

“睡美人”转座系统研究进展

丁昇 吴晓暉*

(复旦大学发育生物学基地, 上海 200433)

摘要 “睡美人 (*Sleeping Beauty*, SB) ”转座系统是 *Tc1/mariner* 转座因子超家族中的一员, 是目前唯一取材于脊椎动物的具有活性的转座系统. 对近年来有关“睡美人”的研究进展作一个综述, 并针对存在的问题提出相应的解决方案.

关键词 睡美人 (*Sleeping Beauty*), 转座, 转基因

学科分类号 Q344

转座系统在当代遗传学和分子生物学研究中有很大的应用价值. 它不仅能作为反向遗传学的工具, 将外源基因转入受体基因组而使受体产生表型的变化, 同时, 其又因为能通过随机的插入突变使内源基因失活而成为正向遗传学研究中的新宠. 线虫中的 *Tc1* 转座系统和果蝇中的 *P* 因子就是最好的例子. 但由于种属特异性的存在, 长期以来转座系统在脊椎动物中的应用仍是一片空白. 1997年, Ivics 等^[1,2] 基于积累的系统发生数据, 利用生物信息学的手段, 对存在于鱼类中的一个已失活的转座系统进行了分子水平的重建, 唤醒了其转座活性. 由于该转座系统已在水下“沉睡”了一千万年, 因而被取名为“睡美人” (*Sleeping Beauty*, SB). 近几年的研究表明, “睡美人”不仅能在体外培养的鱼、小鼠和人等大多数脊椎动物的细胞中起作用, 而且能在小鼠体细胞内介导转基因稳定整合和长期表达^[3,4]. 另外, 整合的“睡美人”还能通过生殖细胞传递给后代, 在后代中继续发生转座^[5~7]. 因此, 在不久的将来“睡美人”很有希望成为功能基因组研究中的一个强有力的遗传学和分子生物学工具, 在脊椎动物, 尤其是哺乳动物的转基因和大规模遗传筛选研究中大显身手.

1 “睡美人”的分子重建

1.1 策略与对象的选择

要得到一个有活性的转座系统的策略只有两个: 找到一个或构建一个. 与当初通过败育现象在果蝇基因组中找到具有转座活性的 *P* 因子不同, 由于脊椎动物中已知的所有转座系统都因突变的长期积累而失活了, 所以唯一的选择就是将失活的基因加以重建, 以期恢复其活性. 从事鱼类研究的 Ivics 等^[1] 将对象锁定在了 *Tc1/mariner* 转座因子

超家族鲑鱼科亚家族上, 其原因之一是通过分子进化分析可知该亚家族是整个超家族最年青的一个分支, 失活的时间相对较短而积累的突变相对较少, 容易被重建. 更为重要的是该类转座系统通过水平转移已扩散到很多不同的鱼类品种中, 经历了相互独立的进化过程, 这为生物信息学方法的应用创造了先决条件.

1.2 “睡美人”转座酶基因的重建

Ivics 等^[1] 收集了来自 8 个不同鱼类品种中的 12 个失活的鲑鱼科亚家族 *Tc1* 类转座酶基因的序列, 经过多重序列比对 (multiple sequence alignment), 找到了 5 个同源性高的保守基序 (motif). 在这些区域的平均变异率要低于其他区段, 说明这些基序对其原有的转座功能可能是必需的, 且在转座酶失去转座功能以前一直受到选择的压力. 因此, 恢复这 5 个基序的庐山真面目是使“睡美人”“苏醒”的关键.

重建的第一步是通过将两个不同来源的片段互补拼接在一起, 从而恢复一个完整的开放阅读框 (图 1a, SB1~ SB3), 在此过程中还要消除中间未成熟的终止子和移码突变. 由于 SB3 的序列还有 24 个位置上与保守区域的一致序列不符, 因此它还不具备转座能力. 然后, 通过 PCR 引入定点突变的方法, 构建了一系列序列 (图 1a, SB4~ SB10), 它们一个比一个更接近一致序列, 它们所编码蛋白的功能也逐项显现了出来: SB4 基因的一部分作为核定位序列能使与之相连的 *LacZ* 基因在核中表达; 通过泳动延滞分析 (mobility shift assay) 证明 SB8 基因编码的 N 端 123 个氨基酸组

* 通讯联系人.

