

真核生物的转座子

- 种类繁多，包括DNA转座子及逆转录转座子两大类。
- 大多数真核生物转座子已经丧失了自发转座能力。在哺乳动物中，目前仅发现少量DNA转座子可以转座，在人类中，只有一类逆转录转座子可以转座。

(一) 真核生物的DNA转座子

- 种类繁多，根据转座子的结构特征和序列相似度分为hAT, *Mariner/Tc1*, *P*, *MuDR*, *EnSpm/CACTA*, *PiggyBac*, *Merlin*, *Harbinger*, *Transib*, *Kolobok*, *ISL2EU*, *Sola*, *Zator*, *Ginger1*, *Ginger2/TDD*, *IS3EU*, *Dada*, *Helitron*, *Polinton*, *Crypton*, *Academ*, *Novosib*和*Zisupton*共23个超家族。
- DNA转座子约占人类基因组的3%，但人类的DNA转座子均已经丧失了转座能力。在哺乳动物中，目前仅发现褐蝙蝠*PiggyBac*家族的转座子能够自发转座。

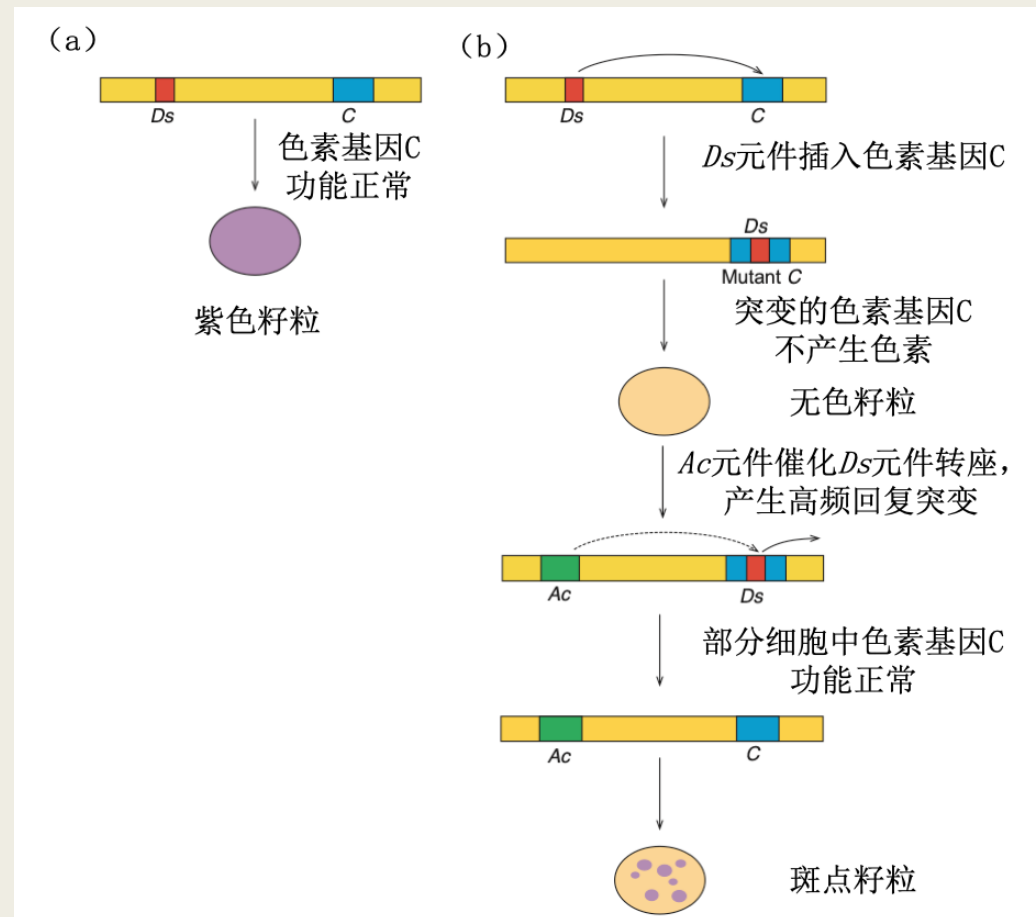
DNA转座子——玉米的Ac/Ds元件



芭芭拉·麦克林托克



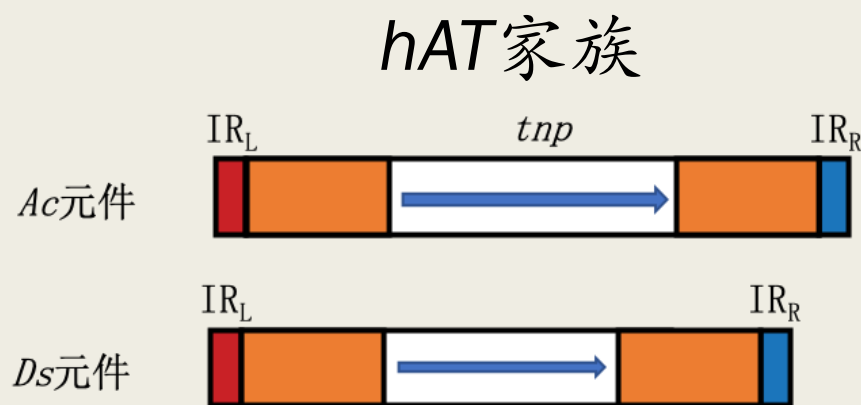
印第安彩虹玉米



Ac/Ds元件

Ac, activation; Ds, dissociation

DNA转座子——玉米的Ac/Ds元件



■ Ac (Activator) 元件

全长4563 bp, 两端各有一个长度为11 bp的IR序列, 中间包含转座酶编码序列(*tnp*), 在IR序列和*tnp*之间为约200 bp的亚末端重复序列。

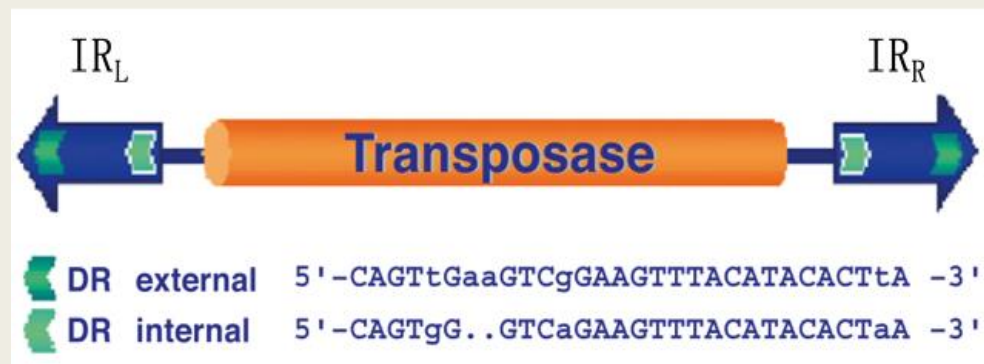
■ Ds (Dissociation) 元件

Ds元件有多种, 均为Ac元件的截短突变体, 由于缺失了部分或全部*tnp*编码序列, 只有在Ac元件的辅助下才能转座。

■ 转座机制

剪切-粘贴机制, 产生8 bp的TSD。

DNA转座子——*Mariner/Tc1*家族



睡美人转座子
(Sleeping beauty)

