

遗传学（第3版）

第13章 转座因子的结构与功能

1. 转座因子的发现与分类
2. 原核生物中的转座子
3. 真核生物中的转座子
4. 转座作用的分子机制
5. 转座因子的遗传学效应及其应用





13.1 转座因子的发现与分类

13.1.1 转座因子的发现

1914年 Emerson 研究玉米果皮色素遗传,发现一种花斑果皮的突变类型→多次回复突变→产生宽窄不同、红白相间的花斑。

这种花斑的产生在于突变基因的不稳定性,但**如何不稳定?**他不清楚

1938年 Rhoades 研究玉米籽粒糊粉层色素遗传,发现修饰的孟德尔分离比:有色:斑点:白色=12:3:1,而这两个基因是不连锁的。他认为

基因 A_1 (控制色素) $\Rightarrow a_1$ 表现为无色;

另一基因 D_+ (斑点) 表型为有色斑点。

这样原品系的基因型为 $A_1 A_1 d_+ d_+$, 突变后产生了 $A_1 a_1 D_+ d_+$ 的植株, 这种双突变植株自交就产生了上述比例 (图13-1)



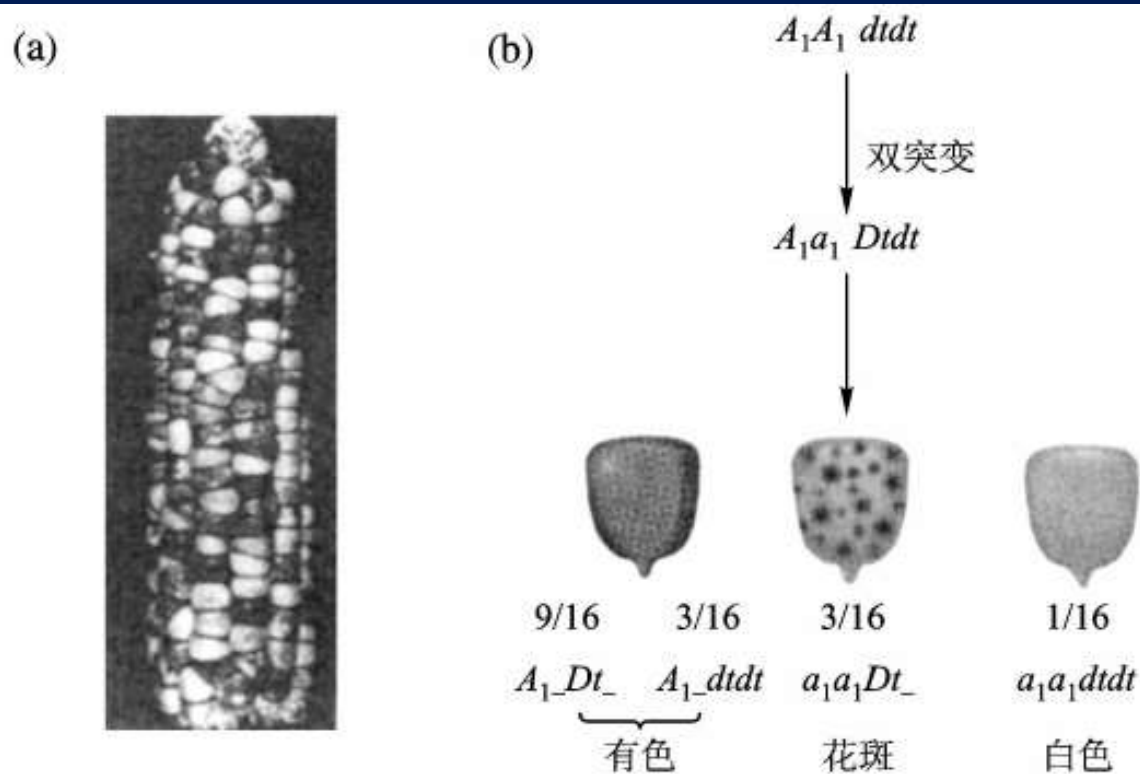


图 13-1 玉米花斑表型的遗传学解释

(a) 由于转座因子而引起的玉米籽粒花斑的表型 (b) 最初认为玉米籽粒花斑形成是双突变的结果



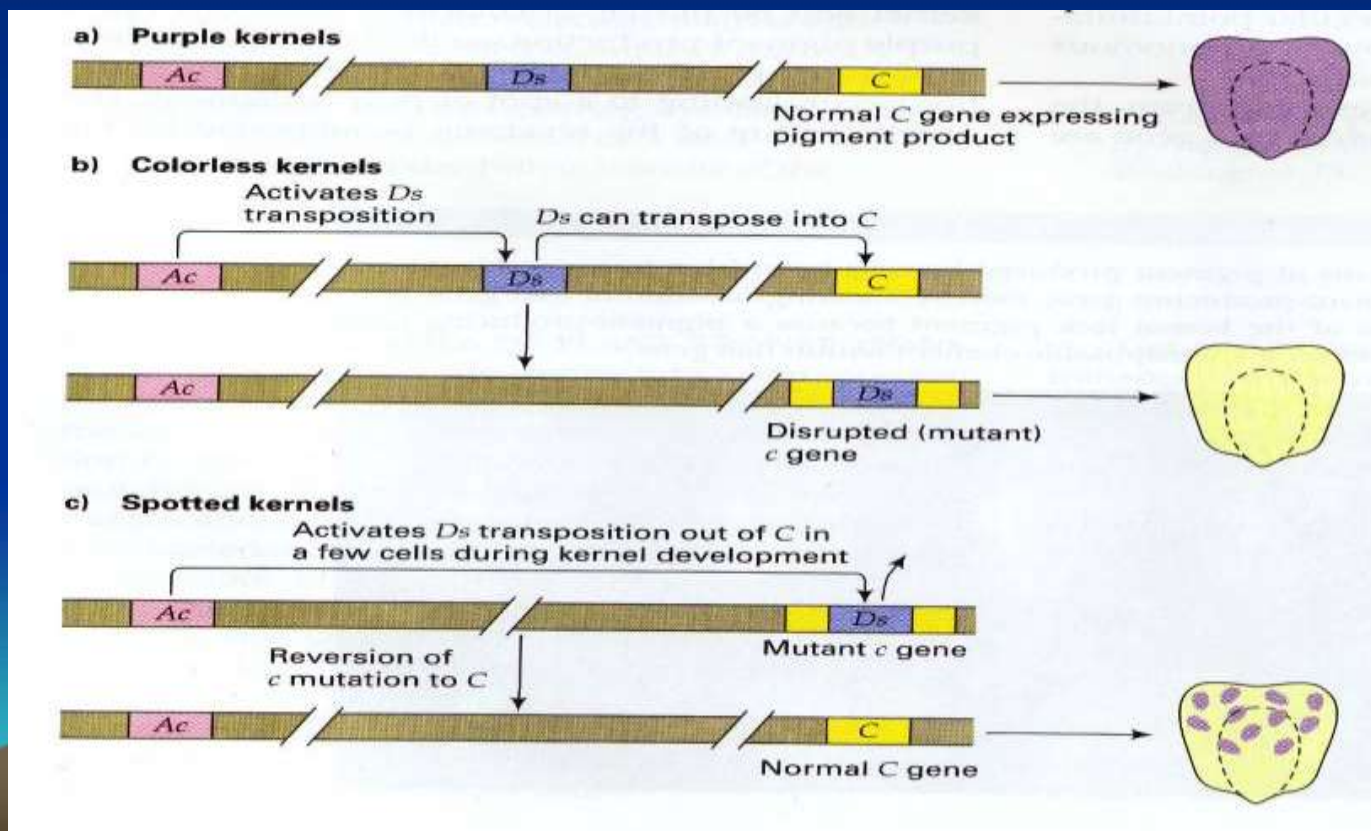
但是**什么因素导致或产生花斑呢**？一种可能是在体细胞中产生了回复突变 $a_1 \rightarrow A_1$ ，但大量的斑点需要很高频率的回复突变。

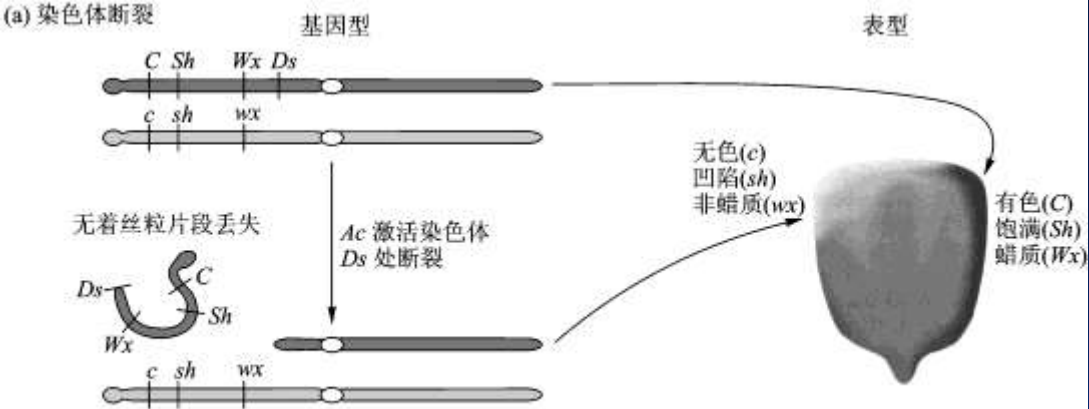
Rhoades用 $a_1 a_1 D_{t-}$ （花斑）特殊无性生殖植物与 $a_1 a_1$ 的植株测交，结果有的后代完全是有颜色的。表明在亲本中每个斑点实际上是回复突变的表型效应。

a_1 成为首次发现的不稳定突变等位基因的例子，而这种等位基因的不稳定性取决于不连锁 D_t 基因的存在。一旦回复突变发生，它们就变得稳定了；即 D_t 基因能离开 A_1 基因，这时 A_1 的表型不再改变。因而 D_t 的缺乏使得表型保持稳定。**Rhoades发现了某些基因的不稳定性，而且这种不稳定性是由另一个独立的因子所控制。**但仍未揭示这种不稳定性的遗传学机制，也缺乏实验证据

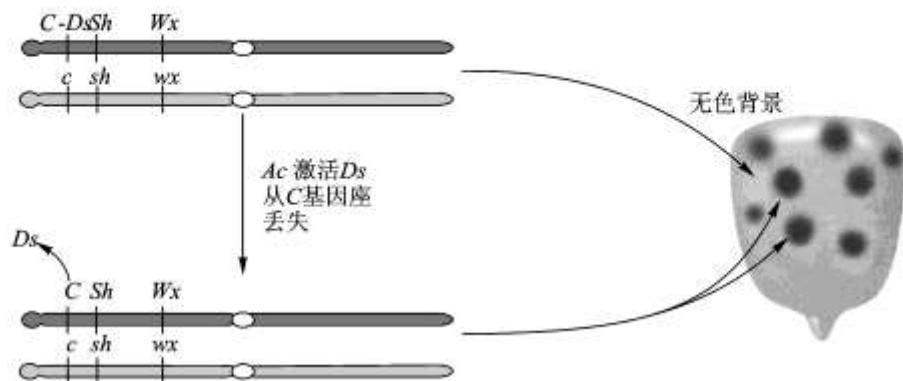


1940年至1950年 McClintock 研究了玉米胚乳的紫色、白色以及白色背景上带有紫色斑点这些表型之间的相互关系。她发现花斑表型是不稳定的，并根据自己的遗传学和细胞学研究结果推断“花斑”这种表型并不是一般的基因突变产生的，而是由于一种控制因子的存在所导致的。





(b) 新的不稳定等位基因



染色体的断裂结果还会形成断裂融合桥 (breakage-fusion-bridge)

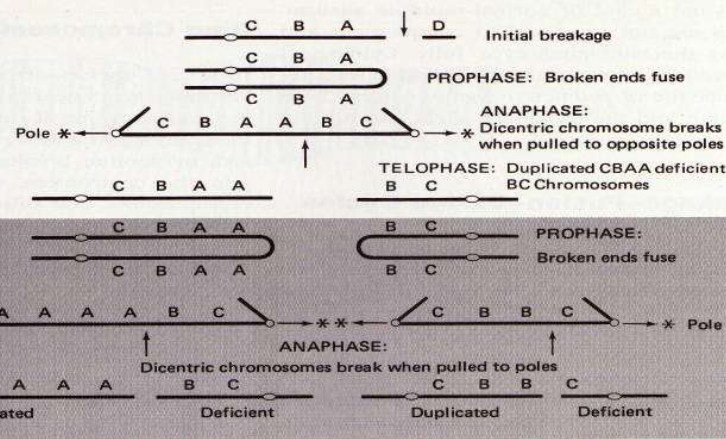
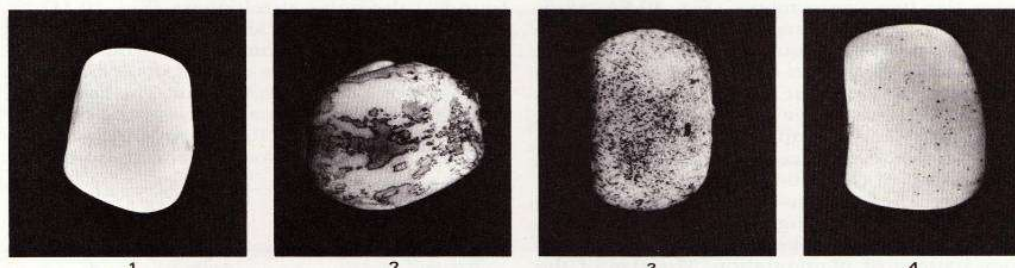


图13-2 Ds转座导致籽粒斑点表型

(a) Ds 位于杂合的一条同源染色体的着丝粒与一系列显性标记之间，另一条同源染色体缺乏 Ds 而有几个隐性基因，在 Ds 处断裂产生携带显性标记的无着丝粒片段而丢失，故同源染色体上隐性基因表达，籽粒产生无色扇形；

(b) Ds 插入 C 基因后籽粒为无色；当 Ac 激活 Ds 从 C 基因切离后，则籽粒出现有色斑点，而且斑点的大小取决于产生它的细胞的分裂次数

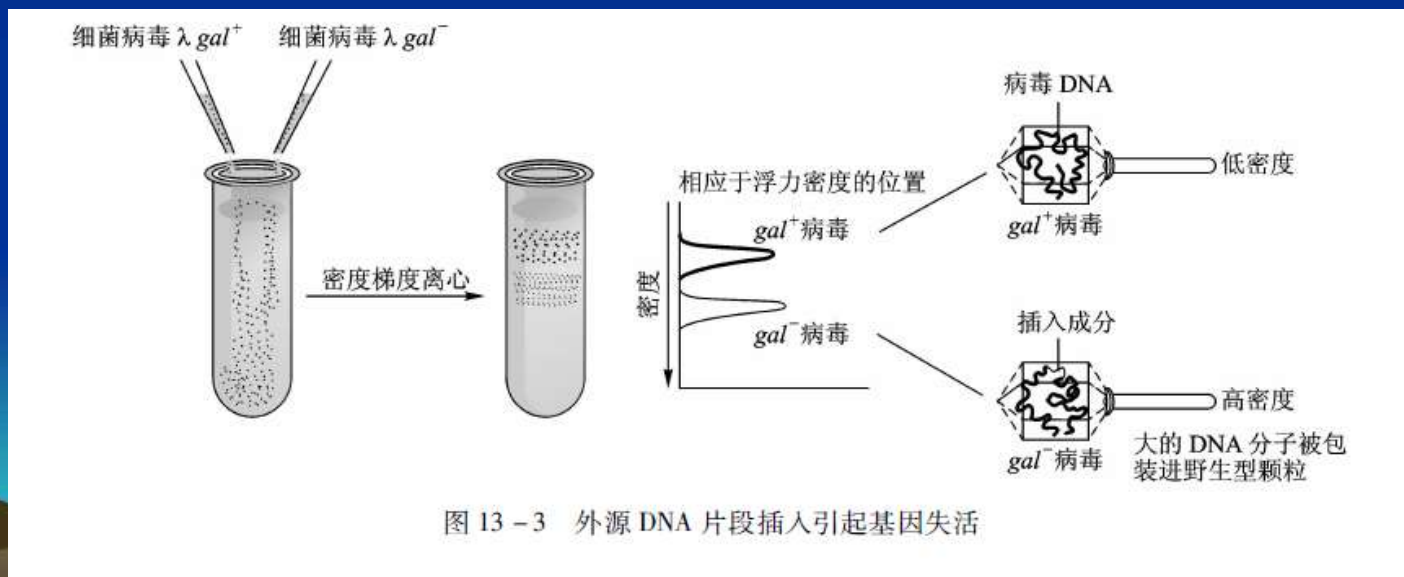


根据大量遗传学和细胞学研究结果, McClintock于1951年提出了“**转座因子**”学说。

这些转座因子既可以沿染色体移动, 也可以在不同染色体之间跳跃——又称为**跳跃基因 (jumping gene)**。

这是遗传学发展史中划时代的重大发现, 将基因概念向前推进了一大步。但这项划时代的成果并未受到当时同行们重视。

直到20世纪60年代Jacob和Monod的乳糖操纵子模型和基因调控理论发表后, 特别是Shapiro在细菌中也发现了可转座的遗传因子后 (图13-3), 这一成果才被接受。



13.1.2 DNA转座

可移动位置的遗传因子包括两种类型：

一种类型是**直接以DNA序列某些区段作为转座成分的**，如在玉米中发现的转座因子Ac/Ds和在大肠杆菌中发现的IS都是涉及DNA直接转座，因而称为DNA转座；

另一种类型是**以RNA介导进行转座**，称为反（逆）转座子（retroposon）或反转录转座子（retrotransposon）。

首先讨论DNA转座，根据转座机制，DNA转座可分为两种类型：
复制型转座（Replicative transposition）

非复制型转座（Nonreplicative transposition）



(1) 复制转座

复制转座:转座子在转座过程中被完整地复制。在新的位点上的转座子是供体转座子的一个完整的拷贝。

即：一个拷贝仍在原位，而另一个同样的拷贝插入新的位点。

在该转座过程中，伴随着转座子的拷贝数的增加。

复制转座由两种酶催化：**转座酶** 作用于原转座子的末端

解离酶 对已复制的拷贝起作用

TnA等一组相关的转座子的移动仅由复制转座机理而进行。

图-复制转座模型

RR





(2) 非复制型转座

非复制转座：转座因子作为一个物质实体 (Physical entity) 直接从一个位点移向另一个位点 (site)，在供体位点留下一个双链打断的断口。若不被修复则将对供体基因组是致死的，插入序列和复合转座子 Tn10 及 Tn5 使用以上机制转座，仅要求转座酶。

保守转座 (Conservative transposition)：是另一类非复制转座事件，在该过程中转座因子从供体位点切除然后插入靶位点，在一系列的过程中每一核苷酸键都被保留。

某些转座子只采用一种转座机制，而另一些转座子可能具备两种途径。如IS1和IS903就采用复制和非复制两种途径。Mu噬菌体能从共同的中间体转变为两种途径中的一种。还有些转座子存在着不同的转座机制交替使用，据此特征难以划分它们的类型。

DNA区段无论以什么机制进行转座，其转座过程有别于同源重组过程。这些DNA序列的转座往往发生在非同源序列之间，也不需要像细菌同源重组中RecA等蛋白质的参与，只依赖于转座区域DNA复制、转座酶和特定的 *IR* 等因子而完成重组过程。



13.1.3 反转录转座子

由RNA介导的转座仅发生在真核生物中，它是由反转录病毒（retroviruses）以其RNA基因组的DNA拷贝插入到宿主细胞的染色体中而产生的。

一些真核细胞的转座子和反转录病毒的前病毒的一般组成相关，并且通过RNA中间体进行转座。这一类转座子称为反转座子（retroposons）和反转录转座子（retrotransposons）。

包括：反转录病毒本身（自由感染宿主细胞）

自己不能转座但可以通过RNA转录进行转座的序列（DNA序列）

RNA→cDNA→dsDNA→整合进染色体DNA上（原病毒/前病毒provirus）

DNA→RNA→cDNA→dsDNA→整合进染色体

(1) 反转录病毒的生活史（[VIIp486, F16.1](#)）

(2) 反转录病毒基因组的结构与功能及转录（[VIIp488, F16.3](#)）

典型的反转录病毒含3—4个“基因”（编码区）

gag: （[Group-specific antigens](#)）nucleocapsid gene 核衣壳蛋白基因
或 internal structural protein 病毒核心蛋白基因

pol: RNA—dependent DNA polymerase 反转录酶基因

env: envelop glycoproteins 外膜糖蛋白基因



此外某些还会有癌基因

非编码区 ([图16.5](#))

R区: 为基因组两端的重复序列, 每个重复单位10—80nt, cDNA(—)的反转录所必需的

PB[±]: 负链cDNA合成的起始位点, 称为引物结合区 (primer binding region) 可与各种tRNA的3'端碱基互补 (16—19nt互补)

U区: U5 80—100nt (单一序列)

U3 含170—1260 (或1350) nt, 含有强的启动子, 从这里开始合成前病毒DNA

(1) 病毒基因组 (RNA)、病毒线形DNA、前病毒DNA

比较:

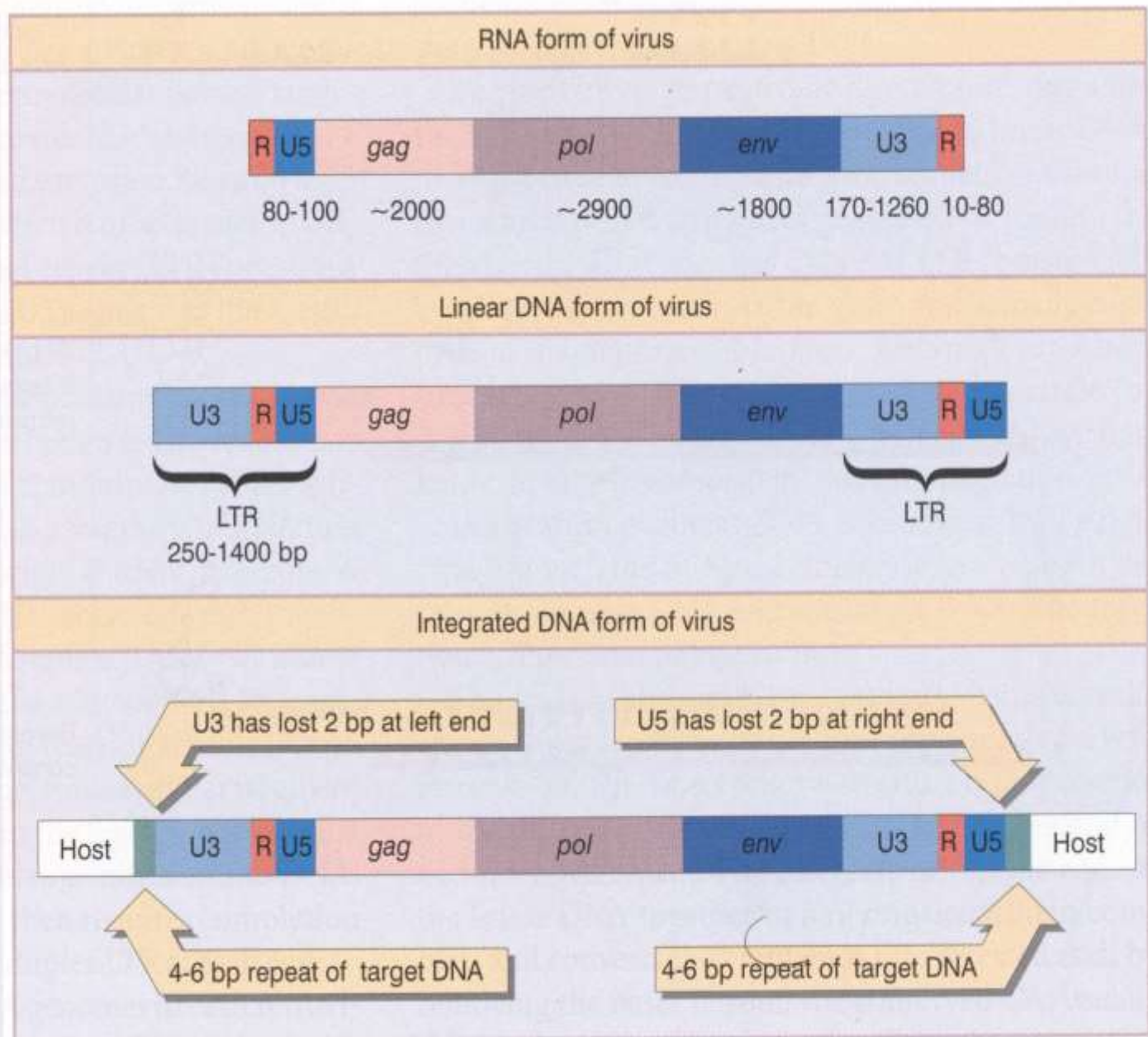
RNA反转录生成前病毒双链DNA并整合到寄主细胞的基因组上, 双链前病毒的长度要比反转录前的单链RNA分子长, 这是病毒DNA分子两端发生重排形成:

LTR(Long terminal repeat)结构: U3-R-U5

这是反转录病毒DNA形成的特有结构

LTR 与病毒DNA整合进寄主染色体有关

Figure 16.5 Retroviral RNA ends in direct repeats (R), the free linear DNA ends in LTRs, and the provirus ends in LTRs that are shortened by two bases each.



