



## 本章重点

1. 基因工程和基因组学的概念和应用；
2. 内切酶的种类和作用；
3. 重组DNA；
4. 基因的分离和鉴定；
5. 基因组图谱的构建和应用；
6. 后基因组学。

## 第十二章 基因工程和基因组学

**转基因烟草**  
来自萤火虫的荧光素酶



### 一、基因工程概述：

**广义遗传工程包括：**

生化工程、蛋白质工程、细胞工程、染色体工程、细胞器工程、**基因工程**及酶工程等。

**狭义遗传工程是指：**

基因工程(重组DNA技术)。



### 第一节 基因工程



**果蝇转基因**  
 $\beta$ -半乳糖苷酶

#### 1. 概念：

**基因工程**：在分子水平上，采取工程建设方式 → 按照预先设计的蓝图 → 借助于实验室技术将某种生物的基因或基因组转移到另一生物中去 → 使后者定向获得新遗传性状的一门技术。

**基因工程技术的建立，使实验生物学领域产生巨大变革。**

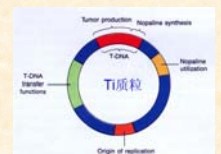


#### 2. 发展：

**1971年** 史密斯 (Smith H. O.) 等人从细菌中分离出的一种**限制性酶** → 酶切病毒DNA分子，标志着DNA重组时代的开始。

**1972年** 伯格 (Berg P.) 等用限制性酶分别**酶切猴病毒**和 **$\lambda$ 噬菌体DNA**，将两种DNA分子用**连接酶**连接起来 → 得到新的DNA分子。

**1973年** 科恩 (Cohen S.) 等进一步将酶切DNA分子与质粒DNA连接起来，并将**重组质粒**转入*E. coli*细胞中。



1982年，美国食品卫生和医药管理局批准，用**基因工程在细菌中生产人的胰岛素**投放市场。

1985年，**转基因植物**获得成功。

1986年，穆勒斯发明了PCR技术，专利转让达3亿美元。

1990年9月14日是基因治疗的诞生日。

在美国马里兰州的一个医疗中心进行第一例基因治疗临床试验。一个4岁的小女孩戴斯瓦(A. DeSilva)，患有严重综合免疫缺失症(SCID)。将腺苷酸脱氨酶(ADA)基因转入骨髓细胞，再送回病人体内，基因治疗SCID获得初步效果。

1994年，延熟保鲜的**转基因番茄**商品生产。



1996年，**威尔穆特**研究小组**克隆羊“多莉”**诞生(利用6岁成年母羊乳腺细胞)。



1997年，**威尔穆特**小组报道用胎儿细胞为核供体，获得了表达治疗人血友病的凝血因子IX转基因克隆羊**“波莉”**

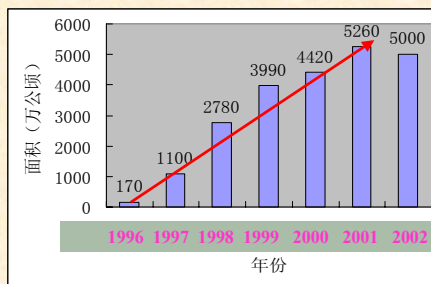


1998年，美国夏威夷大学用一只实验鼠的细胞克隆了3代共50只**实验鼠**。

2000年，美国用无性繁殖技术克隆**猴子**“泰特拉”，意味克隆人体已无技术障碍。

2001年，美国宣布首次克隆成功处于早期阶段的**人体胚胎**。

2002年，12月法国研究者宣布克隆出第一个**女婴**，但一直未获得证实。



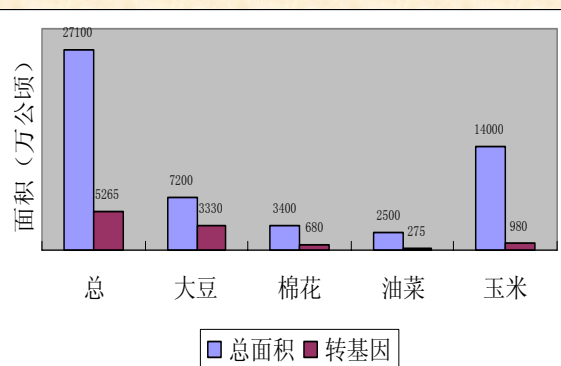
1996年全世界转基因植物种植面积为170万公顷，1997年为1100万公顷，1998年2780万公顷，1999年3990万公顷，2000年4420万公顷，2001年5260万公顷，2002年5000万公顷。