■ 睡美人转座子最为常用的转座子之一是根据来自于大马哈鱼的12个已经失活的Tc1样元件重构而来的能自发转座的转座子。

■ 睡美人转座子的结构

两端的IR序列和中间的转座酶编码序列，如SB10和SB100X。

■ 睡美人转座子的转座机制

剪切-粘贴机制，靶位点产生由5'-TA-3'构成的2 bp的TSD。

(二) 真核生物的逆转录转座子

■ 逆转录转座子仅分布于真核生物中，在玉米中逆转录转座子约占基因组的49%-78%，在人类中逆转录转座子约占基因组的42%。

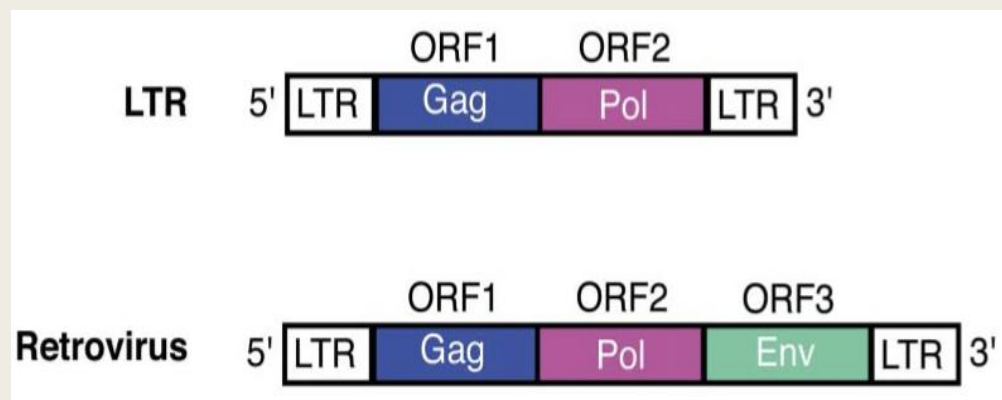
■ 分类

- ┌ LTR (Long terminal repeat) 逆转录转座子
- └ 非LTR逆转录转座子

LTR逆转录转座子

■ 又名病毒样逆转录转座子

结构域逆转录病毒类似，但不能形成病毒颗粒，可能起源于逆转录病毒。



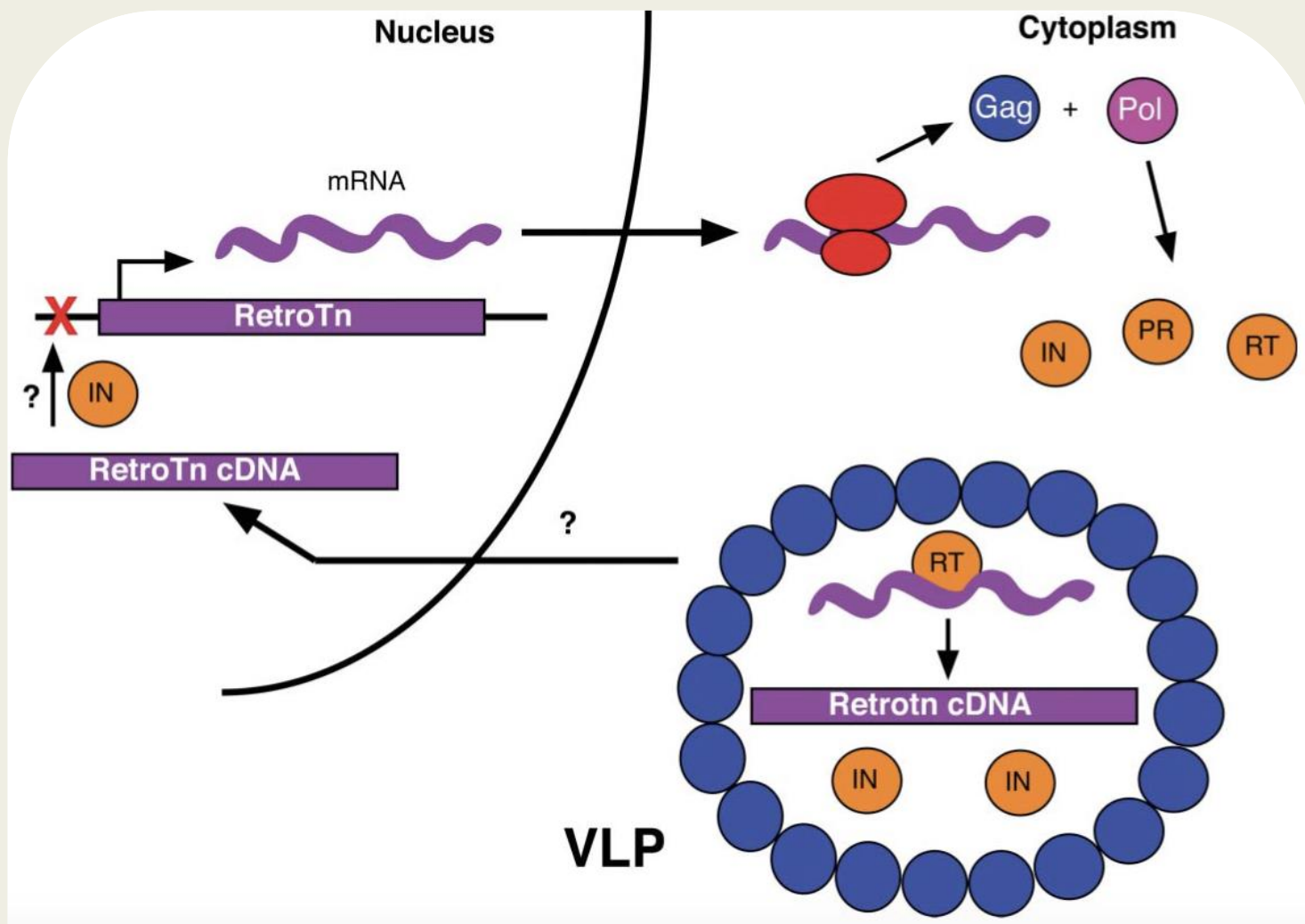
■ LTR逆转录转座子的类型

包括Copia, Gypsy, BEL, DIRS和ERV (endogenous retrovirus) 等5个亚型。ERV又可以分为ERV1, ERV2, ERV3, ERV4和ELV等5类。

■ 人类LTR逆转录转座子

约占人类基因组的8%，其中人类内源逆转录病毒 (HERV) 约占人类基因组的1-2%。

LTR逆转录转座子的转座机制