(2) 反转录病毒基因组的复制及LTR的形成

(F5. 47)

合成起始是由tRNA作为引物结合到引物结合位点 (PBS) 上→反转录酶使该引物向病毒RNA模板5'端延伸, 合成DNA: 合成U5和R →RNase H降解RNA-DNA异源双链中的RNA链→反转录酶经DNA与RNA的杂交序列跳跃到病毒RNA链的3'端→反转录酶延伸引物, RNA-DNA异源双链被RNase H降解→RNA部分引发第二链DNA合成→DNA链延伸, RNA降解→RNase H 降解 tRNA→通过PBS 序列杂交产生第二次跳跃→完成DNA的合成



武汉大学

Wuhan University

反转录病毒基因组的复制

(RNA→cDNA→dsDNA)及LTR的形成



(3) 反转录转座子的类别

在哺乳基因组中有三种类型的反转录转座子：

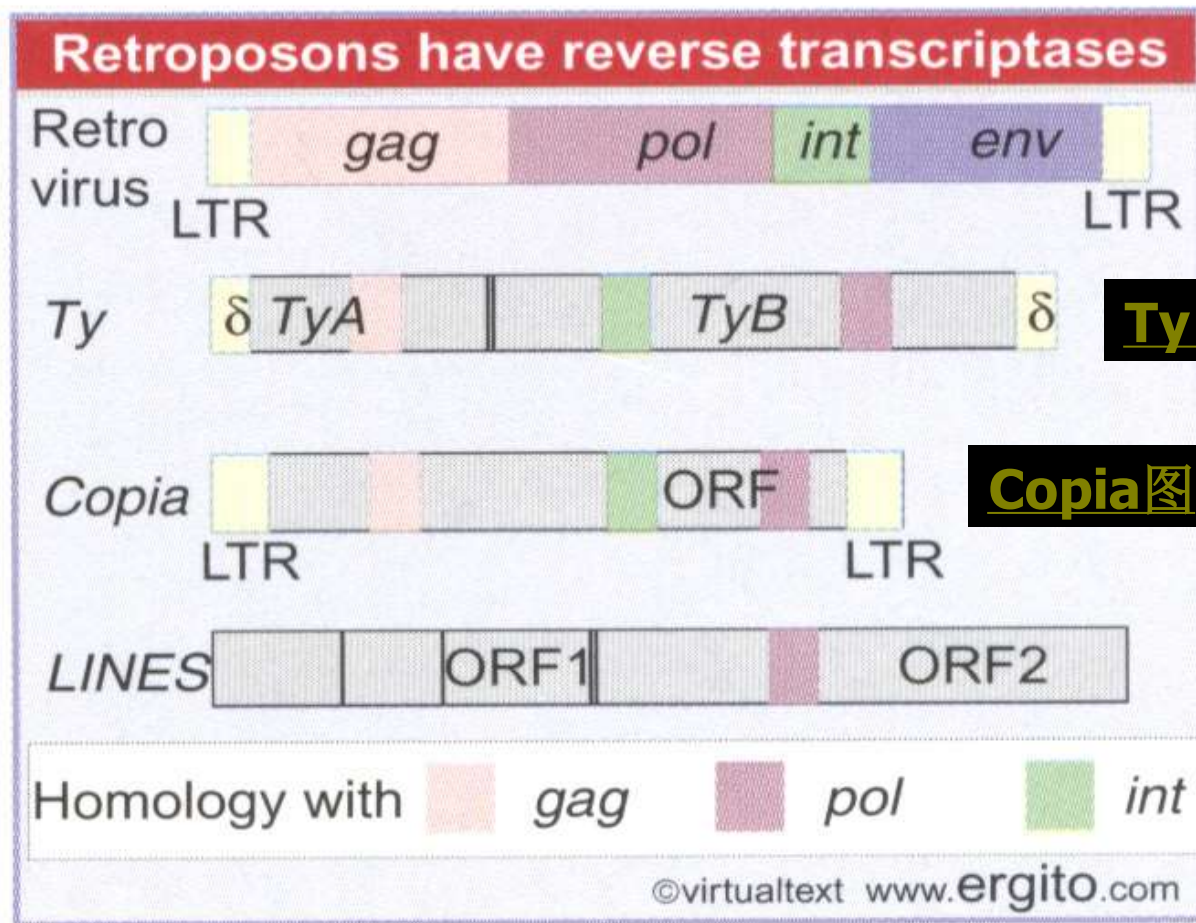
病毒超家族、LINES和非病毒超家族

Mammalian genomes have three types of retroposons

	Viral Superfamily	LINES	Nonviral Superfamily
Common types	<i>Ty</i> (<i>S. cerevisiae</i>) <i>copia</i> (<i>D. melanogaster</i>)	L1 (human) B1, B2 ID, B4 (mouse)	SINES (mammals) Pseudogenes of pol III transcripts
Termini	Long terminal repeats	No repeats	No repeats
Target repeats	4-6 bp	7-21 bp	7-21 bp
Enzyme activities	Reverse transcriptase and/or integrase	Reverse transcriptase /endonuclease	None (or none coding for transposon products)
Organization	May contain introns (removed in subgenomic mRNA)	1 or 2 uninterrupted ORFs	No introns

©virtualtext www.ergito.com

Figure 17.16 Retroposons can be divided into the viral superfamilies that are retrovirus-like or LINES and the nonviral superfamilies that do not have coding functions.



Ty图

Copia图

Figure 17.17 Retroposons that are closely related to retroviruses have a similar organization, but LINEs share only the reverse transcriptase activity.

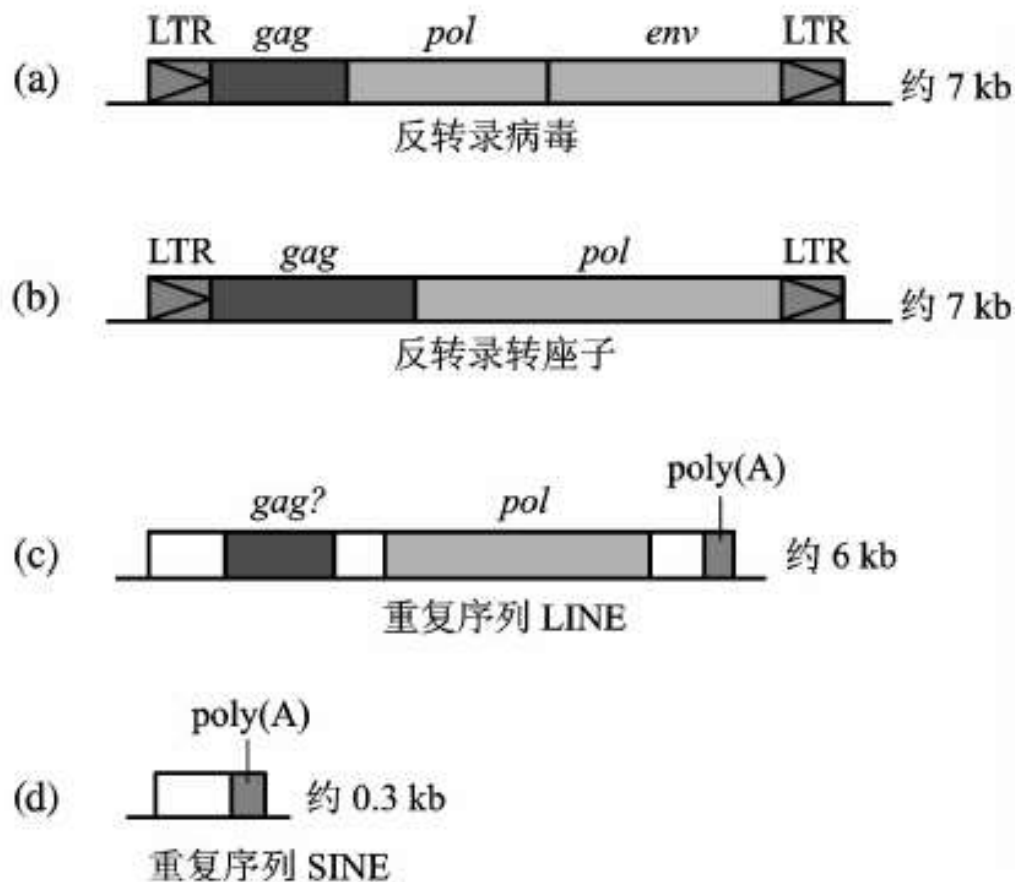


图 13 - 6 4 种反转录转座因子结构比较

反转录病毒(a)和反转录转座子(b)均含有长末端重复序列,归于 LTR 一类,只是后者缺编码外壳蛋白的基因(*env*)。LINE(c)和 SINE(d)为非 LTR 反转录转座子,两者在 3'端均有 poly(A)区



(4)、转座子的类别

原核

IS——两端有IR, 只编码转座酶

类转座因子——结构同IS, 但不能独立存在, 仅作为复合子的两端组件

复合转座子——两端由IS或类IS构成, 可编码抗抗菌素物质

TnA转座子家族——两端为IR, 可编码转座酶、解离酶和抗性物质

真核

转座子 { Ac-Ds双因子系统 (玉米中) 激活——解离因子
P因子——果蝇中父本因子, 在M♀ × P♂ 中导致杂种不育

**反转录
转座子**

反转录病毒: RNA → DNA → 整合宿主靶DNA

病毒超家族 { Ty
Copia
L1NS L1 } { (1) 有长末端重复序列
(2) 编码反转录酶或整合酶
(3) 可含内含子

非病毒超家族 { SINS B1
Alu
假基因 } { (1) 无重复序列
(2) 不编码转座子产物
(3) 无内含子



13.2.1 插入序列 IS (Insertion sequence)

插入序列是最简单的转座因子

IS是细菌染色体和质粒的正常组成部分，可进行同源序列间的重组。*E.coli* K12 chromosome 有8个IS1、5个IS2和IS3。

IS是独立的结构单位: 转座组件 (transposition modules)

自主单位 (autonomous units)

- 仅有其自身转座编码的基因，不含有其它编码蛋白的结构基因
- 结构都较短: ~1000bp (768bp~1500bp)
- 两端含有20-40bp IR (Inverted terminal repeat)末端反向重复序列
- 识别的靶序列5-9bp，转座后，在寄主DNA上产生同向重复。

IS因子插入细菌染色体上可以产生各种效应，这取决于其插入的方向，IS因子虽然不含有其它编码蛋白质的结构基因，但是它们可以含有转录终止子 (transcriptional terminator)或启动子 (promoter)

从分子遗传学角度解释极性效应:

- (1) IS中可能有终止子信号，它的插入会造成mRNA的转录的终止
- (2) 含有无义密码子，造成翻译的终止由此影响到后续基因的翻译

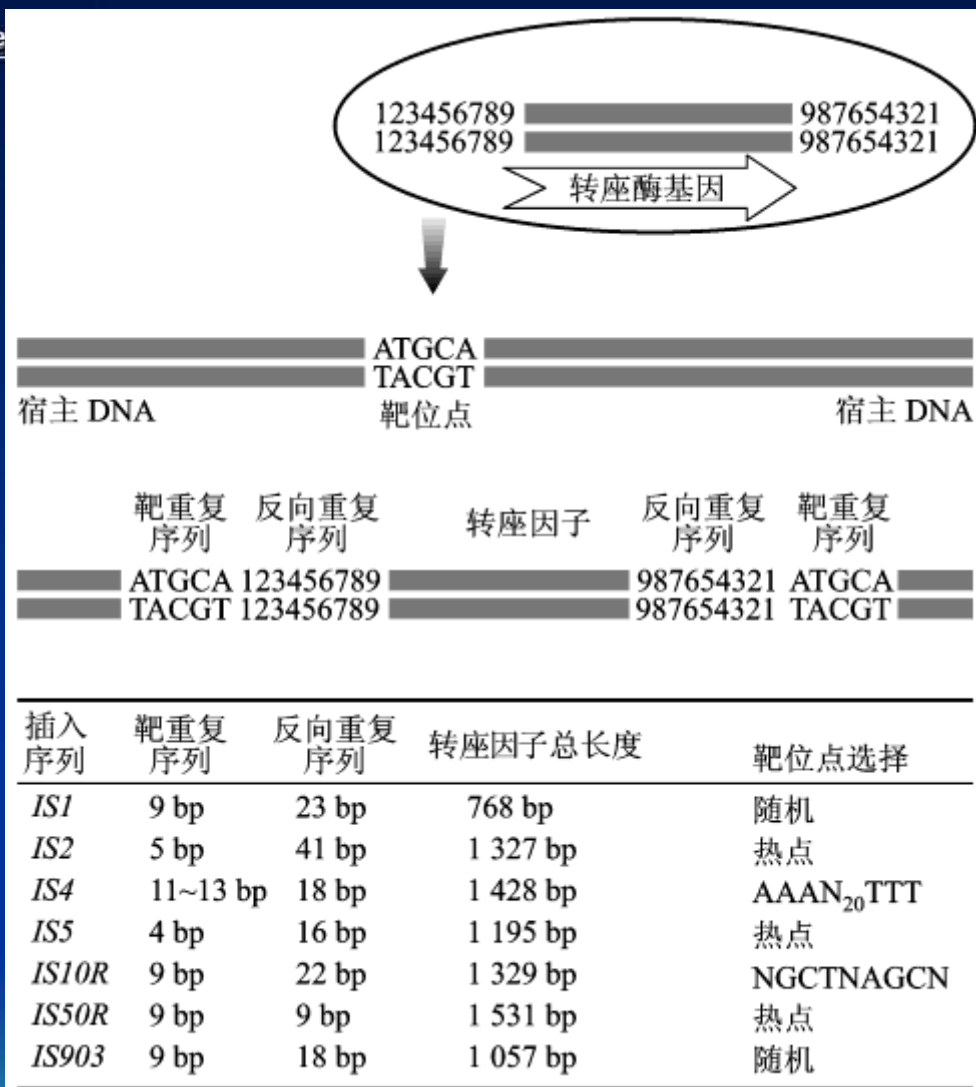


图13—7 IS结构模式图
IS末端的反向重复序列为9 bp，数字1~ 9示碱基序列

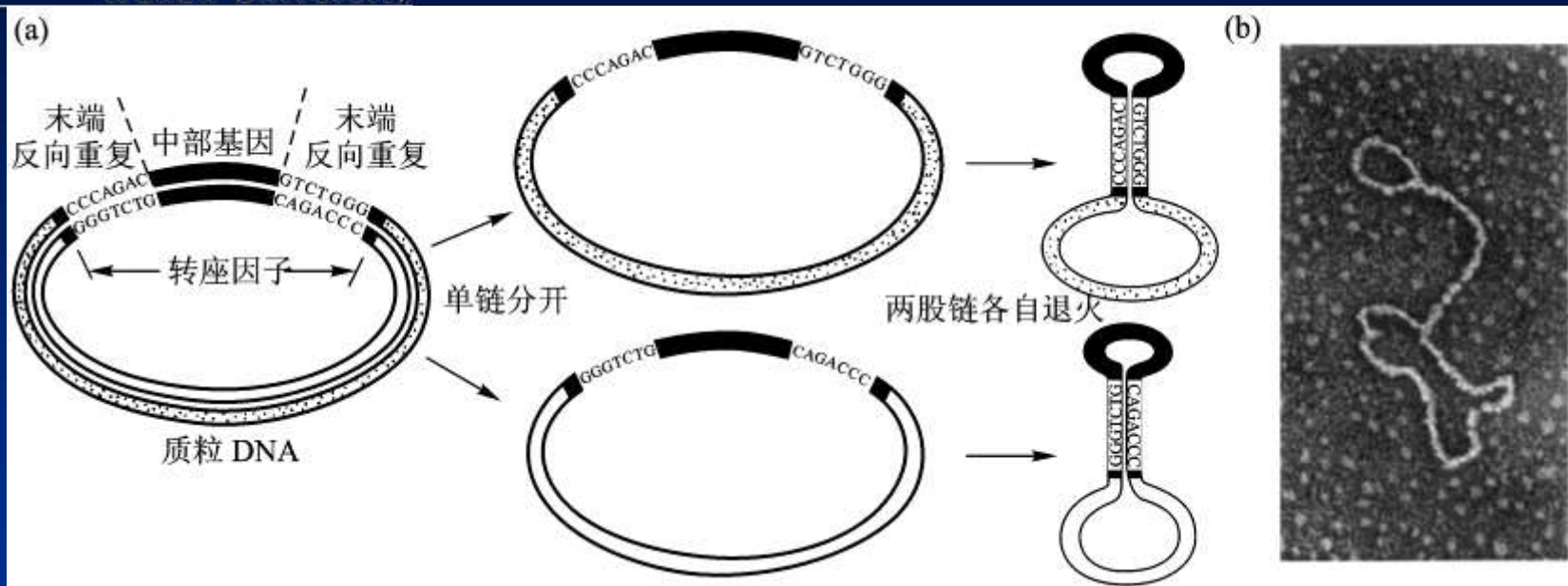


图13—8 颈环结构的形成

含IS质粒经变性复性形成颈环结构（a）及其电镜照片（b），大环是质粒DNA，小环是IS的中间序列，颈的部分是IS的IR

类插入序列（IS-like element）

是指 IS10R、IS50R和IS903,它们的结构和IS相似，但不独立存在。而是作为复合转座子两臂的组件。

13.2.2 转座子

(1) 复合转座子 (composite transposons)

结构：比IS长得多，自身转座基因+其它基因（抗药性基因）

不同的复合转座子的抗性标记不同

复合转座子两端的组件由IS和类IS组成

“两端”我们称为“Arms”，“Arms”可以是同方向的、也可以是反方向的。

一般，当一个复合型转座子两侧组件不同时，该转座子的转座作用主要依靠其中一个组件的功能（Tn10中，IS10R）

维持复合型转座子的一个主要动力是对中心区所携带的标志的选择。一个IS10结构单位自身可以随意地移动，并且比Tn10的移动频率高一个数量级。

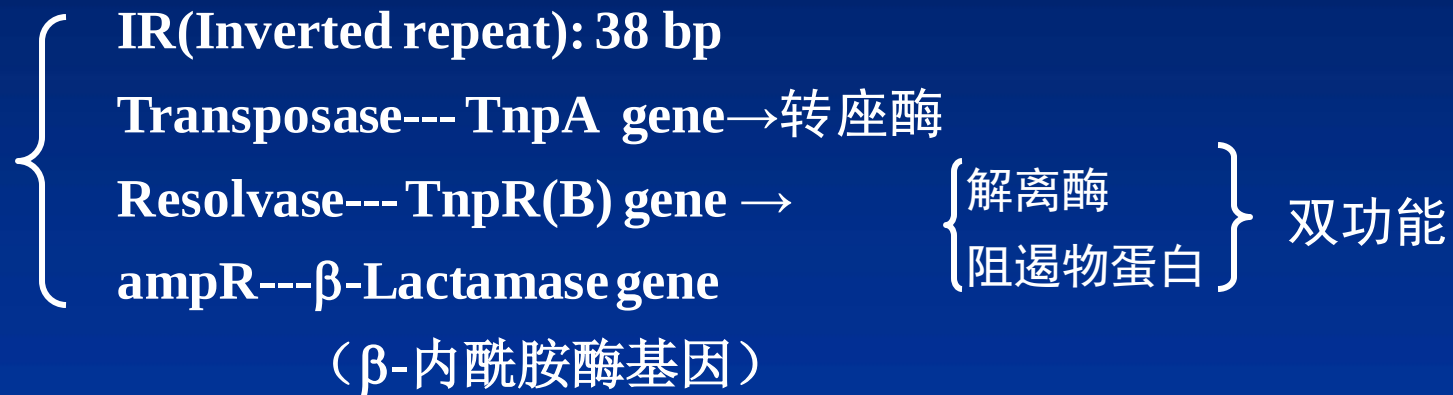
但对Tetr的选择可使Tn10维持在一起，因此在选择的条件下，完整的Tn10的转座频率可明显增加。

IS元件所编码的转座酶的活性负责识别一个靶部位和识别转座子的末端。**只有转座子的末端可以作为转座的底物。**

(2) TnA 转座子家族

TnA家族包括几个相关的转座子，其中Tn3和Tn1000（以前叫 $\gamma\delta$ ）最具代表性
结构特点：比较大——5kb，两端不含IS，通常末端有38bp左右的IR
具有独立的转座酶和解离酶基因，含抗药性基因

功能元件



TnA能和单链DNA结合，并且有反式活性（即能催化其它不产生TnpA的同类型转座子转座），这是不同于IS一型转座过程的。

TnpR有双重作用 一是作为基因表达的阻遏蛋白；另一个作用是解离酶的功能。
TnpR突变，转座频率 $\uparrow$ ，由于TnpR阻遏了TnpA和其自身基因的转录；TnpR蛋白失活，TnpA合成 $\uparrow$ ，转座频率 $\uparrow$ ，这表明TnpA转座酶的数量是转座的限制因子。

TnpA和TnpR基因之间有一个富含A-T的内部顺式控制区。TnpR的两种功能就是通过和此区的结合来实现的。

Res: The site of resolution

解离序列 (res) 是内部的特殊位点, 只有TnA家族才有这一位点。

TnpR解离酶的结合部位共有三个结合位点, 每个长约30—40bp, 结合在每个位点独立地发生。

三个结合位点 I、II、III 具有序列同源性

位点 I: res部位 包括TnpA转录的起始点

位点 II: TnpR转录的起始点

解离是一种非复性反应, 键的断裂和再接不要求能量的输入

①超螺旋发生改变: 由于转座酶的结合, DNA在res位点发生弯曲。

②解离酶在res位点切开一个短的回文, 产生两条双链, 在5'端与解离酶共价结合:



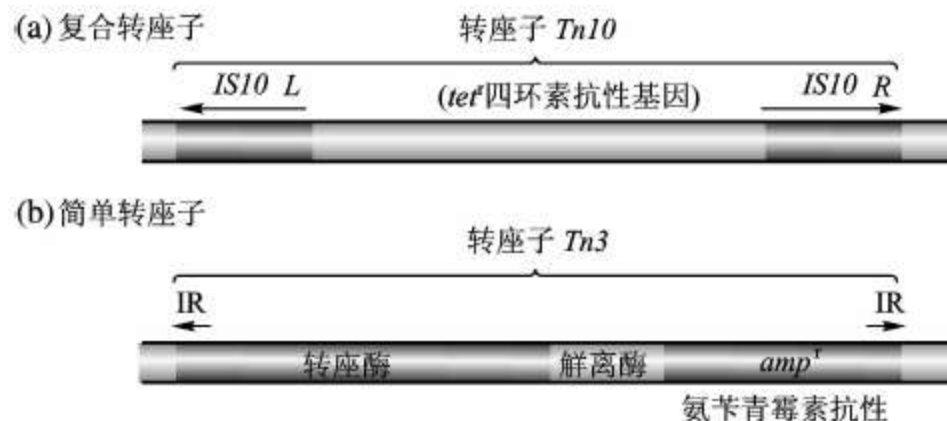


图 13 - 9 复合转座子与简单转座子的结构(引自 Griffiths 等,2012)

- (a) 复合转座子,如 *Tn10*,含转座酶基因 *IS10* 以相反方向插入形成 IR (*IS10L* 不含转座酶基因)
- (b) 简单转座子,如 *Tn3*,有短 IR, *Tn3* 序列编码自己的转座酶,解离酶将负责分离共合体为供体和受体

表 13 - 1 *Tn* 的特征

转座子	抗性标记	长度/bp	反向重复序列 中共同的序列/bp	靶 DNA 中产生的重 复序列的大小/bp
<i>Tn1</i> , <i>Tn2</i> , <i>Tn3</i>	氨苄青霉素	4 975	38	5
<i>Tn4</i>	氨苄青霉素、链霉素、磺胺	205 000	短	含 <i>Tn3</i>
<i>Tn5</i>	卡拉霉素	5 400	8/9	9(<i>Tn5</i> 的每一端都由插 入序列 <i>IS50</i> 按相反方向构成)
<i>Tn6</i>	卡拉霉素	4 200		
<i>Tn7</i>	三甲氧苄二氨嘧啶、链霉素	14 000		
<i>Tn9</i>	氯霉素	2 638	18/23	9(<i>Tn9</i> 的每一端都是同向插 入序列 <i>IS1</i>)
<i>Tn10</i>	四环素	9 300	17/23	9(<i>Tn10</i> 的每一端都按相反方 向构成的插入序列 <i>IS10</i>)

13.2.3 转座噬菌体

1963年 Taylor 发现Mu-phage (Mutator phage)

Mu噬菌体是一种以大肠杆菌为寄主的温和噬菌体，以裂解生长和溶源生长两种方式交替繁衍自己，同时它又能像IS和Tn一样可以在宿主基因组上随机进行转座。即Mu噬菌体具有温和噬菌体和转座因子的双重特性。

Mu-phage：一种DNA噬菌体，38000bp 线状DNA，游离噬菌体和整合状态具相同的基因次序，Mu-DNA不含末端反向重复序列，这是和其它转座子不同的地方。

游离时与整合时的差别，在于两个末端序列的变化

游离：两端连接着一段寄主DNA，左端100bp,右端1500bp

整合：再一次整合时，这两段序列消失



Mu的插入途径:

- (1) 溶源化过程: 插入寄主的任意部位, 造成靶点的倍增5bp
- (2) 裂解生长: 子代Mu-DNA全部随机插入寄主DNA, 能作为转座子再造成其它靶点上的插入。

Mu的复制能力和它的转座能力是密切相关的, Mu的生存依靠转座, 复制转座是其正常生活史中的一种方式。

Mu-phage的另一结构特点:

α 区: 含有包括A、B基因等大多数基因

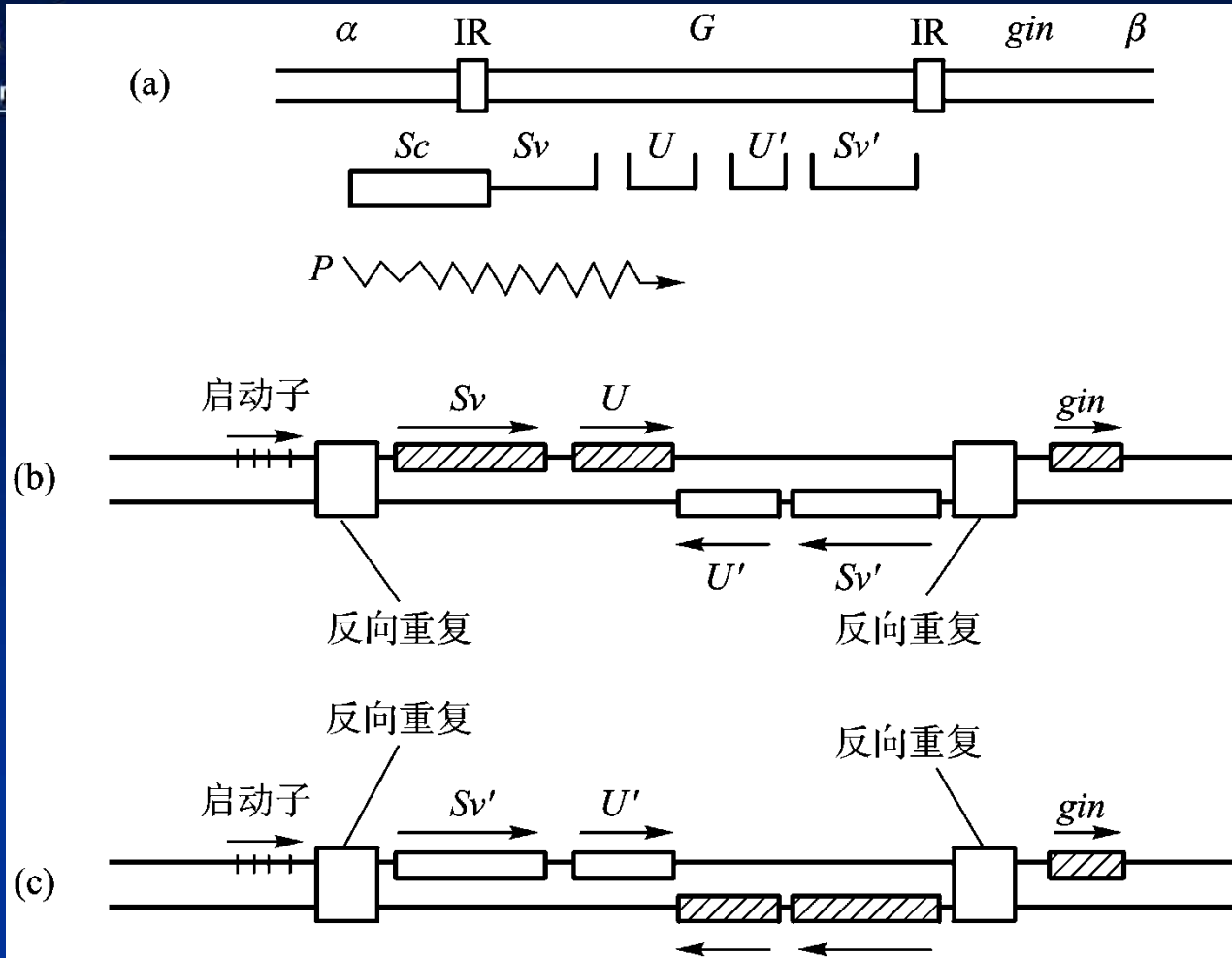
右侧3kb的G区序列: 含Sv、U、U'和Sv'4个基因

β 区: 含gin等两个基因

在转录时G区序列的不同走向导致了不同的寄主特异性:

G (+) $\rightarrow$ Sv和U基因表达 $\rightarrow$ 吸附*E.coli*K12菌株

G (—) $\rightarrow$ Sv'和U'基因表达 $\rightarrow$ *E.coli*C菌株



在取向 (b) 时，可变化区与 Sc 相邻，而在取向 (c) 时，一个不同的可变区 Sv' 在该位置，倒位是由噬菌体的 gin 基因编码产物即转化酶 (invertase) Gin 蛋白来催化的，该蛋白对 G 序列末端的 34bp 反向重复序列起作用。



13.3 真核生物中的转座子

13.3.1 酵母基因组中的转座子

在酵母中研究得较清楚的转座子是 *Ty* (transposon yeast, *Ty*) 系列。

其一般长度约为5900 bp，两端各含有一个 δ 正向长末端重复序列 (LTR)，LTR的长度约340 bp。 δ 因子大约由70%的AT组成，每一个 δ 因子都含有一个启动子和一段被转座酶识别的序列 [图13-12 (a)]。*Ty1*插入酵母染色体后，受体上就会出现5 bp的DR。另外由于 *Ty1*插入后， δ 为DR，所以也有可能发生类似于细菌中复制重组过程形成的小环，丢失一个 δ ，而留在酵母中的 δ 称为Solo δ ，这时酵母细胞表型则恢复正常，而人们可以通过对Solo δ 的分析来判断此处曾插入过 *Ty1* [图13-12 (b)]。



一般认为 *Ty1* 因子转座时，首先以其DNA为模板合成一个拷贝的RNA，然后再通过反转录合成一条新的 *Ty1* 因子，最后这条新的 *Ty1* 转座子再插入到新的位点上（图13-13）。

由于每一个细胞中有多个拷贝的 *Ty1* 因子，所以基因组中不同位置的 *Ty1* 因子也可能发生重组，从而导致染色体的易位、缺失和倒位等结构变异。

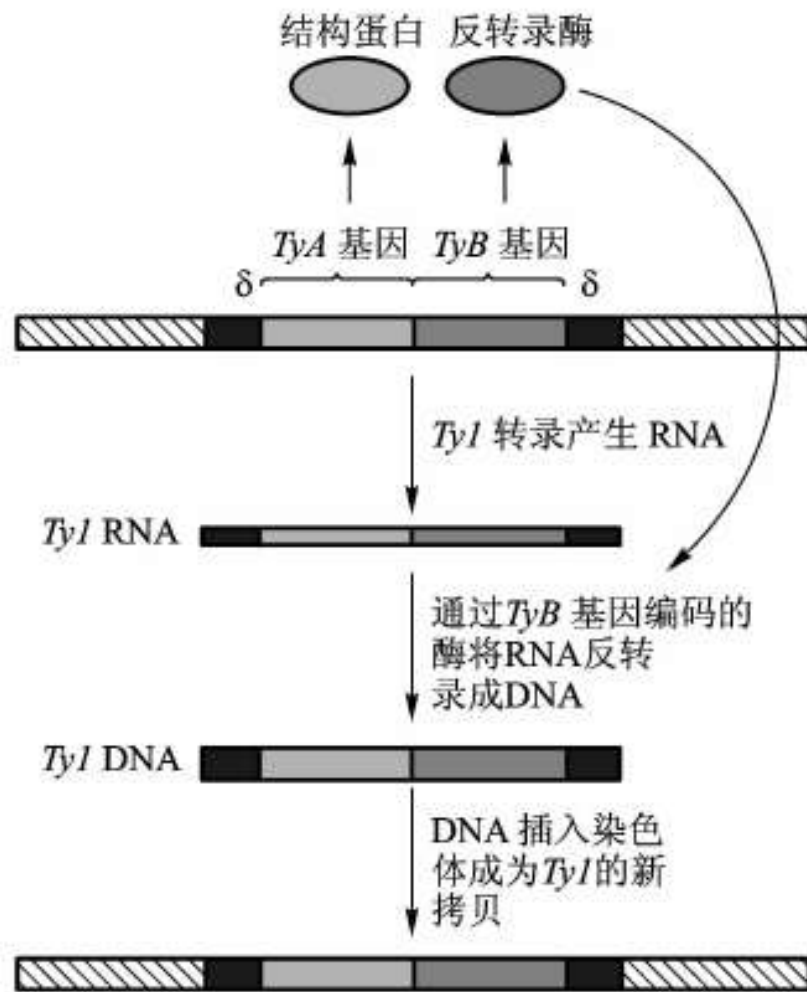


图 13 - 13 酵母 *Ty1* 的转座(引自 Snustad 等,2003)



13.3.2 果蝇基因组中的转座子

P因子

1977年 M.G.Kidwell等报道:

D.melanogaster 某些品系进行杂交 $\rightarrow F_1$ 多

“劣生”

已鉴定出两个系统与杂种不育有关:

第一系统 I(inducer)型

R(

当 I型 ♂ $\times$ R型 ♀ $\rightarrow F_1$ 育性降低 但

第二系统 果蝇被分成两种类型

P型 (父本贡献的 Paternal contributing)

M型 (母本贡献的 Maternal contributing)

P ♀ $\times$ M ♂ $\rightarrow F_1$ 正常 $\rightarrow F_2$ 正常

M ♀ $\times$ P ♂ $\rightarrow F_1$ Hybrid appears normal but is sterile(No progeny)

说明杂种劣生是不对称的, 它可被 $P_{male} \times M_{female}$ 杂交而诱导, 但其反交不能诱导出劣生

		雌性亲本	
		M	P
雄性亲本	M	正常	正常
	P	不育	正常



不育主要是与生殖细胞 (germ cells) 有关的问题:

在P—M系统的杂交后代中 (F_1) 体细胞组织正常, 但性腺不发育。

配子发育中的形态缺陷 (Morphological defect) 从种系细胞分裂时开始

任一 $P_{male} \times M_{female} \rightarrow F_1 \text{ dysgenesis}$

重组染色体的实验表明: 在每条P型雄果蝇染色体中的各个区域也都导致不育, P雄果蝇有很多P因子, 占据许多不同染色体位置。M型果蝇染色体上缺乏P因子。

对杂种劣生的 F_1 中w(white)基因突变的DNA作图表明:

所有的突变都是由于DNA片段插进w位点所引起,

被插入的DNA序列称为Pelement

①结构: P因子在长度上变化, 序列上同源, 两端含31bp反向重复序列, 转座在靶位点产生8bp的同向重复。

最长的P因子2.9kb, 4个开放阅读框: ORF0、ORF1、ORF2、ORF3

最短的P因子: 内部缺失编码区, 丧失产生转座酶的能力, 但仍能被完整的P因子编码酶反式激活



P品系：带有30—50份拷贝的P因子，大约1/3是全长的。

P因子是P品系基因组的惰性成分（inert component），
而当p雄蝇与M雌蝇交配时，它们被激活而转座。

P因子的激活具组织专一性：它仅发生在生殖细胞中，但是P因子在生殖细胞和体细胞中都能被转录。

组织专一性在这里仅指剪接方式里的一种变化，
表现在不同的细胞中转录的方式不同。

②p因子表达的调节机制

P因子：4个外显子，3个内含子。在体细胞中，仅前2个内含子被剪切；
产生 ORF0-ORF1-ORF2编码区→66kDa protein →转座激活的阻遏物
实验证明：第三个内含子的切除是转座作用需要的。第三个内含子的切除和ORF3被拼接成阅读框，这是一个关键的调控作用。这决定P因子能否激活并进行转座。

是什么负责组织的特异拼接？体细胞中含有一种蛋白，它结合在外显子3的序列上，阻止了ORF3的剪接。在生殖细胞中缺乏这个蛋白，因此可以剪接第三个内含子，产生编码转座酶的mRNA



杂种不育依赖于杂交的性别的定向，表明细胞质和P因子同样重要。

细胞质的贡献——细胞型 (cytotype) :表现为一种可遗传的细胞质效应或
细胞质因子

带 P 因子

P细胞型

缺乏P因子

M细胞型

细胞型效应的分子解释 (p 262 图11—15) (GVIIp481 [F15.27](#))

取决于66kDa蛋白抑制转座的能力，这个蛋白质是卵中母体因子产生的
卵 细 胞 中： 66kDa蛋白、母体因子

在P品系中：有充足的66kDa蛋白防止转座的发生。包括在P雌蝇中任何交
配中，该蛋白的存在既阻止转座酶的合成，也阻止转座酶的激活。

在M型雌蝇中：在卵细胞中不存在阻遏蛋白（因为没有P因子），而从雄蝇
亲本中引进的P因子导致种系里转座酶的激活。



果蝇中的转座子除了 P 因子外还有*copia*、*412*、*279*、*Tip*、*FB*等。它们的结构虽有所不同，但两端都有反向重复序列（图13-15）。

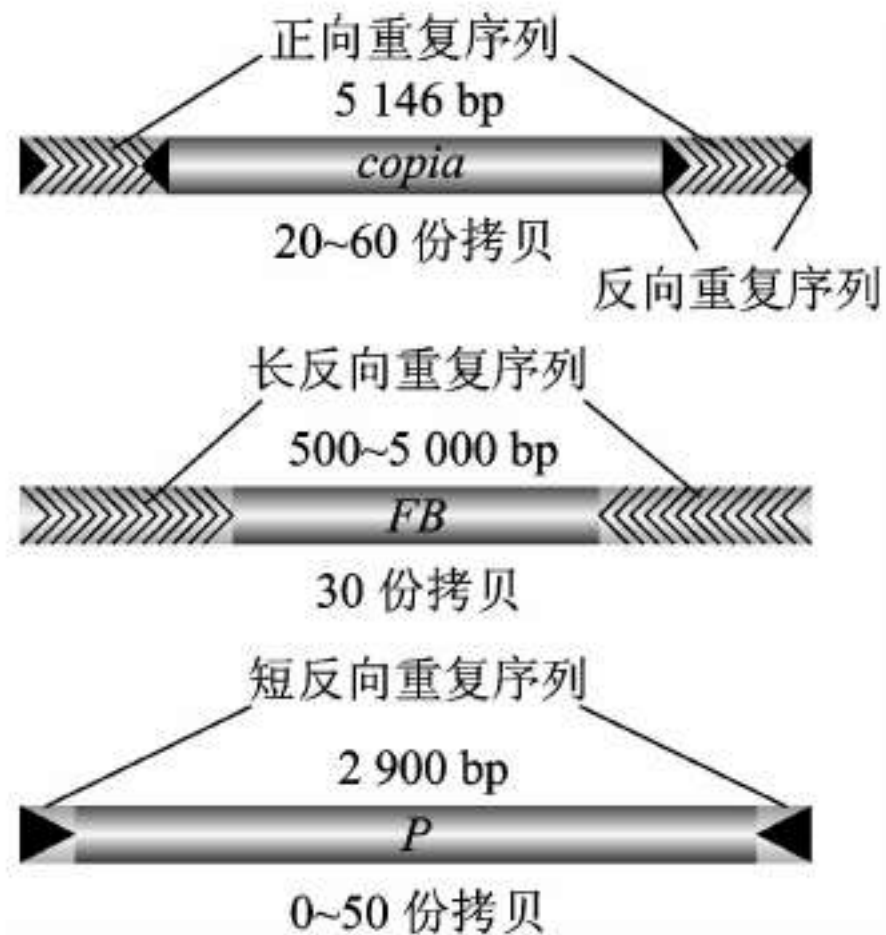


图 13 - 15 果蝇中的 3 种转座子结构



13.3.3 玉米基因组中的转座子

(1) *Ac-Ds* 系统 *Ac-Ds*系统 (*Ac-Ds*)

McClintock 1940-1950年

对玉米进行遗传分析与细胞学观察

玉米籽粒的糊粉层上出现色斑与花纹的品系

9Chr.短臂经常发生断裂，断裂后的染色体接着发生

“Breakage-fusion-bridge Cycle”

其后代籽粒→色斑





武汉大学

Wuhan University

在9Chr.发生断裂的位置称为解离（Dissociation）座位→简称Ds座位

Ds座位→移动到另一个位置→染色体在新的位置上发生断裂

Ds座位的这种移动——**转座**

Ds座位——转座因子 transposable element

McClintock指出，Ds因子使染色体和其自身转座都需要在染色体的另一个位置上有第二个因子存在，使Ds因子活化，此第二个因子称为actiator即Ac因子。Ac因子本身能够转座，但不能使染色体发生断裂，根据它们二者的关系称为**Ac-Ds为双因子系统**





研究表明同一组控制成分的不同成员可以分成两大类：

(1) **自主成分** (Autonomous element) :

具备使自身切除 (excise) 和转座能力。因其含有有活性的自主成分，它在任何位点的插入都可以造成一个不稳定的或可变的 (mutable) 等位基因。自主成分自身的丧失，或其转座能力的丧失后，一个可变的等位基因变成一个稳定的等位基因。

(2) **非自主成分** (Nonautonomous element) :

是稳定的，在一般情况下它们并不自动地转座或经历其它的变化，只有当基因组中其它位置上存在一个相同家族的一个自主成分时，它们才会变得不稳定。**当由一个自主成分提供反式互补时，非自主成分表现出与自主成分相联系的活性状态，包括能够转座到一个新的位置上去。**非自主成分失去转座所需的反式作用功能。



Ac-Ds的分子结构:

Ac element: 4563bp $\xrightarrow{\text{转录拼接}}$ 3500base mRNA (807密码子)

5个外显子构成单一基因 → 转座酶

Ac两端各有11个bp的IR。插入位点的靶序列产生8bp的重复。

Ds element: 长度和序列上都有变化, 但皆与Ac相关:

两端相同的11bp的IR, 它们比Ac短, 很明显Ds因子是Ac因子的缺失变异类型。

缺失使Ds失去转座基因活性 → 失去自身转座的能力

Ac和Ds的分子结构说明:

由于Ac因子编码的转座酶对Ds因子起互补作用, 才使得Ds因子得以转座。因为各种Ds末端的反向重复顺序都存在, 都能被Ac因子的转座酶所识别。

当玉米的Ac因子家族从一个位点转座到另一个位点时, 原来位点上的转座因子都不再保留, 这种转座机制属于Nonreplicative mechanism.



Ds为一种结构和功能都不完整的转座子，只具有与切割有关的识别序列，缺乏与转座有关的功能，所以不能自动转座，容易在基因组中固定下来。

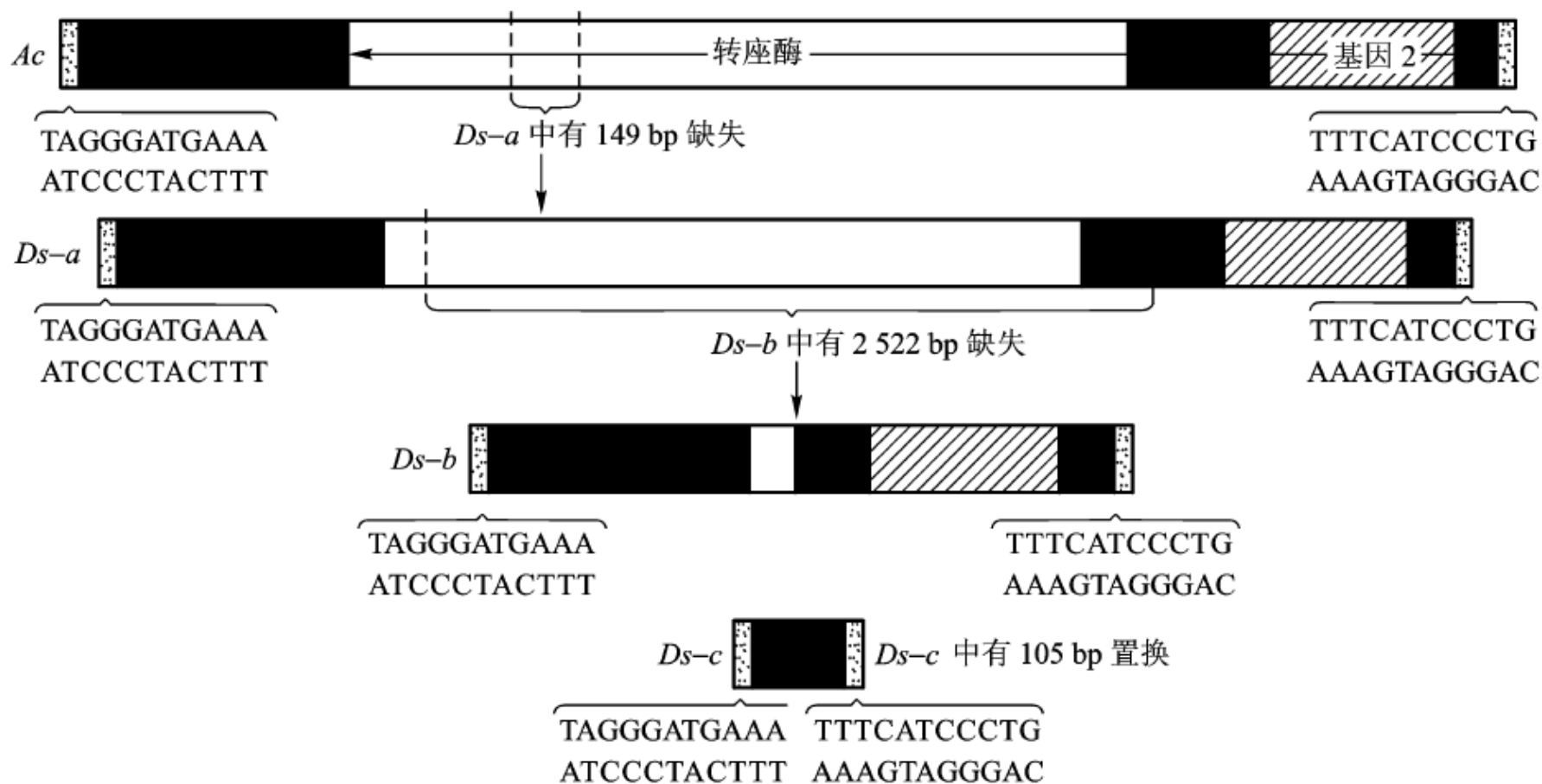


图 13 - 16 玉米中的几个 Ds - Ac 系统



13.3.4 人类基因组中的转座子

已知重复序列占了人类基因组50%以上，其中转座子占重复序列的45%。所有的转座子都是多拷贝的，这些转座子分为4种类型（图13—17）。

			长度	拷贝数	占基因组比例
LINE	自主转座		6~8 kb	850 000	21%
SINE	非自主转座		100~300 bp	1 500 000	13%
反转录病毒类转座子序列	自主转座		6~11 kb	450 000	8%
	非自主转座		1.5~3 kb		
DNA 转座子序列	自主转座		2~3 kb	300 000	3%
	非自主转座		80~3 000 bp		

图13—17 人类基因组中的几种主要转座子



LINE(long interspersed element):长散在元件, 如L1,是唯一在人类和小鼠中都具有活性的一类成员,长为6 500 bp, 含ORF1与ORF2两个阅读框。L1在哺乳动物基因组中的拷贝数可多达十万份。

SINE (short interspersed element):短散在元件, 如Alu元件,长度在100~ 300 bp之间, 是非自主转座子, 其3'端与L1有同源性, 因此能依靠L1进行转座。

Alu family: 大量的长约300 bp的中度重复序列, 广泛地分布在非重复DNA序列之间。用限制性内切酶Alu1, 可将其切割为一个170 bp片段, 和一个130 bp片段, 说明这是一类长度和性质相似的重复序列。*Alu*家族有多个成员, 总拷贝数有30万~ 50万个, 大约平均每6 kbDNA序列中就有1个, 占人体基因组的3%~ 6%。在**小鼠中与Alu序列相关序列称为B1家族**, 长度为130 bp。在**中国仓鼠中则被称为Alu相应家族**(Alu-equivalent family), 该家族也存在于其他哺乳动物的基因组中。



13.4 转座作用的分子机制

13.4.1 DNA转座机制

(1) 复制型转座 (结合 13.1.2部分ppt)

1979 Shapiro **Tn3转座模型**:

① 切开 (cutting): Tn3转座酶识别受体质粒上的靶 (target) 序列以及自身两边的反向重复序列, 并切开。

② 连接 (rejoining): 供体和受体结合成为共联体 (cointegrate)。

③ 复制: 由DNA多聚酶和连接酶参加, 在IS两端形成DR, 一般为5~11 bp。

④ 重组: 在解离位点 (resolution sites, *res*) 特定位点进行重组, 结果共联体分离形成两部分, 一个是原来含有转座子的序列, 另一个是通过转座插入了转座子的序列。

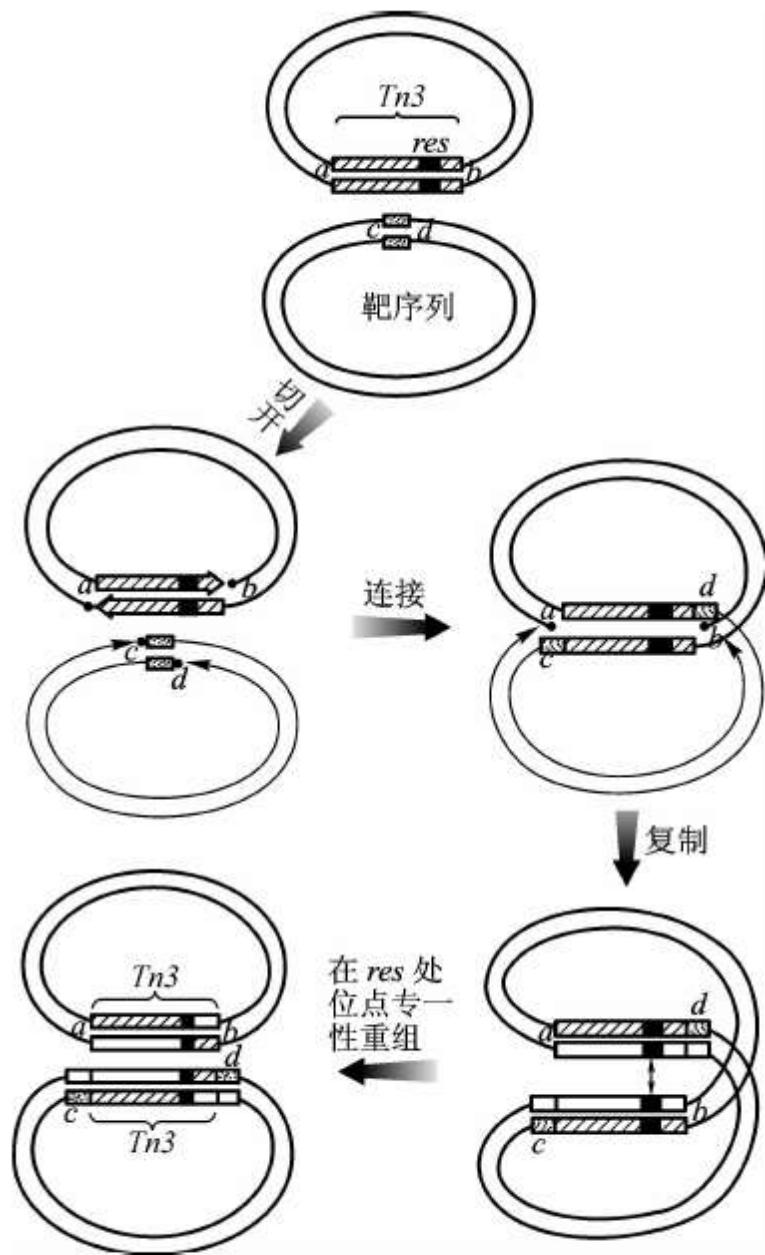


图 13-18 Tn3 复制型转座模型



(2) 非复制型转座

非复制型转座的原理是断裂和重接反应使靶序列重构，只有靶位点发生重连（reunion），而供体链仍保持裂缺，不形成共联体。非复制转座酶也可在靶DNA上产生剪接。转座子两侧产生双链断裂，完整的转座子从供体中释放出来。