Tel: 021-65643718, E-mail: morc@fudan.edu.cn

收稿日期: 2002-07-24, 接受日期: 2002-09-12

成的多肽具有特异性的 DNA 结合活性, 这是转座所必需的; 最终完成的 SB10 基因编码长 340 个氨基酸的蛋白质, 在已有功能的基础上, 其 C 端的 DDE 结构域 (在此结构域中起关键作用的是第 153 位与第 244 位的天冬氨酸和 279 位的谷氨酸) 具有转座催化功能, 抗性筛选实验证实了该蛋白质已具备了转座酶活性.

1.3 “睡美人”转座酶的特异识别序列

鲑鱼科亚家族 Tc1 类转座酶的 DNA 识别序列由两部分组成: 210~ 250 个碱基长的反向重复序列 (IR) 和其内部的顺向重复序列 (DR). 两个这样的

识别序列分列于转座区域的两侧. 通过比对发现不同品种的 IR 差异较大而 DR 的保守性很高. Ivics 等^[1]选择了保守性最好的一个品种 *Tanichthys albonubes* 的 IR/DR 作为“睡美人”转座酶特异识别的 DNA 序列, 称为 T 元件. 它包含 230 bp 的 IR 和 32 bp 的 DR.

带有 T 元件的转座子 (transposon) 和介导其转座的转座酶 (transposase) 一同构成了完整的“睡美人”转座系统 (图 1b). 至此, “睡美人”终于苏醒了.