■ 转座机制与逆转录病毒感染类似。

■ 会在细胞内产生病毒样颗粒 (Virus-like particle, VLP)

非LTR逆转录转座子

- 广泛分布于真核生物中，在多个物种中，非LTR逆转录转座子是基因组中最丰富的序列，约占哺乳动物基因组的30%。
- 很多非LTR逆转录转座子依然能够转座，是人类中唯一能够转座的转座子。

■ 分类

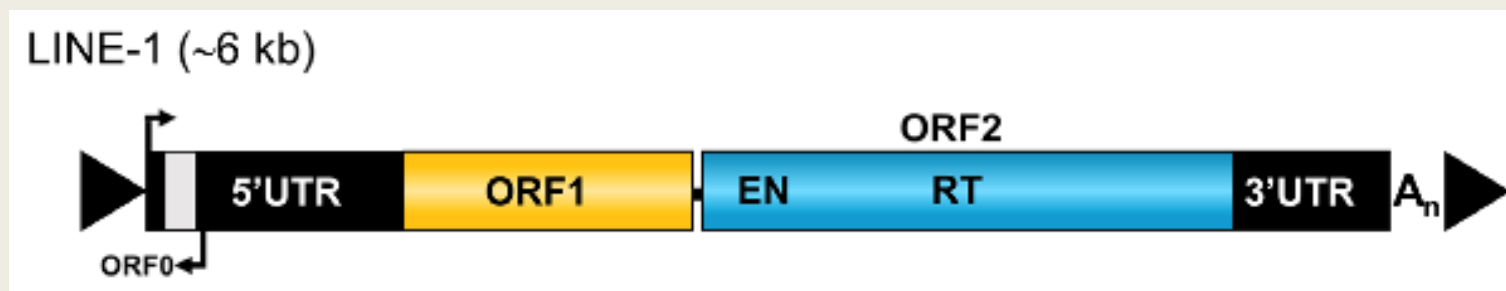
- 自主型：又名长散在核元件（long interspersed elements, LINE)
- 非自主型
 - 短散在核元件(short interspersed element, SINE)
 - 复合非LTR逆转录转座子（composite non-LTR retrotransposon)

非LTR逆转录转座子——LINE

- 包括LINE1 (L1) , LINE2 (L2) , LINE2A (L2A) , LINE2B (L2B) , CRE, NeSL, R1, R2, R4, Hero, Randl/Dualen, Proto1, Proto2, RTE, RTEX, RTETP, I, Nimb, Ingi, Tx1, Vingi, Tad1, Loa, Outcast, Jockey, CR1, Kiri, Rex1, Crack, Daphne, Ambal, 和Penelope等32个亚型。