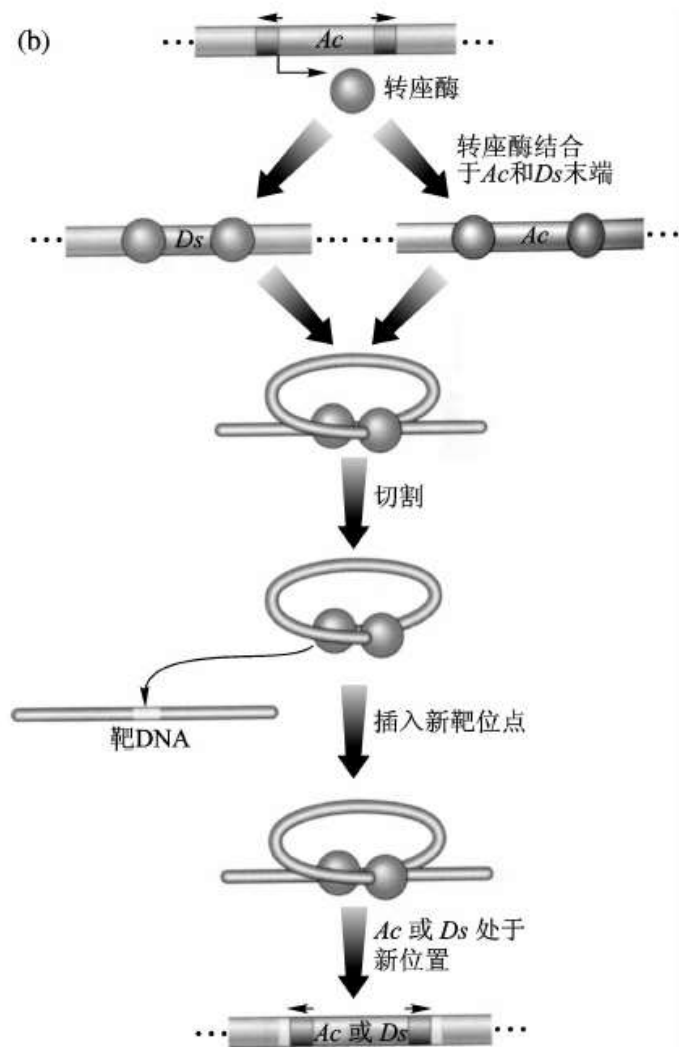
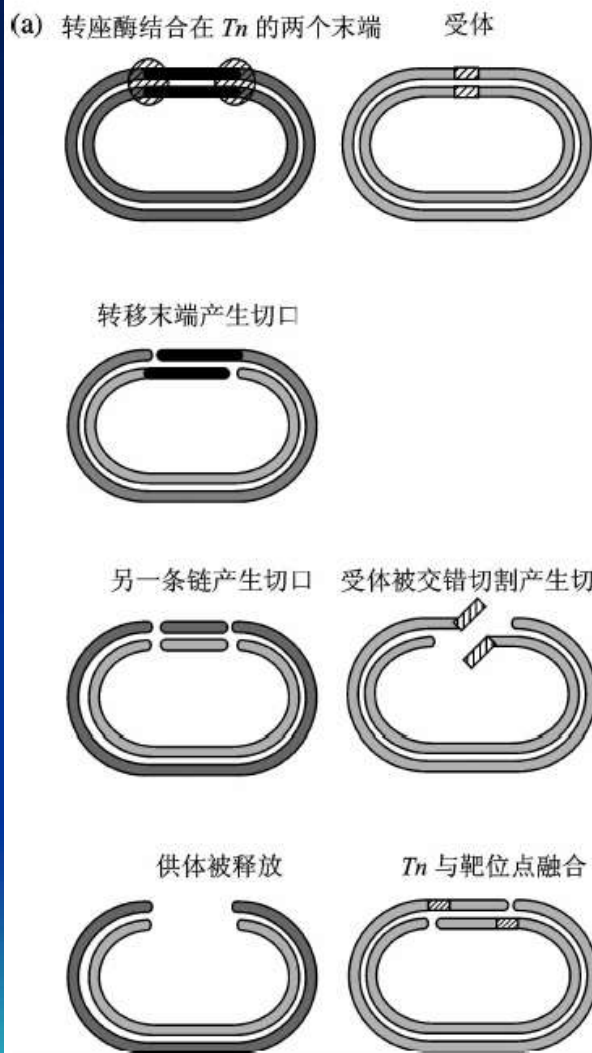


图 13-19 非复制型转座机制

(a) *Tn10* 非复制型转座 (仿自 Lewin, 2011) (b) 玉米中 *Ac-Ds* 非复制型转座 (引自 Griffiths 等, 2012)



13.4.2 反转录转座子的转座机制

(1) 反转录病毒的转座机制

(结合 13.1.3 反转录转座子部分ppt)

(2) *Tyl/copia*类反转录转座子的转座机制

Tyl 和 *copia* 因子是典型的反转录转座子，均由RNA介导转座，只是这类反转录子由于缺少编码外壳蛋白的基因（*env*），因此不能形成病毒颗粒，它们只能在细胞内生活，亦可转录合成RNA。因其基因组中有编码反转录酶的基因（*pol*），所以能将它的RNA反转录为cDNA，并以类似反转录病毒cDNA的整合机制插入到宿主基因组中。



武汉大学

Wuhan University

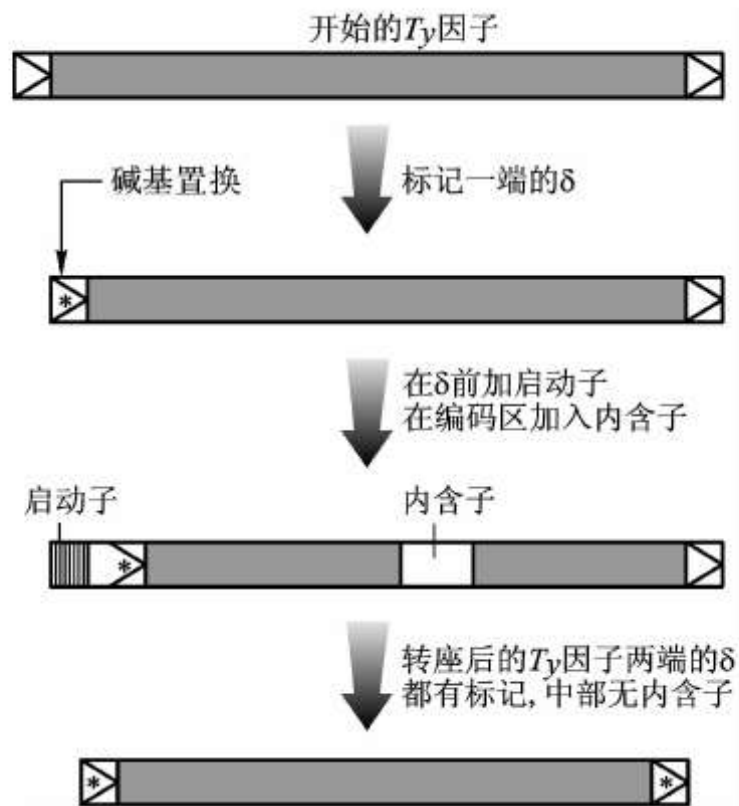


图 13-25 Ty 经 RNA 介导转座的证明
(仿自 Lewin, 2011)

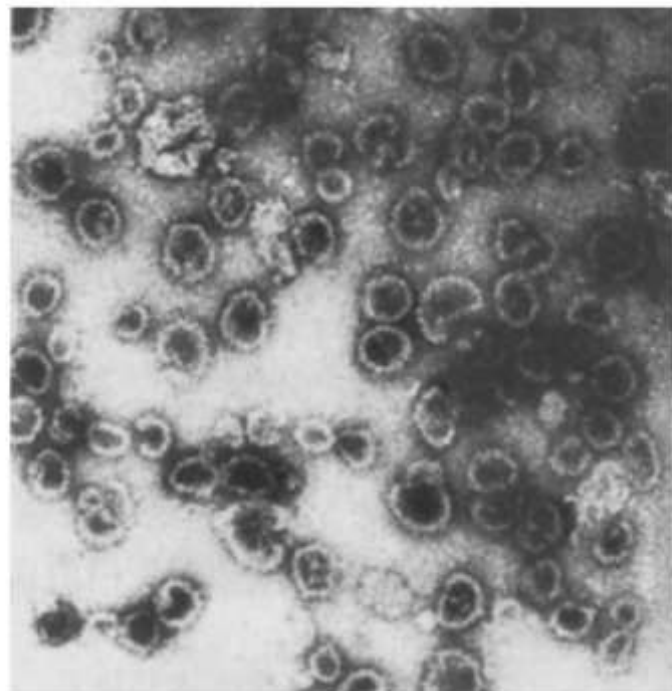


图 13-26 Ty 因子形成病毒样颗粒
(电镜照片)(引自 Lewin, 2011)



13.5 转座因子的遗传学效应及其应用

13.5.1 引起染色体结构变异

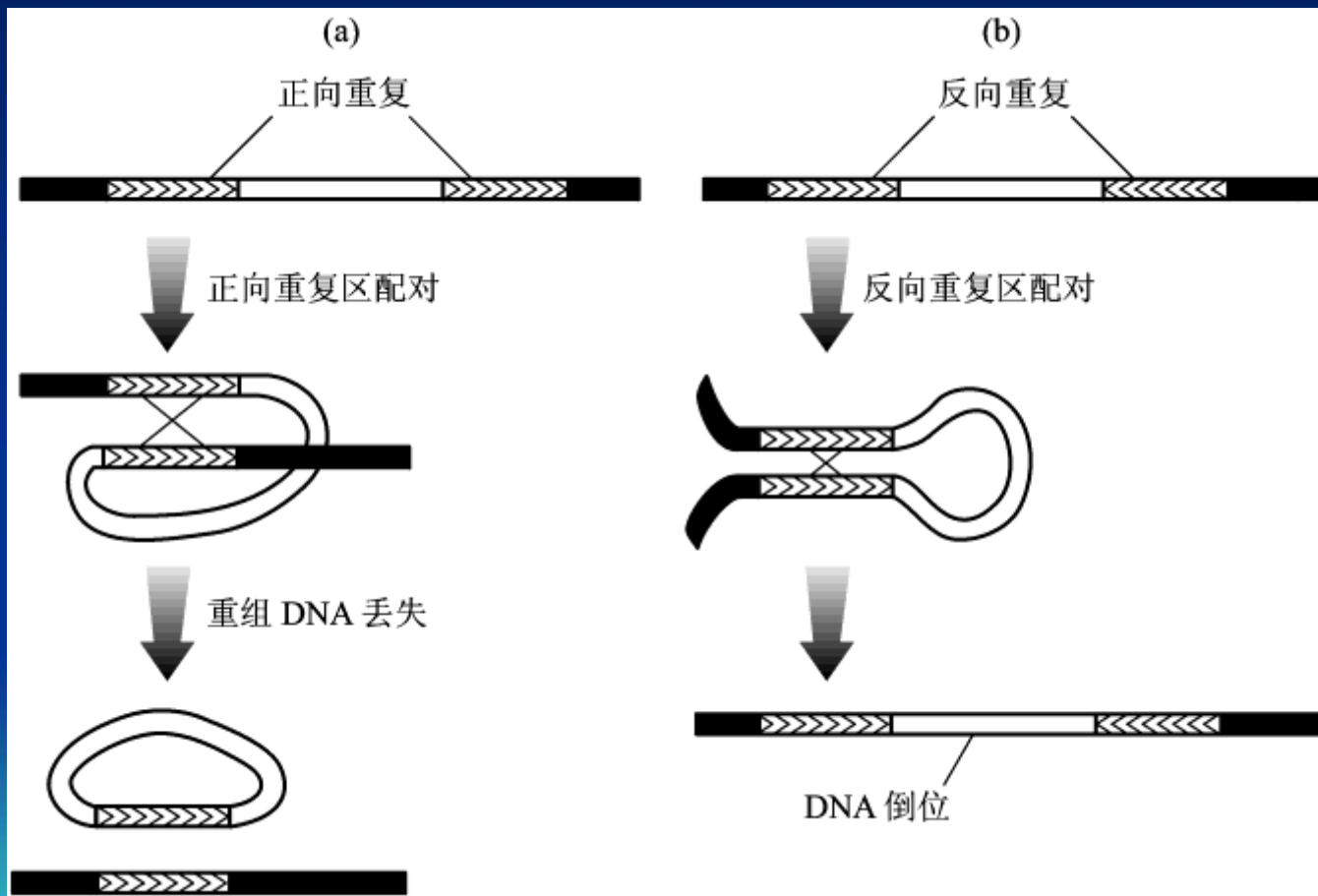


图13-4 由转座子引起的染色体DNA缺失 (a) 或倒位 (b)



13.5.2 诱发基因突变与启动外显子混编

当转座子插入到某个基因中往往导致该基因失活，在某些情况下，插入位点的基因仍然能够正常转录，只是转座子中的插入序列通过转录后的剪接过程而被除掉，因此插入位点的基因仍表现出显性性状，这种现象称为**渗漏突变 (leaky mutation)**。也就是仍有一些残余基因表达的突变。这类基因称为**渗漏基因 (leaky gene)**，又称亚效等位基因 (hypomorph)，即一种突变基因与其野生型有相似的效应，但效应较弱。

当两个转座子被同一转座酶识别而整合到染色体的邻近位置时，则它们之间的DNA将变得易于被转座酶作用而转座。如果它们之间的DNA中含有外显子，则该外显子将被切离，并可能插入另一基因之中。这种效应称为**外显子混编 (exon shuffling)**。所谓外显子混编，即源自一个或几个基因的若干个外显子像“洗牌”那样地进行重排。这也是生物体产生新基因和基因进化多样性的途径之一（图13—27）



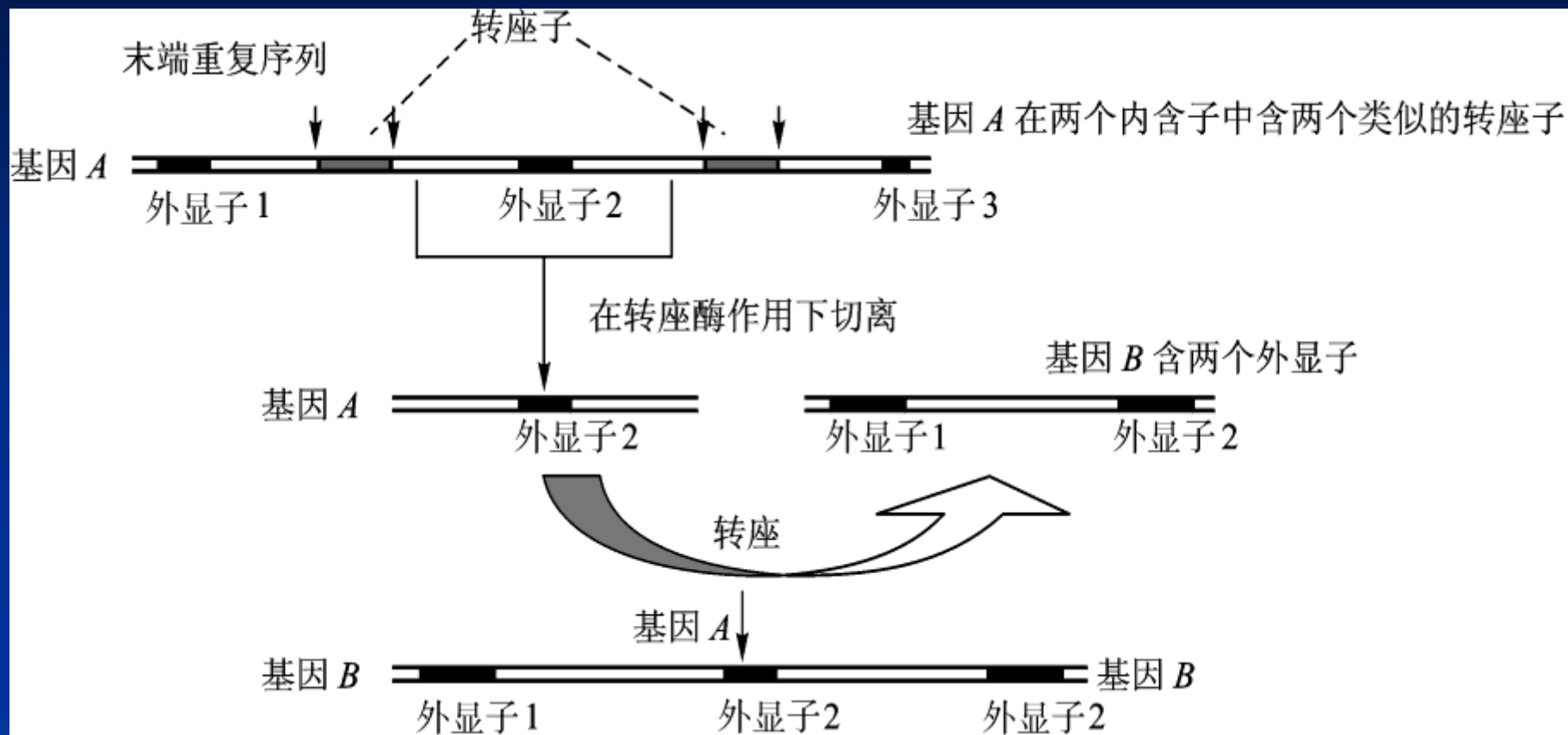


图13—27 双转座启动不同基因间的外显子混编



13.5.3 调节基因表达

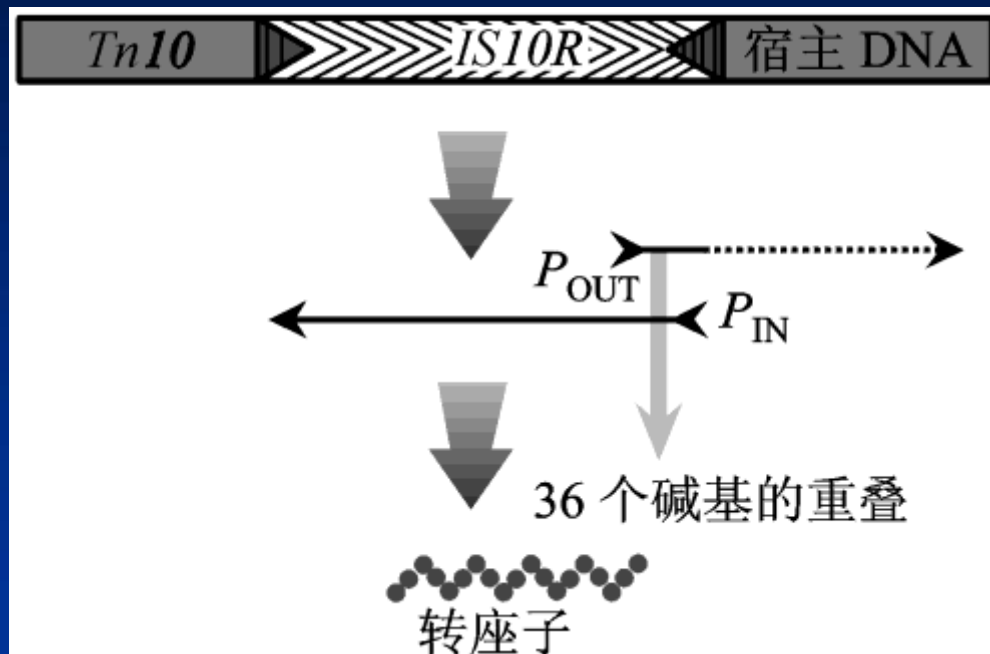


图13—28 转座子调节宿主基因表

IS10R两端具22 bp的反向重复序列。在它的右侧反向重复序列内侧有两个方向相反的启动子，其中 P_{OUT} 是外向启动子，它能激活邻近下游的宿主DNA 中的基因。 P_{IN} 是内向启动子，它们之间有36 bp重叠

转座子除了含有增强子外，有的转座子还含有启动子，也能促进基因的转录活性。如Tn10右侧的IS10R以某一方向插入到由于缺失了启动子而不能表达的argE基因的5'端时，结果使沉默的argE基因重新表达，对该序列分析表明，其末端含有一个外向的启动子，这可能是它启动沉默基因表达的原因（图13—28）。



11.5.4 产生新的变异

由于转座插入位点可能出现新的基因，如像Tn携带的抗药性基因，它的转座不仅造成某个基因的插入突变，同时在此位点上出现一个新的抗药性基因。

由于转座而增加了同源序列的整合，如IS既可插入细菌染色体的不同位置，又可插入质粒中。

在某些情况下，插入到某个基因座位中的转座子也可能失活，转座子失活过程主要与DNA分子的甲基化作用（methylation）有关。

转座子DNA序列的甲基化不仅影响转座子自身的表达，而且还影响邻近基因也呈现出表观遗传变化。





13.5.5 转座子标记目的基因

转座子标签 (transposon tagging) 方法:转座子不仅能在本基因组中转座, 也能转入其它植物的基因组中。转座的结果是使插入位置的基因失活, 从而有效地诱导产生表型突变株。然后构建一个对应于突变株的基因组DNA文库, 用作标签的转座子作为探针从基因库中筛选相对应的克隆, 分离得到相对应于变异的基因, 这就是转座子标签克隆基因。
用转座子给未知的目的基因加以标签, 便于对该基因的识别与分离。由于插入的DNA序列相当于人为地给目的基因加上一段已知的序列标签, 因此DNA插入 (**转座子插入**) 突变分离基因的技术, 又叫做DNA标签法 (DNA-tagging) (**转座子标签法 transposon tagging**)

应用该技术的条件是:

- ① 亲本之一必须含有活跃的转座子。
- ② 所用的转座子必须已经分离和鉴定, 否则无法对其进行标记作为探针
- ③ 转座子插入后要有目的性状的突变型。



13.5.6

作为基因工程的载体

利用P因子作为载体，将外源基因转移到果蝇胚胎种系细胞中，对果蝇进行遗传操作。

玫瑰色眼基因 ry^+ 被整合在细菌的带有缺失的P因子中，同时与携带有完整的P因子的辅助质粒注入 ry^- 果蝇胚胎，结果在其后代中出现转基因 ry^+ / ry^- 子代，原位杂交证明，P因子介导 ry^+ 转移到一些果蝇的种系细胞的染色体中

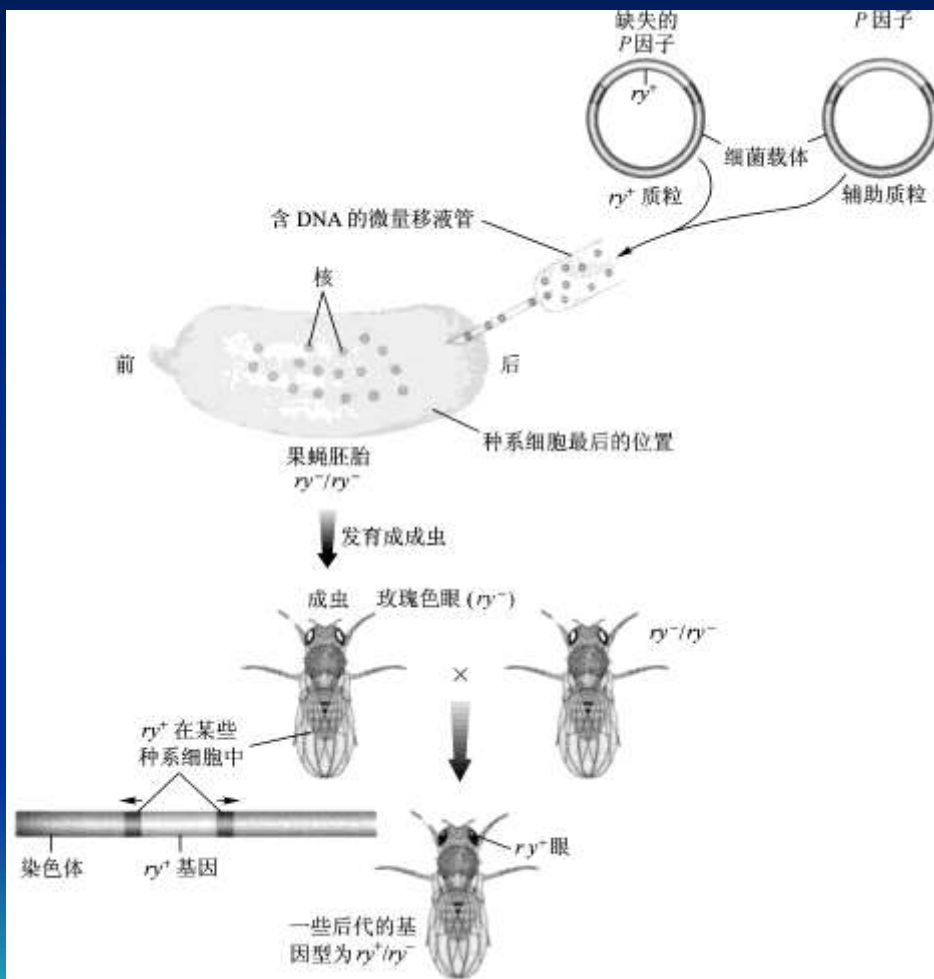
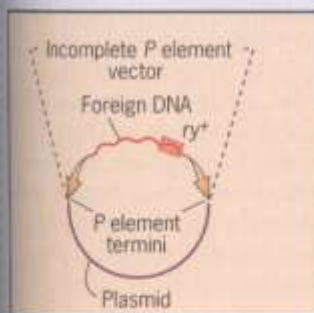
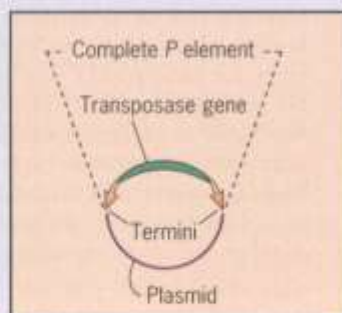


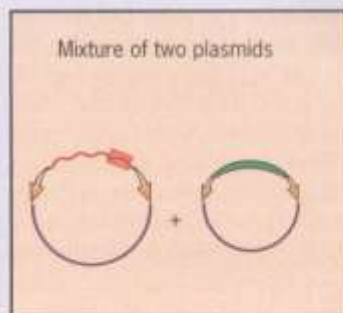
图13- 34 果蝇中P因子介导基因转移



1 Foreign DNA is inserted into an incomplete P element in a plasmid. The insert also contains an eye color gene (ry^+) as a marker. In flies, this gene produces red eyes.



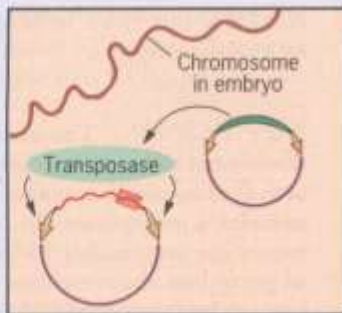
2 A complete P element is inserted into a different plasmid.



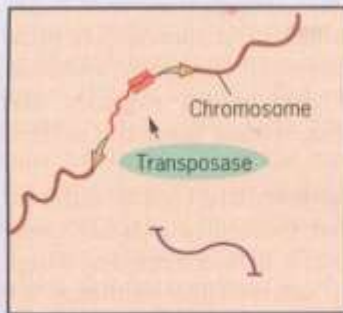
3 The two plasmids are mixed in solution.



4 The plasmid mixture is microinjected into a ry^- mutant *Drosophila* embryo. Flies with the ry^- mutation have brown (rosy) eyes.