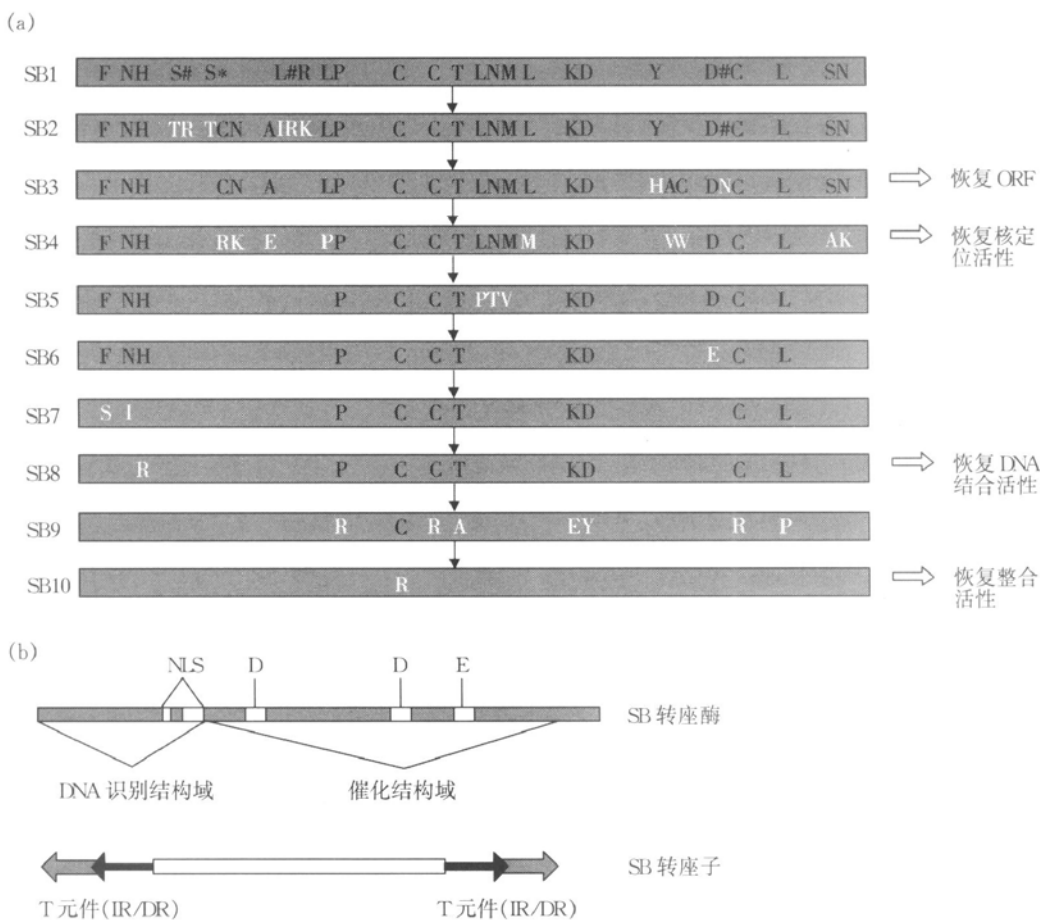


Fig. 1 The reconstruction of *Sleeping Beauty*^[1]

图 1 “睡美人”转座系统的重建过程^[1]

(a) “睡美人”转座酶基因的重建过程. 与一致序列不同的氨基酸残基用黑色字母表示, 经突变成一致序列的氨基酸残基用白色字母表示. “*”表示未成熟的终止密码子, “#”表示移码突变. 右侧箭头旁的说明文字显示在重建的特定阶段 SB 转座酶的结构或功能的恢复情况. (b) “睡美人”转座酶基因各结构域的示意图和“睡美人”转座子结构示意图. “睡美人”转座酶与“睡美人”转座子共同组成“睡美人”转座系统.

2 对“睡美人”转座效率的研究

2.1 检测“睡美人”转座的分子生物学方法

在“睡美人”重新获得转座活性之后, 来自不同国家的几个研究小组, 纷纷投入到对“睡美人”

转座系统的转座效率和转座机理的研究之中. 分子生物学方法的应用使他们的研究深入到分子水平. 表 1 对检测“睡美人”转座的分子生物学方法作一概述.

Table 1 Methods for detecting the transposition of SB

表 1 检测“睡美人”转座的分子生物学方法

方法	简介	用途	相关文献	
抗生素筛选	将抗生素抗性基因导入转座子中, 观察在有无转座酶的情况下经抗性培养基筛选后得到的阳性克隆数的变化	测定转座效率	[1, 3, 5, 8]	
转入报告基因	将 GFP 或半乳糖苷酶基因导入转座子中, 通过 FISH 或蓝白斑筛选观察报告基因的表达	测定转座效率	[3, 7]	
DNA 印迹	以转座子中的特异片段作探针进行杂交, 观察转座前后被标记的限制性酶切片数量和长度的变化	测定转座效率和拷贝数	[1, 5~7]	
PCR	常规 PCR	以转座子外侧的已知序列为引物, 检测转座前后 PCR 产物长度的变化	测定转座效率	[7]
	RT-PCR	检测转座酶在样品中的表达情况	表征转座酶的表达	[8]
	反向-PCR	经过酶切和环化, 以两侧的 T 元件序列为引物扩增出转座子新整合位点两侧的基因组序列	研究转座机制	[8]
	Splinkerette PCR	经过酶切和加接头, 分别以接头中的序列和一侧的 T 元件序列为引物扩增出转座子新整合位点两侧的基因组序列	研究转座机制	[6, 9]

2.2 “睡美人”在不同实验条件下的转座效率

由于实验条件 (包括物种、细胞系、导入转座系统的方法及检测方法等) 的不同, 不同科研小组

得出的“睡美人”转座效率的数据有很大差异. 表 2 结合相应的实验条件总结了各组的实验结果, 以便读者比较.

