



第二章 基因工程工具酶和载体

湖南师范大学生命科学学院
袁婁洲



本章目录

- 2.1 工具酶
- 2.2 克隆载体
- 2.3 表达载体



第二节 克隆载体

- 质粒
- 噬菌体
- 柯斯质粒
- 人工微小染色体



克隆载体应具备的条件

- 具有针对受体细胞的亲缘性或亲和性（可转移性）
- 具有与特定受体细胞相适应的复制位点整合位点
- 具有较高的外源DNA的装载能力
- 具有多种单一的核酸内切酶识别切割位点
- 具有合适的筛选标记



一、质粒

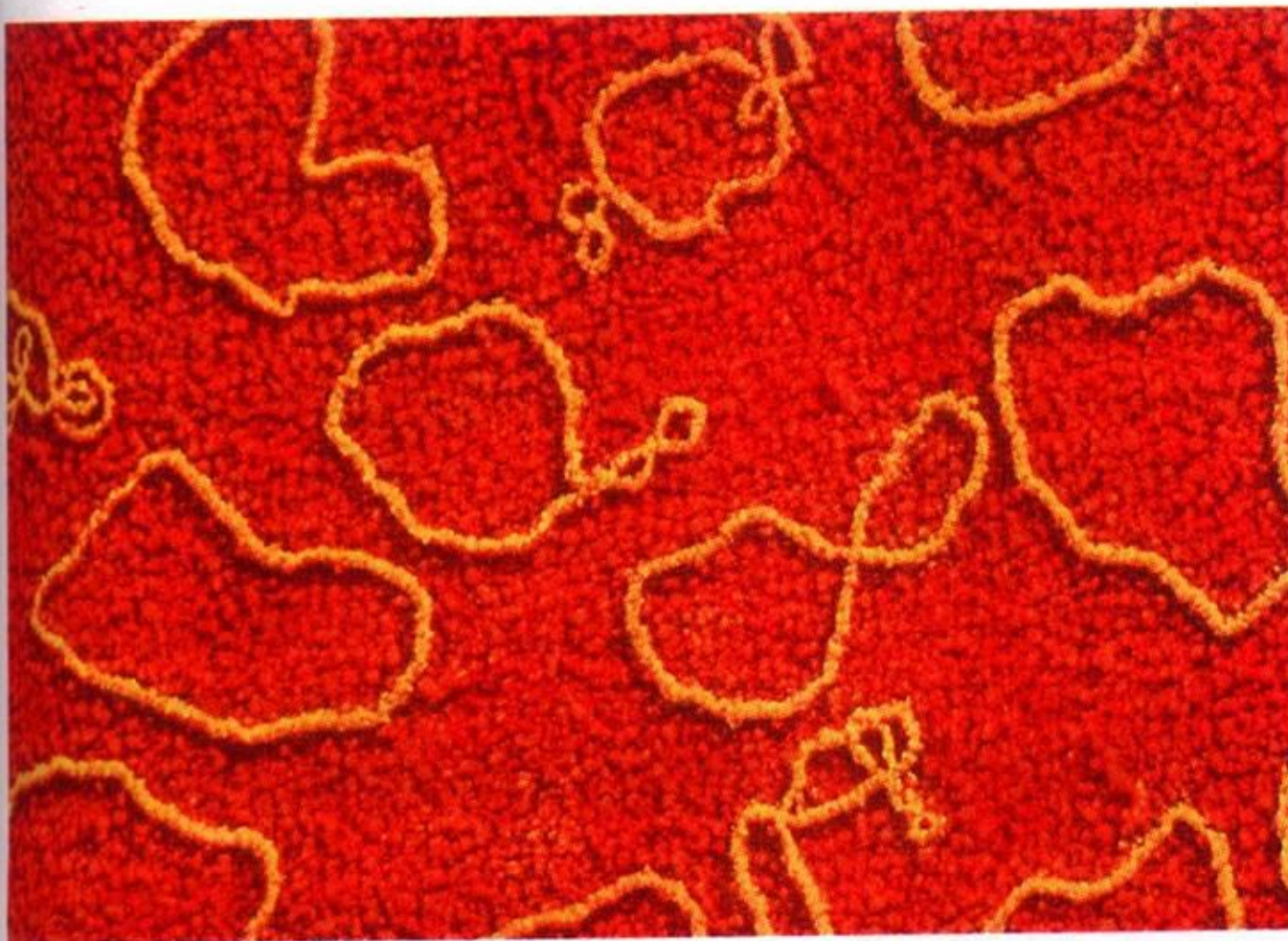
- 细菌染色体外的遗传因子，闭合环状双链DNA
- 大小：1-200kb
- 能自主复制，但要利用寄主细胞复制染色体的同一组酶系
- 复制分松弛型和严谨型两种

松弛型：10-200 拷贝（加氯霉素扩增）

严谨型：1-10 拷贝

- 有某些基因，如抗药性基因，对寄主的生长是有利的

天然质粒



质粒DNA的存在状态

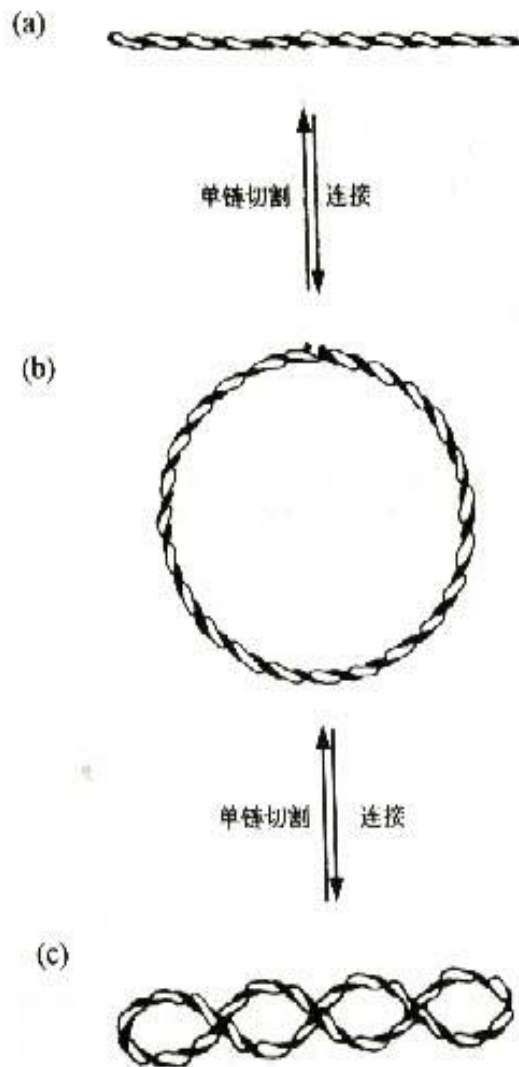


图 4-2 质粒 DNA 的分子构型

(a) 松弛线性的 L 构型; (b) 松弛开环的 OC 构型;
(c) 超螺旋的 SC 构型

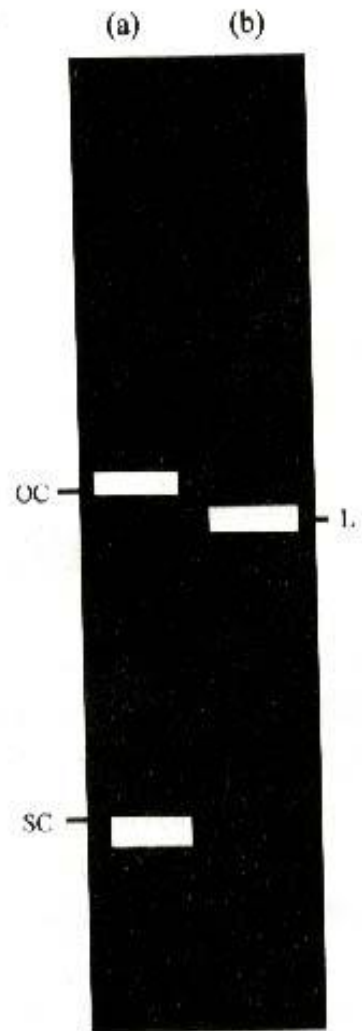


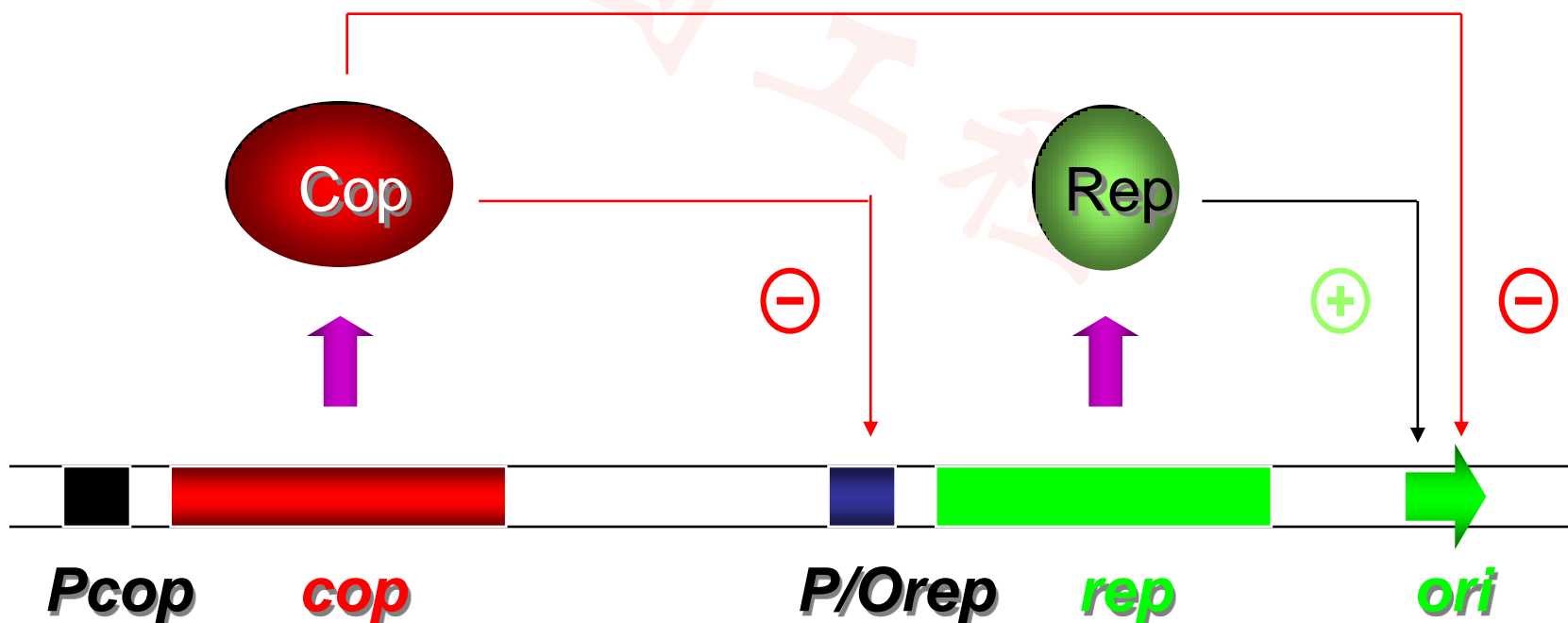
图 4-3 质粒 DNA 琼脂糖凝胶电泳模式图
由于琼脂糖中加有嵌入型染料溴化乙锭,因此,在紫外线照射下 DNA 电泳条带呈橘黄色。
(a) 道中的 SC DNA 走在凝胶的最前沿,OC DNA 则位于凝胶的最后边;(b) 道中的 L DNA 是经核酸内切限制酶切割质粒之后产生的,它在凝胶中的位置介于 OC DNA 和 SC DNA 之间

1、质粒的基本特征

(1) 质粒的自主复制性：

ColE1、pMB1、pSC101等不同启动子

pMB1质粒DNA复制启动控制模型





1、质粒的基本特征

(2) 质粒的不相容性

任何两种含有相似复制子结构的不同质粒，不能同时存在于一个细胞中，这种现象称为质粒的不相容性。

以大肠杆菌的质粒为例：

- ColE1、pMB1 拥有相似的复制子结构，彼此不相容
- pSC101、F、RP4 拥有相似的复制子结构，彼此不相容



1、质粒的基本特征

(3) 携带特殊的遗传标记：抗生素标记基因

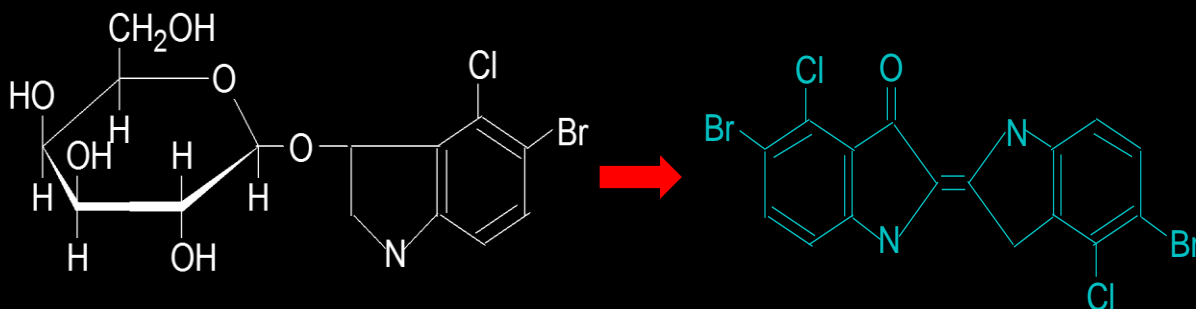
- 1) 氨苄青霉素抗性基因
- 2) 四环素抗性基因
- 3) 氯霉素抗性基因
- 4) 卡那霉素和新霉素抗性基因
- 5) lacZ'基因

Lac Z基因的显色原理:

IPTG: 异丙基- β -D-硫代半乳糖苷
诱导物

X- gal: 5-溴-4氯-3吡啶- β -D-半乳糖苷
底物 (生色剂)