**2001年种植面积>100万公顷的转基因作物:**

**大豆** (3330万 $\text{hm}^2$ ，占全世界转基因作物的63%，均为抗除草剂大豆)、**玉米** (980万 $\text{hm}^2$ ，占19%)、**棉花** (680万 $\text{hm}^2$ ，占13%)、**油菜** (270万 $\text{hm}^2$ ，占5%)；

**其它还有**水稻、小麦、花生、向日葵、亚麻、甘蓝、马铃薯等，番茄、烟草、南瓜和木瓜等50多种转基因作物已培育成功。

**主要分布**在美国 (3570万 $\text{hm}^2$ )、阿根廷 (1180万 $\text{hm}^2$ )、加拿大 (320万 $\text{hm}^2$ ) 和中国 (150万 $\text{hm}^2$ ) 等国。



**2001年全世界转基因作物占相应作物种植总面积的比较**

2001年我国的转基因农作物和林木已达22种，其中转基因棉花、大豆、马铃薯、烟草、玉米、花生、菠菜、甜椒、小麦等进行了田间试验，转基因棉花已经大规模商品化生产。



浙江大学



遗传学第十二章

### 3. 内容:

- ①. 从细胞和组织中分离DNA;
- ②. 限制性内切酶切DNA分子, 制备DNA片段;
- ③. 将酶切DNA分子与载体DNA连接 → 构建能在宿主细胞内自我复制的重组DNA分子;
- ④. 把重组DNA分子引入宿主受体细胞 → 复制;
- ⑤. 重组DNA随宿主细胞的分裂而分配到子细胞 → 建立无性繁殖系(clone)或发育成个体;
- ⑥. 从细胞群体中选出所需要的无性繁殖系 → 并使外源基因在受体细胞中正常表达, 翻译成蛋白质等基因产物、回收; 或筛选出获得定向性状变异的个体。

浙江大学

遗传学第十二章

## 二、限制性内切核酸酶:



### 1. 限制性内切酶(restriction enzyme):

一种水解DNA的磷酸二酯酶, 遗传工程中重要工具。

细菌细胞中存在限制修饰系统:

**限制:** 降解外源DNA, 防御异源遗传信息进入的手段。

**修饰:** 修饰外源DNA片段后, 保留在新细胞中。



遗传学第十二章

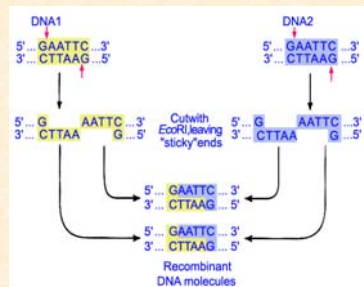
### ① 限制性内切酶如: *EcoR I*、*Hind III*:

#### 酶切方式:

以交错方式切断DNA双链, 产生二个相同单链粘性末端。

#### 重组过程:

两片片段在适宜条件下, 可经磷酸二酯键, 连接成重组DNA分子。



浙江大学

遗传学第十二章

### (1). 限制性内切酶的命名:

根据其来自的生物名称, 用英文字母和数字表示:

- ①. *EcoR I* 来自大肠杆菌 (*Escherichia coli*);
- ②. *Hind III* 来自嗜血杆菌 (*Haemophilus influenzae*).

### (2). 限制性内切酶的类型:

#### 第 I 类酶 (切割部位无特异性):

如 *EcoB* (大肠杆菌 B 株)、*EcoK* (大肠杆菌 K 株) 分子量较大 (约 300000), 作用时需 ATP、 $Mg^{++}$  等辅助因子。

#### 第 II 类酶 (切割部位有特异性):

如 *EcoR I* (大肠杆菌)、*Hind III* (嗜血杆菌), 分子量较小 (约 20000~100000), 作用时需  $Mg^{++}$  存在。

浙江大学

遗传学第十二章

### 第 II 类酶特点:

- ✦ 切割产生平齐末端 (blunt ends), 如 *Sma I*:

Examples of how restriction enzymes cleave DNA: (a) *SmaI* results in blunt ends; (b) *BamHI* results in overhanging (sticky) 5' ends; (c) *PstI* results in overhanging (sticky) 3' ends.