Table 2 SB transposition efficiencies under different experiment conditions

表 2 “睡美人”在不同实验条件下的转座效率

实验系统	类别	物种	细胞系	导入转座子拷贝数	导入方法	检测方法	转座方向	转座效率 ¹⁾	参考文献	
体外系统	鱼类	鲤鱼	EPC	> 1	转染	新霉素筛选	质粒至染色体	1×10^{-4}	[1, 2]	
	两栖类	爪蟾	A6	> 1	转染	新霉素筛选	质粒至染色体	3×10^{-4}	[2]	
	哺乳类	仓鼠	K1		> 1	转染	新霉素筛选	质粒至染色体	8%	[2]
		奶牛	MDBK		> 1	转染	新霉素筛选	质粒至染色体	4×10^{-3}	[2]
		猴	Cos-7		> 1	转染	新霉素筛选	质粒至染色体	1.8×10^{-3}	[2]
		小鼠	LMTK		> 1	转染	新霉素筛选	质粒至染色体	8×10^{-4}	[1, 2]
			3T3		> 1	转染	新霉素筛选	质粒至染色体	9×10^{-4}	[2]
			胚胎干细胞	1	电穿孔	嘌呤霉素筛选	质粒至染色体	1×10^{-6}	[8]	
			胚胎干细胞	1	电穿孔	嘌呤霉素筛选	染色体之间	3.5×10^{-5}	[8]	
	人	HeLa		> 1	转染	新霉素筛选	质粒至染色体	8.7×10^{-3}	[1, 2]	
体内系统	哺乳类	小鼠 C57/6 ²⁾	肝细胞	> 1	尾部静脉注射	蓝白斑筛选	质粒至染色体	5% ~ 6%	[3]	
		小鼠 FVB/n ²⁾	生殖细胞	1	显微注射	DNA 印迹	染色体之间	20.4%	[5]	
		小鼠 BCF1 ²⁾	生殖细胞	20	显微注射	PCR	染色体之间	67%	[7]	
		小鼠 FVB/n ²⁾	生殖细胞	> 1	显微注射	DNA 印迹	染色体之间	2 ³⁾	[6]	
		小鼠 FVB/n ²⁾	单细胞胚胎	> 1	显微注射	DNA 印迹	质粒至染色体	16%	[4]	

¹⁾除特别说明外, 转座效率是指由于 SB 转座酶的作用, 而发生转座的研究对象 (细胞或个体) 的数量, 占有研究对象数量的百分比 (不考虑在同一研究对象中发生转座的次数). ²⁾表示品系. ³⁾这里的转座效率是指每个个体中由于 SB 转座酶的作用而发生转座的平均次数.

2.3 影响“睡美人”转座效率的因素

2.3.1 实验体系: 从表 2 发现,“睡美人”在小鼠体内系统中的转座效率明显高于在体外培养的小鼠细胞系中的转座效率. 这可能是由于小鼠的某个发育阶段(很可能是配子发育阶段)有利于“睡美人”的转座^[7].

2.3.2 物种和细胞系:“睡美人”能在脊椎动物的许多物种中发生转座,这说明它的种属特异性不强.但在不同物种中或同一物种的不同品系或细胞系中它的转座效率有明显不同,说明“睡美人”的转座可能受到某些物种或细胞特异性分子的影响.

2.3.3 转座子的长度和拷贝数: Izsvak 等^[2]通过使用 HeLa 细胞的实验发现在 2.2 kb 至 10.3 kb 的范围内,“睡美人”转座子的长度每增加 1 kb,其转座效率就会降低 30%.最近, Karsi 等^[10]用小鼠的 3T3 细胞进行的研究也得出了类似的结果.另外,从表 2 中可以看出,在相同条件下,转座子多拷贝比单拷贝更有利于转座的发生.

2.3.4 转座子外侧的序列: Luo 等^[8]在研究“睡美人”在小鼠胚胎干细胞的转座时发现,同样的转座子从质粒转到染色体上的效率和在染色体之间的转座效率有明显的区别(表 2).这很可能是转座子外侧序列的不同(包括长度、结构、原核和真核的不同)对“睡美人”的转座效率产生了影响.

2.3.5 转座酶的表达水平: Yant 等^[3]通过体内实验表明,低的转座酶表达水平对“睡美人”转座更有效.但 Izsvak 等^[2]通过细胞培养却得出了相反的结论.实验条件的不同可能是产生矛盾的原因.因此,要得到更有说服力的结果必须进行更系统更有针对性的研究.