β -半乳糖苷酶分解x-gal 形成蓝色产物

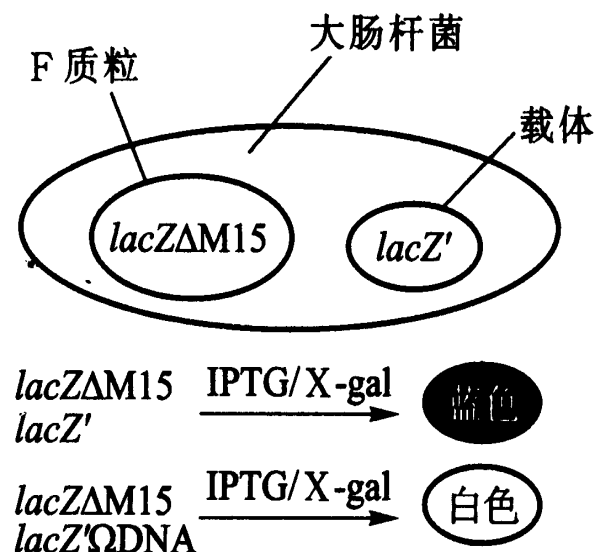
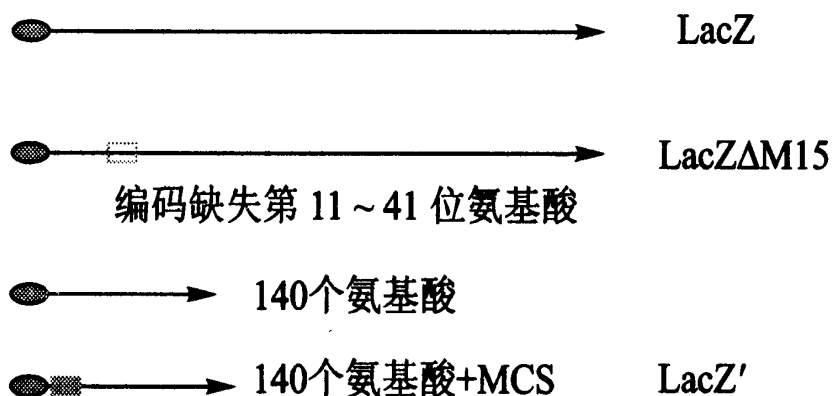


lacZ' 基因 α 互补显色的原理

载体上 lacZ' 基因，产物为正常 β -半乳糖苷酶的N端140aa，也被称为 α 片段

宿主染色体或F因子上带有N端突变的 lacZ Δ M15 的 ω 或 β 片段，能与 α 片段互补显色。

互补宿主有 JM101 JM103, JM105, JM109, NM522 等变异的大肠杆菌)



LacZ基因表达的蓝白斑菌落



空转
化子：
蓝色

重组
子：
白色



2、常用基因工程质粒：pBR322

F Bolivar & R L Rodriguez构建

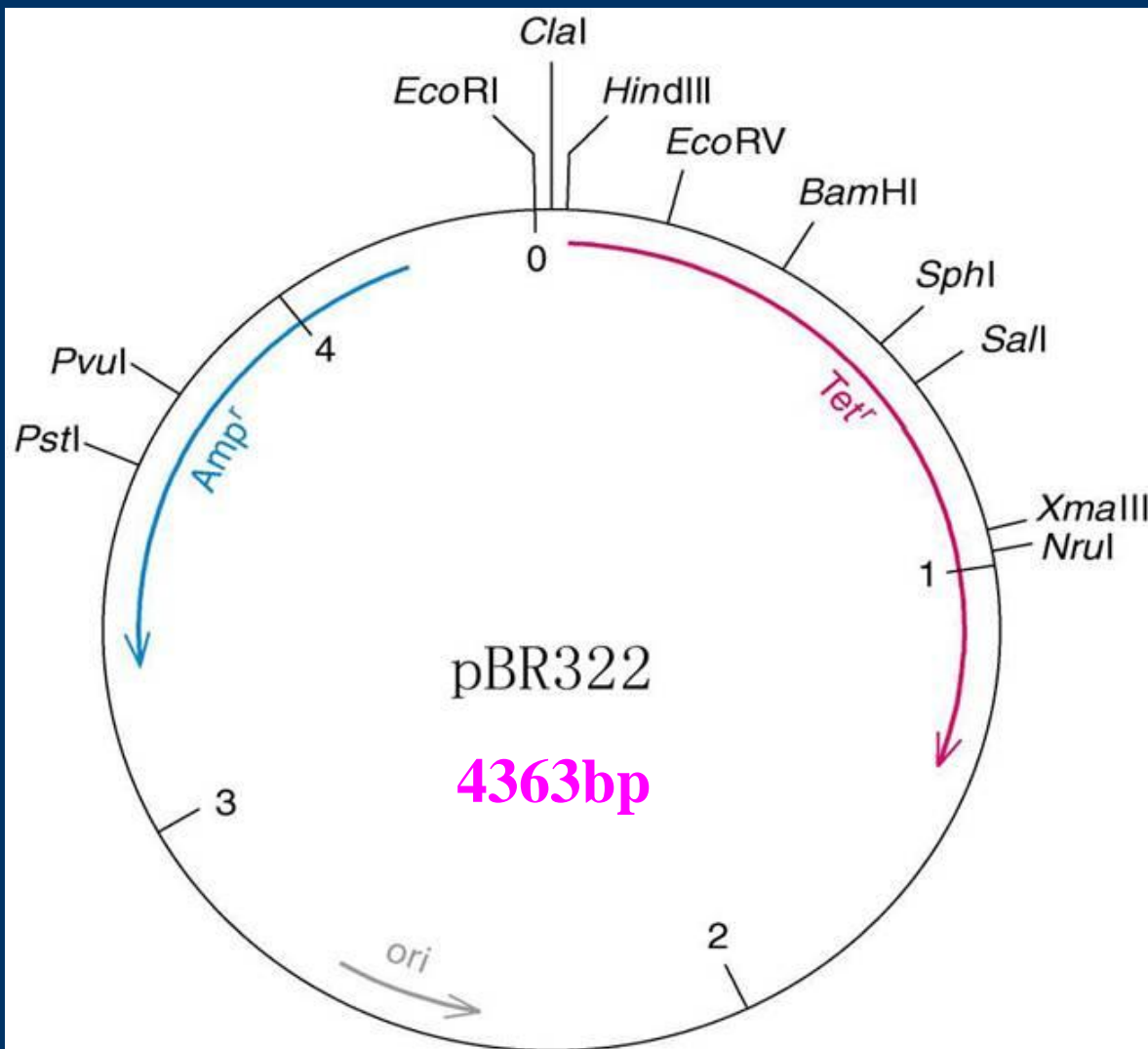
松弛型质粒

Amp^r, Tet^r

多种限制性酶切位点

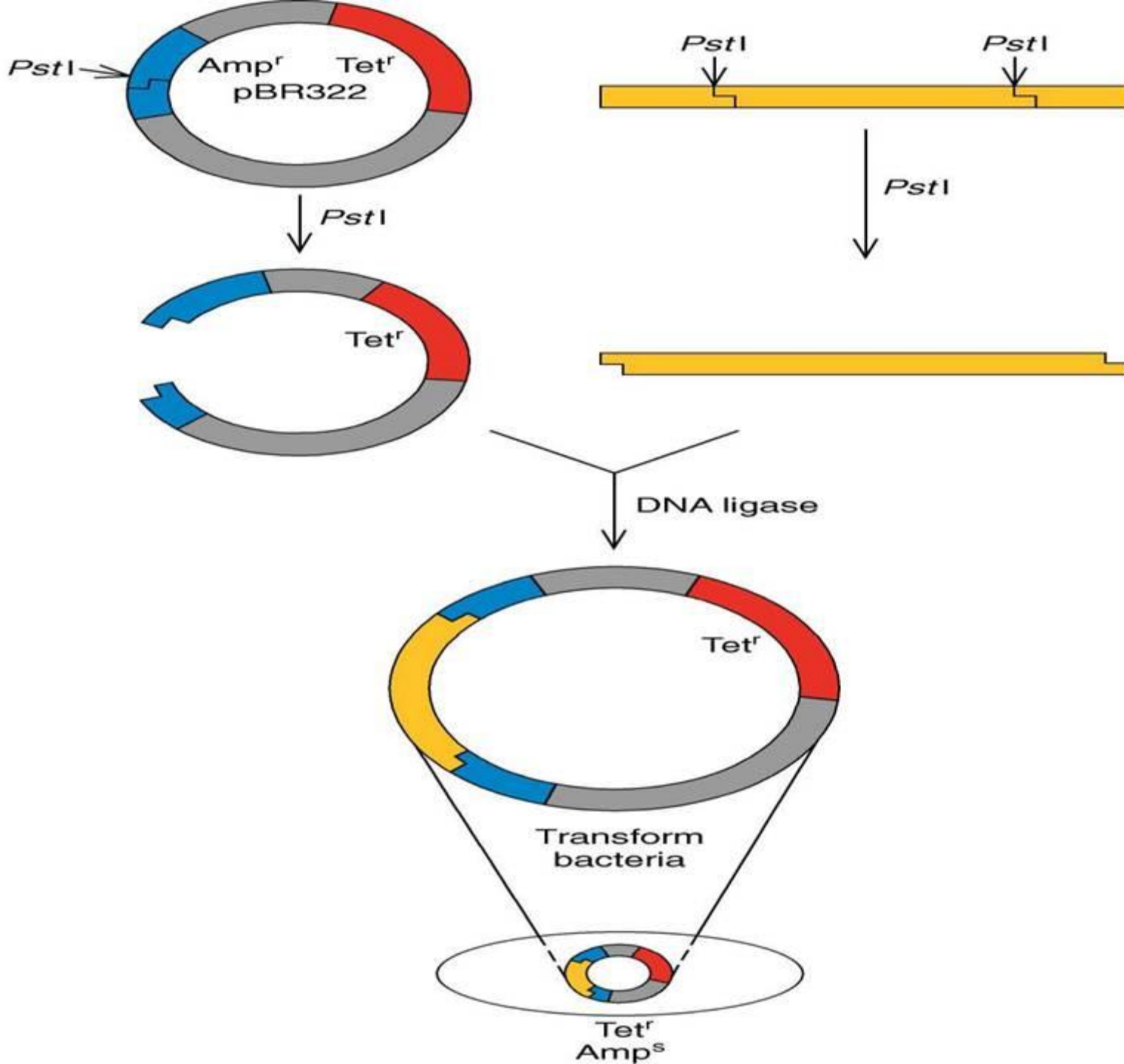
全长 4362 bp

pBR322结构图





pBR322 质粒克隆外源基因的过程



程



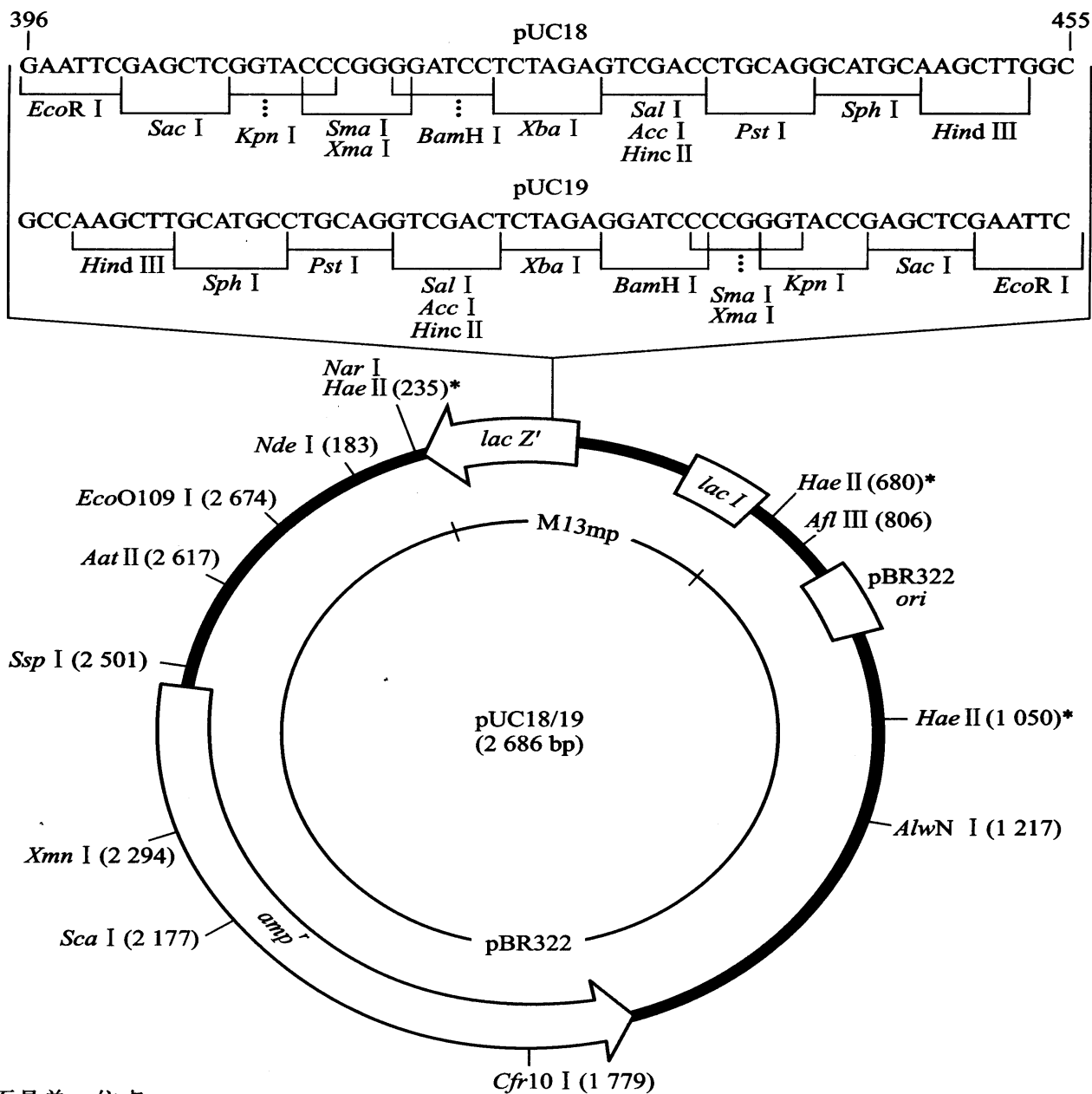
2、常用基因工程质粒： pUC 质粒

- pUC 质粒 是由大肠杆菌 pBR322 质粒 与M13 噬菌体改建而成的双链DNA质粒载体
- pUC 18/pUC 19 全长 2674 bp ， 有 Amp 抗性基因， 带有lac Z'基因
- 一个复制起点*ori*。因在复制起点*ori*区域内缺失 *rop* 基因（改进的pMB1复制子）， 细胞即使生长在无氯霉素环境中也可形成高拷贝数（500–700）。
- 宿主菌携带 β 一半乳糖苷酶C端(ω 片段)序列。重组子菌落是白色，空转化子是蓝色。

