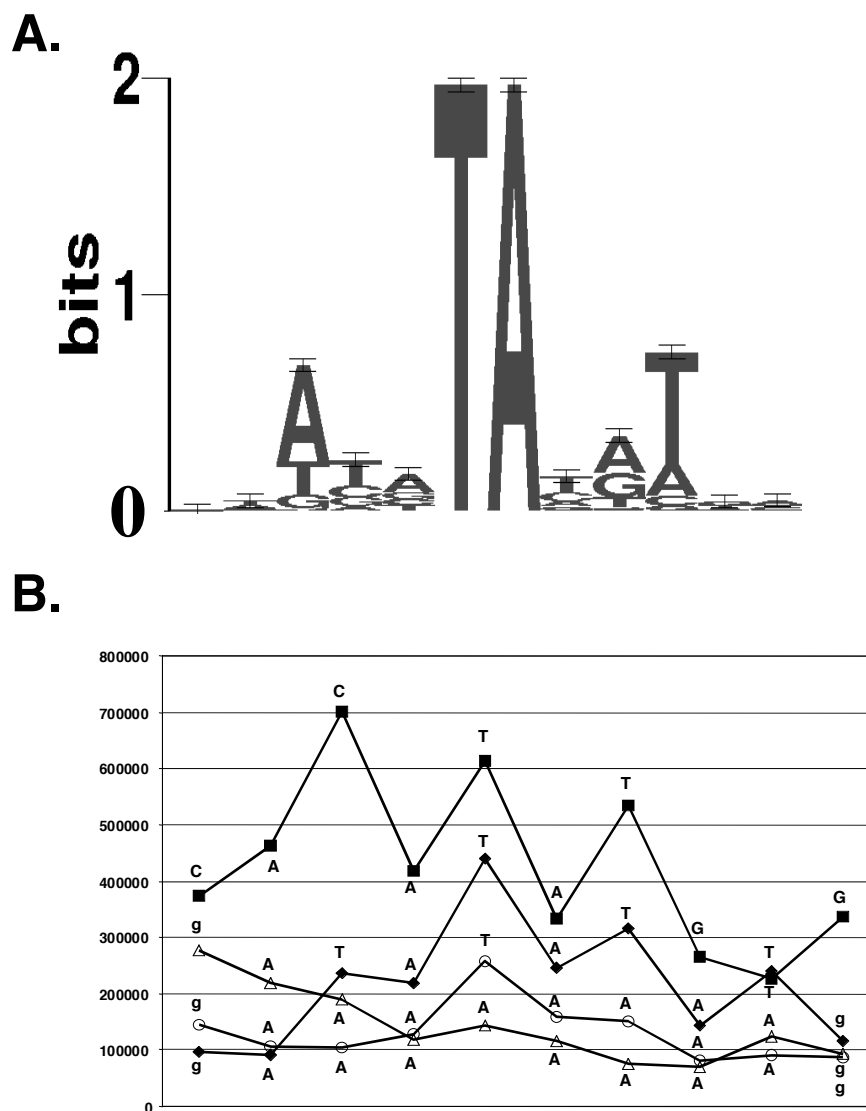


Figure 3. The preferred insertion site of *Sleeping Beauty* is a bendable AT-repeat. (A) Consensus sequence of insertion sites. Seqlogo analysis of five base pairs upstream and downstream of TA target sites. The y-axis represents the strength of the information, with 2 bits being the maximum for a DNA sequence. (B) Bendability of *Sleeping Beauty* and Tc1 target sites. DNase I digestion of radioactively labeled, 32-bp oligos containing either the 8-bp SB (Vigdal *et al.*, 2002) or 10-bp Tc1 consensus insertion sequences (van Luenen and Plasterk, 1994), or two different 8-bp sequences with low predicted bendability was performed. The labeled bands in were quantitated with a PhosphorImager and graphed relative to each other. The uppercase letters indicate the core target sequence, while the lowercase letters indicate the identical flanking sequence. (◆) SB consensus, (■) Tc1 consensus, (△) Bad Bender, (○) Bad Bender2.

two IRs (Figure 2). However, as mentioned above, physical proximity of the DRs is not sufficient for SB transposition; a highly specific configuration of functional DNA elements within the inverted repeats has a critical importance. As mentioned earlier, SB transposase preferentially binds the inner DRs within the transposon inverted repeats. It was also found that HMGB1 enhances transposase binding to both DRs, but its effect is significantly more pronounced at the inner sites. It appears, therefore, that the order of events that take place during the very early steps of transposition is binding of transposase molecules first to the inner sites, and then to the outer sites. The pronounced effect of HMGB1 on binding of the transposase to the inner sites

suggests that HMGB1 enforces ordered assembly of a catalytically active synaptic complex (Figure 2). Indeed, interference with this sequence of events by replacing the outer transposase binding sites with the inner sites abolishes SB transposition (Cui *et al.*, 2002). This ordered assembly process probably controls that cleavage at the outer sites occurs only if all the previous requirements had been fulfilled. An assembly pathway similar to the one we propose for *Sleeping Beauty* has been described for bacteriophage λ (Richet *et al.*, 1986).

The IR/DR-type organization of inverted repeats introduces a higher level regulation into the transposition process. The repeated transposase binding sites, their

dissimilar affinity for the transposase, and the effect of HMGB1 to differentially enhance transposase binding to the inner sites are all important for a geometrically and timely orchestrated formation of synaptic complexes, which is a strict requirement for the subsequent catalytic steps of transposition.