3 “睡美人”的转座原理

使用各种 PCR(表 1)结合测序的方法对“睡美人”的整合位点在转座前后序列上的变化进行了研究^[3,5,7,8].这些研究结果表明“睡美人”的转座采用经典的切割和粘贴(cut-and-paste)机制(图 2).首先,“睡美人”转座酶特异识别转座子两端的 T 元件,将转座子从原整合位点处切下,并在切点处留下原本属于 T 元件末端序列的 3 个碱基,在转座子的两侧形成 5' 端突出.同时,转座酶特异切割基因组其他位置上的 TA 二核苷酸,产生两个 3' 突出的 TA 末端,作为新的整合位点.然后,切离的转座子转移至新的位点处并重新拼接入基因组.最后,利用细胞内的 DNA 修复机制将

所有剩下的缺口补平.因此,一旦“睡美人”转座事件发生,在理论上应留下以下痕迹:原整合位点处 3 个碱基长的转座标签(transposon-tag)和新整合位点两侧的 TA 二碱基正向重复.

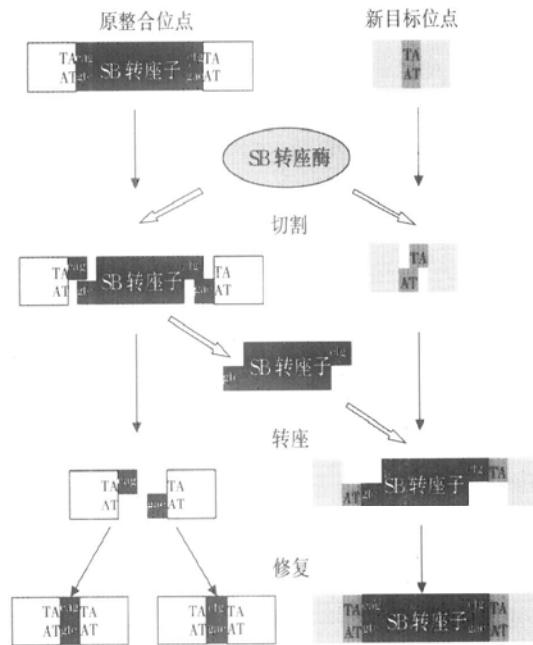


Fig. 2 The transposition mechanism of SB^[8]
图 2 “睡美人”的转座原理^[8]

4 “睡美人”转座系统的应用前景

“睡美人”转座系统才重建不久,目前的研究还集中在其转座功能的本身,但其巨大的应用潜力已经展示在我们的面前.

4.1 在遗传筛选方面的应用前景

正向遗传学的主要任务是从表型出发研究基因的功能,其常用的手段是进行基因组范围的突变筛选.在线虫和果蝇中,基于 DNA 转座系统的大规模突变筛选的策略已成功建立.而在小鼠中由于缺乏有活性的转座系统,相似的策略迟迟未能形成.“睡美人”转座系统在小鼠生殖细胞和体细胞中的成功转座将有助于填补这一空白.与目前普遍使用的化学诱变剂相比,转座系统的优势在于其在引入插入突变的位置上留有标签,便于研究者定位.“睡美人”的另一优势是其转座子结构有很大的改造空间. Dupuy 等^[6]将“睡美人”加以改造,使其具有了捕捉多聚腺苷酸尾(polyA-trapping)的功能,从而能更有针对性地研究未知基因的功能.另外, Fisher 等^[5,8]报道,“睡美人”有在同一个染色体上发生邻近转座的偏好,这有利于对连锁座位的突变分析.