#### a) Cut with *SmaI*



浙江大学

遗传学第十二章

- 有的产生粘性末端 (sticky ends), 如 *Bam*H I、*Pst* I;
- 识别特定碱基顺序, 如回文对称序列 (palindrome, 又称反向重复序列, 从两个方向阅读其序列相同的序列)。

### b) Cut with *Bam*HI



### c) Cut with *Pst*I



部分常用的限制性内切酶

### Characteristics of Some Restriction Enzymes

Enzymes with 6-bp Recognition Sequences	BamHI "bam-H-one"	Bacillus amyloliquefaciens H	5'GGATCC3' 3'CTAAGG5'
EcoRI "eco-R-one"	E. coli RY13	GAATTC CTTAAG	
HindIII "hin-D-three"	Haemophilus influenzae Rd	AGCTTT TTCGAA	
PstI "P-S-T-one"	Providencia stuartii	CTGCAG GACGTC	
SmaI "sma-one"	Serratia marcescens	CCCGGG GGGCCC	
Enzymes with 4-bp Recognition Sequences	HaeIII "hay-three"	Haemophilus eggplant	GGCC CCGG
HpaII "hepa-two"	Haemophilus parainfluenzae	CCGG GGCC	
Sau3A "saw-three-A"	Staphylococcus aureus 3A	GATC CTAG	
Enzymes with 8-bp Recognition Sequences	NotI "not-one"	Nocardia ciftidis-civianum	GCGGCCGC CGCCGGCG

\*In this column the two strands of DNAs are shown with the sites of cleavage indicated by arrows. Since the sequence is symmetrical about a center point in each recognition sequence, the DNA molecules resulting from the cleavage are symmetrical.  
Key R:guanine, Y:pyrimidine.  
Adapted from R.J. Roberts, 1978, CRC Critical Reviews in Biochemistry 4:123 Copyright ©1978 CRC Press, Inc., Boca Raton, FL. Reprinted with permission.



## 三、载体 (vector) :

**运载工具:** 将“目的”基因导入受体细胞的运载工具。

DNA片段 + 适合的载体DNA → 重组DNA → 在载体DNA的运载下, 高效率地进入宿主细胞, 并在其中进行复制。

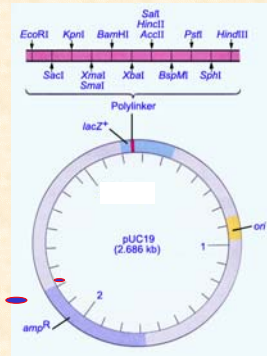
**DNA载体:** 质粒、噬菌体、病毒、细菌或酵母菌人工染色体等。



### 载体的条件:

- ①. 具有复制原点, 能自我复制;
- ②. 具多克隆位点 (multiple cloning site, MCS 或 polylinker region), 即有多种限制酶的切点;
- ③. 选择时的遗传标记, 如抗生素基因;
- ④. 易从宿主细胞中回收。

pUC<sub>19</sub>质粒



### (-)、细菌质粒:

**质粒**是细菌细胞内独立于细菌染色体而自然存在的、能自我复制、易分离和导入的环状双链DNA分子。

质粒具有**重组表型检测标记**, 检测是否携带外源DNA片段。

#### 在细胞内的复制程度:

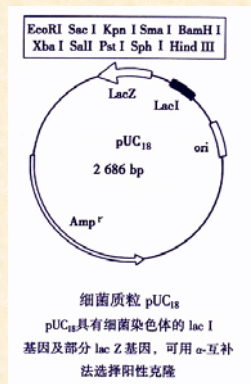
**严紧型:** 一个细菌细胞内质粒数量有1~2个;

**松弛型:** 每个细胞内有的20~60个。



### 如pUC<sub>18</sub>质粒具有以下特点:

- ①. **分子量小**, 可接受较大外源片段;
- ②. **拷贝数多**, 500个/细胞;
- ③. **克隆位点的酶切位点多**, 克隆方便;
- ④. 具有用于检测重组质粒的**选择标记** (α-互补的显色表型)。