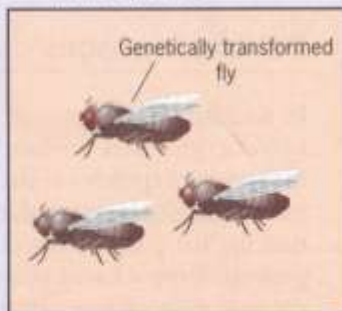
5 In the embryo's germ line, transposase from the complete P element catalyzes excision of the incomplete P element from its plasmid.



6 The excised P element is inserted into a chromosome in the embryo's germ line.



7 The embryo becomes an adult.



8 When the adult fly reproduces, the progeny are screened for incorporation of the incomplete P element with its ry^+ marker gene.



9 Genetically transformed progeny, recognized by their red eyes, are propagated.

Figure 1 Genetic transformation of *Drosophila* using P element vectors. Foreign DNA inserted between P element termini is integrated into the genome through the action of a transposase encoded by the complete P element. Flies with this DNA in their genomes can be propagated in laboratory cultures.



武汉大学

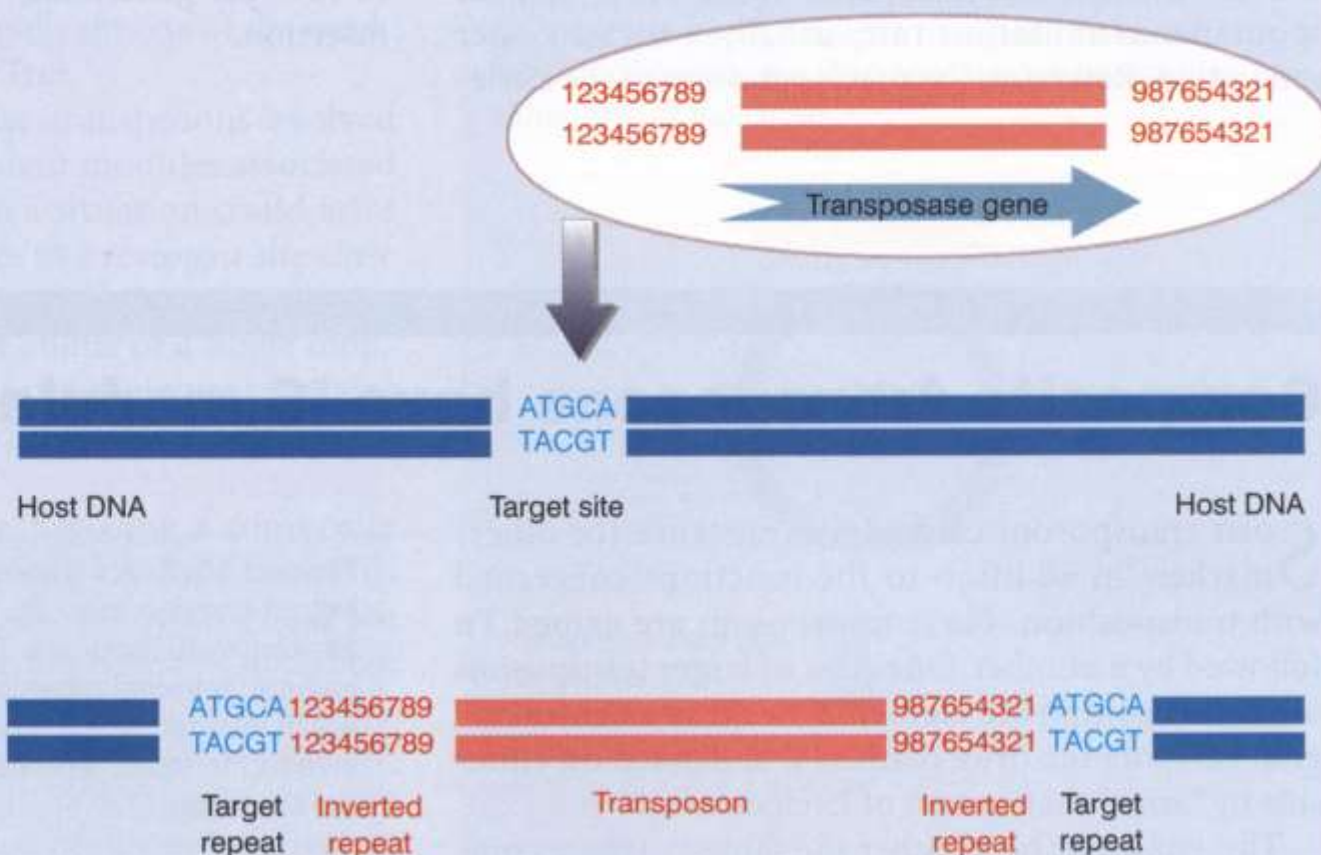
Wuhan University

本章结束

谢谢！



Figure 15.1 Overview: transposons have inverted terminal repeats and generate direct repeats of flanking DNA at the target site. In this example, the target is a 5 bp sequence. The ends of the transposon consist of inverted repeats of 9 bp, where the numbers 1 through 9 indicate a sequence of base pairs.



	Target repeat	Inverted repeat	Transposon	
	ATGCA	123456789	987654321	ATGCA
	TACGT	123456789	987654321	TACGT
	Target repeat	Inverted repeat	Transposon	Target repeat
			Overall length	Target selection
IS1	9 bp	23 bp	768 bp	random
IS2	5 bp	41 bp	1327 bp	hotspots
IS4	11-13 bp	18 bp	1428 bp	AAAN ₂₀ TTT
IS5	4 bp	16 bp	1195 bp	hotspots
IS10R	9 bp	22 bp	1329 bp	NGCTNAGCN
IS50R	9 bp	9 bp	1531 bp	hotspots
IS903	9 bp	18 bp	1057 bp	random

[返回](#)

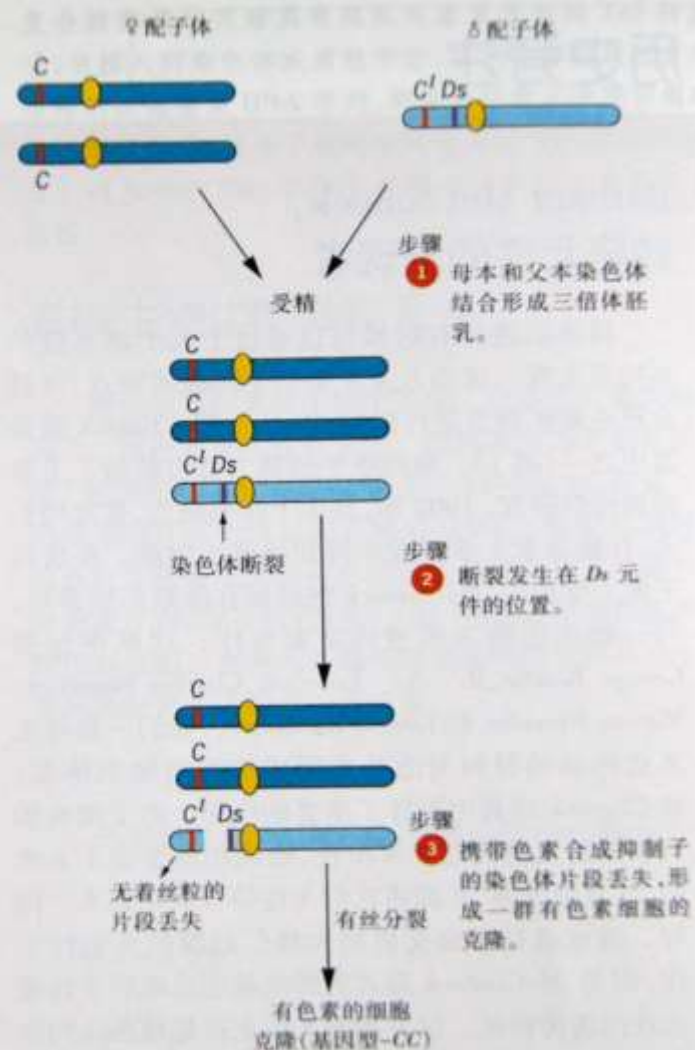
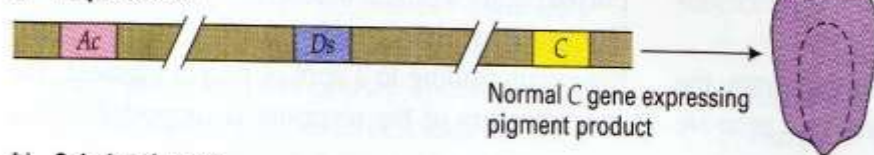
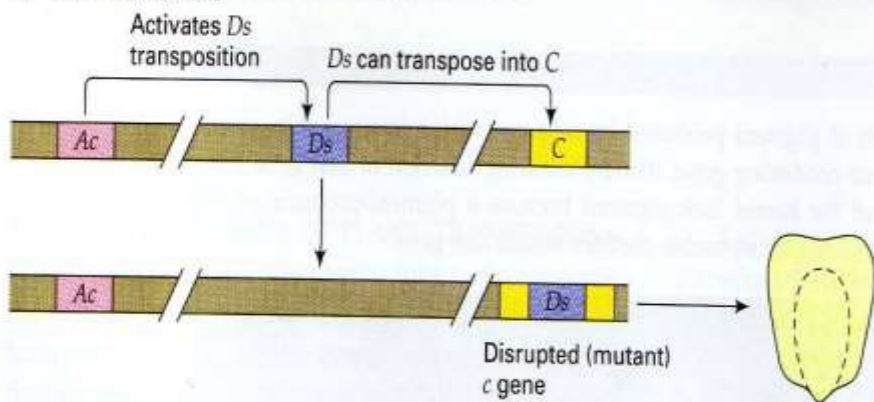


图 18.9 玉米中由 *Ds* 转座元件引起的染色体断裂。在 9 号染色体短臂上的等位基因 *C* 能在糊粉中产生正常的色素；而等位基因 *C'* 抑制这种色素形成。

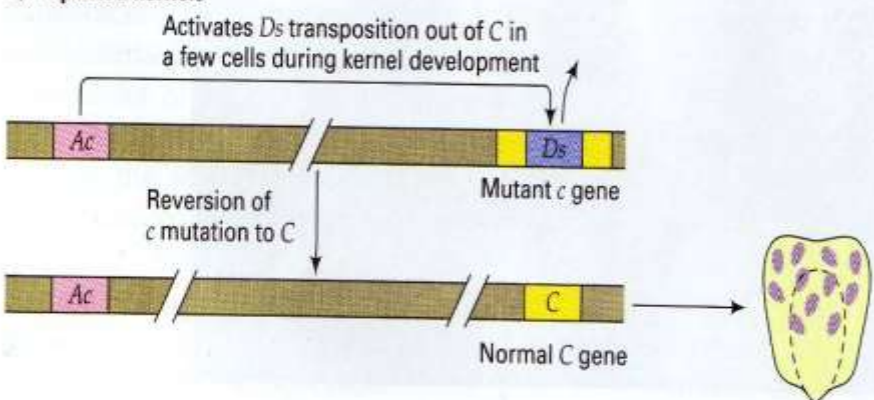
a) Purple kernels



b) Colorless kernels



c) Spotted kernels



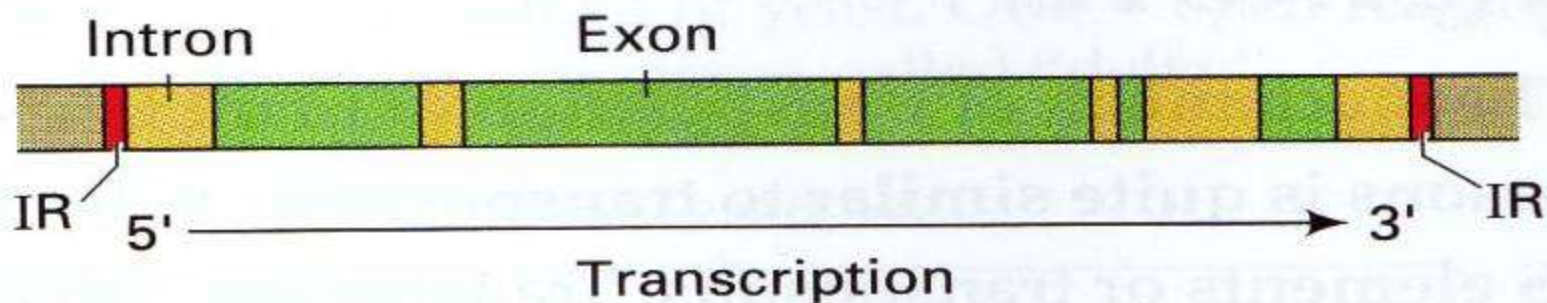
DA

返回1

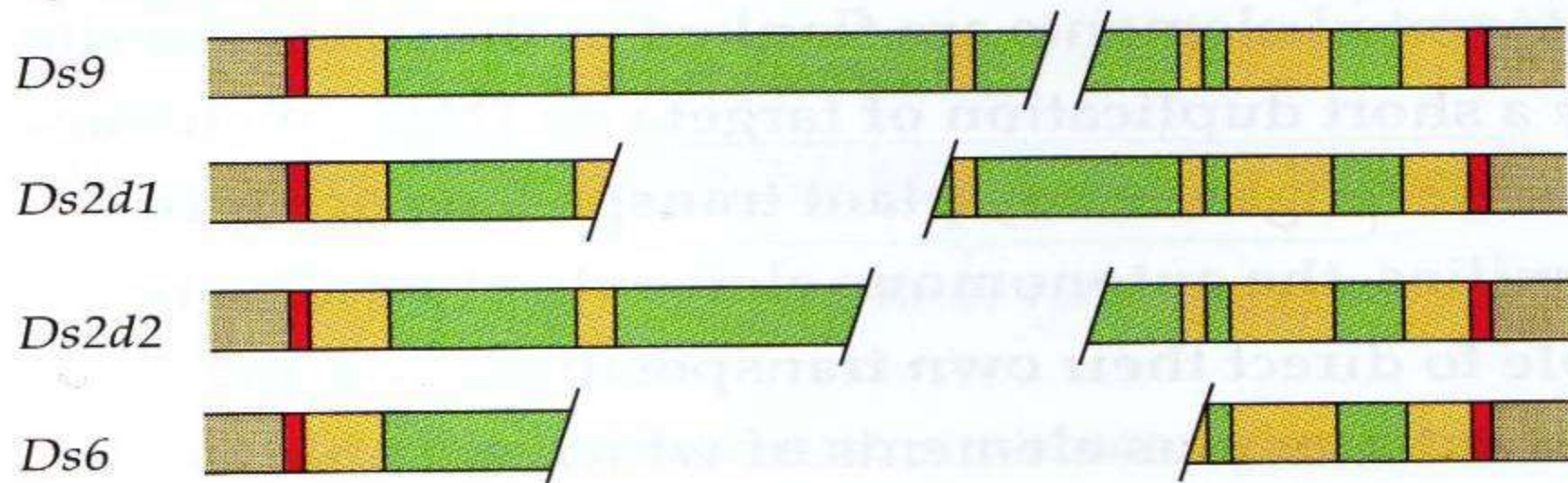
返回



a) Activator element (*Ac*)



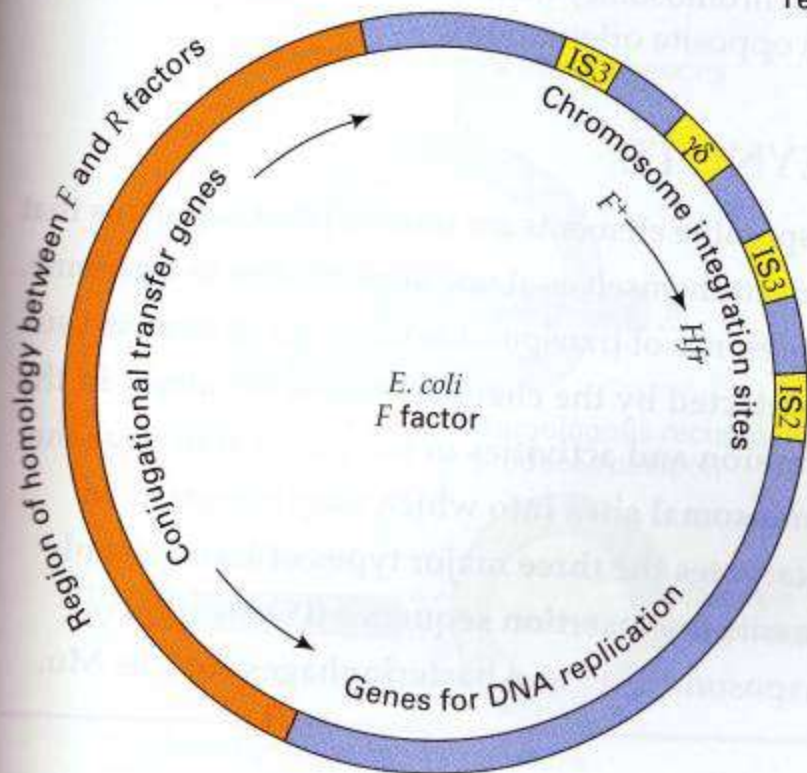
b) Dissociation elements (*Ds*)



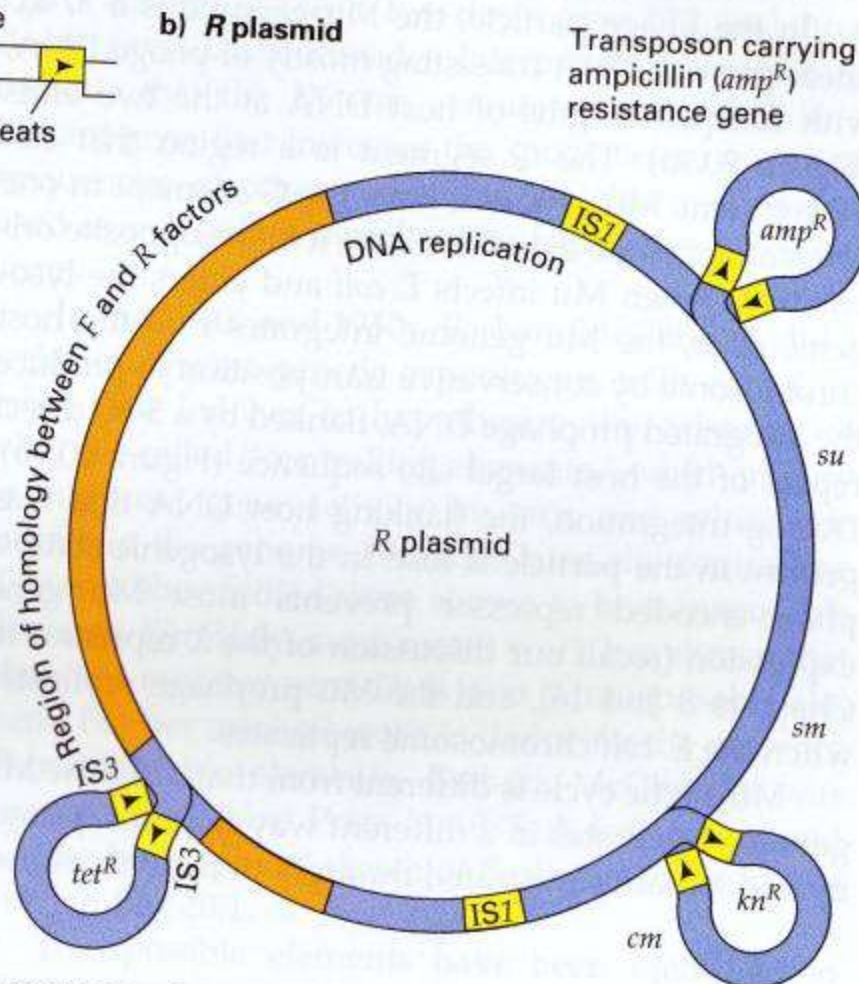


General transposon structure

a) *E. coli*
F factor



b) R plasmid



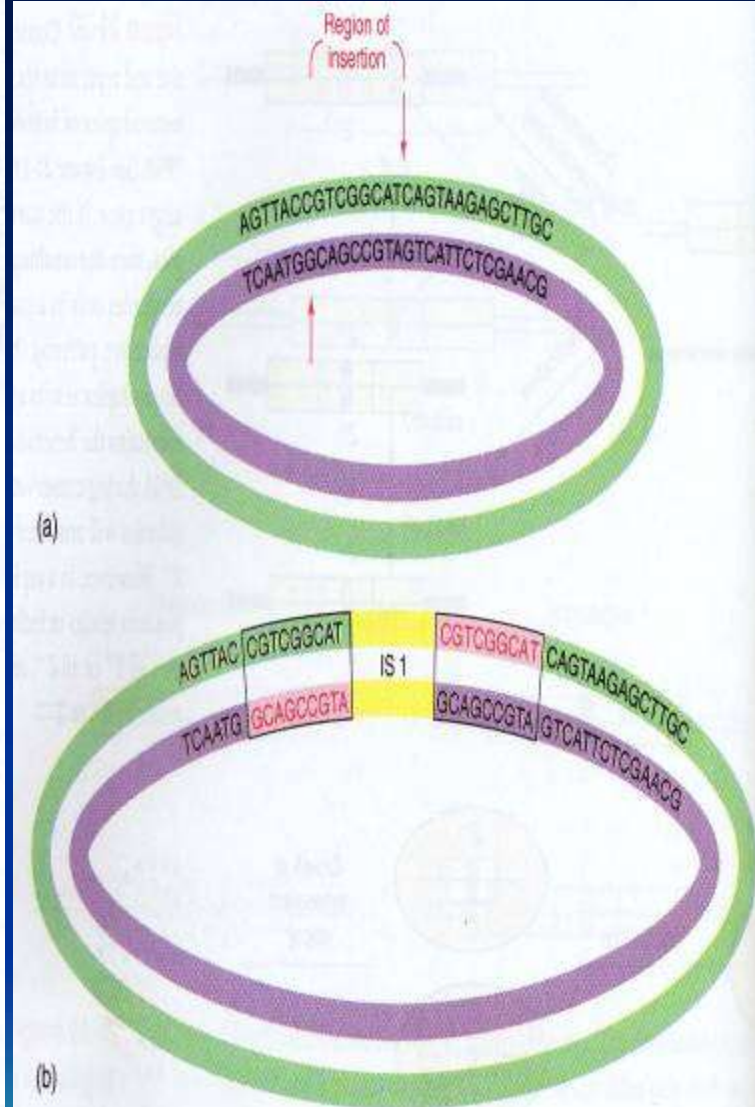
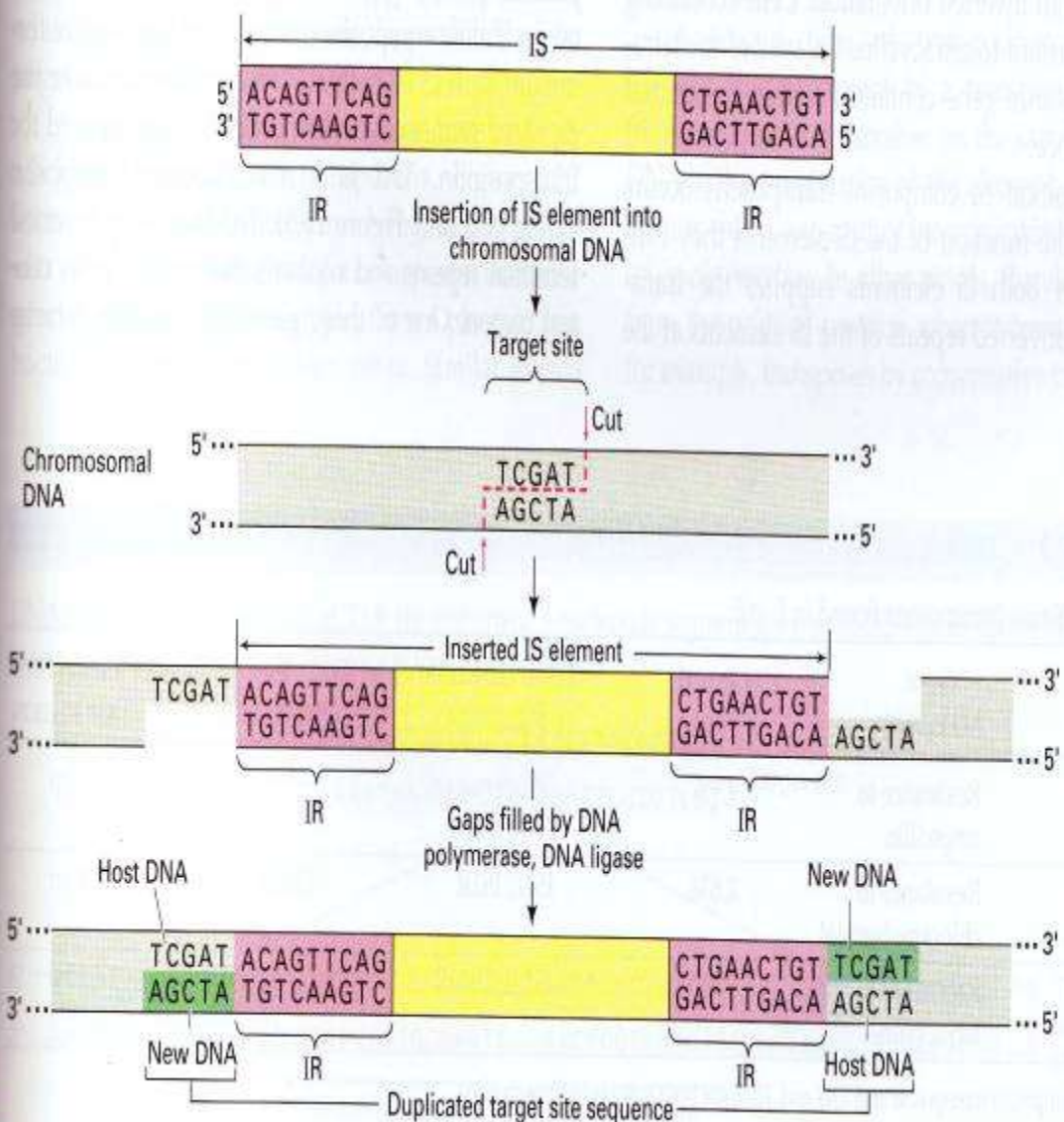
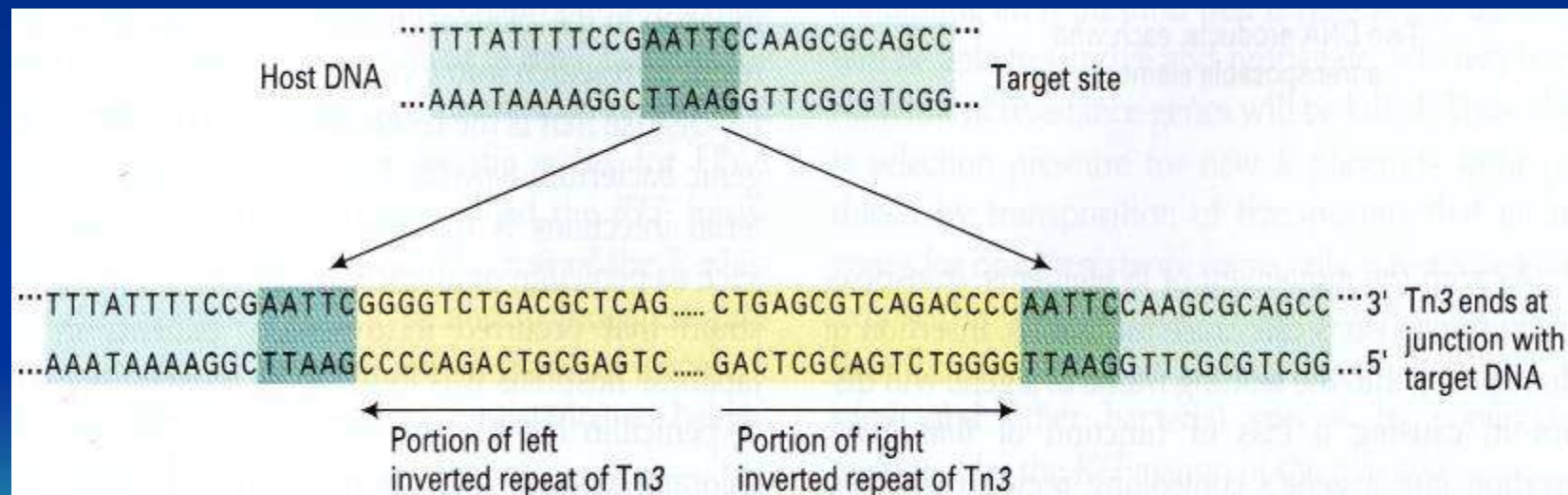
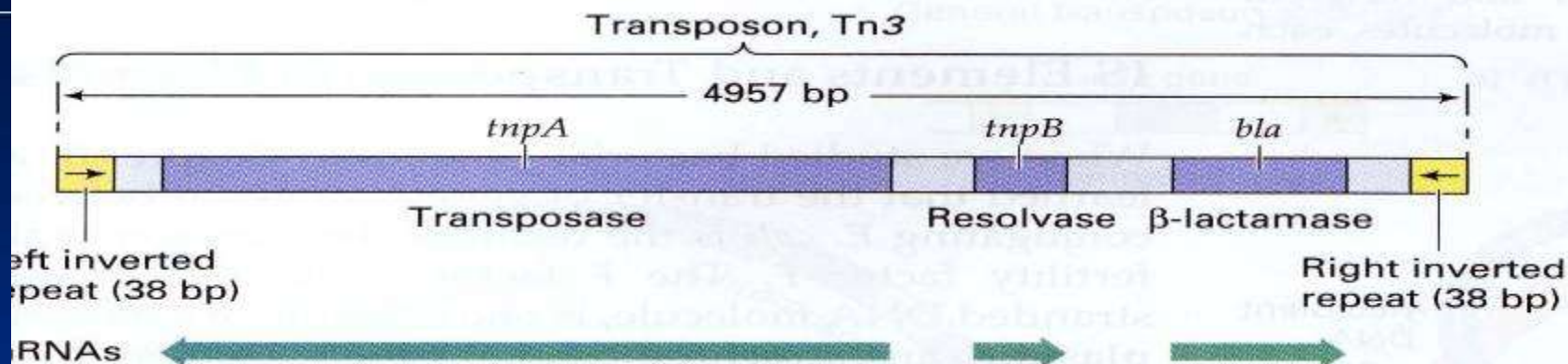