Transposon Excision

All the DDE recombinases catalyze similar chemical reactions (Craig, 1995), which begin with a single-strand nick that generates a free 3'-OH group. To catalyze second strand cleavage, DDE recombinases developed versatile strategies (Turlan and Chandler, 2000). The position of 5'-cleavage of the second strand required for the liberation of the element occurs directly opposite to the 3'-cleavage site in V(D)J recombination (Gellert, 2002) and for the bacterial Tn10 element (Kennedy *et al.*, 1998) (thereby generating blunt ended products), but it is staggered inwards the element by three nucleotides for *Sleeping Beauty* (Luo *et al.*, 1998) and by two nucleotides for the Tc1 and Tc3 elements (van Luenen *et al.*, 1994). Thus, transposon excision leaves behind three-nucleotide-long 3'-overhangs in SB transposition. DNA repair of the broken DNA ends generates transposon "footprints" that are therefore identical to the first or last three nucleotides of the transposon (Luo *et al.*, 1998). In V(D)J recombination, the single-strand nick is converted into a double-strand break by a transesterification reaction in which the free 3'-OH attacks the opposite strand, thereby creating a hairpin intermediate (van Gent *et al.*, 1996; Gellert, 2002). Tn5 and Tn10 transposons also transpose via a hairpin intermediate, with the difference that the hairpin is on the transposon and not on flanking DNA (Kennedy *et al.*, 1998; Bhasin *et al.*, 1999). In contrast to V(D)J recombination, the excision sites do not have a hairpin structure in SB transposition (unpublished results). Whether second-strand cleavage occurs by transesterification or by hydrolysis in SB transposition needs to be investigated.

Transposon Integration: Target Site Selection of *Sleeping Beauty*

Most transposable elements do not integrate randomly into target DNA, and display some degree of specificity in target site utilization (Craig, 1997). Target selection may depend on primary DNA sequence and chromatin structure, which can influence target site utilization by modulating the accessibility of DNA. In some systems, including the bacterial transposon Tn10 and the Tc1 and Tc3 transposons in *Caenorhabditis elegans*, target site selection is primarily determined by the transposase itself (van Luenen and Plasterk, 1994; Junop and Haniford, 1997). Sequences responsible for target site selection of Tn10 and retroviruses have been mapped to the core catalytic domain of the transposase (or integrase) (Katzman and Sudol, 1995; Junop and Haniford, 1997), containing the DDE signature.

Several members of the DDE domain recombinase family integrate fairly randomly, yet not all possible sites are utilized within a genome with equal frequencies. Despite the implication that the conserved catalytic domain is responsible for locating the target site, no common pattern

of integration can be recognized on the sequence level. Therefore, assuming that there might be common features of target selection in the DDE family, it is an attractive hypothesis that structural properties of the target DNA will be among them. Because of the great potential of SB in genetic applications, determination of the factors affecting specificity of SB's target site selection is of great importance.

Sleeping Beauty Displays a Random Pattern of Integration in the Human Genome

In order to analyze SB's insertion profile on the genomic level, transposon insertions were generated in human HeLa cells using the *in vivo* transposition assay shown in Figure 1B. 138 insertion sites were identified and mapped on human chromosomes, using NCBI's human genome BLAST service. Although some chromosomes were hit more frequently than others, no clear preference was apparent for any chromosome, or for certain subchromosomal regions (Vigdal *et al.*, 2002). This observation indicates that most (if not all) chromosomes can serve as good targets for transposition. 35% of the transposition events occurred in transcribed regions. Because about one third of the human genome is estimated to be transcribed (I. H. G. S. C., 2001), this frequency suggests no preference for or against insertion into genes. Taken together, these results indicate a fairly random pattern of integration of SB elements in human chromosomes.

Sleeping Beauty Prefers a Palindromic AT-repeat for Insertion

Sleeping Beauty, like all other Tc1/*mariner* elements, integrates at TA dinucleotides (Plasterk *et al.*, 1999), which occur approximately once every 20 basepairs, on average, in vertebrate genomes. We investigated whether all TAs are equally good targets, or if other sequence determinants exist that influence SB's target site selection. Chromosomal sequences flanking integrated transposons were used to determine the DNA sequence of a consensus target site. We found six bases directly surrounding the insertion site forming a short, palindromic AT-repeat: ATATATAT (Vigdal *et al.*, 2002), in which the central underlined TA is the insertion site (Figure 3A).

Both major steps of DNA transposition, i. e. excision and integration, are catalyzed by the same catalytic domain of the transposase. We investigated whether excision and integration are under the same rules with respect to primary sequence of the DNA that serves either as a target for transposition or as a donor site for element excision. If that were the case, an element that had transposed into the consensus target would excise more efficiently out of this site than from others. A collection of SB elements flanked by different base compositions were tested for their respective transpositional efficiencies. No dramatic effects on transposition were observed; the transposition efficiency ranged from a low of 50% to a high of 128% (Vigdal *et al.*, 2002). In general, homopolymeric stretches of nucleotides make poor excision substrates, whereas alternating purine and pyrimidine bases flanking the element are more efficient donor substrates. Moreover, the preferred

integration target sequence does not constitute an efficient donor site. Apparently, SB is only moderately sensitive to the base composition of sequences immediately flanking the transposon at the excision step of transposition.

Structural Properties of DNA at Sleeping Beauty Insertion Sites

A preference for inserting into a particular sequence raises the possibility that integration sites might have common structural features. We found three properties of DNA that together define preferred sites for integration of *Sleeping Beauty* and other Tc1/*mariner* transposons. These are bendability (Figure 3B), A-philicity and a symmetrical pattern of hydrogen bonding sites in the major groove of the target DNA (Vigdal *et al.*, 2002). DNA bending can lead to changes in the width and depth of the major and minor grooves, affecting a protein's access to bases of the DNA (Brukner *et al.*, 1995). A bendable structure may allow transposase and/or auxiliary host factors, to deform the bound DNA into a spatial optimum for strand transfer. A-philicity represents the propensity of DNA to form an A-DNA like double helix (Ivanov and Minchenkova, 1995). A-DNA has a wide and shallow minor groove that is believed to provide proteins easier access to form hydrogen bonds with bases within the DNA helix. We have detected a 10-bp palindromic pattern of hydrogen bonding for both *Sleeping Beauty* and Tc1 genomic insertions (Vigdal *et al.*, 2002). Such palindromic pattern and the symmetry of the consensus target site sequence (Figure 3A) together indicate that the target DNA is recognized by a dimeric or multimeric form of the transposase, consistent with our earlier finding that SB transposase forms tetramers in solution (Izsvák *et al.*, 2002).