4.2 在转基因方面的应用前景

转基因是从序列出发研究基因功能的重要手段。从“睡美人”转座系统将 LacZ 和 GFP 等报告基因整合入多种物种和细胞系中，并产生稳定表达的实验结果，可以看出其作为转基因载体的潜力。最近，Dupuy 等^[4]对 SB 介导的转基因作了专门研究，他们将带有 GFP 或 agodi 基因（决定小鼠毛色）的转座子线性化后，与体外转录的编码 SB10 转座酶的 mRNA 一同注射入小鼠的单细胞胚胎，结果表明 SB 转座酶的存在提高了转基因的效率，而且转入的外源基因能通过生殖细胞传递给后代。与另一类常用的转基因载体——病毒载体相比较，“睡美人”转座系统具有以下优势：a. 结构简单。与外源基因共同整合入受体基因组的只是两侧各 250 bp 的 IR/DR 序列，对外源基因的影响较小。b. 与反转录病毒载体相比，对插入片段的长度限制较小。c. “睡美人”转座酶的表达水平可影响转座，因此可尝试利用不同的启动子调控转座酶的表达量，从而进一步实现对外源基因表达量的控制。另外，也可尝试在“睡美人”转座酶的编码序列前带上组织特异性或底物诱导的启动子，实现外源基因的时空特异性表达。

4.3 在基因治疗方面的应用前景

Yant 等^[3]证明了“睡美人”能在成年小鼠的体细胞中发生转座，并通过尾静脉注射的方法，将插入了人类凝血因子 IX 的“睡美人”转座子导入患有 B 型血友病的小鼠中，发现小鼠体内的凝血因子 IX 的表达量提高了 80 倍，并有 5 个月以上的持续表达，同时伴随着小鼠凝血时间的明显缩短。另外，利用“睡美人”转座系统在造血干细胞中表达外源基因，以用于基因治疗的尝试也正在进行之中^[11]。虽然转座的随意性和安全性等问题有待解决，但这使“睡美人”转座系统用于体细胞的基因治疗成为可能。

5 存在的问题与解决方案

5.1 “睡美人”转座效率偏低的问题

从表 2 的整体看来，“睡美人”的转座效率还偏低，尤其不能满足应用于大规模突变筛选所要达到的高诱变率的要求。解决这个问题可以从治标和治本两方面考虑。

5.1.1 治标之策：根据前面提及的影响“睡美人”转座效率的因素，尽可能在设计实验时选择有利于提高效率的因素，如摸索合适的转座酶浓度、控制转座子的长度和选择在染色体间转座等。另外，还

可以引入多个转座子，这样可能会增加转座酶的识别几率，而且不同转座子间可能会发生协同作用而便于转座。

5.1.2 治本之策：转座效率不高也可能是转座系统本身不够完善所致。其实，利用同源序列比对的方法重建“睡美人”是存在着一定缺陷的。该方法的前提是“睡美人”在各鱼类品种中水平转移之前具有完全的活性，只是在转移之后才逐渐失活。但也存在着另一种可能，即“睡美人”在第一次水平转移之前就由于突变而失活或活性下降，而这种突变在每一个序列中都存在，通过比对是发现不了的。这至少给我们以启示：对现有的“睡美人”转座系统本身进一步加以改造，可能会提高其转座的效率。据 Cui 等^[12]最新报道，通过改变 T 元件内的 4 个特定碱基可使 SB 在 HeLa 细胞中的转座效率提高 3 倍。采用体外对“睡美人”转座酶基因和 T 元件进行随机的点突变，并结合抗生素筛选检测转座效率的方法可能寻找到新的“睡美人”改进型。

5.2 位置效应引起的转基因静默现象

Horie 等^[7]在利用“睡美人”转座系统将 GFP 基因转入小鼠的实验中发现，虽然用 DNA 印迹法检测到基因已转入了小鼠基因组但用 FISH 检测不到其表达。除了可能是实验方法灵敏度不够以外，很可能是由于转入的基因插到了基因组中的异染色质区，影响了表达。这种现象在其他转基因实验中也时常发生，被称为位置效应引起的转基因静默现象。