Figure 21-22 Duplication of a short sequence of nucleotides in the recipient DNA is associated with

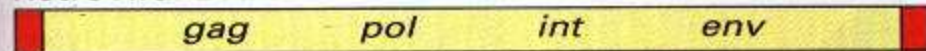
IS

返回

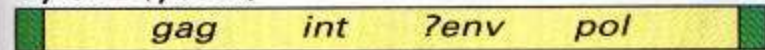


(b) Genes in retroviral DNA and viral retrotransposons

Retroviral DNA



Ty 912 (yeast)



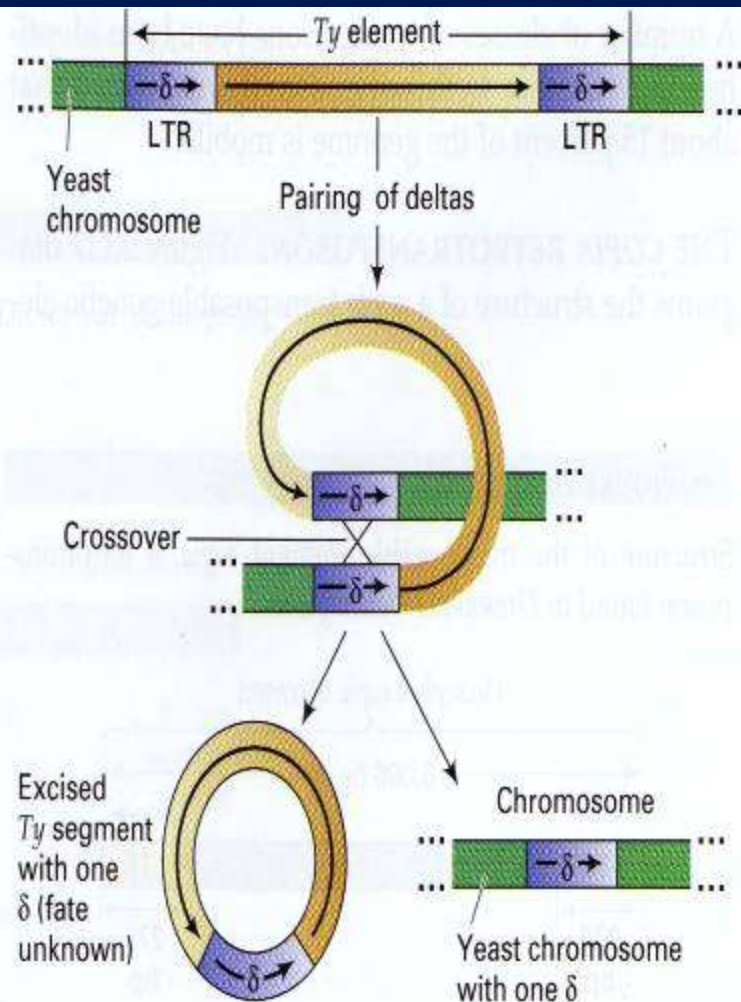
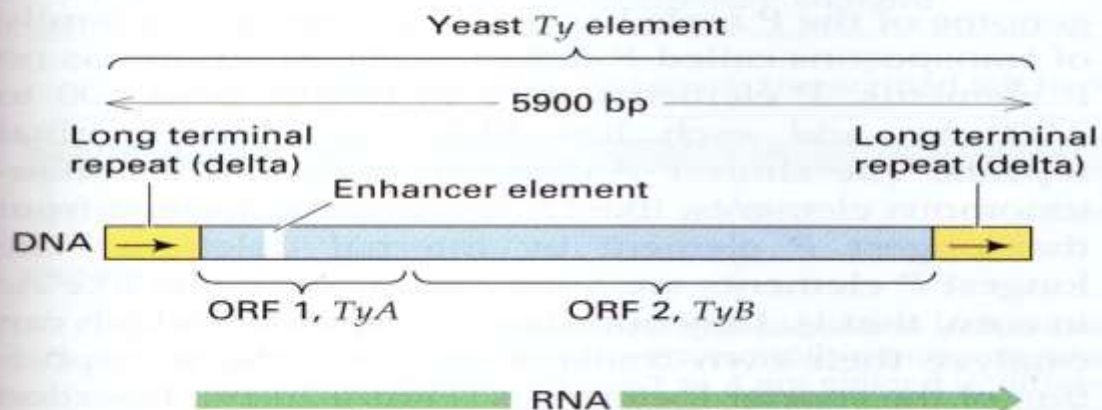
Copia (Drosophila)



Figure 21-32 A comparison of the genes of integrated retrovirus DNA and the yeast Ty elements and *Drosophila copia* elements. The

~ FIGURE 20.15

The Ty-transposable element of yeast. ORF = open reading frame. LTR = long terminal repeat, called "delta."

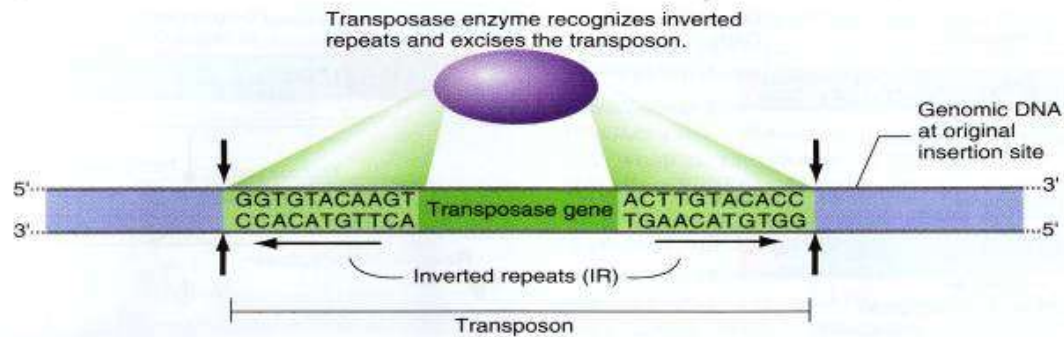


Ty

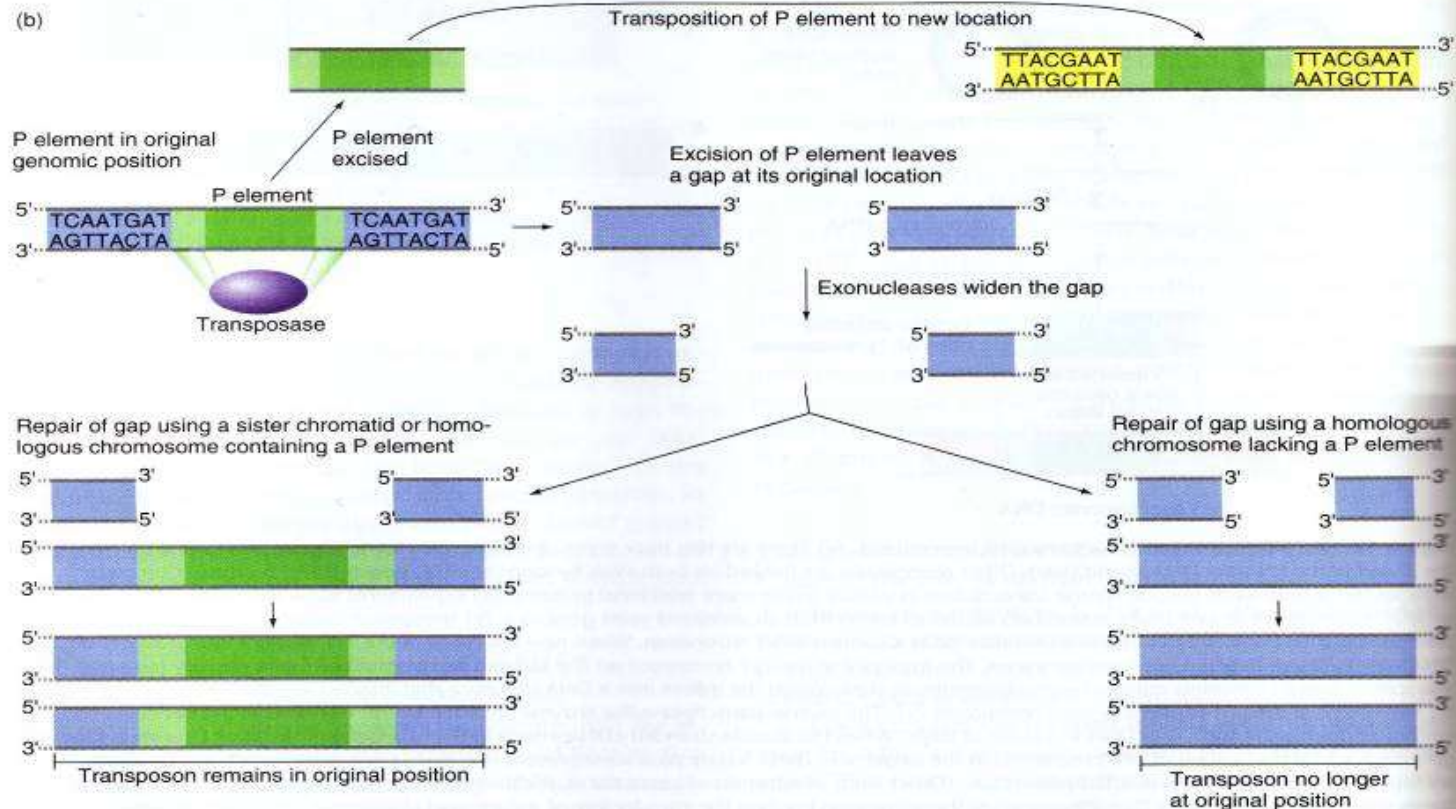
返回



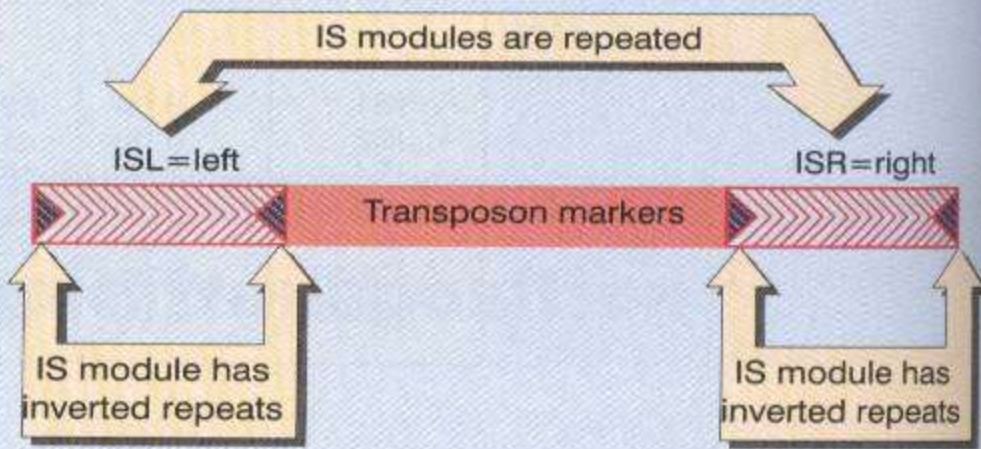
(a)



(b)

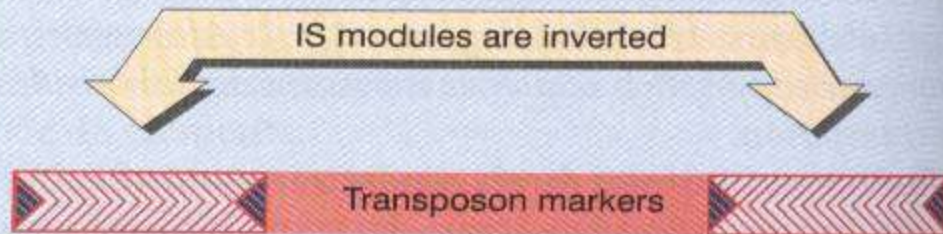


返回



Example

Tn9 IS1 *cam*^R IS modules identical both functional



Transposon	Left end	Markers	Right end
Tn903	IS903	<i>kan</i> ^R	both IS ends functional
Tn10	IS10L nonfunctional	<i>tet</i> ^R	IS10R functional
Tn5	IS50L nonfunctional	<i>kan</i> ^R	IS50R functional

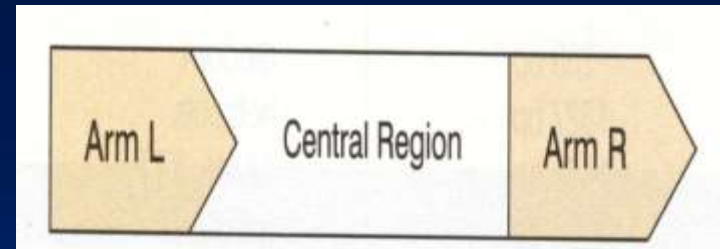
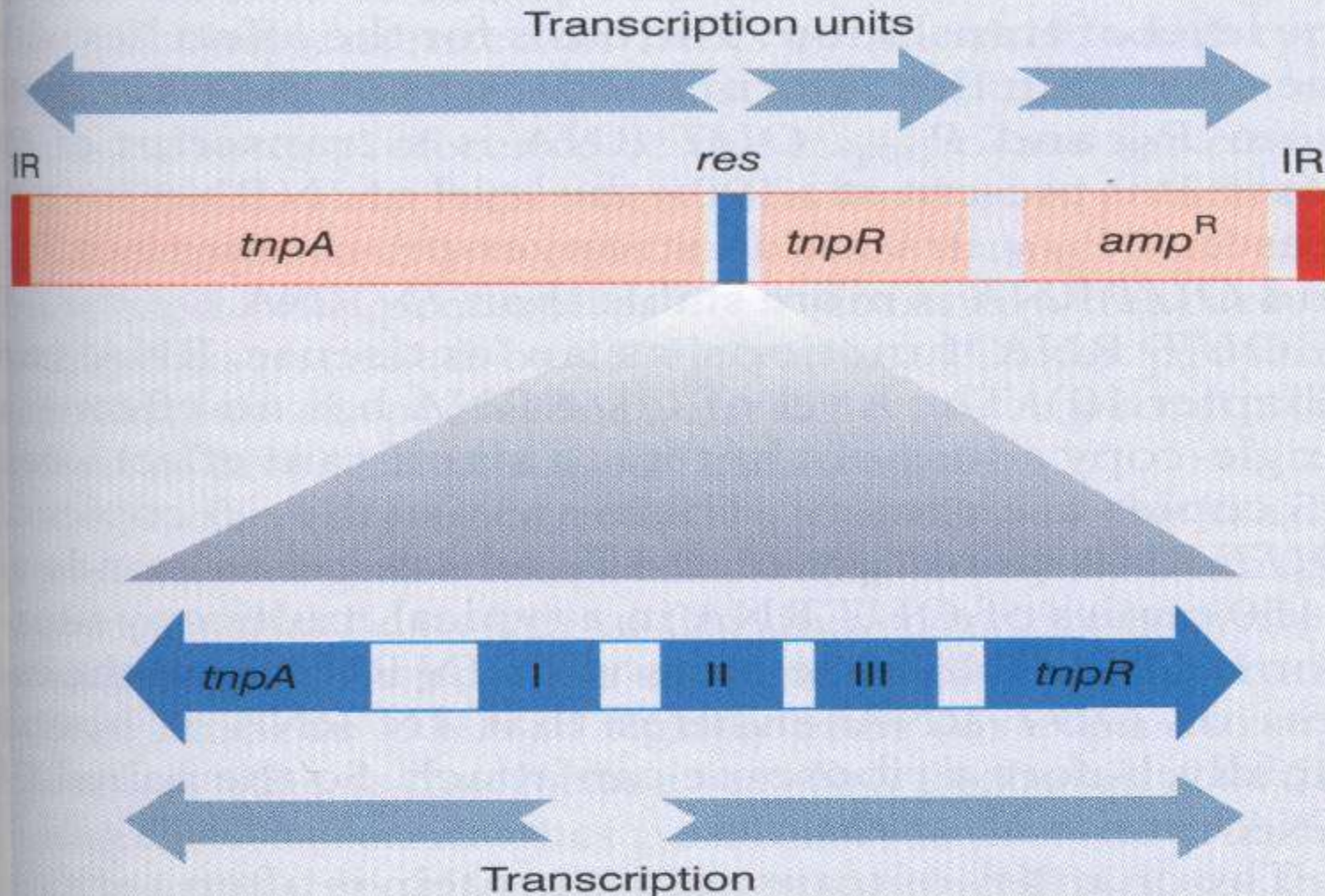


Figure 15.2 A composite transposon has a central region carrying markers (such as drug resistance) flanked by IS modules. The modules have short inverted terminal repeats. If the modules themselves are in inverted orientation (as drawn), the short inverted terminal repeats at the ends of the transposon are identical.

[返回](#)

Figure 15.16 Transposons of the TnA family have inverted terminal repeats, an internal *res* site, and three known genes.



返回

Figure 15.4 The direct repeats of target DNA flanking a transposon are generated by the introduction of staggered cuts whose protruding ends are linked to the transposon.

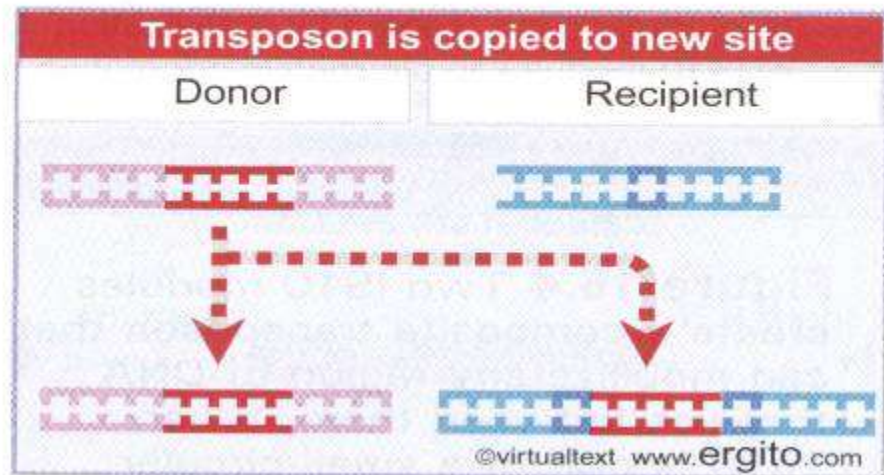
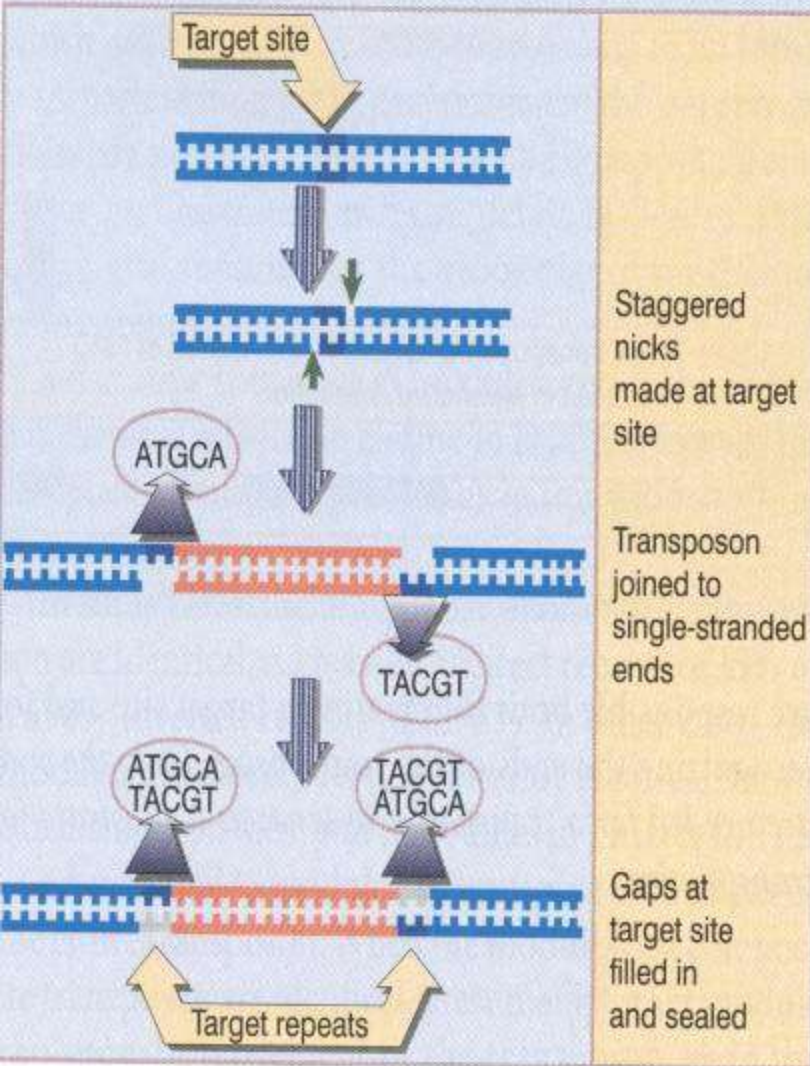
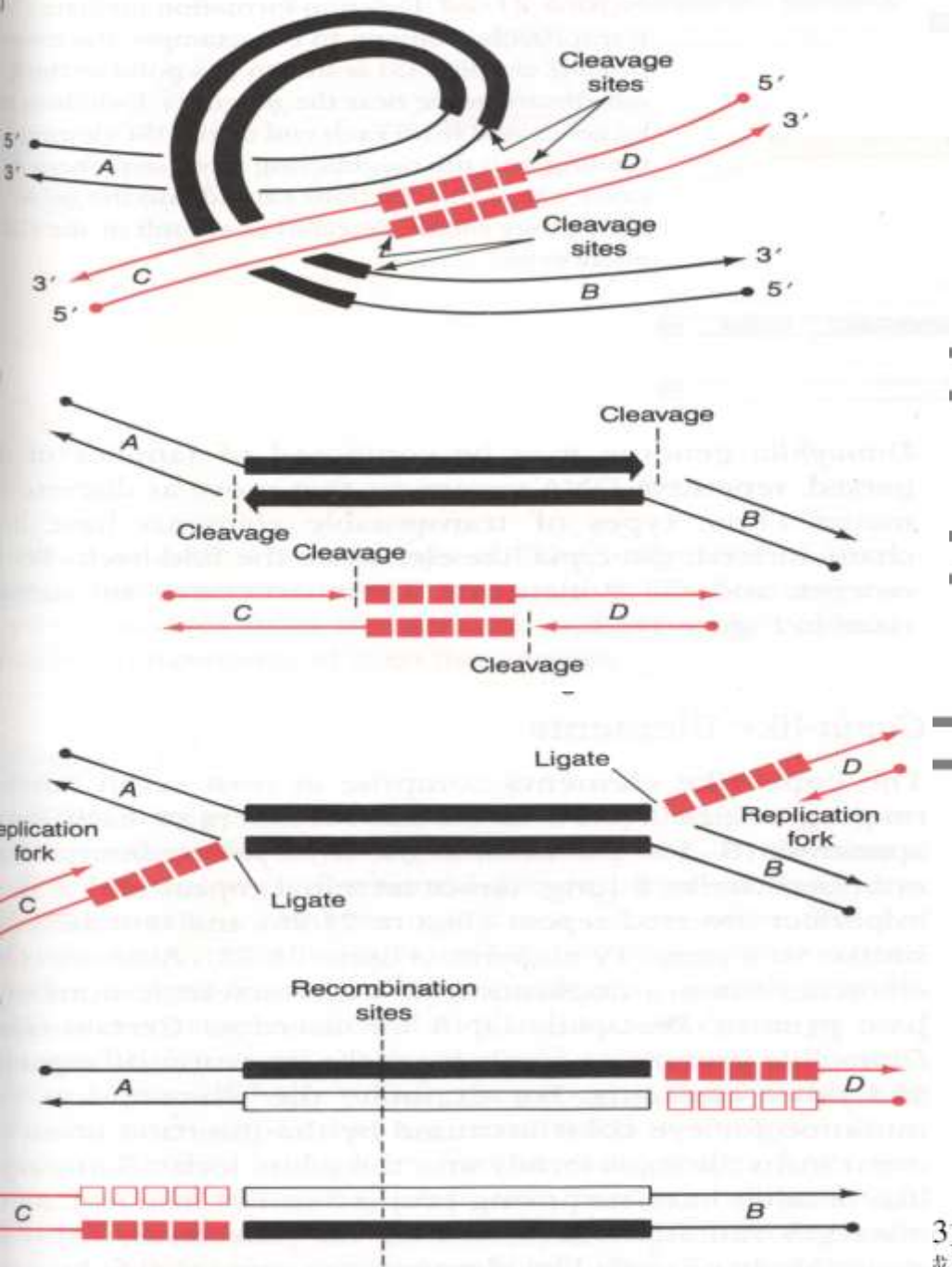
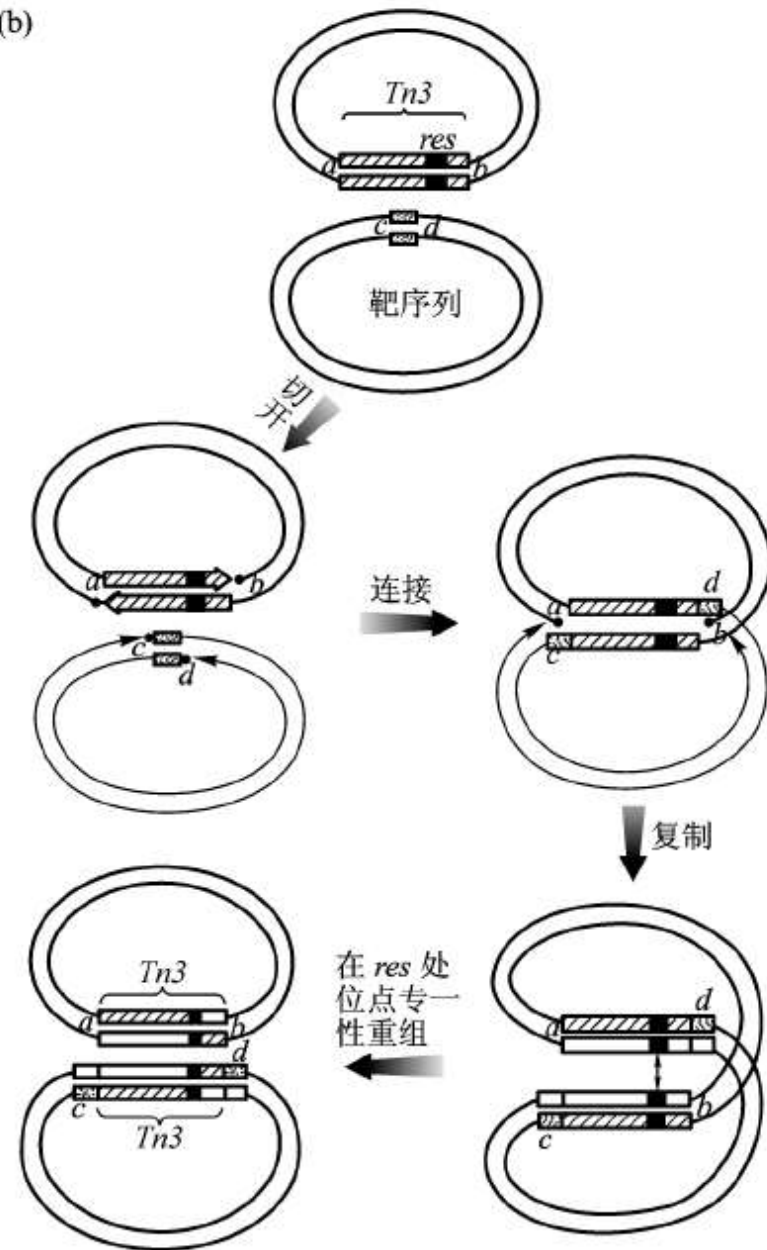