Alltogether the data indicate that, similar to some other transposable elements, target site selection of SB shows considerable specificity, and that it is primarily determined on the DNA structural rather than on the sequence level. The results indicate that a combination of particular physical properties generate a spatial optimum of the DNA for transposase interaction, and the ability of the transposase polypeptide to efficiently interact with such sequences specifies the sites where insertions occur.

Regulation of *Sleeping Beauty* Transposition by Intrinsic Factors

Transposon movement is usually restricted in genomes to minimize the mutational damage inflicted on the host cell (Hartl *et al.*, 1997). Consequently, naturally occurring transposons may not be the most active ones. Transposon size is among the factors that can affect the efficiency of transposition. Indeed, deleted, inactive transposon copies frequently accumulate in genomes, probably because the shorter versions transpose better. Small size may not be an absolute requirement for mobilization, but it has been observed for different transposons, including SB, that longer elements tend to transpose less efficiently (Lampe *et al.*, 1998; Izsvák *et al.*, 2000). The efficiency of *Sleeping Beauty* transposition is higher when the donor transposon is present on a plasmid as opposed to a chromosome (Luo *et al.*, 1998). Furthermore, reducing the outer distance

between the two transposon ends in a plasmid-based vector enhances transposition (Izsvák *et al.*, 2000). However, when the transposon ends are less than 300 bp apart from each other, transposition rates drop. These observations collectively indicate that physical distance between the transposon inverted repeats is an important determinant of the transposition process, perhaps by affecting synaptic complex assembly or stability.

Transposition of *mariner* elements was shown to be regulated by "overproduction inhibition"; an increase in the amount of transposase, either by increasing gene dosage or by increasing expression of a single gene, results in a decrease in transpositional activity (Lohe and Hartl, 1996). We have found in tissue culture experiments that the efficiency of SB transposition positively correlates with the number of available transposase and transposon molecules within a well-defined range (Izsvák *et al.*, 2000). However, high level SB transposase gene dosage in living mice resulted in repressed transpositional activity, suggesting the existence of a regulatory mechanism similar to overproduction inhibition (Yant *et al.*, 2000). Future work will have to address the issue of regulation of SB transposition by transposase overexpression.

Links between DNA Transposition, Retroviral Integration and V(D)J Recombination

It has been proposed that there are mechanistic links between different recombination reactions such as the cut-and-paste transposition used by Tc1/*mariner* elements, bacteriophage Mu transposition, and retroviral integration (Craig, 1995). Likewise, V(D)J recombination is a transposition-like reaction (van Gent *et al.*, 1996). The recombinases responsible for these reactions all contain DDE domains (Plasterk *et al.*, 1999). The similarity between V(D)J recombination and transposition was further supported by the discovery that the RAG proteins can mediate transposition of DNA flanked by recombination signal sequences (Agrawal *et al.*, 1998; Hiom *et al.*, 1998), suggesting that the V(D)J recombination machinery evolved from an ancient RAG transposon.

The similarities are further emphasised in requirements for host factors. HMGB1/2 were shown to enhance V(D)J recombination, by enhancing binding of the RAG proteins to recombination signal sequences, and enforcing the cleavage to occur only at a defined combination of different recombination signals, which is termed the 12/23 rule (van Gent *et al.*, 1997; Aidinis *et al.*, 1999). HMGA1, that belongs to another family of HMG proteins, is required for retroviral integration (Hindmarsh *et al.*, 1999; Li *et al.*, 2000). In addition to the HMG proteins, the requirement for cellular proteins involved in double-strand DNS break repair provides an addition to the several links between DNA transposition, V(D)J recombination (Gellert, 2002) and retroviral integration (Daniel *et al.*, 1999) that exist on evolutionary, structural and functional levels.

Genetic Applications

Sleeping Beauty Transposition in Somatic Tissues of the Mouse

In vitro induction of transposition and selection of cultured cells that harbor integrated transposons in their chromosomes can be useful when combined with embryonic stem cell technology (Luo *et al.*, 1998). However, there is considerable interest in technologies that allow the delivery and expression of genes in certain tissues or organs *in vivo*, for the correction of genetic diseases.

Evidence that the SB system can potentially be developed as a useful vector for gene therapy came from experiments in which the two components of the transposon system were administered into living mice by tail vein injection (Table 2) (Yant *et al.*, 2000). Using this simple technology, about 5% of hepatocytes of the experimental animals expressed a foreign marker protein, β -galactosidase, from the *lacZ* gene within the transposon vector. Taking into account that the transposon vector cannot infect cells (thus active cellular uptake is not promoted), a 5% transformation efficiency is a significant result, because other integrating and infectious vectors such as retroviruses and adeno-associated virus vectors also transform hepatocytes *in vivo* with similar efficiencies. Thus, SB can mediate efficient chromosomal integration of transgene constructs *in vivo* in a mammalian model

system. For gene therapy, chromosomal integration of transgene constructs itself does not solve the problem, because in many cases the transferred gene has to be expressed for a prolonged period of time, and it has to express the gene product at a level that will have a therapeutic effect. The following four studies clearly demonstrated that the SB transposon can fulfill both requirements (Table 2). First, transgenic mice generated with an SB vector containing the human α -1-antitrypsin (hAAT) cDNA expressed hAAT in their blood for more than 6 months (Yant *et al.*, 2000). Second, transfer of an SB vector containing a human Factor IX (FIX) expression cassette resulted in partial correction of the bleeding disorder in haemophilic mice (Yant *et al.*, 2000), and sustained the production of biologically active FIX at levels which would convert a severely affected patient with haemophilia B to one with a much milder phenotype. Third, SB-mediated gene therapy in fumarylacetoacetate hydrolase (FAH)-deficient mice has been shown to correct hereditary tyrosinemia type 1 in 62 % of the animals receiving a FAH-expressing transposon construct and the transposase (Montini *et al.*, 2002). This latter study demonstrated an average transposon copy number of 1/ diploid cellular genome in the liver, and a long-lasting transgene expression even after serial transplantation of hepatocytes. Finally, the applicability of transposon-mediated gene transfer in human tissues was demonstrated in an *ex vivo* study, in which skin tissues of

Table 2. *In vivo* gene transfer in vertebrates in somatic tissues and in the germline using *Sleeping Beauty*.