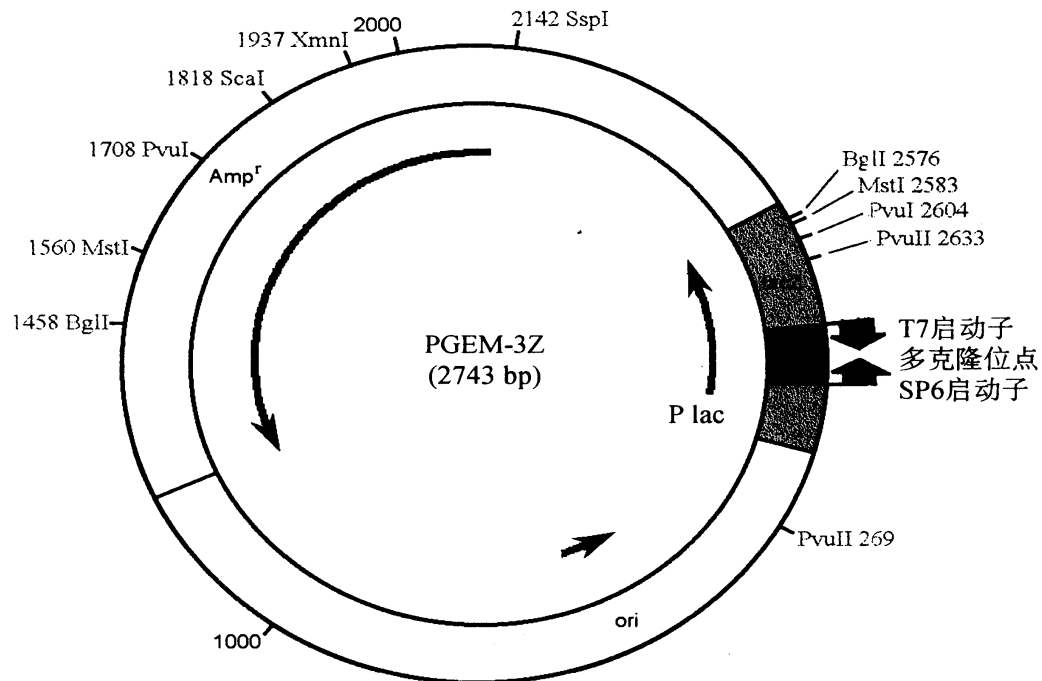


pUC 18/pUC 19 质粒结构

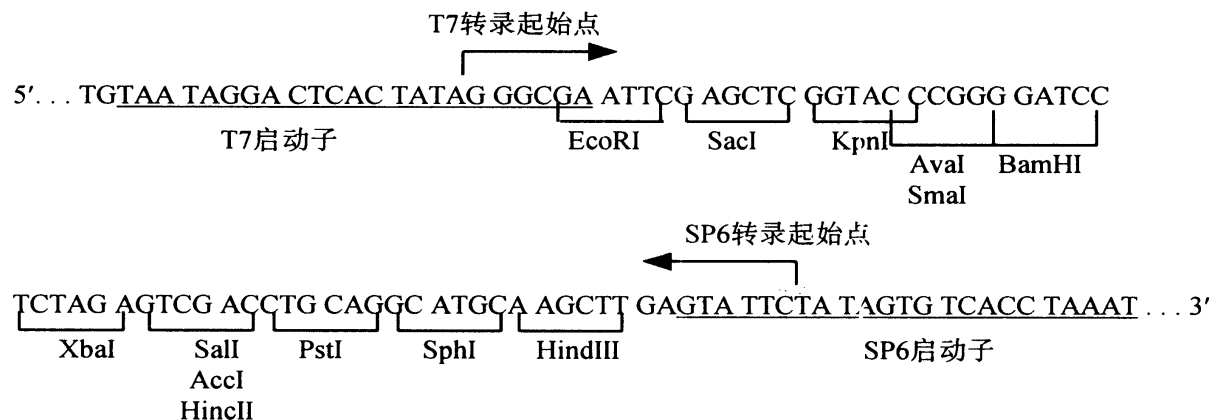
基因工程



* 不是单一位点



多克隆位点
PGEM-3Z





改建质粒的要求

- 尽可能使质粒缩小，以便插入更大的外源DNA片段；
- 增加多种酶切位点；
- 增加标记基因；
- 提高外源DNA的表达水平；
- 增加质粒在受体细胞中的拷贝数



二、 λ 噬菌体

- λ 噬菌体DNA 是线性双链DNA分子, 48.5kb
- 噬菌体 λ 粒子, DNA线性双链, 带有12bp的粘性末端, 称为cos位点。

噬菌体感染细菌以后, 双链DNA分子通过COS位点成环状

- Λ 基因组三个区域:

左侧: 包括外壳蛋白基因, 20kb

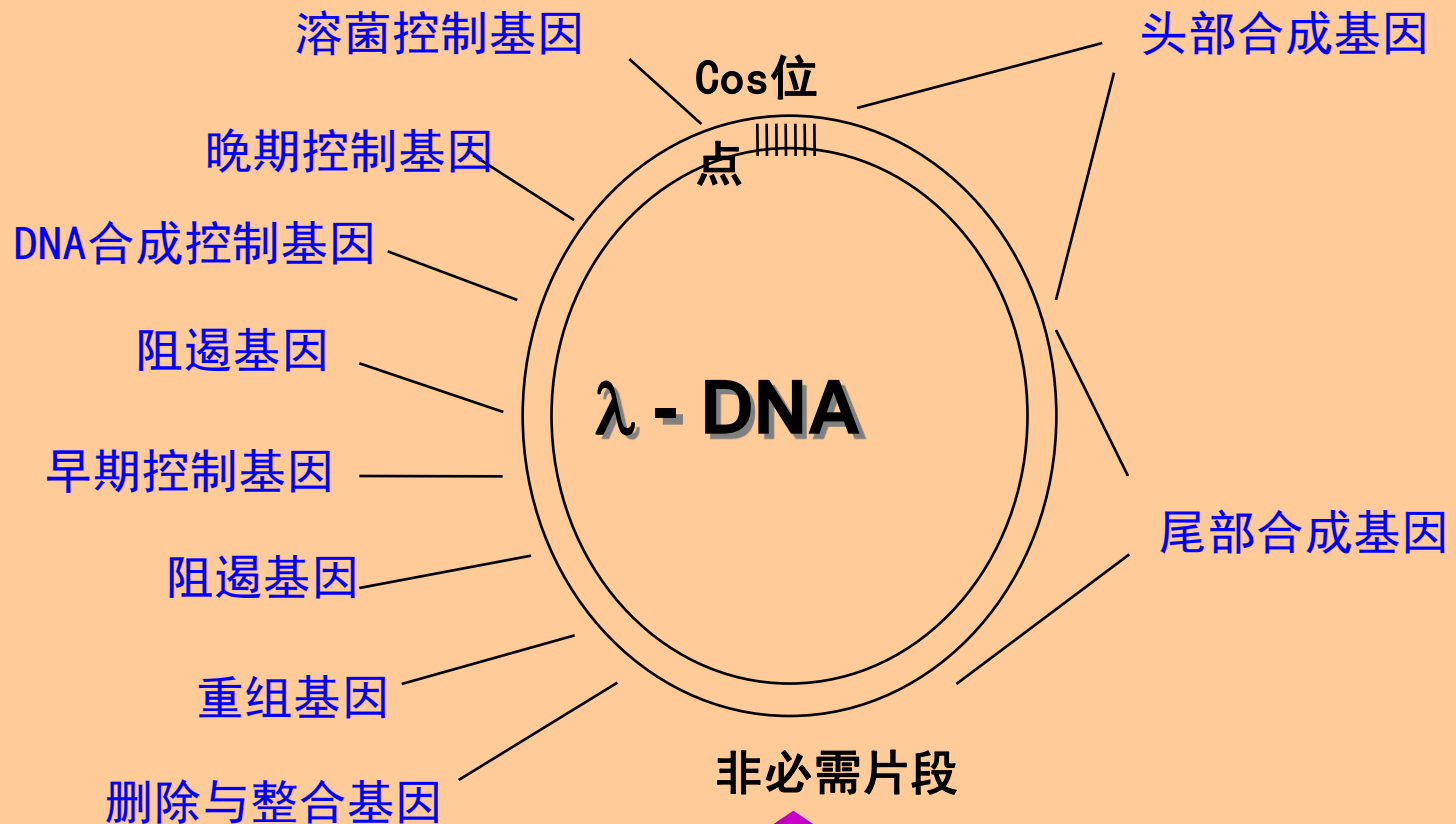
中间区: 18kb

右侧: 复制及裂解基因, 12kb

λ DNA+外源DNA之和必须在39-53kb之间, 所以可插入7-21kb外源片段



λ 噬菌体的基因组结构



COS序列 3' **CCCGCCGCTGGA** 5'
5' **TCCAGCGGCGGGG** 3' **COS序列**

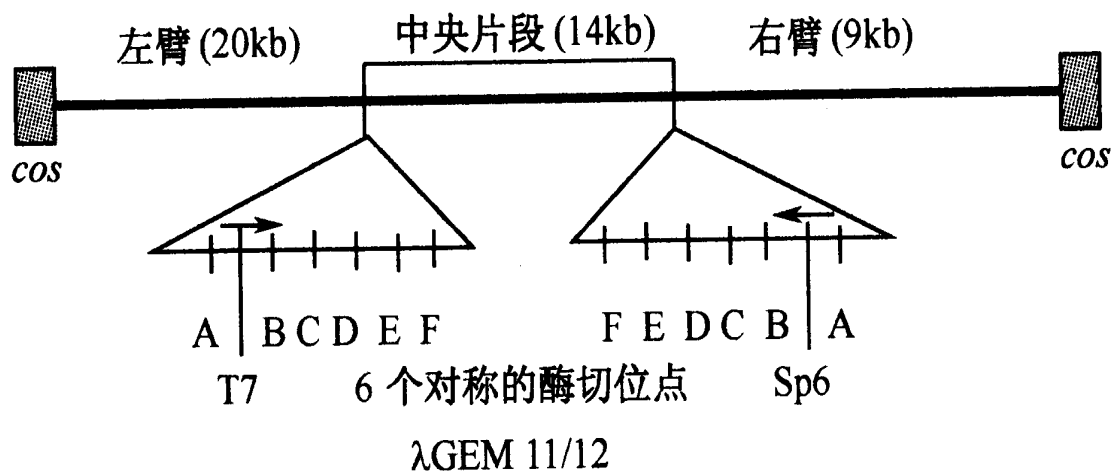
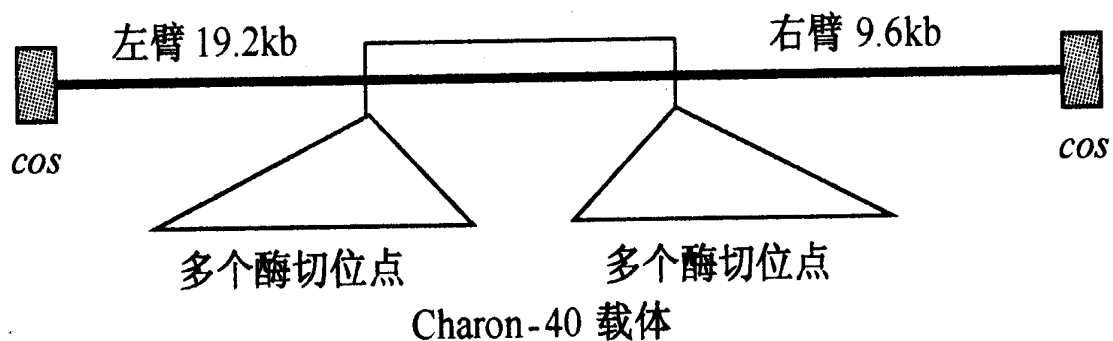


λ -DNA载体的构建：删除重复的酶切位点

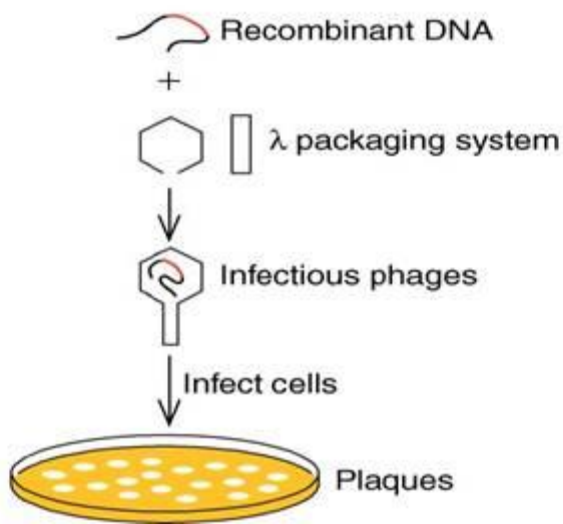
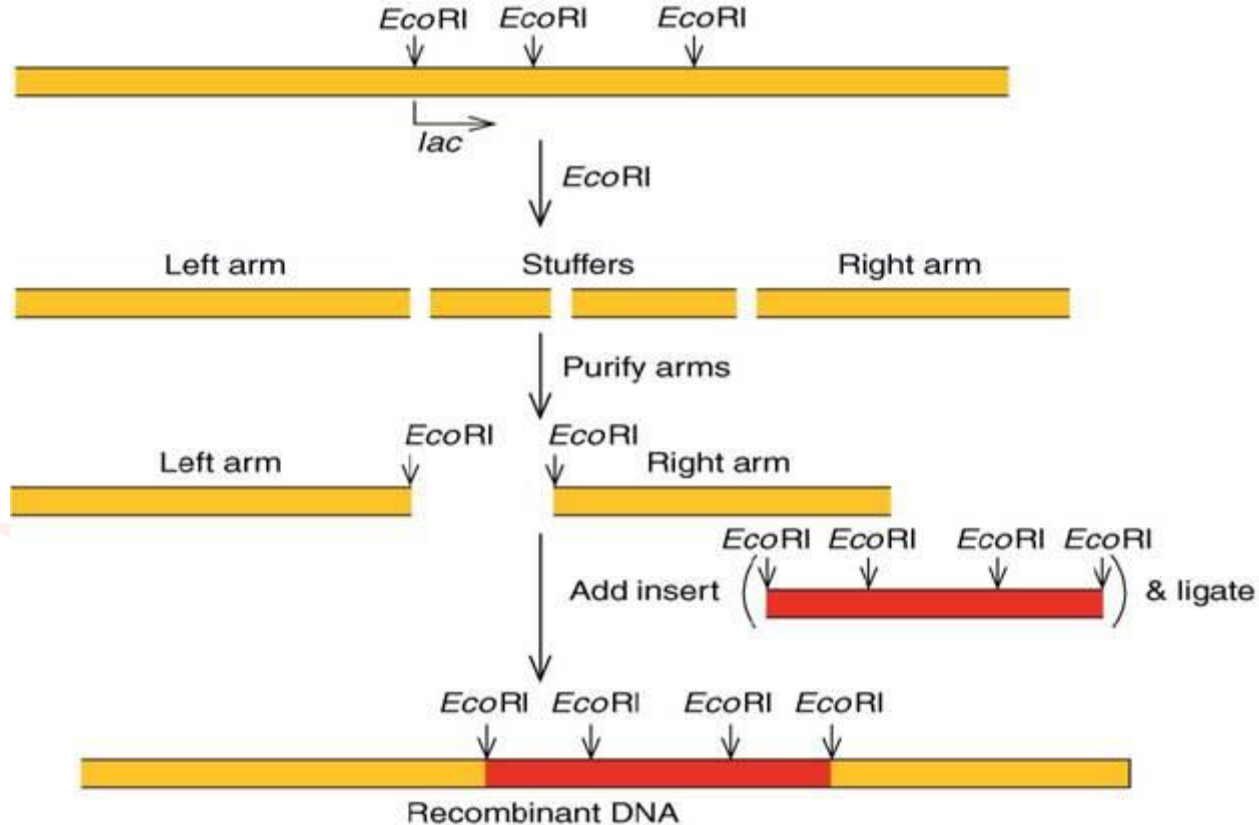
- 野生型的 λ -DNA链上有5个EcoRI位点和7个HindIII位点，不利于重组操作，必须删除至1 - 2个
- 同时，为了便于各种来源的DNA片段的克隆，还需要增加一些单一的酶切位点。如利用酶切和采用定点突变技术去除或增加酶位点
- 增加标记基因lacZ