### (二)、λ噬菌体(温和型):

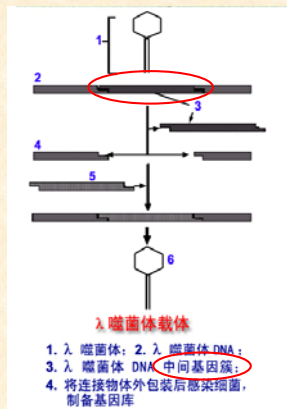
基因组全长为49kb。

噬菌体DNA中间约2/3的序列为**中间基因簇**，两端为DNA左、右臂。**中间基因簇**可被外源DNA替代而不影响侵染细菌能力。

能接受**15kb外源DNA**片段作为cDNA或核DNA克隆载体。

#### 优点:

- 不易引起生物危害，有助于“目的”基因进入细胞并增殖;
- 携带**大片段**外源DNA分子，占总量25%时仍不失活。



### (三)、柯斯质粒(cosmid):

噬菌体DNA + 部分细菌质粒DNA序列 → 柯斯质粒。

有噬菌体cos序列、细菌质粒复制原点、抗生素抗性标记。

这种质粒分子量较小,但可接受**长达50kb**的外源DNA片段 → 克隆真核生物基因。

一个长片段DNA可能具有真核生物基因的编码序列及其它调控序列。



### (四)、穿梭载体(shuttle vectors):

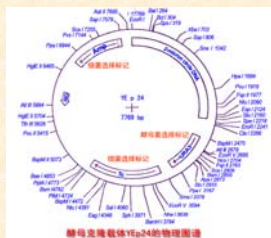
指能在**两种不同**生物中复制的载体。

如能在**原核生物**(如*E. coli*)、**真核细胞**(如酵母)中复制的载体。

穿梭载体需**具有**细菌质粒复制原点、真核生物自主复制序列(Auto-nomously replicating sequence, ARS)以及两者的选择标记。

穿梭载体在细菌中用于克隆、扩增基因,在酵母重用与基因表达分析。

酵母菌的**YEp**(yeast episoma plasmid)等**系列载体**均是穿梭载体。



### (五)、细菌人工染色体(bacterial artificial chromosome, BAC):

**BAC载体**一般可携带**大于50kb**的外源DNA片段。F因子改造造成BAC载体,甚至可用于克隆100kb以上的DNA片段。

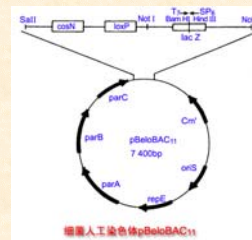
#### 特点:

带有外源DNA的BAC载体在细胞中是单拷贝的;

载体分子量很小(7.4kb);

选择标记:**氯霉素抗性基因**;

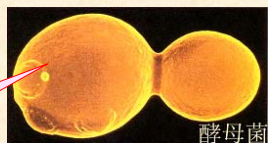
多克隆位点。



### (六)、酵母人工染色体(yeast artificial chromosome, YAC):

**YAC**具有自主复制序列、克隆位点和可在细菌和酵母菌中选择的标记基因;还具有酵母菌染色体一些特点;可接受**100~1000kb**的外源DNA片段。

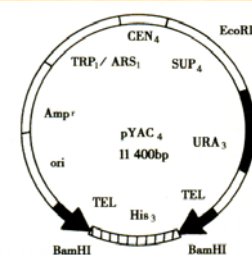
YAC已成为人类基因组计划和图位克隆基因的重要工具;并促进了**人类人工染色体**(human artificial chromosome, HAC)的研究。



1996年完成了酵母菌全基因组序列的测定。

### pYAC<sub>4</sub>特点:

- 两个可在酵母菌中利用的**选择基因**, URA<sub>3</sub>和TRP<sub>1</sub>(色氨酸合成基因);
- 酵母菌着丝粒**序列(CEN<sub>4</sub>);
- 一个**自主复制序列**(ARS<sub>1</sub>);
- 两个嗜热四膜虫**末端**重复序列(TEL) → 保持重组YAC为线状结构;
- 在两个末端序列中间,有一段**填充序列**(His<sub>3</sub>) → 使pYAC<sub>4</sub>在细菌细胞中能稳定扩增;
- Amp**抗性及细菌质粒复制原点;
- 一个EcoRI**克隆位点**,位于酵母菌Sup<sub>4</sub> tRNA基因内。



**酵母菌人工染色体pYAC4**在克隆外源DNA时,用BamHI和EcoRI双酶切,得到二个人工染色体臂,与EcoRI酶切的外源DNA片段连接,构成重组的酵母人工染色体,用于转化酵母菌。

**(4)、Ti质粒及其衍生载体:**

适合于植物的载体系统 → 将重组DNA运送到植物细胞中 → 目的基因表达。

**Ti质粒(细菌质粒):** 自然存在于**土壤农杆菌(Agrobacterium tumefaciens**, 革兰氏阴菌)细胞 → 可诱导植物产生瘤细胞(tumor-inducing, **Ti**) → 冠瘿瘤(grown galltumors)。

Ti质粒有一部分**转移DNA** (transfer DNA, **T-DNA**) → 当农杆菌感染植物时, **T-DNA**便转移到植物染色体上, 诱导**冠瘿瘤**, 并能合成**冠瘿碱** (opine), 作为农杆菌碳源和氮源。