<i>Vector</i>	<i>Organism</i>	<i>Target</i>	<i>Method of transfer</i>	<i>Transferred gene</i>	<i>Aim</i>	<i>Comment</i>	<i>Ref.</i>
plasmid	mouse	hepatocyte	hydrodynamic tail vein injection	<i>lacZ</i> α -antitrypsin Factor IX	gene therapy	5-6% integration/ long term expr.	Yant <i>et al.</i> , 2000
adonovirus/ transposon hybrid	mouse	hepatocyte	tail vein injection/ infection	<i>lacZ</i> Factor IX	gene therapy	therapeutic level/ long term expr.	Yant <i>et al.</i> , 2002
plasmid	human	keratinocyte	transfection	LAMB3	gene therapy	17% integration/ therapeutic level/ long term expr.	Ortiz-Urda <i>et al.</i> , 2003
plasmid	mouse	hepatocyte	hydrodynamic tail vein injection	FAH	gene therapy	4% integration/ long term expr.	Montini <i>et al.</i> , 2002
linearized plasmid+ mRNA	mouse	germ line	pronuclear injection	GFP	germ line transgenesis	transposon array, 3 transposition/ embryo, long term expr.	Dupuy <i>et al.</i> , 2002
plasmid+ mRNA	zebrafish	germ line	microinjection	GFP	germ line transgenesis	long term expr.	Dupuy <i>et al.</i> , 2002
linearized DNA	mouse	germ line	oocyte injection	neomycin	mutagenesis	0.2 transposition/ sperm	Fischer <i>et al.</i> , 2001
linearized gene trap plasmid+ mRNA	mouse	germ line	pronuclear injection	GFP	mutagenesis	2 transposition/ sperm	Dupuy <i>et al.</i> , 2001
linearized DNA	mouse	germ line	pronuclear injection	GFP	mutagenesis	1 transposition/ sperm	Horie <i>et al.</i> , 2001

junctional epidermolysis bullosa patients were transfected with an SB vector expressing laminin (Ortiz-Urda *et al.*, 2003). Transformed cells were shown to regenerate healthy human skin with normal laminin expression, and thus to enable genetic correction of a life-threatening disease (Ortiz-Urda *et al.*, 2003).

Together, the above studies have demonstrated the potential usefulness of *Sleeping Beauty* for the correction of human genetic diseases. One problem with respect to further applications is that the efficiency of *in vivo* gene transfer into many types of tissue with naked DNA constructs is rather low, therefore the overall transformation rates with plasmid-borne SB vectors can be insufficient in clinical applications. A second problem is that the hydrodynamic injection method used in the above studies (injection of a large volume of DNA solution into the bloodstream in a couple of seconds) is hardly applicable in humans. A potential solution to both problems was offered by engineering an adenovirus/SB hybrid vector (Table 2) (Yant *et al.*, 2002). Adenovirus vectors are very

efficient at infecting cells, but transgene expression from these vectors is transient due to the lack of stable genomic integration. Repeated administration of adenovirus vectors can induce an immune response against viral proteins and the elimination of transduced cells. The adenovirus/SB hybrid combines the advantages of the two systems: high efficiency gene transfer and stable transgene integration and expression.

Sleeping Beauty Transposition in the Vertebrate Germline

A requirement for the generation of stable transgenic stocks and for insertional inactivation of endogenous genes is that genetic changes are directed to the germline so that mutations will be passed on to the next generation. Classical methods to express foreign genes in vertebrate animals rely on injection of nucleic acids into oocytes or fertilized eggs. These techniques are relatively simple, but because plasmid-borne genes are not equipped to promote chromosomal integration, their presence and expression is usually transient and mosaic, and very rarely results in