λ噬菌体置换型载体的结构

(LacO+ 腺病毒 DNA 235bp 重复序列)



利用λ噬菌体载体克隆外源DNA的过程

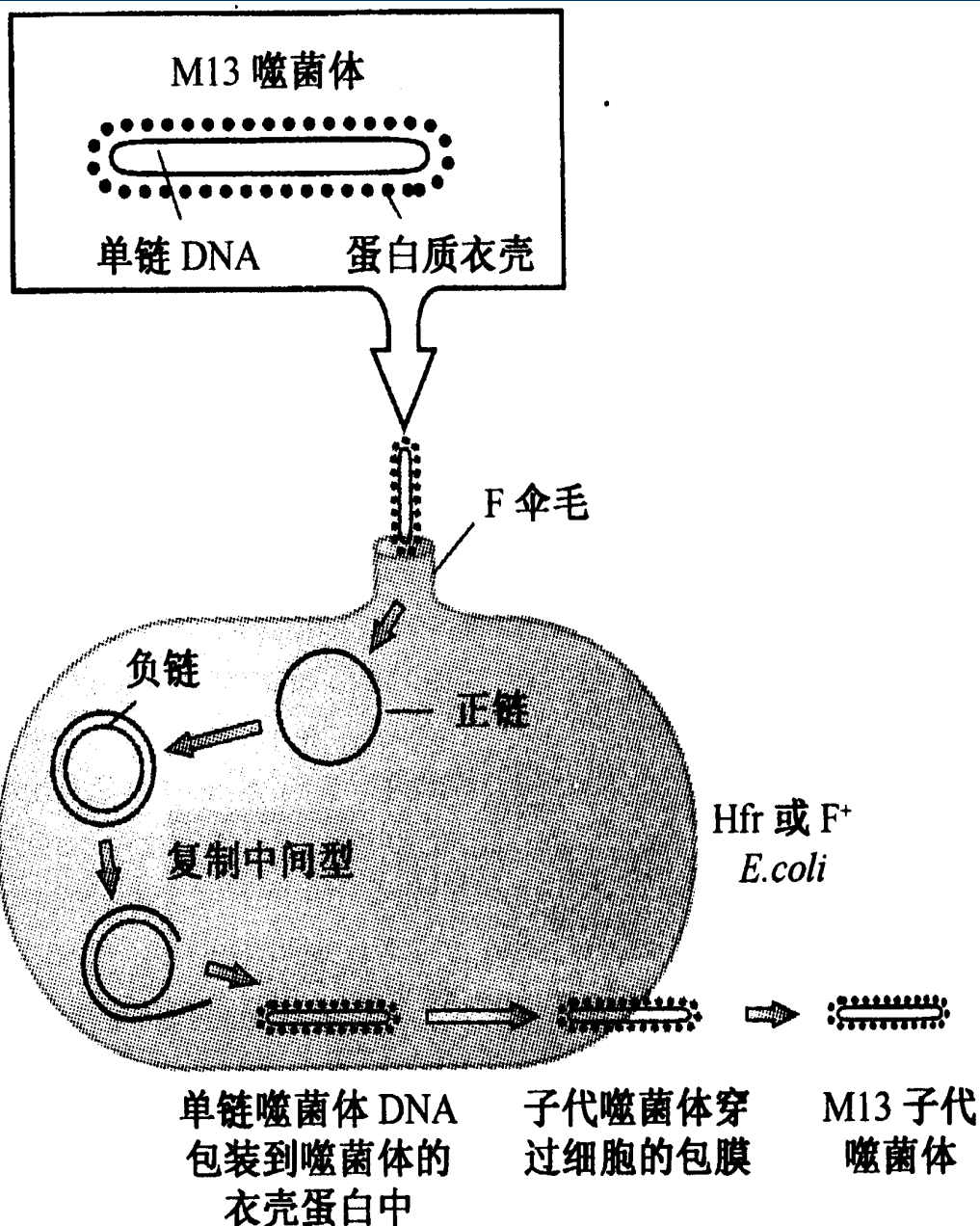




三、M13单链噬菌体

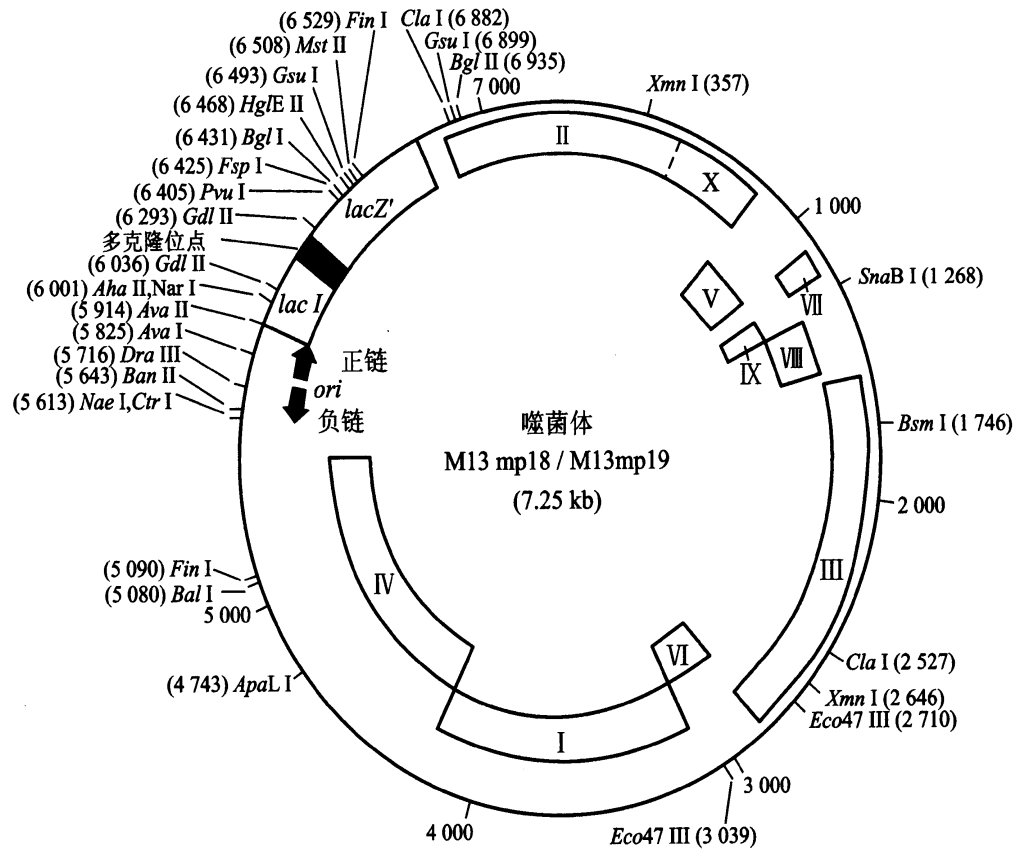
- **M13噬菌体的生物学特性： 生物结构**
- M13噬菌体的外型呈丝状
- M13 噬菌体由外壳包装蛋白和正链DNA组成
- M13 DNA全长6407个核苷酸
- M13 DNA上至少有10个基因
- M13 噬菌体不裂解宿主细胞， 但抑制其生长

M13 噬菌体的生活周期





M13 噬菌体载体的结构



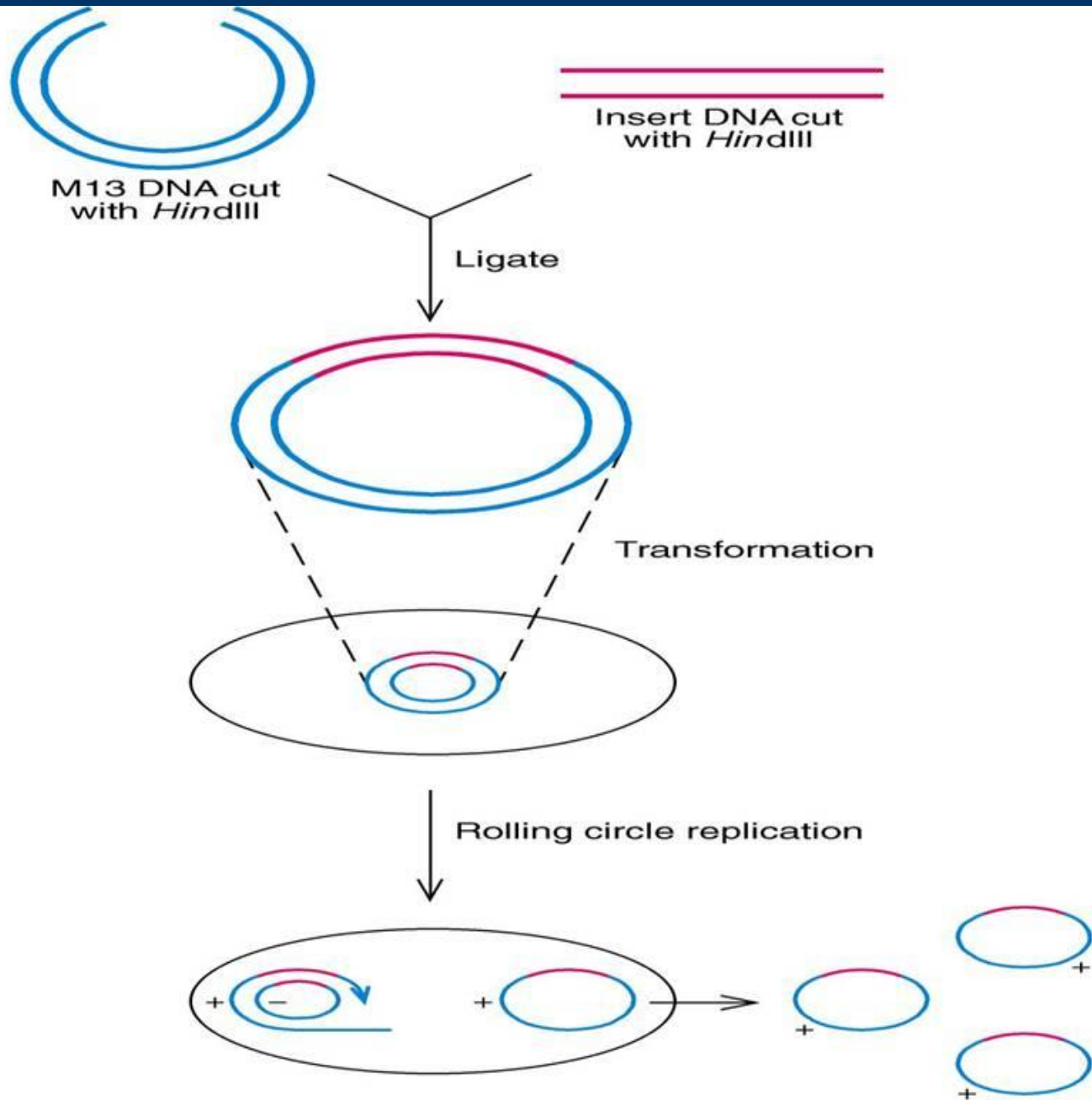
多克隆位点
M13 mp18

1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	7	8	
Thr	Met	Ile	Thr	Asn	Ser	Ser	Ser	Val	Pro	Gly	Asp	Pro	Leu	Glu	Ser	Thr	Cys	Arg	His	Ala	Ser	Leu	Ala	Leu	Ala	
ATG	ACC	ATG	ATT	ACG	AAT	TCG	AGC	TCG	GTA	CCC	GGG	GAT	CCT	CTA	GAG	TCG	ACC	TGC	AGG	CAT	GCA	AGC	TTG	GCA	CTG	GCC
				EcoR I		Sac I		Kpn I		Sma I Xma I		BamH I		Xba I		Sal I Acc I Hinc II		Pst I		Spl I		Hind III				

M13 mp19

1	2	3	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	5	6	7	8	
Thr	Met	Ile	Thr	Pro	Ser	Leu	His	Ala	Cys	Arg	Ser	Thr	Leu	Glu	Asp	Pro	Arg	Val	Pro	Ser	Ser	Ans	Ser	Leu	Ala	
ATG	ACC	ATG	ATT	ACG	CCA	AGC	TTG	CAT	GCC	TGC	AGG	TCG	ACT	CTA	GAG	GAT	CCC	CGG	GTA	CCG	AGC	TCG	AAT	TCA	CTG	GCC
					Hind III		Spl I			Pst I		Sal I		Xba I		BamH I			Sma I		Kpn I		Sac I		EcoR I	