■ 人类的L1

长度约为6 kb , 大约有50万个拷贝, 约占人类基因组的17%。



非LTR逆转录转座子——SINE

■ 包括SINE1/7SL, SINE2/tRNA, SINE3/5S, SINEU和SINE4等5类。

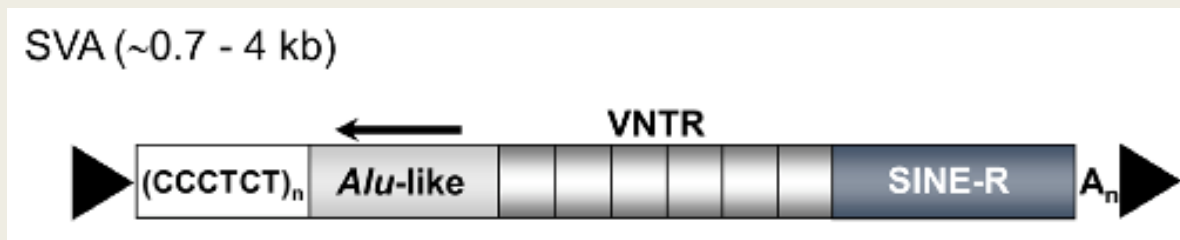
■ 人类的SINE

因其含有限制性内切酶AluI的识别序列而得名，其长度约为280-300 bp，在人基因组中约含有100万个拷贝。

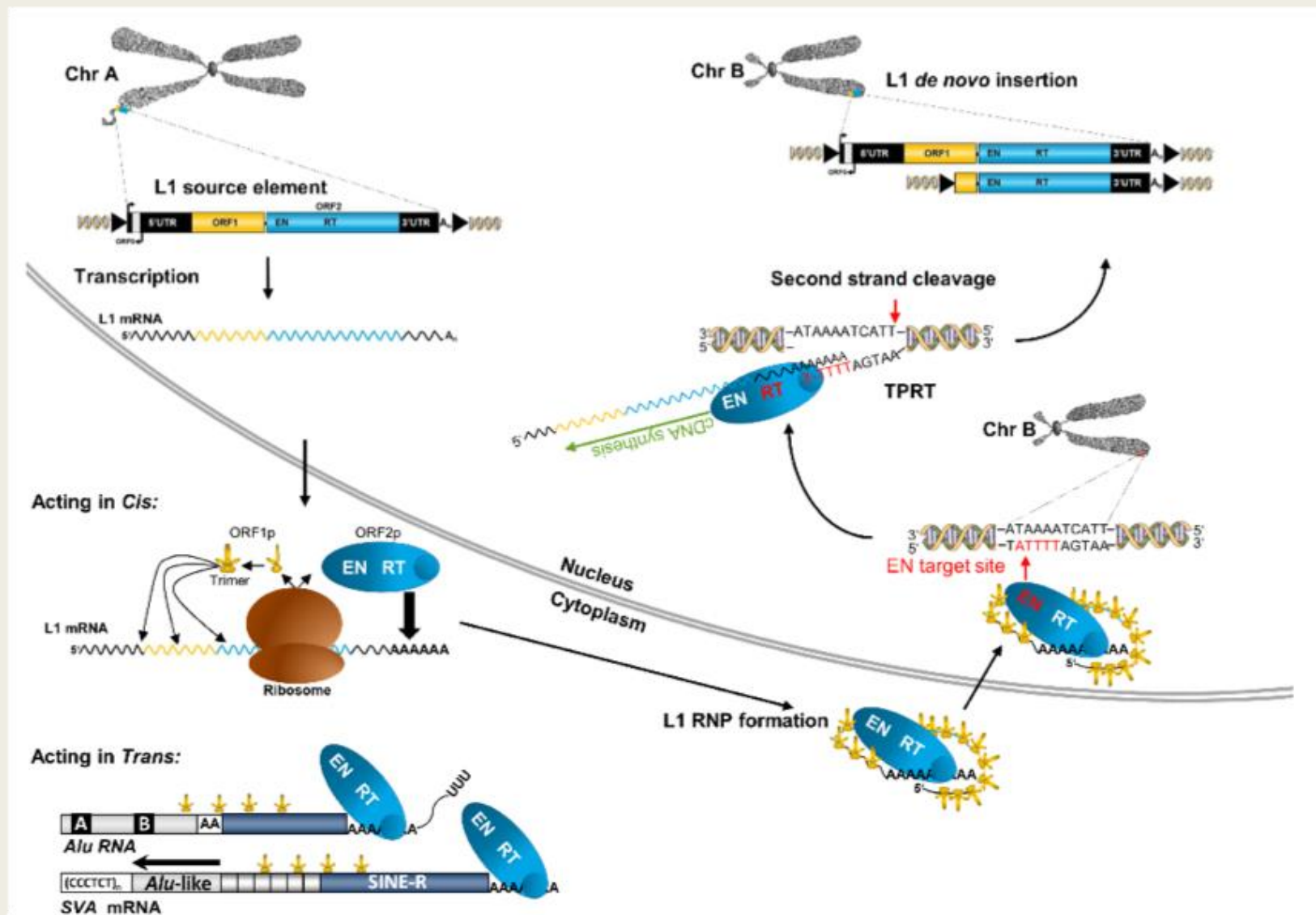


非LTR逆转录转座子——复合非LTR逆转录转座子