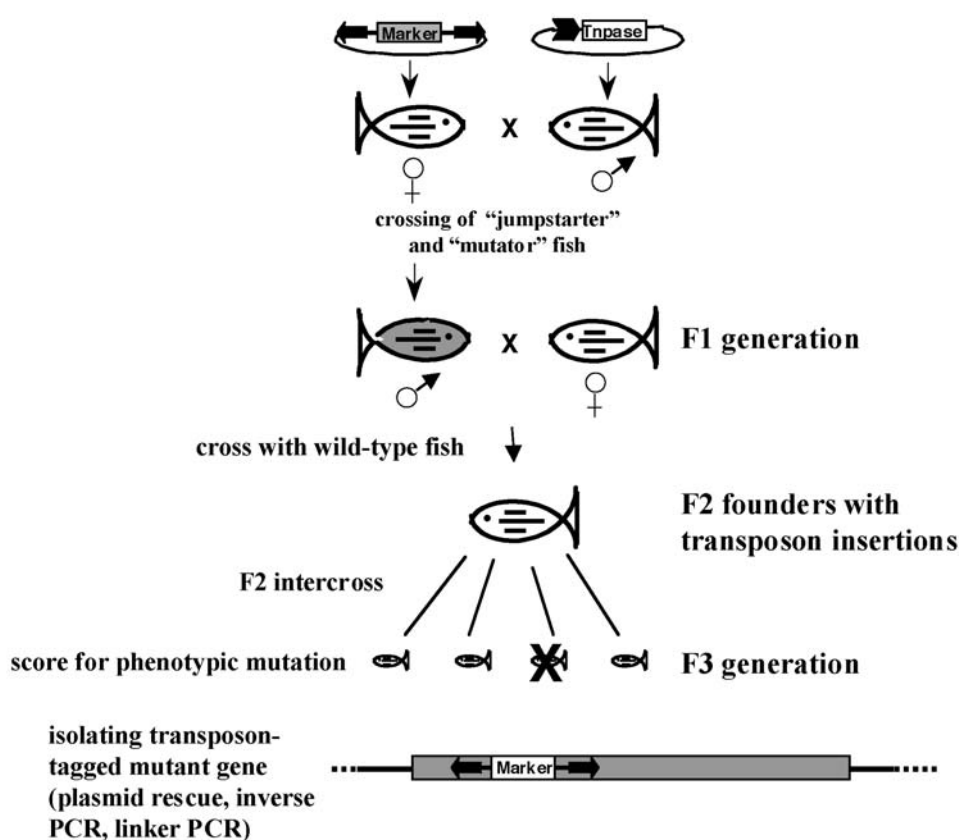


Figure 4. Experimental strategies to induce transposition in the vertebrate germline. The figure shows a possible strategy for the zebrafish, but some of these methods can be adapted to other vertebrate species. The "jumpstarter" stock expresses the transposase in the male germline, and the "mutator" stock contains the mutagenizing transposon, preferably equipped with a gene trap marker. The two components are brought together in a hybrid after breeding of the two lines. Transposition events will occur in the germline of males of F1 heterozygotes. These animals are crossed with wild-type females to segregate the insertion events in individuals of the F2 generation. F2 intercross yields F3 progeny, 25% of which is expected to contain the transposon insertion allele in a homozygous form. Phenotypic changes can be observed in F3. Mutant genes can easily be cloned by different PCR methods making use of the inserted transposon as a unique sequence tag.

genomic integration. An improvement of germline transgenesis by *Sleeping Beauty* transposition has been demonstrated in zebrafish (Nasevicius and Ekker, 2000) and in the mouse (Dupuy *et al.*, 2002) (Table 2). The mouse experiments were done by pronuclear microinjection of a GFP-marked transposon together with SB transposase mRNA synthesized *in vitro* (Dupuy *et al.*, 2002). A transgenic frequency of 45% was established (as opposed to 29% in the absence of transposase) (Table 2). Integrated transposons undergo germline transmission, and are expressed in the offspring. These results demonstrate the usefulness of transposition for the generation of stable transgenic lines in vertebrate models.

Breeding of "Jumpstarter" and "Mutator" Stocks to Induce Transposition in the Germline of the Hybrid

This is most likely the preferred method for generating large numbers of transposon insertions for insertional mutagenesis, a method that cannot be applied for retrovirus vectors. In this experimental setup, two transgenic lines need to be generated first; a "jumpstarter" stock expressing the transposase in the male germline, and a "mutator" stock containing the transposon to be mobilized (Figure 4). These two stocks are crossed to bring the two components of the transposon system together, and transposition is expected to occur in the sperm cells of males of the heterozygous hybrids. Such males would be crossed to wild type females to segregate the different insertion events in the genomes of their sperm cells into separate animals (Figure 4).

Recent evidence suggests that such a scheme might be useful for generating insertional mutations in vertebrate model systems (Fischer *et al.*, 2001). Two separate transgenic mouse lines have been established: one expressing the SB transposase from the protamine 1 promoter which is active during spermiogenesis, and the other containing an integrated *neo*-marked SB element. In 20% of the offspring of double-transgenic males the transposon jumped to different genomic locations, and transposon insertions are stably transmitted in the absence of the transposase (Fischer *et al.*, 2001) (Table 2). Subsequent to this study, two other papers describing a similar experimental design and even more encouraging results were published (Dupuy *et al.*, 2001; Horie *et al.*, 2001). Both studies employed a ubiquitous promoter to drive the expression of the Sleeping Beauty transposase in transgenic mice, and a multicopy array of transposons as donors for transposition. Horie *et al.* (2001) have found up to 80% of the progeny of a double-transgenic male to contain transposition events, and estimated the frequency of germline transposition to be about one event per gamete, whereas Dupuy *et al.* (2001) estimate that, on average, their double-transgenic males carry about two new transposon insertions per sperm cell. Although further studies are required to establish a protocol for inducing large numbers of transposon insertions in the mouse germ line, it appears that promoter choice and a chromosomal pool of transposons with a sufficiently large number of elements available for mobilization will be among the important parameters.

What frequency of transposition is required for optimal mutagenesis? Insertional mutagenesis is less efficient than chemical mutagenesis. However, on average, the

phenotypic effect of a transposon insertion is more dramatic than that of a single nucleotide substitution. Although relatively random insertion of transposon vectors can be a clear advantage for gene identification through insertional mutagenesis, a limiting factor can be the overall frequency of transposition. A critical parameter contributing to the success of insertional mutagenesis with transposable elements will be if multiple transposon insertions per gamete can be generated. The relative inefficiency of gene inactivation with insertional mutagens can be counterbalanced by the fecundity of model species such as fish or frogs, where several thousand embryos can be generated in a single mating.

Preferably, the transposon would contain a gene trap construct to allow for selection of transposon insertions into genes. In the mutator stock, the gene trap must not be expressed. Offspring of double-transgenic males can be examined for expression of the gene trap marker such as the green fluorescent protein (GFP), which is indicative of transposition of the marker elements into expressed genes. Usually, mutations can only be observed when products of both copies of a gene are inactivated. By bringing the insertions identified in the founder animals to homozygosity, animals that contain two mutant copies of the affected gene can be generated. The spatial/temporal expression of the transposon marker and possible phenotypic effects (mutations) can be examined over the course of embryonic development, and the affected tissues/organs/developmental pathways can be colocalized with the marker.