M13 噬菌体载体的应用

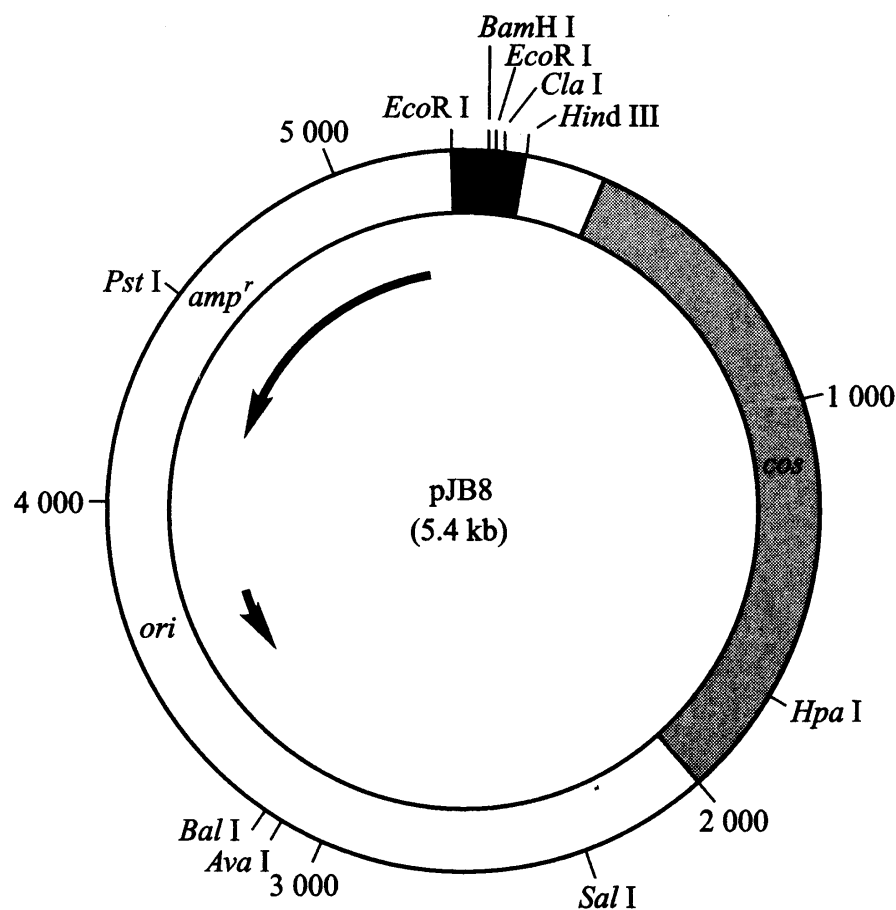




四、柯斯质粒cosmid

- 1978年由Collins和 Hohn改建
- 正常的质粒与噬菌体的cos位点构成
- 有Amp^r, Tet^r 抗性基因
- Cos位点可识别噬菌体的外壳蛋白, 可被包装
- 包装能力: 30-40kb

柯斯质粒 (cosmid) 的结构

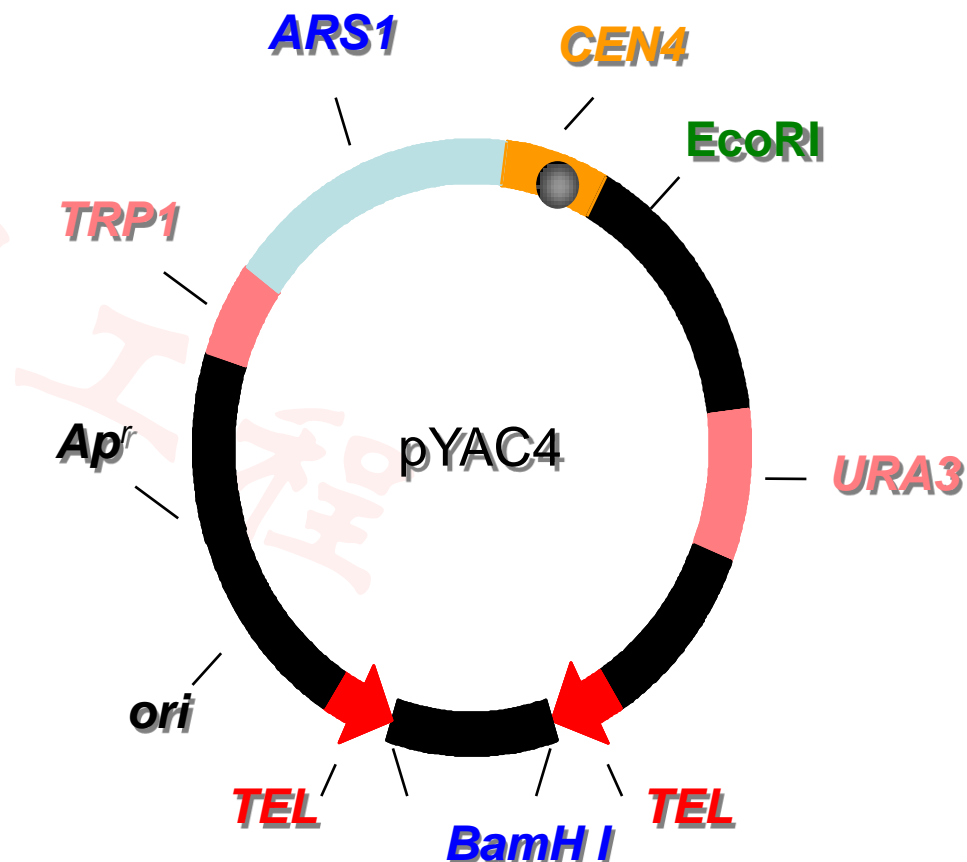


五、人工微小染色体

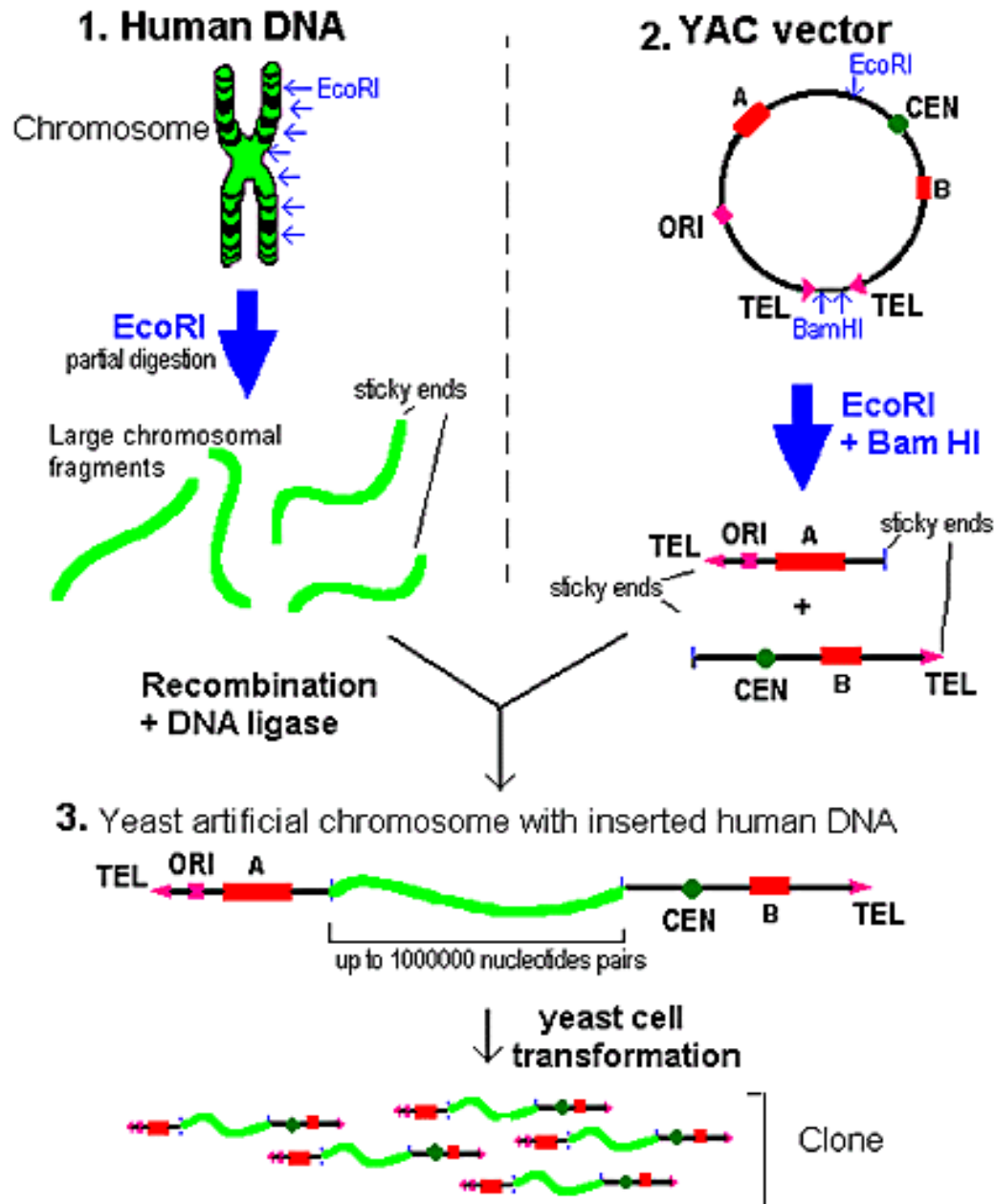
- YAC, yeast artificial chromosome
- BAC, bacteria artificial chromosome
- PAC, P1 artificial chromosome
- MAC, mammal cell artificial chromosome
- Fosmid, F因子 改建 而来

酵母人工染色体 (YAC) 结构

- YAC载体应含有下列元件：
- 酵母染色体的端粒序列
- 酵母染色体的复制子
- 酵母染色体的中心粒序列
- 酵母系统的选择标记
- 大肠杆菌的复制子
- 大肠杆菌的选择标记
- YAC载体的装载量为800 - 1000 kb



酵母人工染色体的应用



Cloning into a Yeast Artificial Chromosome (YAC)



克隆载体的发展历史

- 第一代：环状质粒，装载8-10个kb片段；
- 第二代：经过改建的病毒，装载能力20kb左右；
- 第三代：cosmid, 装载能力30-40kb；
- 第四代：YAC, 装载能力1-2 Mb；
- 第五代：新型载体：BAC, PAC, MAC, Fosmid等, 装载能力80-200kb



第三节 表达载体

原核表达载体

真核表达载体

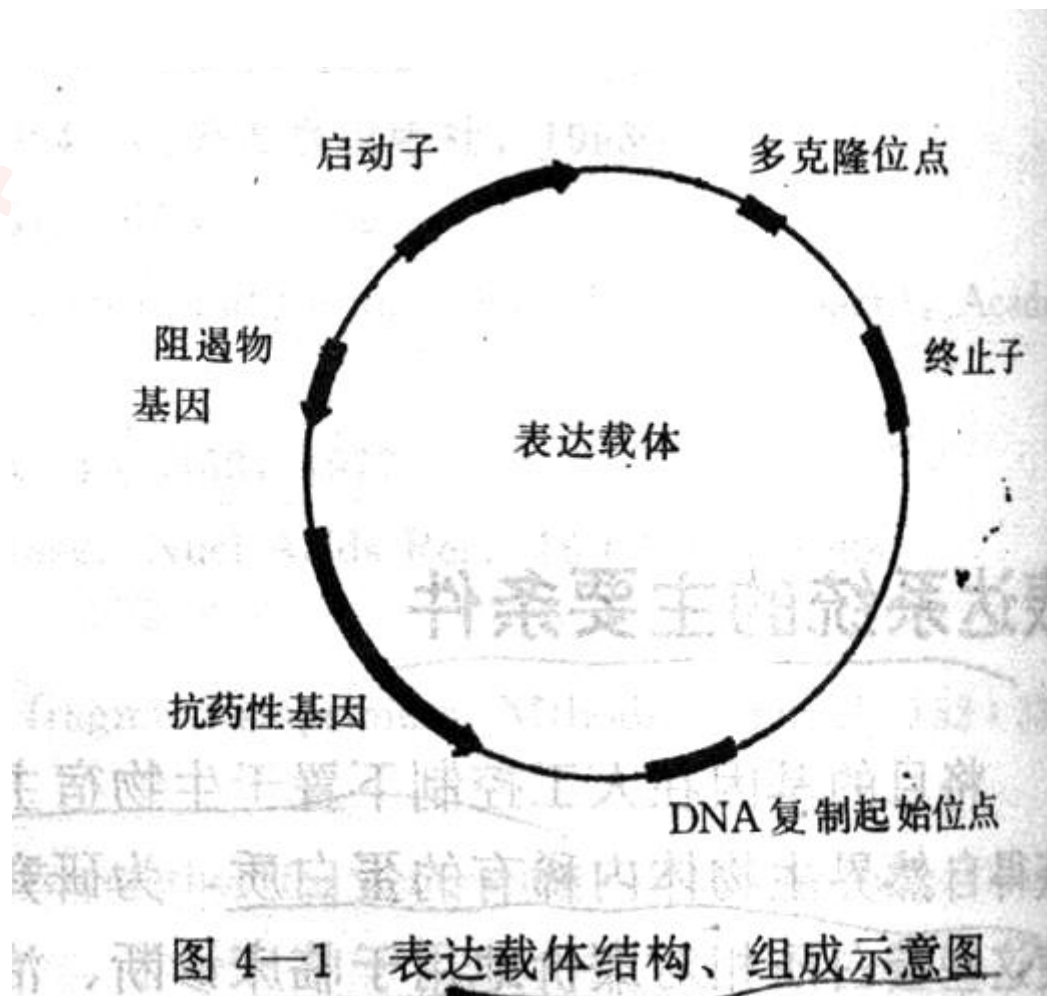


表达载体的组成

- 表达载体是将目的基因在人工控制下置于生物宿主中大量生产的载体. **表达载体三个系统:**
 1. DNA复制及质粒DNA的筛选: 有DNA复制起点ori, 及Amp, Tet抗性基因
 2. 目的基因的转录: 这一系统包括启动子, 抑制物基因和转录终止子. 启动子位于目的基因的上游, 常用的如Placz等.
 3. 蛋白质的翻译: 包括核糖体识别位点SD, 翻译起始密码子和终止密码子.