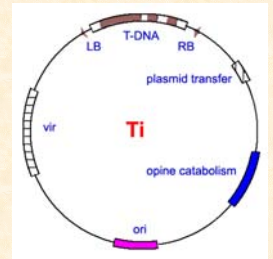
农杆菌感染产生冠瘿瘤



**○Ti质粒特点:**

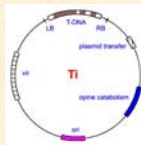
1. 具有**五个主要功能区域:**

- ①. **T-DNA:** LB与RB间T-DNA序列整合到植物基因组;
- ②. **质粒转移区(PT);**
- ③. **冠瘿碱(opine)代谢区;**
- ④. **复制原点;**
- ⑤. **毒性区。**



2. T-DNA中直接参与**转移并整合**的序列: T-DNA两端与其它序列交界处的**25bp**不完全直接重复, 右端的这段序列称为**右界(RB)**, 左端的序列称为**左界(LB)**。

3. **V区的毒性基因:** T-DNA转移所必需, 毒性基因可以顺式及反式两种方式控制T-DNA转移。



**○Ti质粒是200~500kb环状DNA分子, 主要有两套载体系统:**

(1). **二元载体(binary vectors):** 该系统具有两个质粒:

- ①. **用于克隆外源基因片段的克隆质粒:** 在*E. coli*和农杆菌中复制, 容易操作, 并可在二者间转移, 是一种穿梭质粒;
- ②. **非致病质粒:** 具有毒性基因, 没有T-DNA序列。

将两个质粒分别导入进农杆菌中, 经农杆菌介导, 克隆质粒中的外源DNA转移到植物染色体中。

(2). **共整合载体(integrated vectors):**

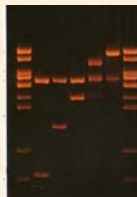
这个系统包括两个质粒, 一个用作克隆外源基因, 另一个具有毒性基因, 但**这两个质粒具有一段同源序列**, 在农杆菌中重组整合成一个载体。

**四、基因的分离与鉴定:**

一般而言, 一个基因 → 是编码一条多肽链的一个DNA片段, 包括启动子、终止子及内含子等。

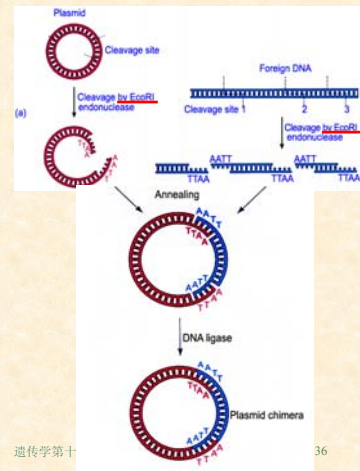
**在DNA重组技术出现之前,** 由于DNA分子量大而结构单一, DNA是细胞中最难分析的大分子。

DNA重组技术发展成功**之后,** DNA已成为细胞中最易操作分析的大分子。



**DNA的体外重组:**

- 1. 体外通过**限制性内切酶**的修剪和**DNA连接酶**的连接;
- 2. **目标DNA分子 + 载体** DNA, 共价连接;
- 3. 获得**重组DNA分子**。



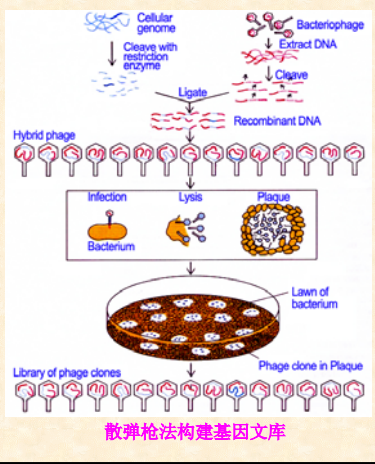
(一)、从基因库中分离基因:

1. 基因库 (library) :

是一组DNA和cDNA序列克隆的**集合体**。从基因库中分离基因, 首先要构建基因库。

根据克隆核酸序列、来源, **基因库**可分为:

- 核基因库、染色体库、cDNA库、线粒体库等。



①. 核基因库 :

**核基因库(genomic library):** 将某生物全部基因组DNA酶切后与载体连接构建而成。

**构建文库可采用不同的载体:**

**质粒载体:** 重组质粒导入感受态细胞 → 收集菌落。

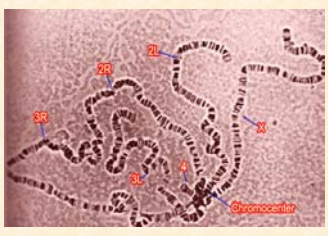
**噬菌体或柯斯质粒 (cosmid)载体:** 将重组DNA包装进噬菌体 → 感染细菌 → 收集噬菌斑。