Acknowledgements

We thank D. Fiedler, E. Stüwe and A. Katzer for their technical assistance. Work in this laboratory is supported by EU grant QLG2-CT-2000-00821.

References

- Agrawal, A., Q.M. Eastman, and D.G. Schatz. 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature*. 394: 744-751.
- Aidinis, V., T. Bonaldi, M. Beltrame, S. Santagata, M.E. Bianchi, and E. Spanopoulou. 1999. The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. *Mol. Cell. Biol.* 19: 6532-6542.
- Ariyoshi, M., D.G. Vassilyev, H. Iwasaki, H. Nakamura, H. Shinagawa, and K. Morikawa. 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. *Cell*. 78: 1063-1072.
- Bhasin, A., I.Y. Goryshin, and W.S. Reznikoff. 1999. Hairpin formation in Tn5 transposition. *J. Biol. Chem.* 274: 37021-37029.
- Breitling, R. and J.-K. Gerber. 2000. Origin of the paired domain. *Dev. Genes. Evol.* 210: 644-650.
- Brukner, I., R. Sanchez, D. Suck, and S. Pongor. 1995. Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J.* 14: 1812-1818.

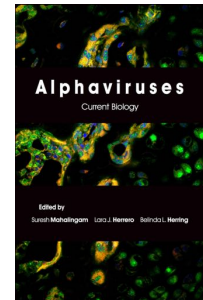
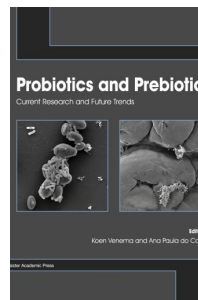
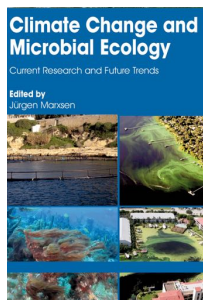
- Bustin, M. 1999. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell. Biol.* 19: 5237-5246.
- Capy, P., R. Vitalis, T. Langin, D. Higuier, and C. Bazin. 1996. Relationships between transposable elements based upon the integrase-transposase domains: is there a common ancestor? *J. Mol. Evol.* 42: 359-368.
- Craig, N.L. 1995. Unity in transposition reactions. *Science*. 270: 253-254.
- Craig, N.L. 1997. Target site selection in transposition. *Annu. Rev. Biochem.* 66: 437-474.
- Cui, Z., A.M. Geurts, G. Liu, C.D. Kaufman, and P.B. Hackett. 2002. Structure-function analysis of the inverted terminal repeats of the *Sleeping Beauty* transposon. *J. Mol. Biol.* 318: 1221-1235.
- Czerny, T., G. Schaffner, and M. Busslinger. 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* 7: 2048-2061.
- Daniel, R., R.A. Katz, and A.M. Skalka. 1999. A role for DNA-PK in retroviral DNA integration. *Science*. 284: 644-647.
- Davies, D.R., I.Y. Goryshin, W.S. Reznikoff, and I. Rayment. 2000. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*. 289: 77-85.
- Doak, T.G., F.P. Doerder, C.L. Jahn, and G. Herrick. 1994. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc. Natl. Acad. Sci. U.S.A.* 91: 942-946.
- Dupuy, A.J., K. Clark, C.M. Carlson, S. Fritz, A.E. Davidson, K.M. Markley, K. Finley, C.F. Fletcher, S.C. Ekker, P.B. Hackett, S. Horn, and D.A. Largaespada. 2002. Mammalian germ-line transgenesis by transposition. *Proc Natl Acad Sci U S A* 99: 4495-4499.
- Dupuy, A.J., S. Fritz, and D.A. Largaespada. 2001. Transposition and gene disruption in the male germline of the mouse. *Genesis*. 30: 82-88.
- Dyda, F., A.B. Hickman, T.M. Jenkins, A. Engelman, R. Craigie, and D.R. Davies. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*. 266: 1981-1986.
- Feng, J.A., R.C. Johnson, and R.E. Dickerson. 1994. Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science*. 263: 348-355.
- Fischer, S.E., E. Wienholds, and R.H. Plasterk. 2001. Regulated transposition of a fish transposon in the mouse germ line. *Proc. Natl. Acad. Sci. U.S.A.* 98: 6759-6764.
- Franz, G., T.G. Loukeris, G. Dialektaki, C.R. Thompson, and C. Savakis. 1994. Mobile *Minos* elements from *Drosophila hydei* encode a two-exon transposase with similarity to the paired DNA-binding domain. *Proc. Natl. Acad. Sci. U.S.A.* 91: 4746-4750.
- Franz, G. and C. Savakis. 1991. *Minos*, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucleic Acids Res.* 19: 6646.
- Gellert, M. 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem.* 71: 101-132.
- Goodier, J.L. and W.S. Davidson. 1994. Tc1 transposon-like sequences are widely distributed in salmonids. *J. Mol. Biol.* 241: 26-34.
- Haring, M.A., C.M. Rommens, H.J. Nijkamp, and J. Hille. 1991. The use of transgenic plants to understand transposition mechanisms and to develop transposon tagging strategies. *Plant Mol Biol* 16: 449-461.
- Hartl, D.L., E.R. Lozovskaya, D.I. Nurminsky, and A.R. Lohe. 1997. What restricts the activity of *mariner*-like transposable elements? *Trends Genet.* 13: 197-201.
- Hindmarsh, P., T. Ridky, R. Reeves, M. Andrade, A.M. Skalka, and J. Leis. 1999. HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration *in vitro*. *J. Virol.* 73: 2994-3003.
- Hiom, K., M. Melek, and M. Gellert. 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell*. 94: 463-470.
- Horie, K., A. Kuroiwa, M. Ikawa, M. Okabe, G. Kondoh, Y. Matsuda, and J. Takeda. 2001. Efficient chromosomal transposition of a Tc1/*mariner*-like transposon *Sleeping Beauty* in mice. *Proc. Natl. Acad. Sci. U.S.A.* 98: 9191-9196.
- International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature*. 409: 860-921.
- Ivanov, V.I. and L.E. Minchenkova. 1995. The A-form of DNA: in search of the biological role. *Mol. Biol.* 28: 1258-1271.
- Ivics, Z., P.B. Hackett, R.H. Plasterk, and Z. Izsvák. 1997. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*. 91: 501-510.
- Ivics, Z., Z. Izsvák, A. Minter, and P.B. Hackett. 1996. Identification of functional domains and evolution of Tc1-like transposable elements. *Proc. Natl. Acad. Sci. U.S.A.* 93: 5008-5013.
- Izsvák, Z., Z. Ivics, and P.B. Hackett. 1995. Characterization of a Tc1-like transposable element in zebrafish (*Danio rerio*). *Mol. Gen. Genet.* 247: 312-322.
- Izsvák, Z., Z. Ivics, and R.H. Plasterk. 2000. *Sleeping Beauty*, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* 302: 93-102.
- Izsvák, Z., D. Khare, J. Behlke, U. Heinemann, R.H. Plasterk, and Z. Ivics. 2002. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in *Sleeping Beauty* transposition. *J. Biol. Chem.* 277: 34581-34588.
- Junop, M.S. and D.B. Haniford. 1997. Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. *EMBO J.* 16: 2646-2655.
- Katayanagi, K., M. Miyagawa, M. Matsushima, M. Ishikawa, S. Kanaya, M. Ikehara, T. Matsuzaki, and K. Morikawa. 1990. Three-dimensional structure of ribonuclease H from *E. coli*. *Nature*. 347: 306-309.
- Katzman, M. and M. Sudol. 1995. Mapping domains of retroviral integrase responsible for viral DNA specificity and target site selection by analysis of chimeras between human immunodeficiency virus type 1 and visna virus integrases. *J. Virol.* 69: 5687-5696.
- Ke, Z., G.L. Grossman, A.J. Cornel, and F.H. Collins. 1996. *Quetzal*: A transposon of the Tc1 family in the mosquito *Anopheles albimanus*. *Genetica*. 98: 141-147.