一、原核表达载体

- 启动子
- 核糖体结合位点
- 密码子
- 终止子



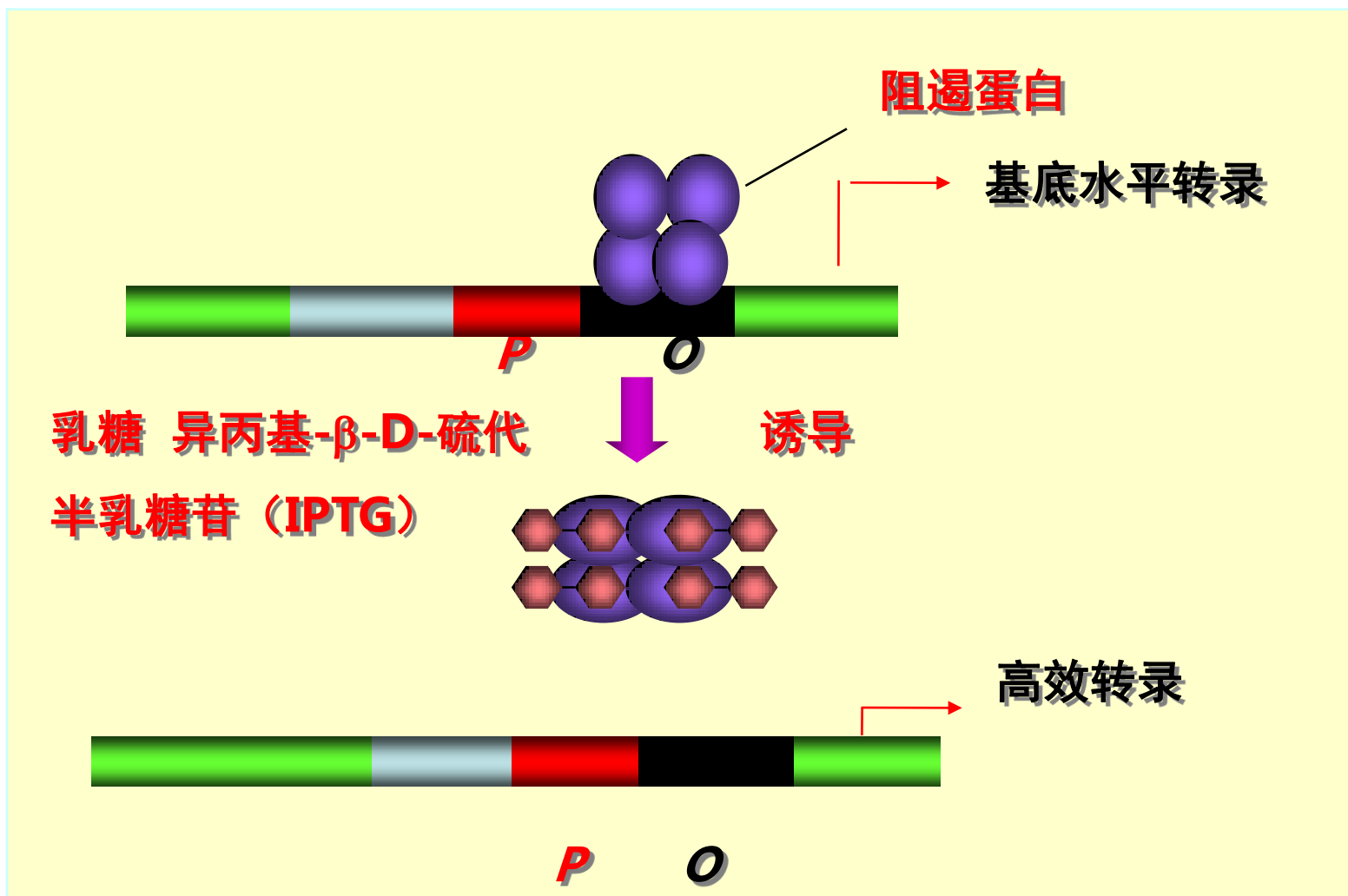


1. 启动子

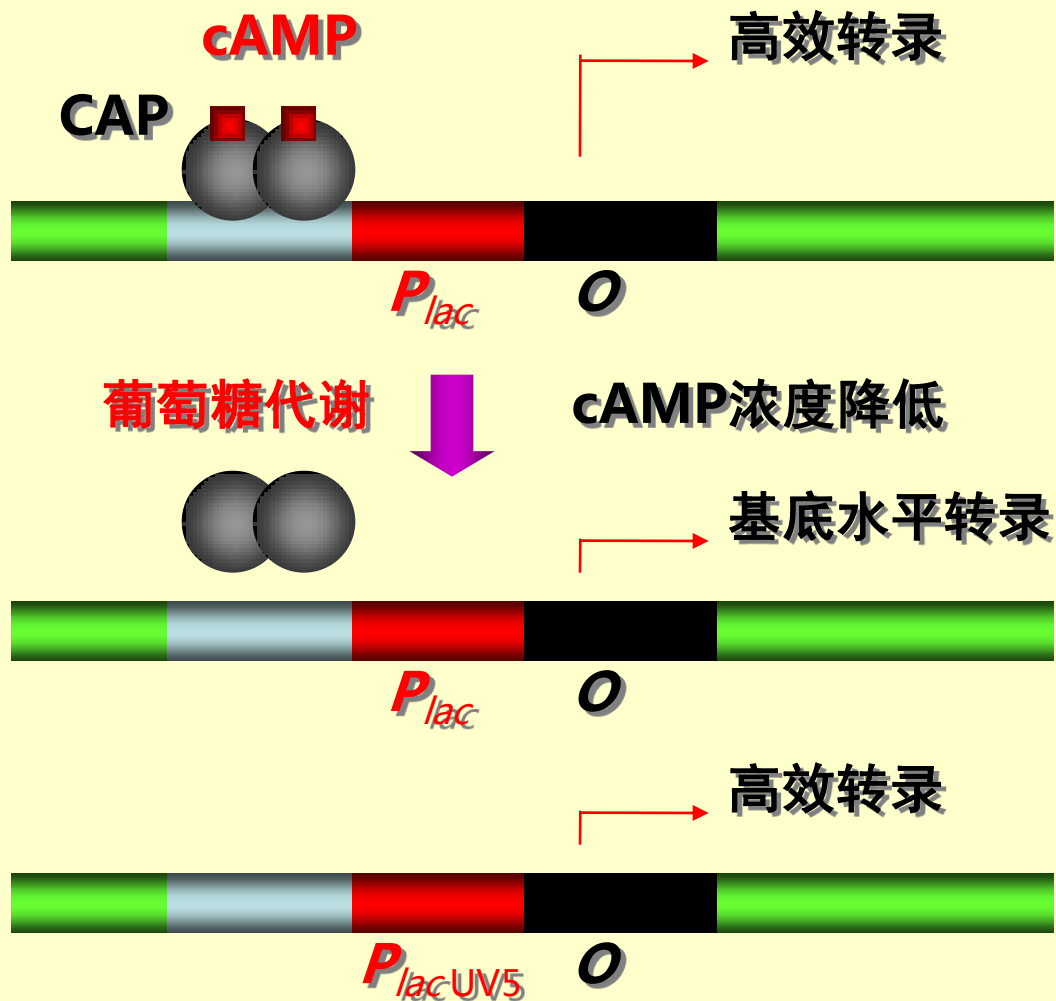
原核表达载体常见的启动子

启动子	-35 区序列	-10 区序列
<i>P_λL</i>	T T G A C A	G A T A C T
<i>P_{recA}</i>	T T G A T A	T A T A A T
<i>P_{traA}</i>	T A G A C A	T A A T G T
<i>P_{trp}</i>	T T G A C A	T T A A C T
<i>P_{lac}</i>	T T T A C A	T A T A A T
<i>P_{tac}</i>	T T G A C A	T A T A A T

*PlacZ*启动子的可控性



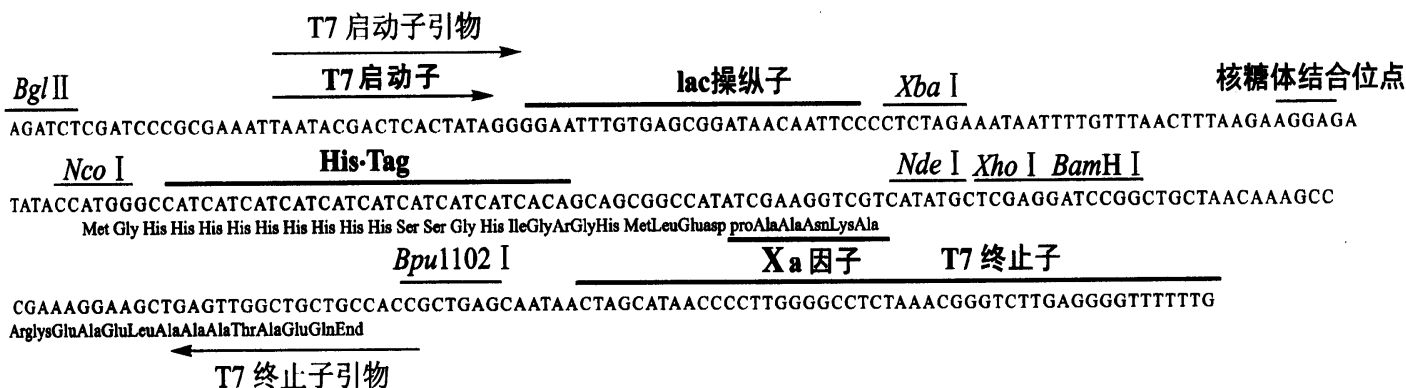
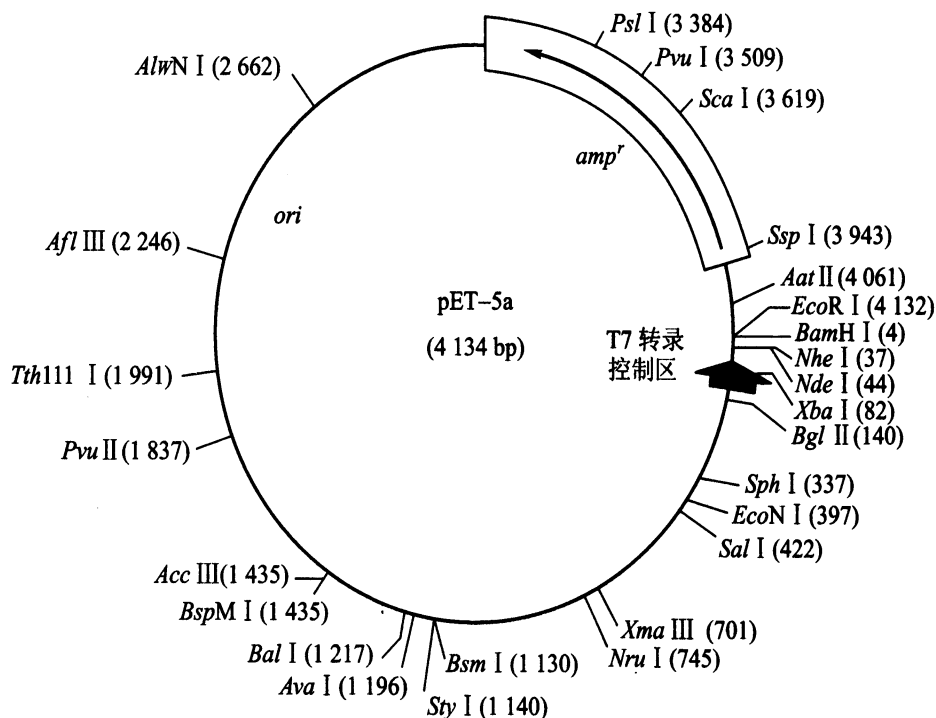
*P_{lac}Z*启动子的可控性



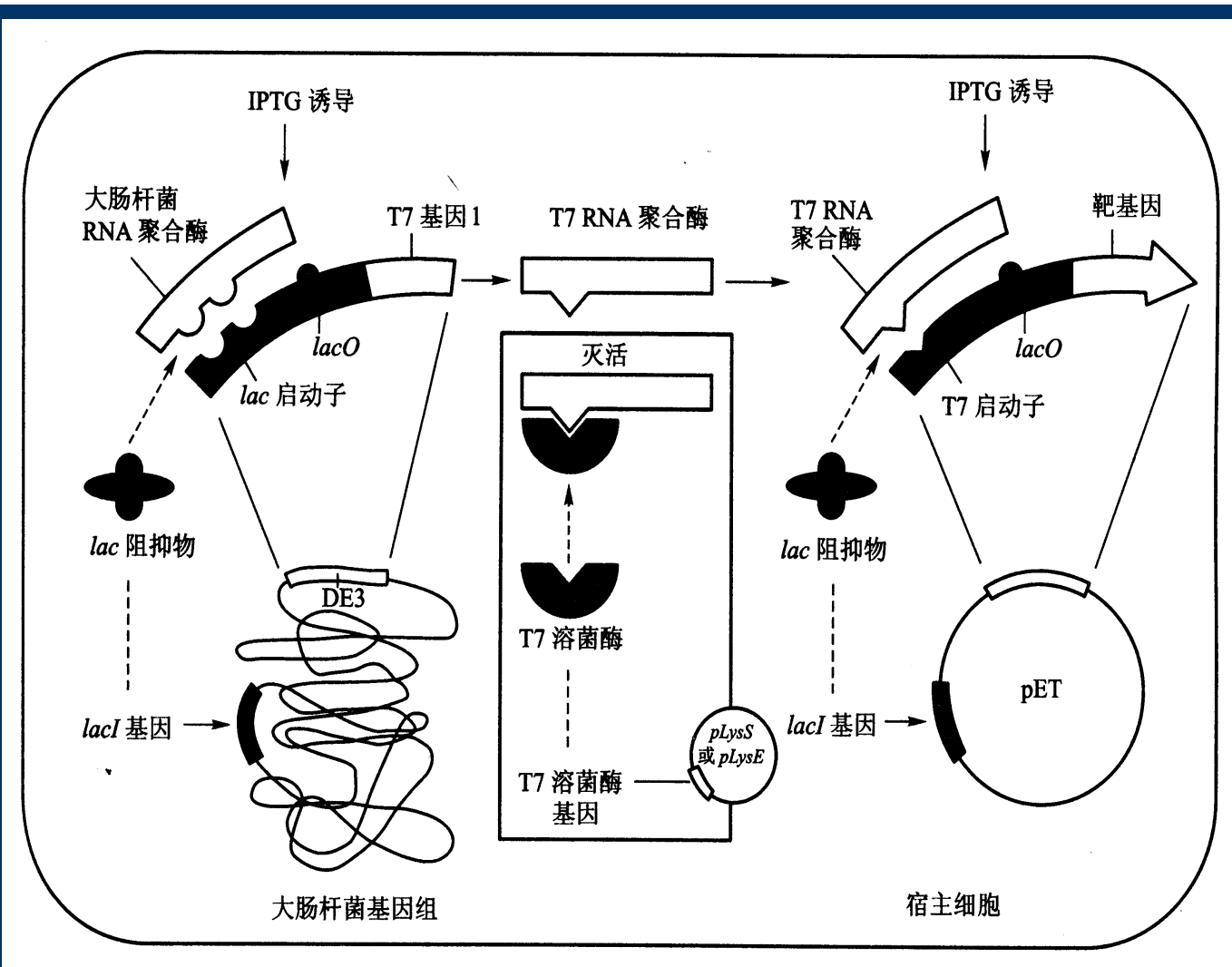


高效表达的T7启动子pET 5a 、 pET 16b 结构

工程



T7表达载体与受体系统



大肠杆菌菌株：BL21 (DE3) 和BL21 (DE3) pLysS



T7噬菌体启动子系统的优点

- 1) T7噬菌体RNA聚合酶合成RNA的速度高于大肠杆菌5倍；
- 2) T7噬菌体RNA聚合酶只识别自己的启动子序列，不启动大肠杆菌DNA任何序列的转录；
- 3) T7噬菌体RNA聚合酶对抑制大肠杆菌RNA聚合酶的抗生素有抗性；
- 4) T7噬菌体RNA聚合酶/启动子系统在一定条件下，基因表达产物可占细胞总蛋白的25%以上。