**BAC(细菌人工染色体)或YAC(酵母人工染色体)载体:** 重组人工染色体 → 导入宿主细胞 → 收集细胞。

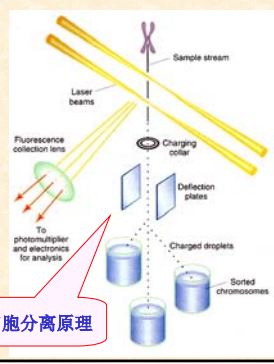
②. 染色体基因库:

将染色体用来构建基因库 → 可选择特异基因和分析染色体结构和组织。

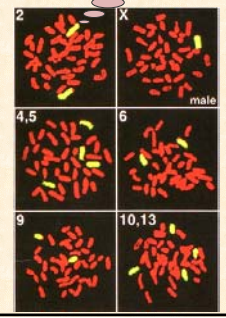
如果蝇的**多线染色体**: 对染色体进行**微切割**, 可以构建染色体区段的**基因文库**。



人类基因组项目研究中, 利用**流动细胞分离 (flow cytometry)**技术将人类染色体分开 → **构建单个染色体基因库** → 加速人类基因组作图和分析。



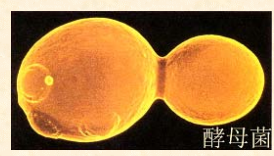
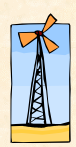
分离染色体进行涂色验证



流式细胞分离原理

**酵母菌:**

利用改良的**脉冲电泳方法(pulsed field gel electrophoresis, PFGE)**, 将酵母菌16根染色体**依据其分子量大小分开**后, 用于构建染色体库, 在基因组的分析中发挥作用。

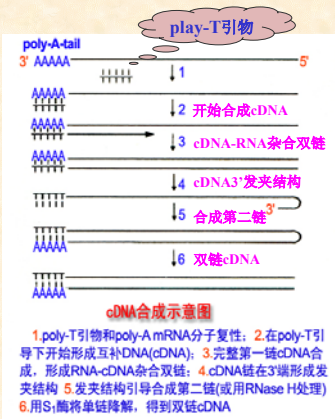


③. cDNA库:

以**mRNA**为模板 → 反转录酶合成**互补DNA** → 构建基因库。  
真核生物mRNA 3'端有一段多聚A尾端 → 双链DNA分子经两端补齐 → 以带限制性酶切点的人工接头连接 → 酶切后与载体(通常是噬菌体)连接 → 制备cDNA库。

**cDNA库与核DNA库不同:**

**cDNA库仅具有细胞或组织内表达基因的mRNA序列** → 仅包括基因组的**部分基因序列**。

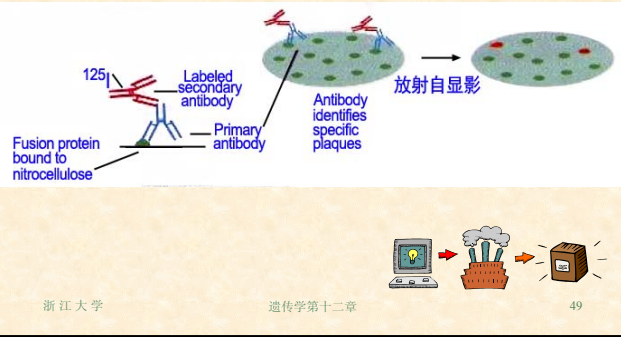






### ③. Western 杂交分析:

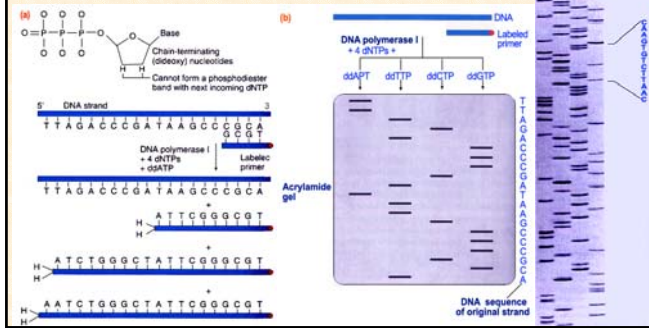
用于蛋白质的分析。



### (3). 核酸序列测定:

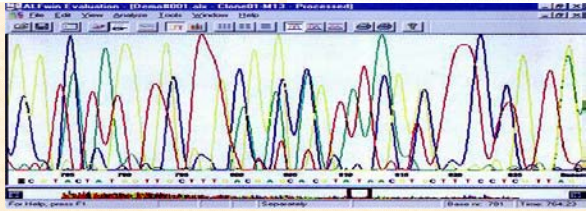
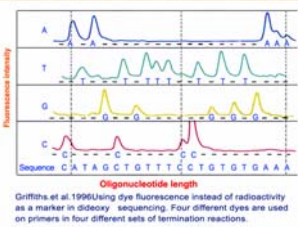
测定克隆后的DNA片段核酸序列。