Figure 16.6 Replicative transposition creates a copy of the transposon, which inserts at a recipient site. The donor site remains unchanged, so both donor and recipient have a copy of the transposon.

返回



(b)



返回

3复制型转座模型
成机制 (b) Tn3转座模型

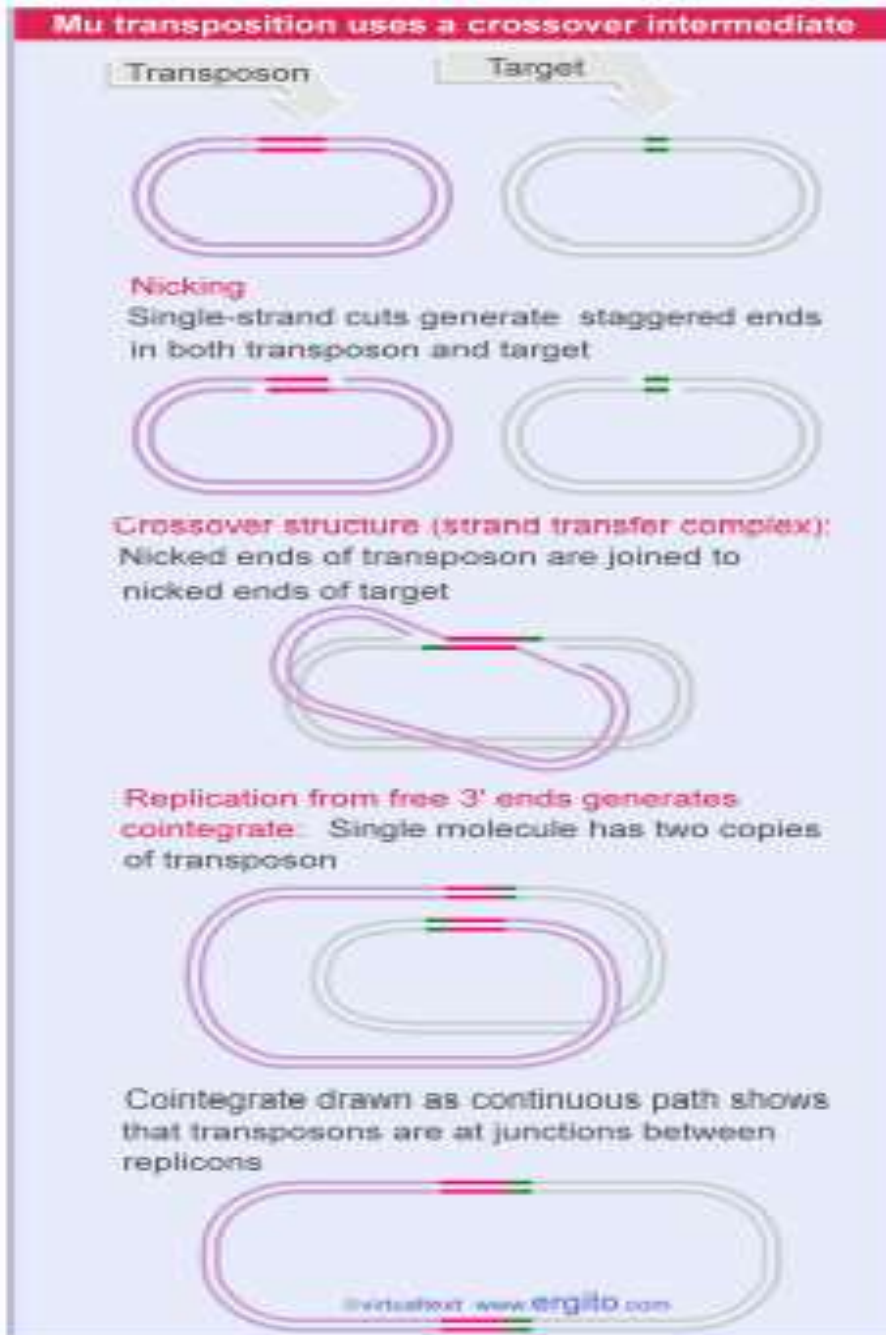


Figure 16.14 Mu transposition generates a crossover structure, which is converted by replication into a cointegrate.

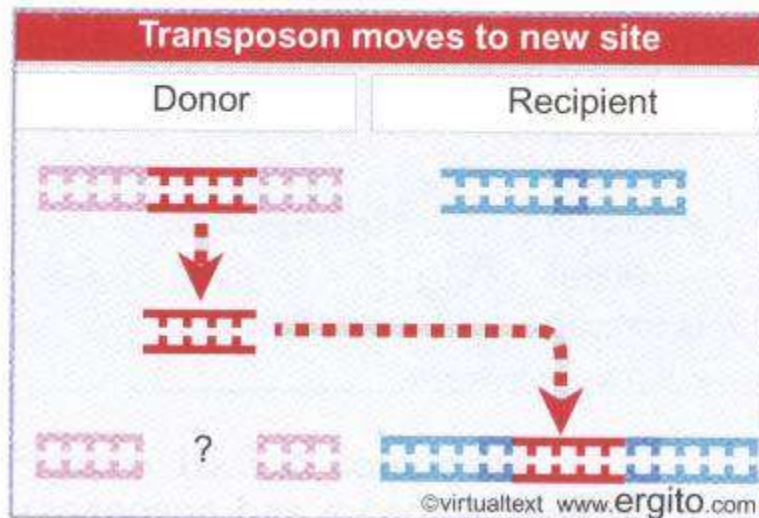


Figure 16.7 Nonreplicative transposition allows a transposon to move as a physical entity from a donor to a recipient site. This leaves a break at the donor site, which is lethal unless it can be repaired.

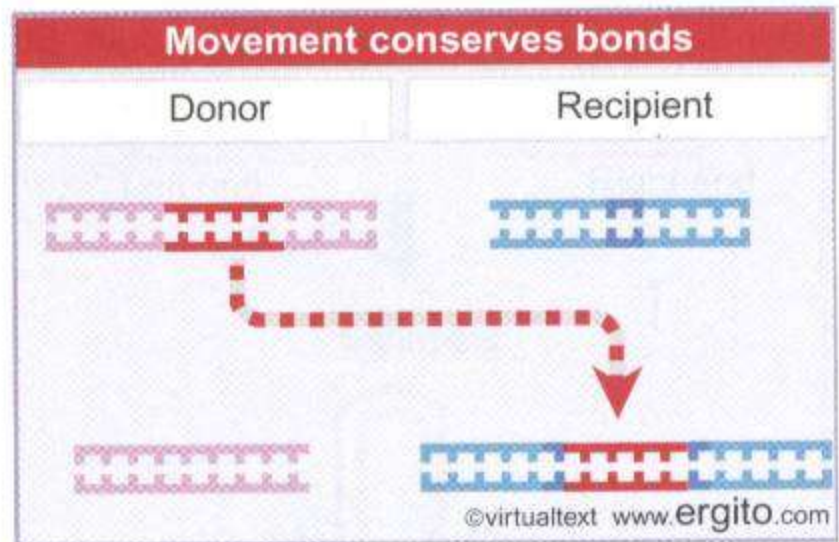
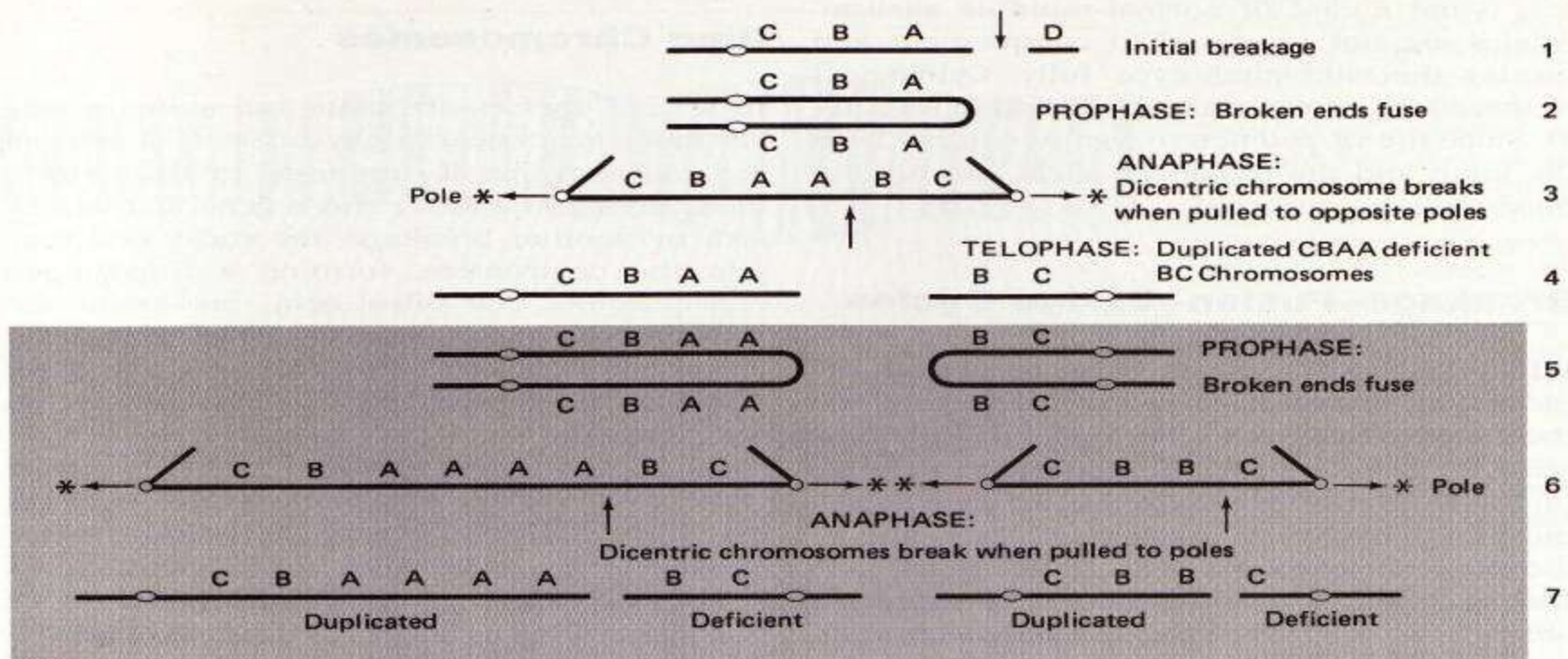
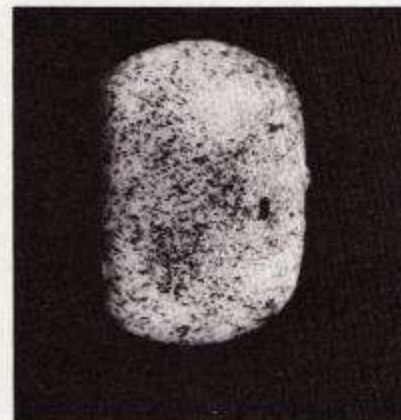
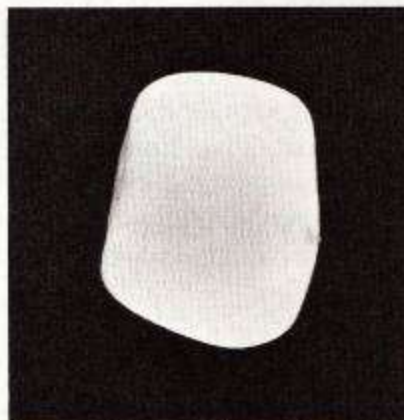


Figure 16.8 Conservative transposition involves direct movement with no loss of nucleotide bonds; compare with lambda integration and excision.



A

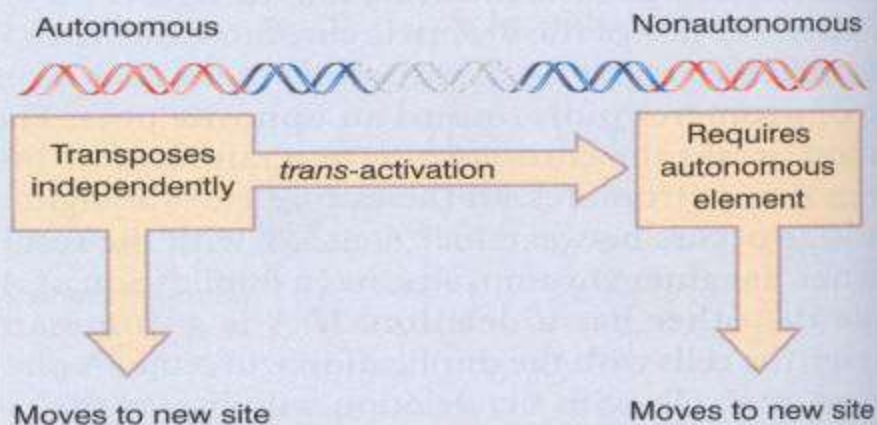


返回

返回1

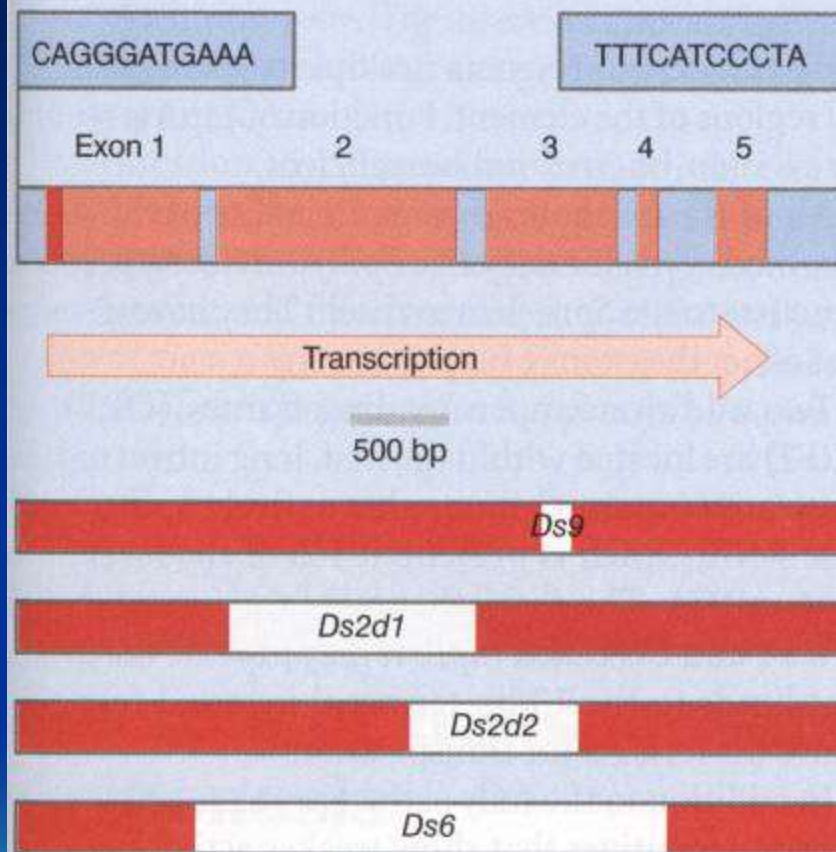


Figure 15.22 Each controlling element family has both autonomous and nonautonomous members. Autonomous elements are capable of transposition. Nonautonomous elements are deficient in transposition. Pairs of autonomous and nonautonomous elements can be classified in >4 families.



<i>Ac</i> (activator) <i>Mp</i> (modulator)	<i>Ds</i> (dissociation)
<i>Spm</i> (suppressor-mutator) <i>En</i> (enhancer)	<i>dSpm</i> (defective <i>Spm</i>) <i>I</i> (inhibitor)
Dotted	Unnamed
<i>Mu</i> (mutator)	Not known

Figure 15.23 The *Ac* element has two open reading frames; *Ds* elements have internal deletions.



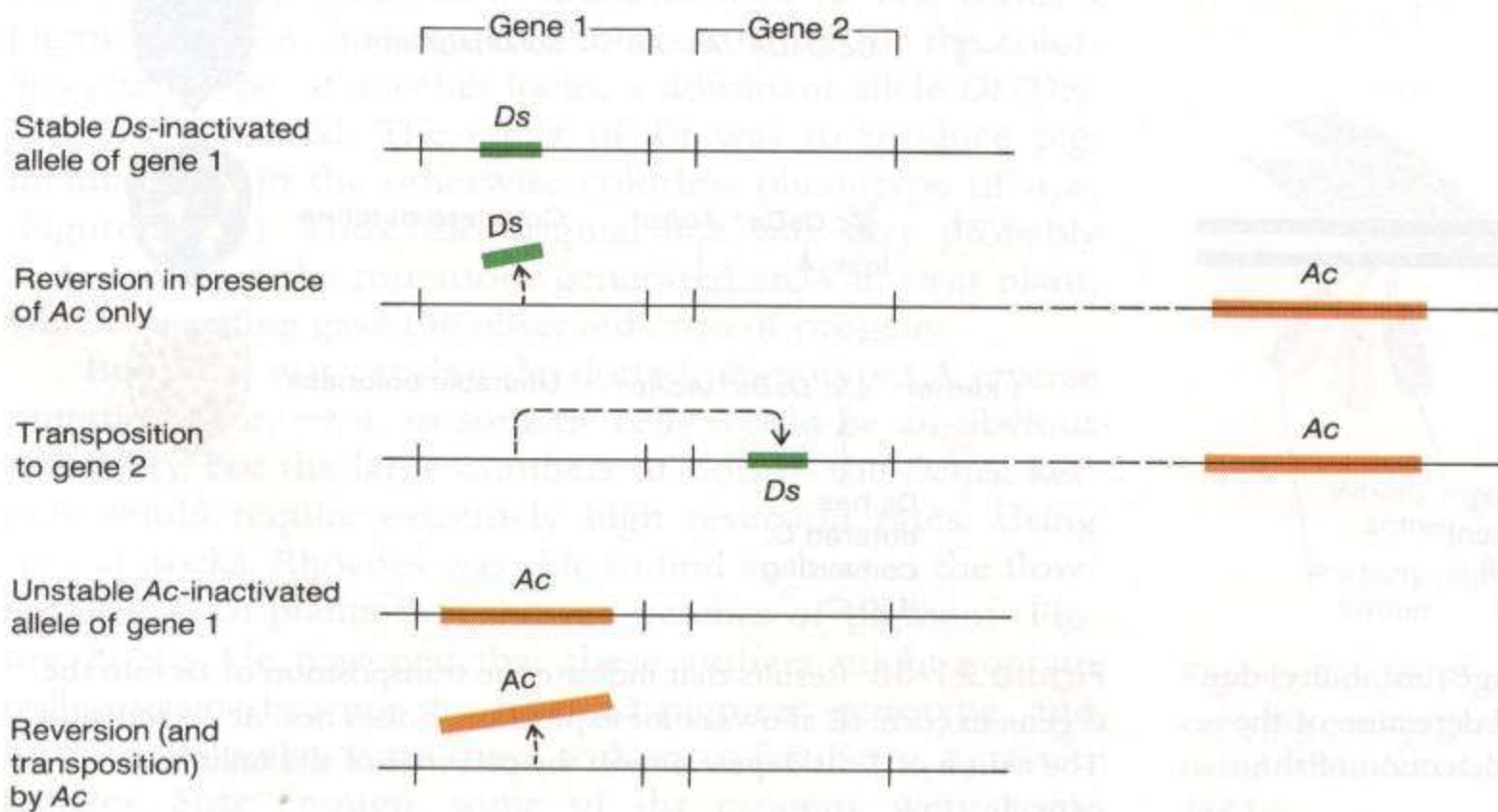




Figure 15.24 *Spm/En* has two genes. *tnpA* consists of 11 exons that are transcribed into a spliced 2500 base mRNA. *tnpB* may consist of a 6000 base mRNA containing ORF1 + ORF2.

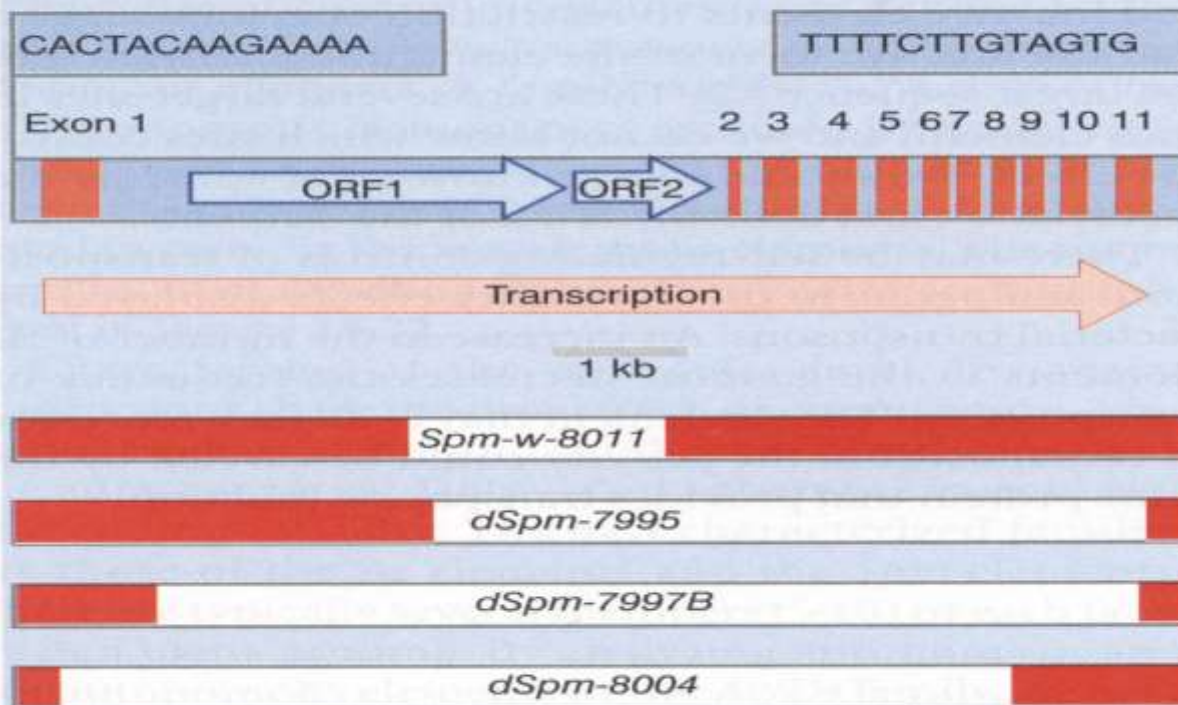




Figure 15.25 Hybrid dysgenesis is asymmetrical; it is induced by P male x M female crosses, but not by M male x P female crosses.