2. 核糖体结合位点

大肠杆菌核糖体结合位点Shine-Dalgarno (SD) 序列包括下列四个特征结构：

- ① 序列组成为**5' UAAGGAGG 3'**，与核糖体小亚基中的16S rRNA 3'端区域**3' AUUCCUCC 5'**互补并与之专一性结合；
- ② 翻译起始密码子**AUG**；
- ③ SD序列与翻译起始密码子**AUG**之间的距离(6-8bp)及碱基组成；
- ④ 基因编码区5' 端若干密码子的碱基序列



3、 融合表达载体

- 蛋白质融合表达是指**外源基因与载体已有的担体蛋白的编码基因拼接在一起**，并作为一个新的开放阅读框进行表达。
- 通过人工设计引入的蛋白酶切割位点或化学试剂特异性断裂位点，可以在体外从纯化的融合蛋白分子中释放回收目的蛋白。



融合蛋白表达的优缺点

- ① 与高表达的担体蛋白共同表达，可以保证目的蛋白**表达率高，稳定性好**；
- ② 担体蛋白常常是结构和功能研究得比较清楚的蛋白质，利用免疫亲和层析很**容易对担体蛋白进行分离和纯化**；
- ③ 某些担体蛋白能很成功地生成折叠正确、有生物活性的蛋白质，**可以帮助目的蛋白形成正确的空间结构**。
- ④ 通过化学裂解和蛋白酶很**容易将担体蛋白从目的蛋白中裂解出来**。



融合蛋白表达的优缺点

用于融合蛋白构建的担体蛋白类型和优点

谷胱甘肽转移酶 (GST) 维持良好空间构象

金黄色葡萄球菌蛋白A (SAPA) 免疫亲和层析 pRIT2T

硫氧化还原蛋白 (TrxA) 维持良好空间构象 pTrxFus

b-半乳糖苷酶 (LacZ) 免疫亲和层析

泛素蛋白 (Ubi) 维持良好空间构象



融合蛋白中目的蛋白的回收

化学断裂法:

如溴化氰 (CNBr) 它与多肽链中的甲硫氨酸残基侧链的硫醚基反应, 生成溴化亚氨内酯, 后者不稳定, 在水的作用下肽键断裂, 形成两个多肽降解片段。

酶促裂解法

单残基位点与多残基位点断裂法



酶促裂解法：单残基位点

每种蛋白酶均具有相应的**断裂位点决定簇**

单残基位点断裂的蛋白酶

如在**精氨酸、谷氨酸、赖氨酸**

残基处切开酰胺键，使目的蛋白分离

蛋白内切酶

切割位点

梭菌蛋白酶

Arg-

葡萄球菌蛋白酶

Glu-C

假单孢菌蛋白酶

Lys-C

猪胰蛋白酶

Arg-C

Lys-C

担体蛋白

目的蛋白

N ■■■■■ Arg C N ■■■■■ C

酶促裂解法：多残基位点

启动子 担体基因 接头 目的基因

表达

Met

Stop

Ile-Glu-gly-Arg

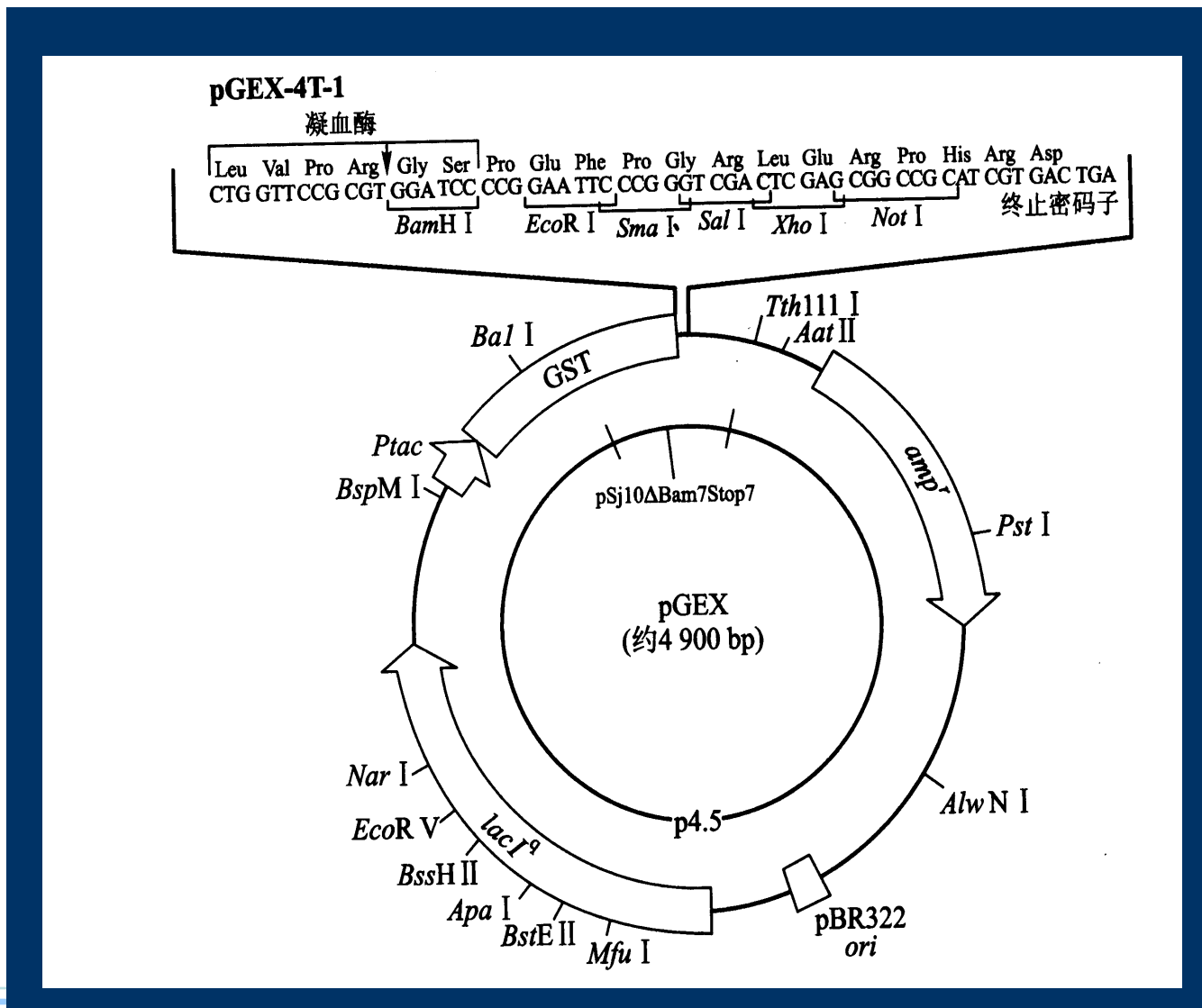
凝血因子Xa的识别、作用序列

亲和层析

酶解回收

Ile-Glu-gly-Arg

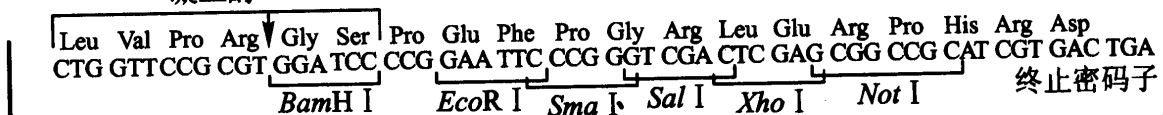
谷胱甘肽S-转移酶 (GST) 融合蛋白表达系统



融合蛋白表达载体移码问题

pGEX-4T-1

凝血酶



GST融合表达载体目的基因移码分析

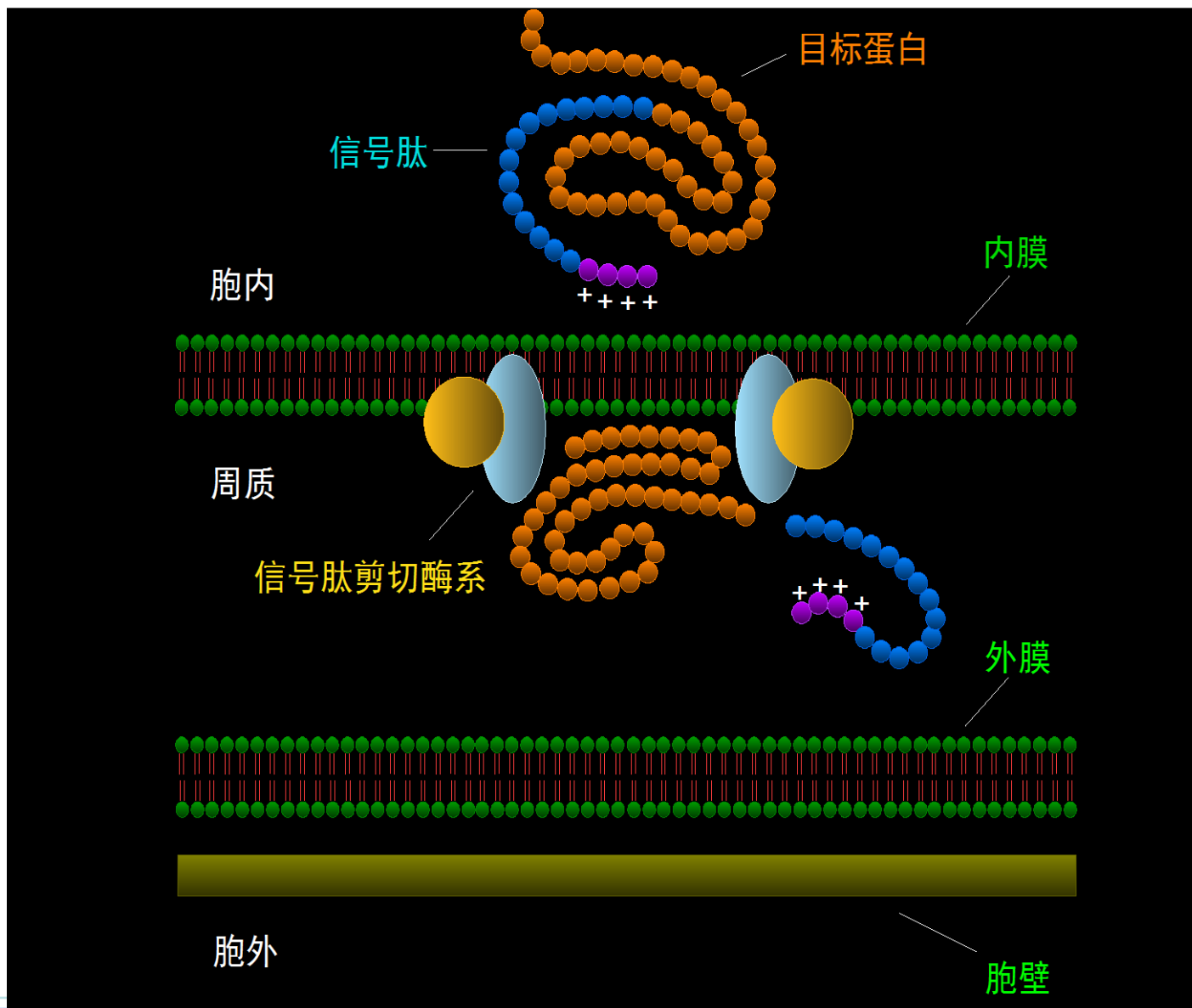
- Leu Val Pro Arg Gly Ser Pro Glu Phe Pro Gly ...
- CTG GTT CCG CGT GGA TCC CCG G ↓ AA TTC CC ↓ G GGT... GST载体
- 凝血酶 BamH I EcoR I Sma I
- CC G ↓ AA TTC ATG CCA GAC GGG CCC AAT ACC ... 目的基因EcoR I切
- CTG GTT CCG CGT GGA TCC CCG GAA TTC ATG CCA GAC GGG...不移码
- 
- CC CCC ↓ GGG ATG CCA GAC GGG CCC AAT ACC ... 目的基因Sma I切
- CTG GTT CCG CGT GGA TCC CCG GAA TTC CCG GGA TGC CAG AC.. 移码



4、 分泌型表达载体

- 分泌表达是指周质蛋白、细胞内膜与外膜蛋白等以带有**N端信号肽的前体形式**首先在细胞质中合成，随后穿膜运送到周质或定位到内、外膜上的分泌过程。穿膜过程中，信号肽被信号肽酶切除。将目的蛋白的基因置于原核蛋白信号肽的序列的下游可能实现分泌型表达。