- 包括SVA (*SINE-R/VNTR/Alu*), LAVA (*L1/Alu/VNTR/Alu*), PVA(*PTGR2/VNTR/Alu*), FVA(*FRAM/VNTR/Alu*)等4类。
- 人类及人类的祖先特有的转座元件——SVA (*SINE-R/VNTR/Alu*)
长度为0.7-4 kb。



非LTR逆转录转座子的转座机制



■ LINE元件可以自发转座。

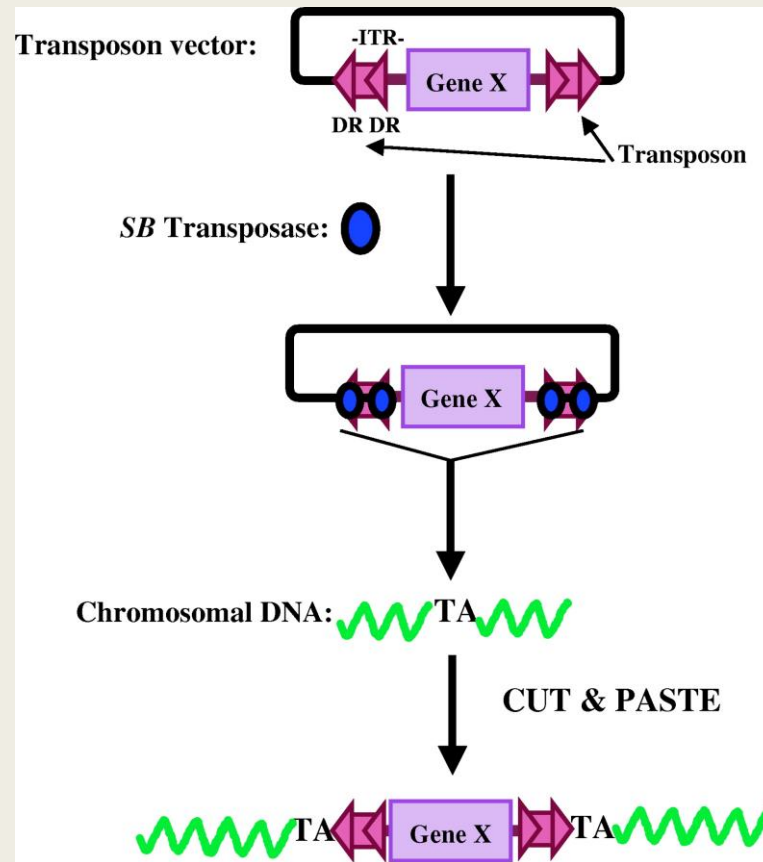
■ SINE和复合非LTR逆转录转座子借助LINE元件编码的蛋白进行转座。

(三) 转座子的应用

基因测序文库构建



基因递送



基因诱捕 (Gene trap)