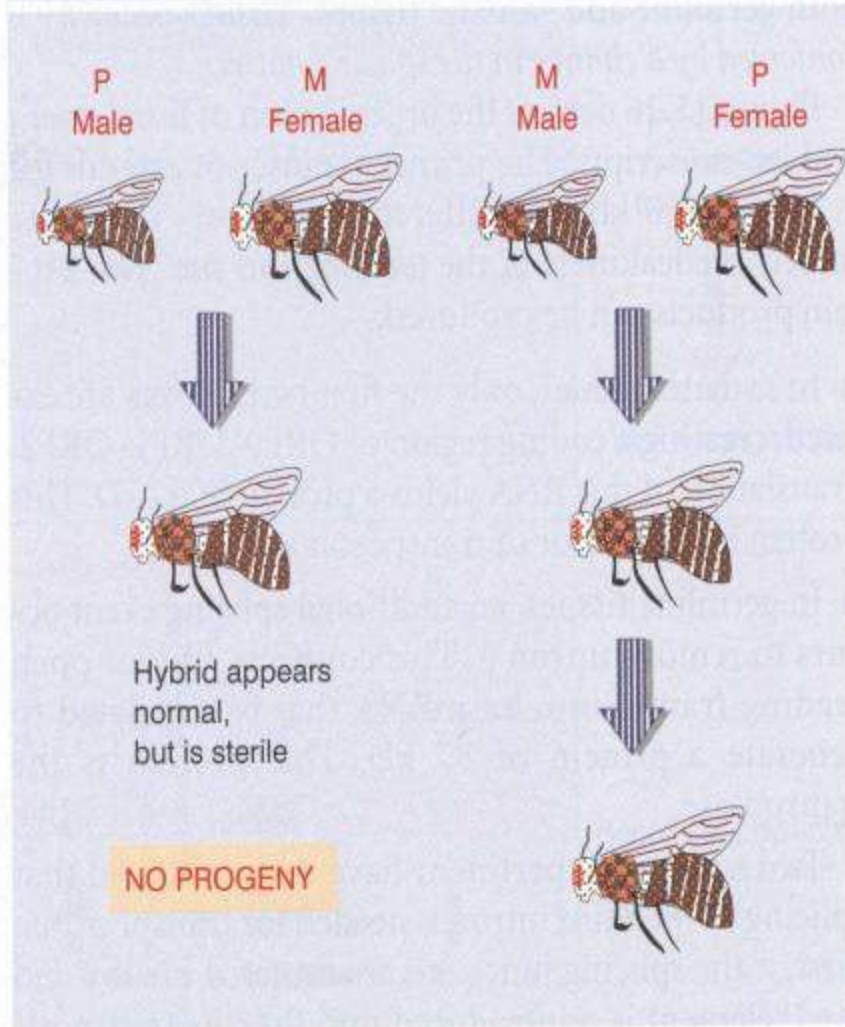
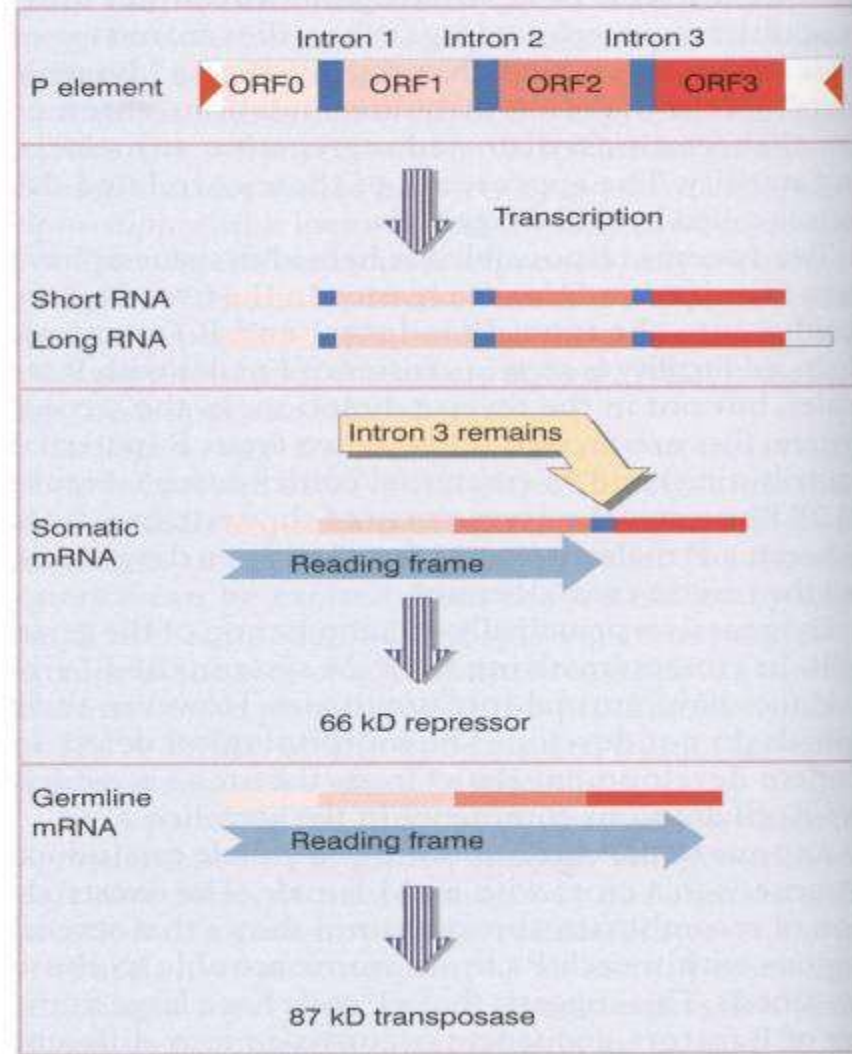


Figure 15.26 The P element has four exons. The first three are spliced together in somatic expression; all four are spliced together in germline expression.



[返回](#)

[返回](#)

[返回2](#)

Figure 15.27 Hybrid dysgenesis is determined by the interactions between P elements in the genome and 66 kD repressor in the cytotype.

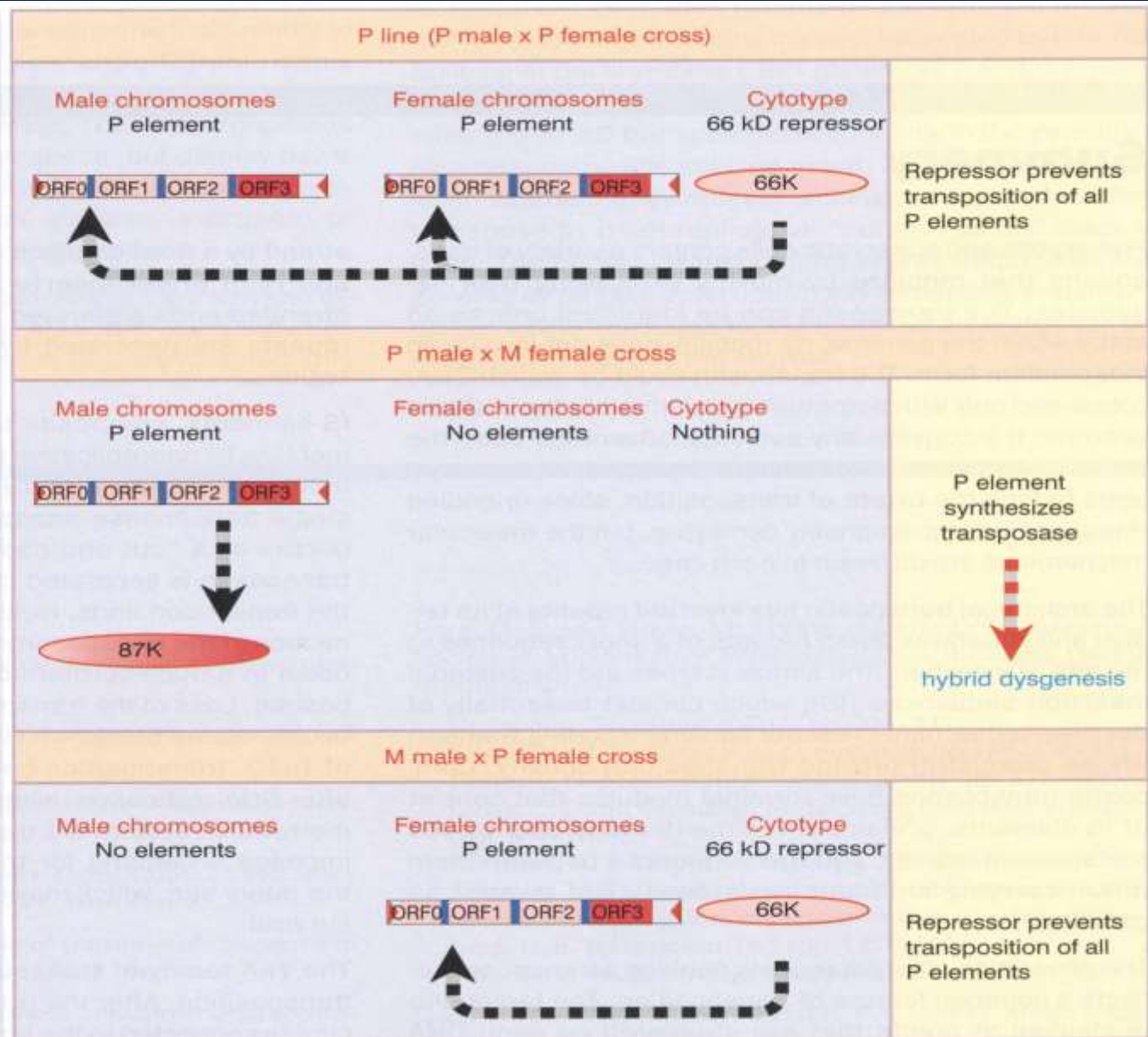


Figure 16.1 Overview: the reproductive cycles of retroviruses and retroposons involve alternation of reverse transcription from RNA to DNA with transcription from DNA to RNA. Only retroviruses can generate infectious particles. Retroposons are confined to an intracellular cycle.

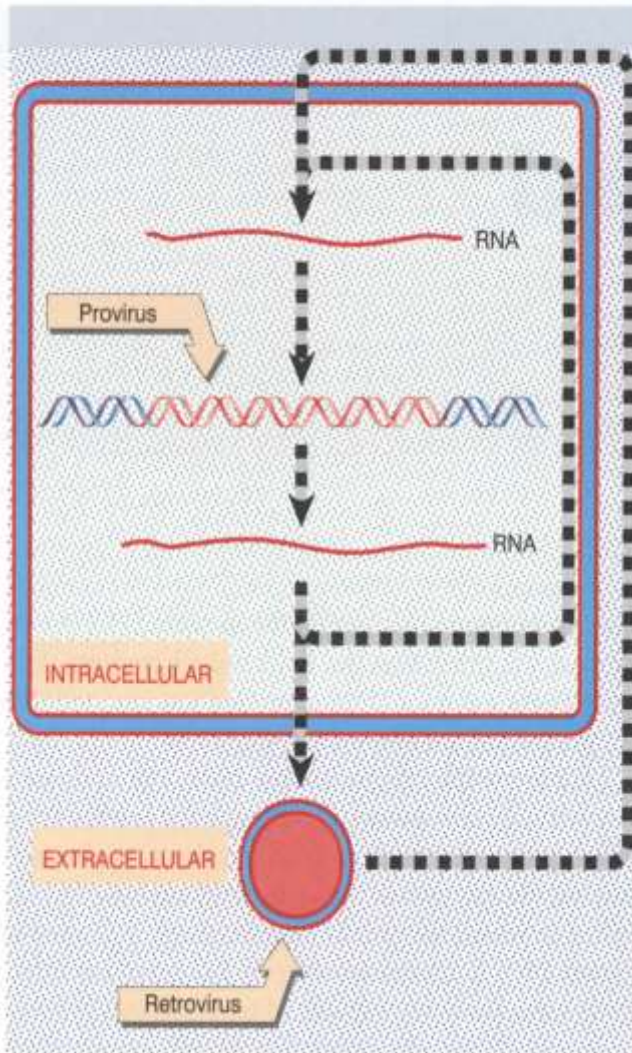
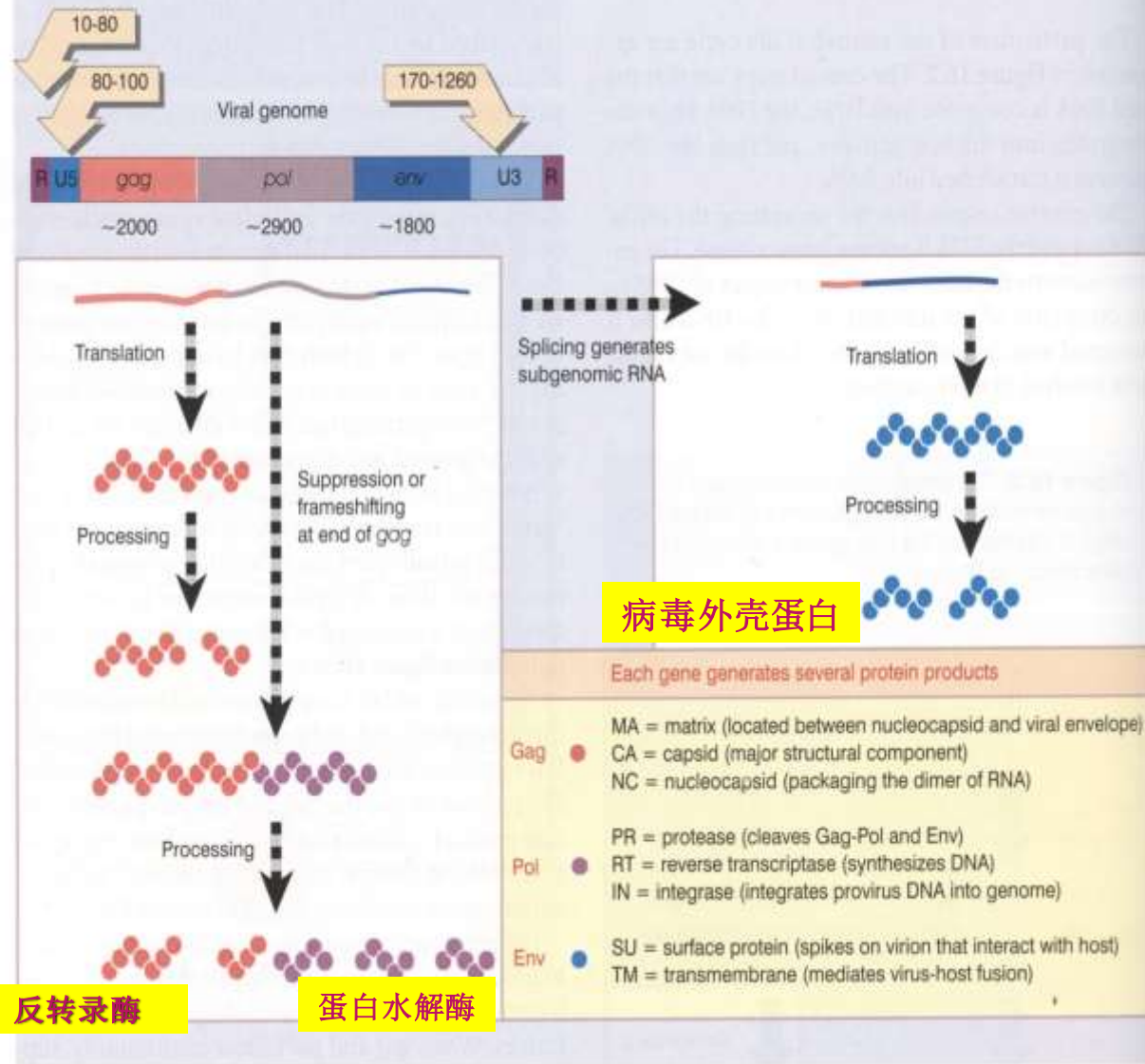


Figure 16.3 The 'genes' of the retrovirus are expressed as polypeptides that are processed into individual products.

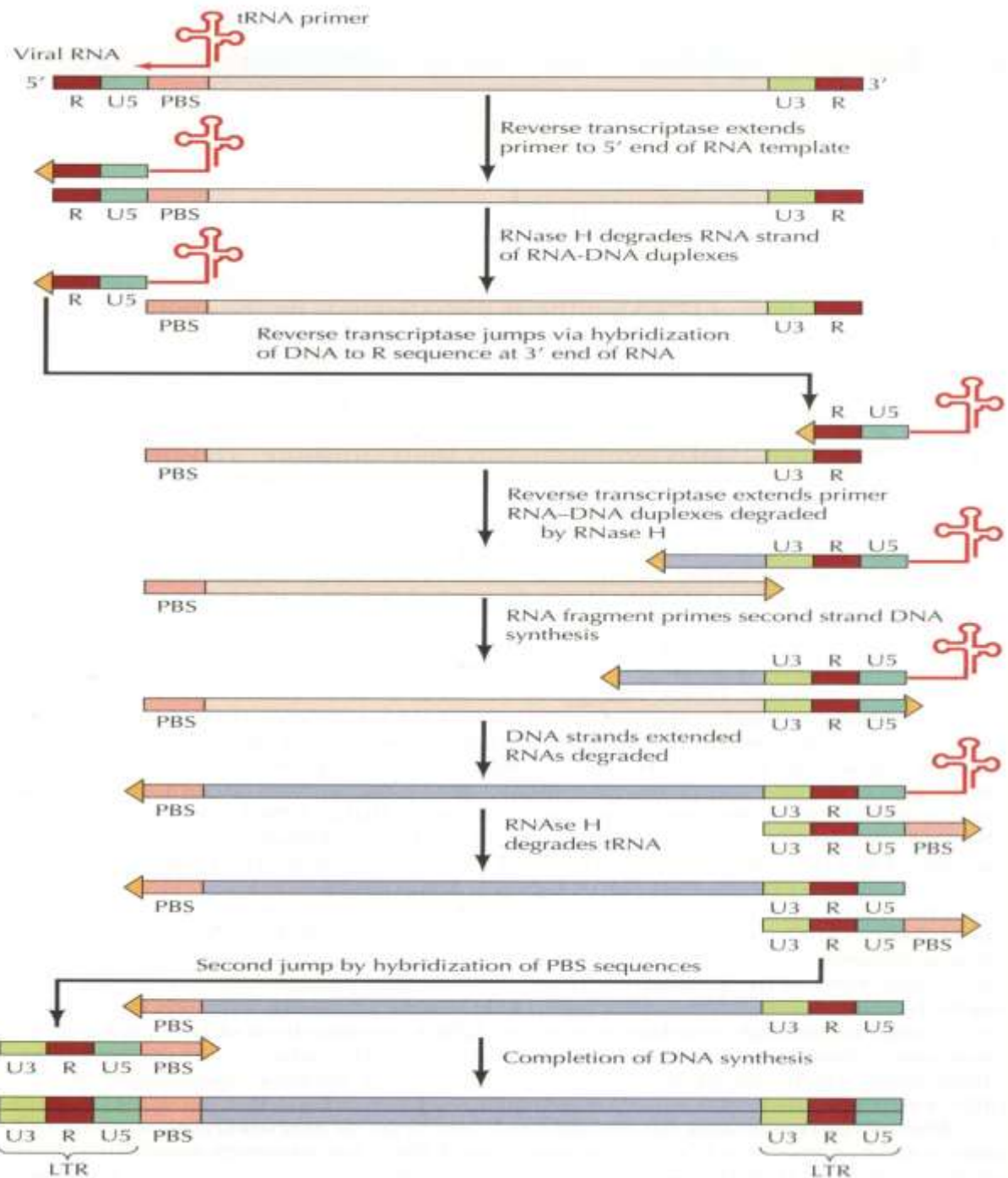


[返回](#)

[返回](#)

Figure 5.47

Generation of LTRs during reverse transcription LTRs consist of three sequence elements: a short repeat sequence (R) of about 20 nucleotides that is present at both ends of the viral RNA; a sequence unique to the 5' end of viral RNA (U5); and a sequence unique to the 3' end of viral RNA (U3). Repeats of these sequences are generated during DNA synthesis as reverse transcriptase jumps twice between the ends of its template. Synthesis is initiated using a tRNA primer bound to a primer-binding site (PBS) adjacent to U5 at the 5' end of the viral RNA. The polymerase copies R, and the RNA strand of the RNA-DNA hybrid is then degraded by RNase H. The polymerase then jumps to the 3' end of the viral RNA in order to synthesize a complete DNA strand complementary to the RNA template. The polymerase jumps again during synthesis of the second strand of DNA, which is also initiated by a primer bound close to the 5' end of its template. The result of these jumps is the formation of LTRs that contain U3-R-U5 sequences.



[返回](#)

Figure 16.11 Ty elements terminate in short direct repeats and are transcribed into two overlapping RNAs. They have two reading frames, with sequences related to the retroviral *gag* and *pol* genes.

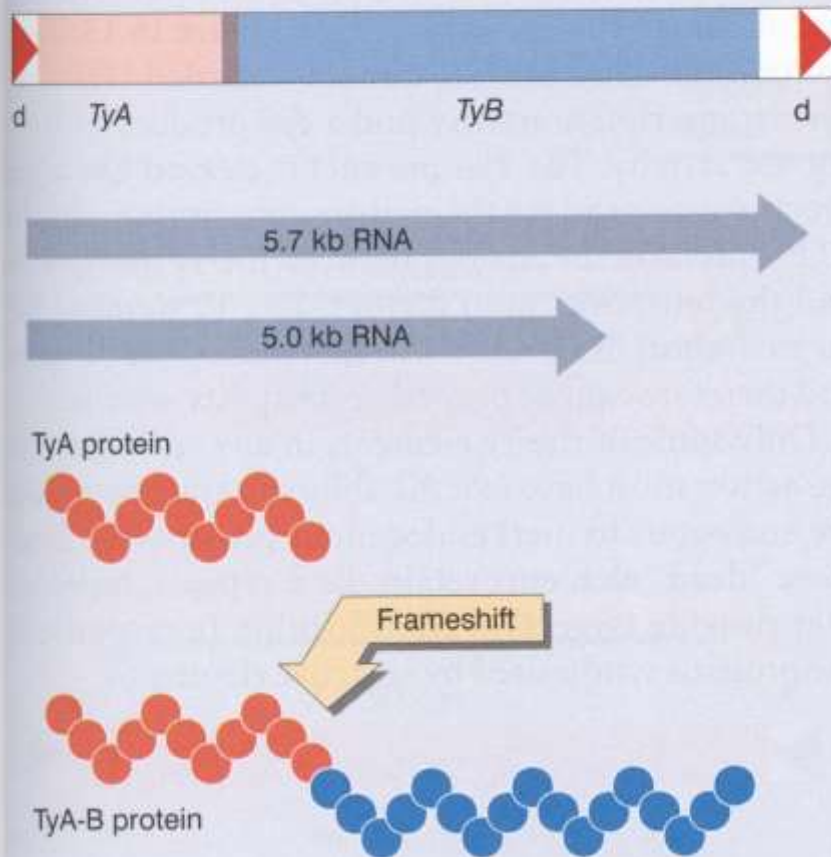


Figure 16.12 A unique Ty element, engineered to contain an intron, transposes to give copies that lack the intron. The copies possess identical terminal repeats, generated from one of the termini of the original Ty element.

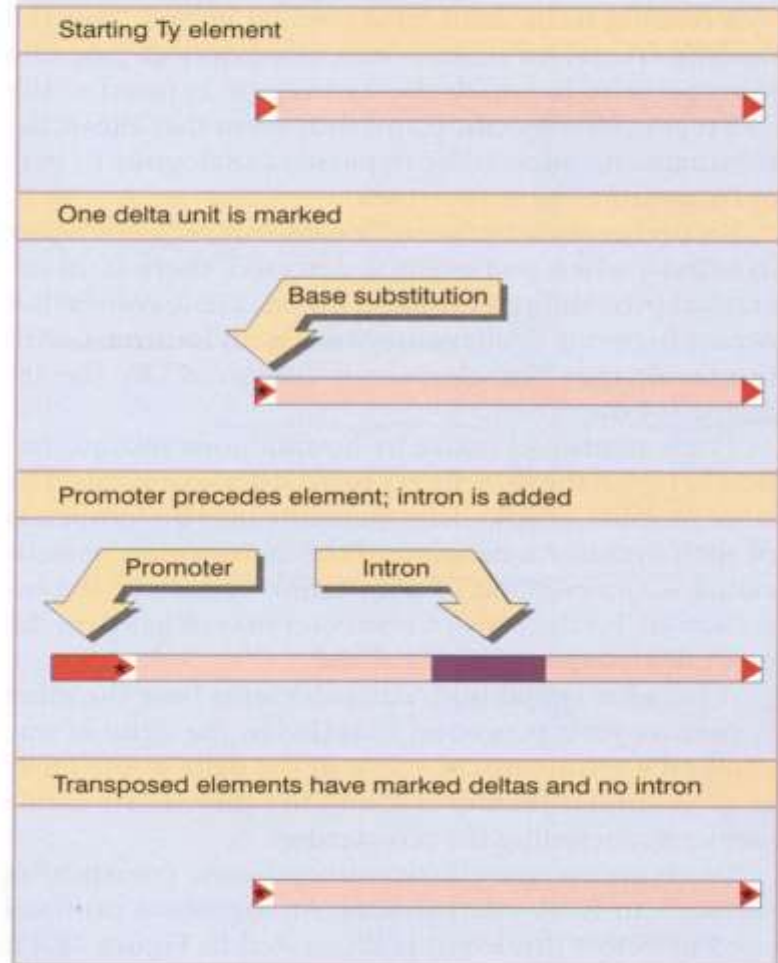




Figure 16.14 Three types of transposable element in *D. melanogaster* have different structures.