分泌表达的原理





分泌表达的优点：

- ①信号肽被切除后的蛋白质N末端的氨基酸序列与天然蛋白是一致的
- ②周质空间中的蛋白酶活性比细胞质中的要低，从而使蛋白质较稳定地存在于周质中
- ③周质中只有少量的细菌蛋白，使分离纯化更容易
- ④周质空间存在一个氧化的环境，使得二硫键更容易形成

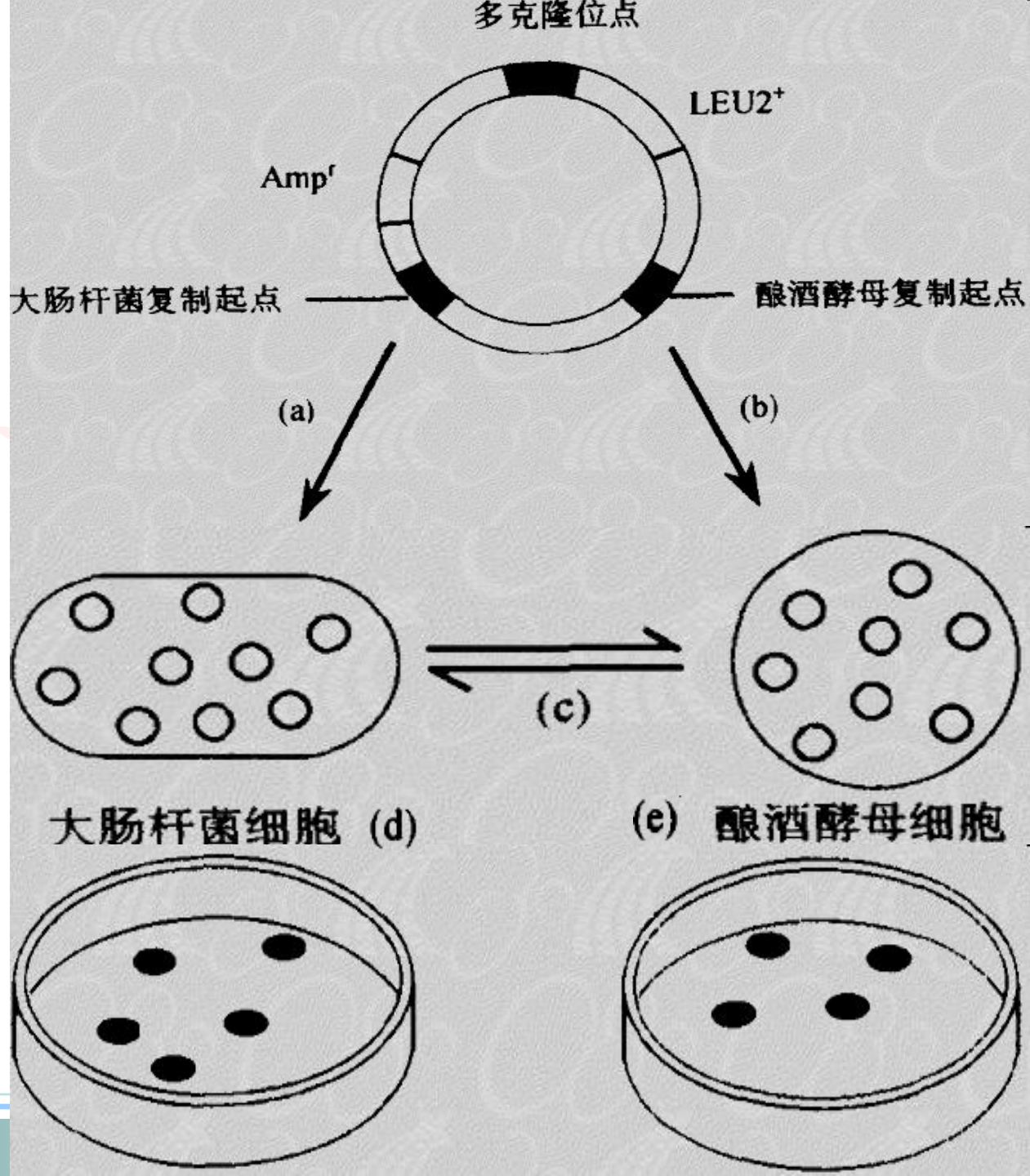


二、真核表达载体

- (1) 原核DNA序列
- (2) 启动子 和增强子
- (3) 剪接信号
- (4) 终止信号和多聚腺苷化的信号
- (5) 遗传标记

穿梭载体

(shuttle vector) 是一类由人工构建的具有两种不同复制起点和选择标记，因而可在两种不同的寄主细胞中存活和复制的质粒载体。

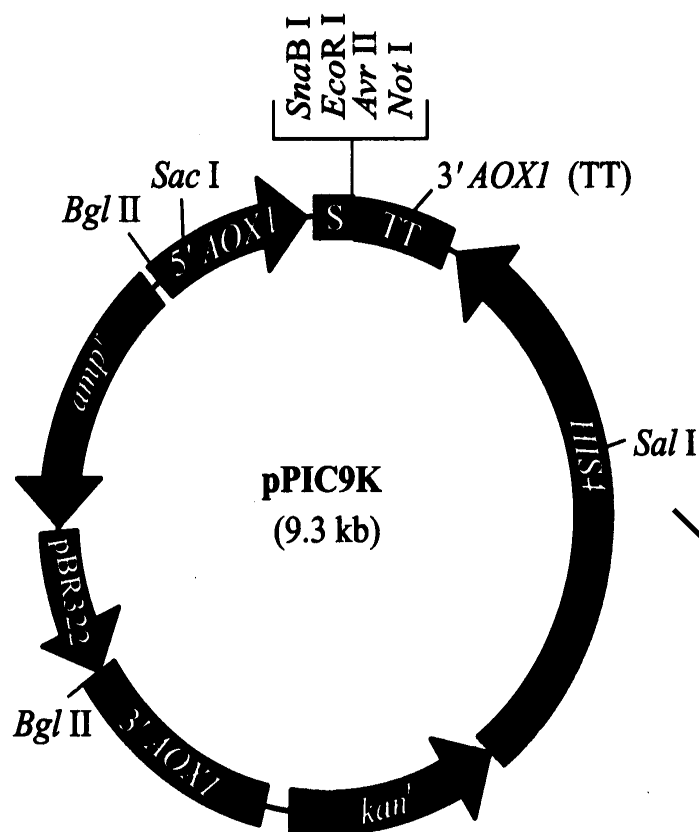




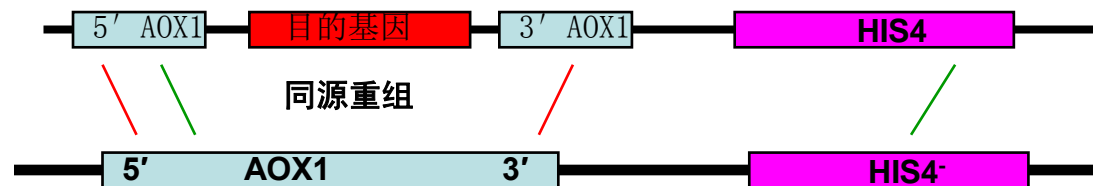
1、甲醇酵母载体

- 利用甲醇作为唯一碳源。乙醇氧化酶将甲醇氧化为甲醛，乙醇氧化酶的启动子是强启动子，编码乙醇氧化酶的基因是AOX1和AOX2. AOX1基因的表达受甲醇严格调控，诱导表达水平可达30%以上。
- 诱导机制：培养基以葡萄糖或甘油为碳源，抑制转录；
- 培养基以甲醇为唯一碳源时，诱导基因转录、蛋白表达
- 野生型Mut⁺， 突变型Mut^s（AOX1基因缺失）

甲醇酵母表达系统的载体与受体菌



- 受体菌：GS115，KM71，SMD1168等
- 含有AOX1和AOX2基因
- 组氨酸缺陷型。如果表达载体上带有组氨酸的基因，可补偿宿主的组氨酸缺陷。



受体菌基因组DNA



2、哺乳动物细胞表达载体

pCMV启动子

EGFP

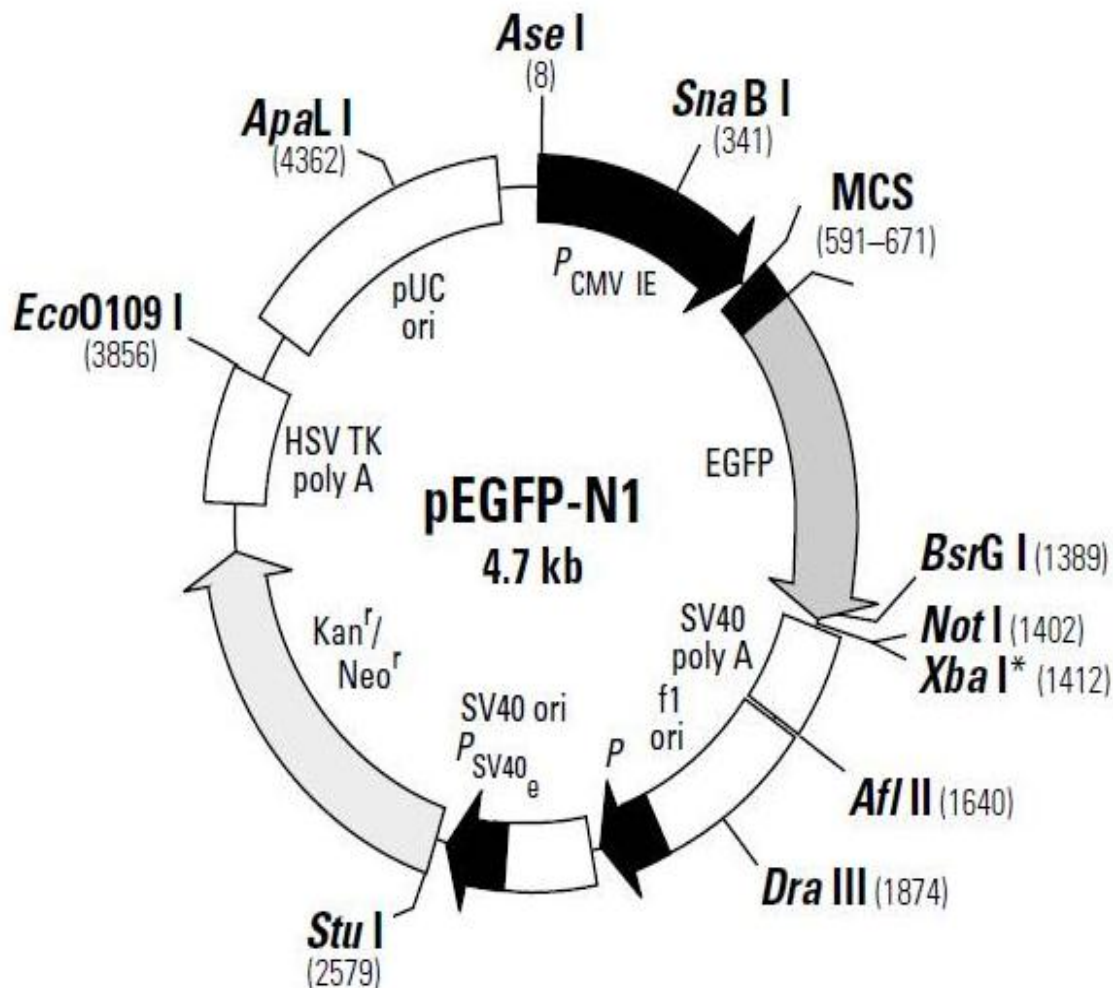
SV40 *ori*

*Neo*抗性盒

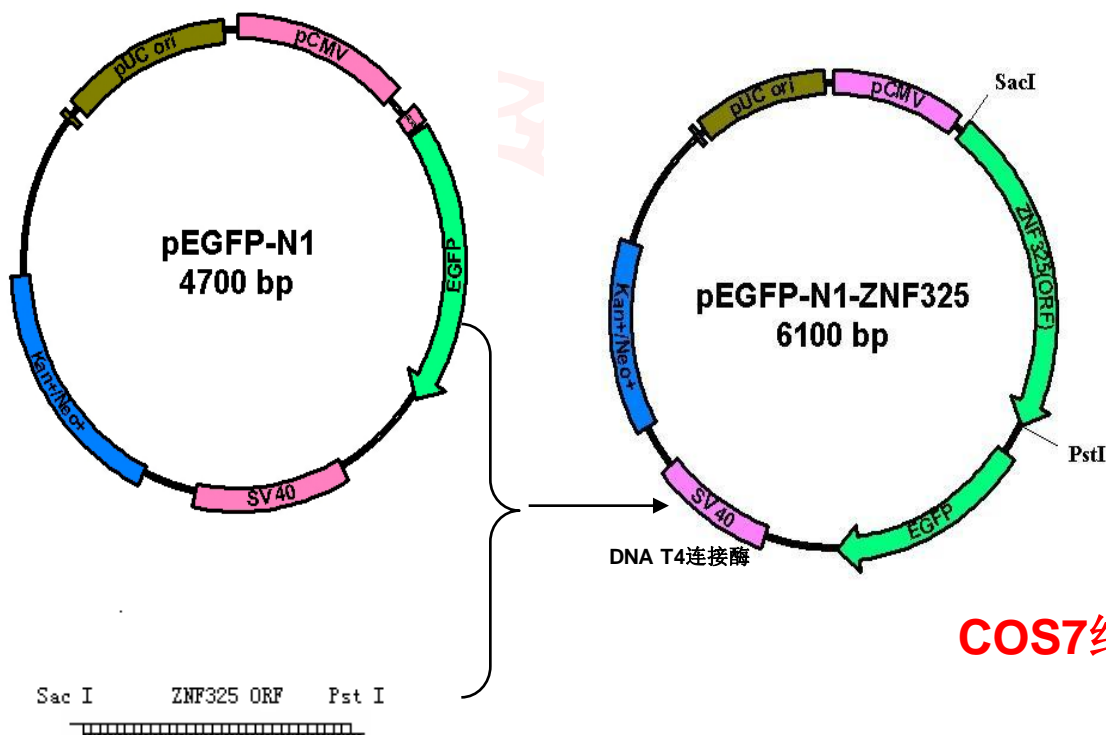
*HSV-TK*基因的多聚腺苷化信号

pUC *ori*盒

*kan*抗性基因



表达载体与目的基因的连接



Sac I 6100bp

Pst I 6100bp

Sac I + Pst I

4700+1400bp

COS7细胞系