要解决这一问题，可尝试在转入基因的两端加上某些顺式作用调控元件，其目的是隔绝周围的染色质状态对转基因表达的影响。有以下元件可供候选^[13, 14]（表 3）。

Table 3 The cis-acting regulatory elements to eliminate the position effect

表 3 可用于消除位置效应的顺式作用调控元件

元件名称	简介
基质附着区和骨架附着区 (matrix attachment region and scaffold attachment region)	能与核骨架结合的 DNA 片段 两个 MAR 或 SAR 间能形成一个相对独立的染色质区域
绝缘子 (insulator)	能通过消除周围染色质的影响，产生相对独立的功能域
座位控制区 (locus control region)	能隔绝周围染色质的影响 有细胞系特异的增强子功能 影响复制时间和复制起点的选择

经过千万年的沉睡, 梦醒时分的“睡美人”正在撩开她神秘的面纱. 在不久的将来, 她动人的全貌必将展现于世.

参 考 文 献

- 1 Ivics Z, Hackett P B, Plasterk R H, *et al.* Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, 1997, **91** (4): 501~ 510
- 2 Izsvak Z, Ivics Z, Plasterk R H. *Sleeping Beauty*, a wide host-range transposon vector for genetic transformation in vertebrates. *J Mol Biol*, 2000, **302** (1): 93~ 102
- 3 Yant S R, Meuse L, Chiu W, *et al.* Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genetics*, 2000, **25** (1): 35~ 41
- 4 Dupuy A J, Clark K, Carlson C M, *et al.* Mammalian germ-line transgenesis by transposition. *Proc Natl Acad Sci USA*, 2002, **99** (7): 4495~ 4499
- 5 Fischer S E J, Wienholds E, Plasterk R H. Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci USA*, 2001, **98** (12): 6759~ 6764
- 6 Dupuy A J, Fritz S, Largaespada D A. Transposition and gene disruption in the male germline of the mouse. *Genesis*, 2001, **30** (2): 82~ 88
- 7 Horie K, Kuroiwa A, Ikawa M, *et al.* Efficient chromosomal transposition of a Tc1/*mariner*-like transposon *Sleeping Beauty* in mice. *Proc Natl Acad Sci USA*, 2001, **98** (16): 9191~ 9196
- 8 Luo G, Ivics Z, Izsvak Z, *et al.* Chromosomal transposition of a Tc1/*mariner*-like element in mouse embryonic stem cells. *Proc Natl Acad Sci USA*, 1998, **95** (18): 10769~ 10773
- 9 Devon R S, Porteous D J, Brookes A J. Splinkerettes improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res*, 1995, **23** (9): 1644~ 1645
- 10 Karsi A, Moav B, Hackett P, *et al.* Effects of insert size on transposition efficiency of the *Sleeping Beauty* transposon in mouse cells. *Marine Biotechnology*, 2001, **3** (3): 241~ 245
- 11 Richardson P D, Augustin L B, Kren B T, *et al.* Gene repair and transposon-mediated gene therapy. *Stem Cell*, 2002, **20** (2): 105~ 118
- 12 Cui Z, Geurts A M, Liu G, *et al.* Structure-function analysis of the inverted terminal repeats of the *Sleeping Beauty* transposon. *J Mol Biol*, 2002, **318** (5): 1221~ 1235
- 13 Bell A C, West A G, Felsenfeld G. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science*, 2001, **291** (5503): 447~ 450
- 14 Li Q, Harju S, Peterson K R. Locus control regions: coming of age at a decade plus. *Trends in Genetics*, 1999, **15** (10): 403~ 408

Sleeping Beauty Transposition System

DING Sheng, WU Xiao-Hui*

(Institute of Development Biology and Molecular Medicine, Fudan University, Shanghai 200433, China)

Abstract *Sleeping Beauty* (SB), a member of the Tc1/*mariner* superfamily of transposable elements, is the only active DNA transposon system from vertebrate. The research progresses of SB in recent years are reviewed and some modifications are proposed to improve the performance of SB as a genetic manipulating system.

Key words *Sleeping Beauty*, transposition, transgene

* Corresponding author. Tel: 86-21-65643718, E-mail: more@fudan.edu.cn

Received: July 24, 2002 Accepted: September 12, 2002