一般采用Sanger(1977)发明的**双脱氧核糖核酸终止法**测定核酸序列。



在Sanger双脱氧法中, 可用荧光标记来代替放射性标记:

ALExpress 全自动激光荧光核酸测序及片段分析系统



### (4). 核酸序列分析:

测定的核酸序列是否为基因、有什么功能, 需用计算机软件或生物学实验进一步分析。

#### ①. 同源序列的分析:

##### ● 同源性比较:

将待测序列在核酸和蛋白质两个水平上**比较基因间的同源性**, 将序列发送到**Blast**等DNA Data数据库进行比较。

##### ● 阅读框架的分析:

一个ORF是一条能编码一条多肽链的DNA序列, 具有翻译起始信号和终止信号等。

cDNA 序列可分析内含子等。



### ②. 无任何同源性的新序列:

进行**基因功能性研究**的方法:

- 将基因导入生物体进行**基因失活**或**过量表达**, 根据**表型**推测基因作用;
- 用**噬菌体显现**(phage display)技术;
- **酵母双杂交**系统进行蛋白质的功能分析。



### (二). PCR扩增基因:

聚合酶链式反应(**polymerase chain reaction, PCR**)可以体外快速扩增DNA。

美国**Mullis(1986)**发明(现代生物学发展史上的**里程碑**)。

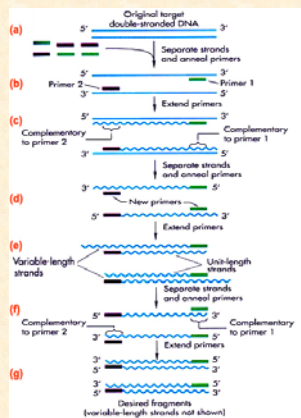


**PCR反应三个步聚(一个循环):**

- 1. 变性:** 94-95°C使模板DNA双链变成单链;
- 2. 复性:** 50-70°C下, 引物分别与互补DNA单链互补配对;
- 3. 延伸:** 在引物的引导和Taq酶作用下, 72°C下合成模板DNA的互补链。

PCR反应通常有25-35个循环, 一般可扩增5kb左右的片段。

由该技术衍生出一些新的技术, 如RAPD和AFLP等技术。

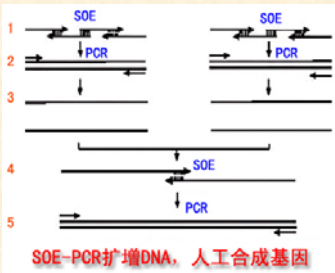


**(二)、人工合成基因:**

根据已知的基因或氨基酸序列, 将化学合成寡核苷酸的方法与酶促合成DNA的方法结合起来, 可很快地人工合成基因。

如SOE-PCR(sequence overlapped extension, SOE)技术, 可扩增出完整的基因序列。

1. 化学合成的寡核苷酸(80~100个核苷酸)通过SOE将单链部分补齐;
2. PCR扩增DNA片段;
3. DNA变性;
4. 两个单链部分经SOE补齐;
5. PCR扩增DNA, 利用多级SOE-PCR扩增出完整的基因。