- Kennedy, A.K., A. Guhathakurta, N. Kleckner, and D.B. Haniford. 1998. Tn10 transposition via a DNA hairpin intermediate. *Cell*. 95: 125-134.
- Kim, D.R., Y. Dai, C.L. Mundy, W. Yang, and M.A. Oettinger. 1999. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev.* 13: 3070-3080.
- Lam, W.L., P. Seo, K. Robison, S. Virk, and W. Gilbert. 1996. Discovery of amphibian Tc1-like transposon families. *J. Mol. Biol.* 257: 359-366.
- Lampe, D.J., M.E. Churchill, and H.M. Robertson. 1996. A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. *EMBO J.* 15: 5470-5479.
- Lampe, D.J., T.E. Grant, and H.M. Robertson. 1998. Factors affecting transposition of the *Himar1 mariner* transposon *in vitro*. *Genetics*. 149: 179-187.
- Landree, M.A., J.A. Wibbenmeyer, and D.B. Roth. 1999. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev.* 13: 3059-3069.
- Leung, P.C., D.B. Teplow, and R.M. Harshey. 1989. Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. *Nature*. 338: 656-658.
- Li, L., K. Yoder, M.S. Hansen, J. Olvera, M.D. Miller, and F.D. Bushman. 2000. Retroviral cDNA integration: stimulation by HMG I family proteins. *J. Virol.* 74: 10965-10974.
- Lohe, A.R. and D.L. Hartl. 1996. Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. *Mol. Biol. Evol.* 13: 549-555.
- Lohe, A.R., E.N. Moriyama, D.A. Lidholm, and D.L. Hartl. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements. *Mol. Biol. Evol.* 12: 62-72.
- Luo, G., Z. Ivics, Z. Izsvák, and A. Bradley. 1998. Chromosomal transposition of a Tc1/*mariner*-like element in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 95: 10769-10773.
- Merriman, P.J., C.D. Grimes, J. Ambroziak, D.A. Hackett, P. Skinner, and M.J. Simmons. 1995. Selements: a family of Tc1-like transposons in the genome of *Drosophila melanogaster*. *Genetics*. 141: 1425-1438.
- Montini, E., P.K. Held, M. Noll, N. Morcinek, M. Al-Dhalimy, M. Finegold, S.R. Yant, M.A. Kay, and M. Grompe. 2002. *In vivo* correction of murine tyrosinemia type I by DNA-mediated transposition. *Mol. Ther.* 6: 759-769.
- Moschetti, R., C. Caggese, P. Barsanti, and R. Caizzi. 1998. Intra- and interspecies variation among *Bari-1* elements of the *melanogaster* species group. *Genetics*. 150: 239-250.
- Nasevicius, A. and S.C. Ekker. 2000. Effective targeted gene "knockdown" in zebrafish. *Nat. Genet.* 26: 216-220.
- Ortiz-Urda, S., Q. Lin, S.R. Yant, D. Keene, M.A. Kay, and P.A. Khavari. 2003. Sustainable correction of junctional epidermolysis bullosa via transposon-mediated nonviral gene transfer. *Gene Ther.* 10: 1099-1104.
- Pellizzari, L., G. Tell, and G. Damante. 1999. Co-operation between the PAI and RED subdomains of Pax-8 in the interaction with the thyroglobulin promoter. *Biochem. J.* 337: 253-262.
- Petrokovski, S. and S. Henikoff. 1997. A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. *Mol. Gen. Genet.* 254: 689-695.
- Plasterk, R.H., Z. Izsvák, and Z. Ivics. 1999. Resident aliens: the Tc1/*mariner* superfamily of transposable elements. *Trends Genet.* 15: 326-332.
- Radice, A.D., B. Bugaj, D.H. Fitch, and S.W. Emmons. 1994. Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. *Mol. Gen. Genet.* 244: 606-612.
- Richet, E., P. Abcarian, and H. Nash. 1986. Synapsis of attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. *Cell*. 46: 1011-1021.
- Rio, D.C., G. Barnes, F.A. Laski, J. Rine, and G.M. Rubin. 1988. Evidence for *Drosophila* P element transposase activity in mammalian cells and yeast. *J. Mol. Biol.* 200: 411-415.
- Spanopoulou, E., F. Zaitseva, F.H. Wang, S. Santagata, D. Baltimore, and G. Panayotou. 1996. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. *Cell*. 87: 263-276.
- Turlan, C. and M. Chandler. 2000. Playing second fiddle: second-strand processing and liberation of transposable elements from donor DNA. *Trends Microbiol.* 8: 268-274.
- van Gent, D.C., K. Hiom, T.T. Paull, and M. Gellert. 1997. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J.* 16: 2665-2670.
- van Gent, D.C., K. Mizuuchi, and M. Gellert. 1996. Similarities between initiation of V(D)J recombination and retroviral integration. *Science*. 271: 1592-1594.
- van Luenen, H.G., S.D. Colloms, and R.H. Plasterk. 1994. The mechanism of transposition of Tc3 in *C. elegans*. *Cell*. 79: 293-301.
- van Luenen, H.G. and R.H. Plasterk. 1994. Target site choice of the related transposable elements Tc1 and Tc3 of *Caenorhabditis elegans*. *Nucleic Acids Res.* 22: 262-269.
- Vigdal, T.J., C.D. Kaufman, Z. Izsvák, D.F. Voytas, and Z. Ivics. 2002. Common physical properties of DNA affecting target site selection of *Sleeping Beauty* and other Tc1/*mariner* transposable elements. *J. Mol. Biol.* 323: 441-452.
- Vos, J.C., I. De Baere, and R.H. Plasterk. 1996. Transposase is the only nematode protein required for *in vitro* transposition of Tc1. *Genes Dev.* 10: 755-761.
- Vos, J.C. and R.H. Plasterk. 1994. Tc1 transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. *EMBO J.* 13: 6125-6132.
- Yant, S.R., A. Ehrhardt, J.G. Mikkelsen, L. Meuse, T. Pham, and M.A. Kay. 2002. Transposition from a gutless adeno-transposon vector stabilizes transgene expression *in vivo*. *Nat Biotechnol* 20: 999-1005.
- Yant, S.R., L. Meuse, W. Chiu, Z. Ivics, Z. Izsvák, and M.A. Kay. 2000. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat. Genet.* 25: 35-41.
- Zayed, H., Izsvák, Z., Khare, D., Heinemann, U. and Z. Ivics. 2003. The DNA-bending protein HMGB1 is a cellular cofactor of *Sleeping Beauty* transposition. *Nucleic Acids Res.* 31: 2313-2322.

Further Reading

Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at [caister.com](http://www.caister.com)

- **MALDI-TOF Mass Spectrometry in Microbiology**
Edited by: M Kostrzewa, S Schubert (2016)
www.caister.com/malдитof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
Edited by: RP Vries, IB Gelber, MR Andersen (2016)
www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
Edited by: RL Dorit, SM Roy, MA Riley (2016)
www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
Edited by: V Bhadauria (2016)
www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
Edited by: R Quatrini, DB Johnson (2016)
www.caister.com/acidophiles
- **Climate Change and Microbial Ecology: Current Research and Future Trends**
Edited by: J Marxsen (2016)
www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
Edited by: G Lear (2016)
www.caister.com/biorem
- **Microalgae: Current Research and Applications**
Edited by: MN Tsaloglou (2016)
www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
Edited by: H Shintani, A Sakudo (2016)
www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
Edited by: N Vasilakis, DJ Gubler (2016)
www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
Edited by: WD Picking, WL Picking (2016)
www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
Edited by: AM Romani, H Guasch, MD Balaguer (2016)
www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
Edited by: S Mahalingam, L Herrero, B Herring (2016)
www.caister.com/alpha
- **Thermophilic Microorganisms**
Edited by: F Li (2015)
www.caister.com/thermophile



- **Flow Cytometry in Microbiology: Technology and Applications**
Edited by: MG Wilkinson (2015)
www.caister.com/flow
- **Probiotics and Prebiotics: Current Research and Future Trends**
Edited by: K Venema, AP Carmo (2015)
www.caister.com/probiotics
- **Epigenetics: Current Research and Emerging Trends**
Edited by: BP Chadwick (2015)
www.caister.com/epigenetics2015
- ***Corynebacterium glutamicum*: From Systems Biology to Biotechnological Applications**
Edited by: A Burkovski (2015)
www.caister.com/cory2
- **Advanced Vaccine Research Methods for the Decade of Vaccines**
Edited by: F Bagnoli, R Rappuoli (2015)
www.caister.com/vaccines
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
Edited by: AT Coste, P Vandeputte (2015)
www.caister.com/antifungals
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
www.caister.com/bacteria-plant
- ***Aeromonas***
Edited by: J Graf (2015)
www.caister.com/aeromonas
- **Antibiotics: Current Innovations and Future Trends**
Edited by: S Sánchez, AL Demain (2015)
www.caister.com/antibiotics
- ***Leishmania*: Current Biology and Control**
Edited by: S Adak, R Datta (2015)
www.caister.com/leish2
- ***Acanthamoeba*: Biology and Pathogenesis (2nd edition)**
Author: NA Khan (2015)
www.caister.com/acanthamoeba2
- **Microarrays: Current Technology, Innovations and Applications**
Edited by: Z He (2014)
www.caister.com/microarrays2
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
Edited by: D Marco (2014)
www.caister.com/n2