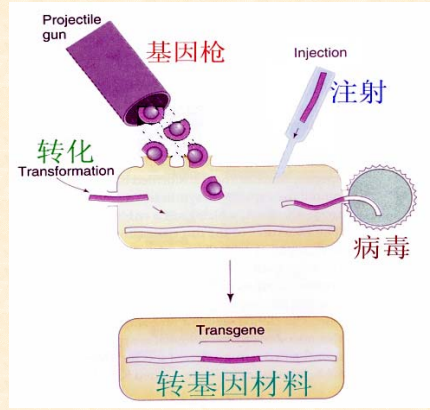
**五、基因工程的应用:**

目前, 基因工程研究发展迅速, 已取得一系列重大突破。基因工程技术已广泛用于工业、农业、畜牧业、医学、法学等领域, 为人类创造了巨大的财富。

具有生长激素的转基因鼠



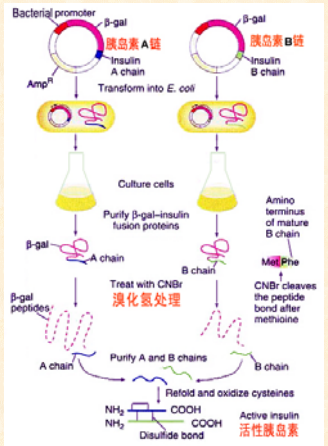
**转基因的方法和技术的:**



**(一)、基因工程工业:**

最早应用基因工程生产人的蛋白质的方法是在细菌中表达人的胰岛素(1982)。

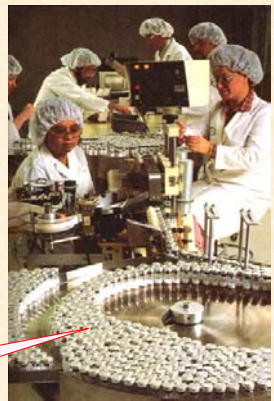
基因工程生产人的胰岛素的方法如图所示。



现已在细菌中生产10多种药品, 例如表皮生长因子、人生长激素因子、干扰素、乙型肝炎工程疫苗等。

目前, 酵母菌、植物悬浮细胞、植株和动物培养细胞均成功地应用于表达外源蛋白。

基因工程应用大肠杆菌生产人生长激素



## （一）、植物基因工程：

**植物基因转化**是指将外源基因转移到植物细胞内、并整合到植物基因组中稳定遗传和表达的过程。

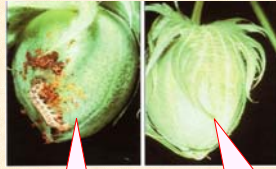
许多植物基因已经被分离、克隆。

**农杆菌转化法**和**基因枪转化法**应用最多。



对照

抗虫稻



对照

转基因抗棉铃虫品种

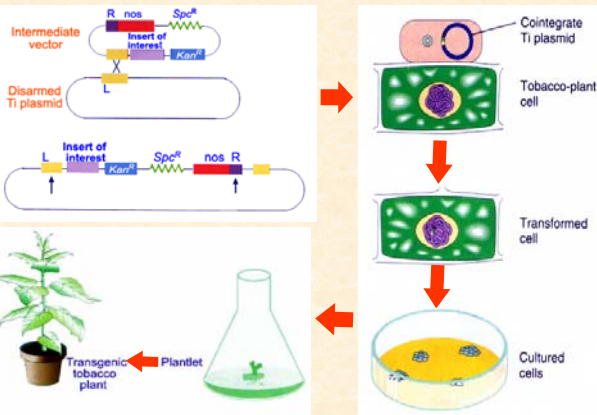
## 1. 根癌农杆菌转化技术：

**根癌农杆菌**介导的植物转化技术应用最早（双子叶植物 → 单子叶植物）。

**过程：**将目的基因与启动子（花椰菜病毒35S）及终止子组成嵌合DNA分子 → 插入到Ti衍生质粒RB与LB内构成**重组质粒** → 转化**农杆菌**细胞 → 利用重组农杆菌去**感染**植物细胞 → 使质粒部分DNA包括目的基因，整合到植物染色体，实现 → **遗传转化**。

利用**抗草甘磷**(glyphosate) *E. coli*中分离克隆的EPSP合成酶基因，已培育出高抗除草剂转基因植物。

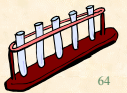
## 农杆菌转化法：



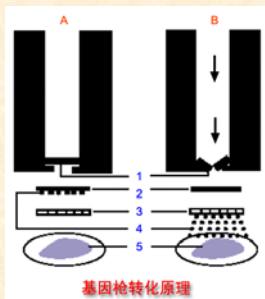
## 2. 基因枪转化技术：

以高压气体为动力，高速**发射包裹有重组DNA的金属颗粒** → 将目的基因直接导入植物细胞 → 整合到染色体上。

转化的载体多以 **pUC系列质粒**为基础构建 → 通常具有**细菌复制原点**及**抗性选择标记**、可在植物中表达的启动子终止子及调控序列、植物抗性选择标记(如除草剂、潮霉素等抗性)。



不同基因枪类型的**共同特点**：是用一种**动力系统**将金属微粒和包被DNA导入**受体细胞或组织**进行转化。



1. 位于高压气体桶底部的**爆破片**；
2. 载有重组DNA和金属微粒的**大载体**；
3. **阻挡板**；
4. 包裹有重组DNA的**金属微粒**；
5. 被轰击**样品**。

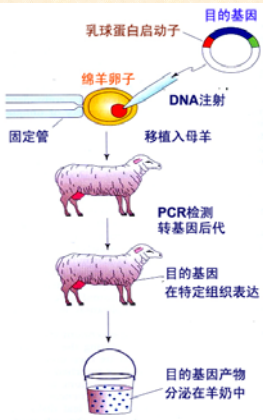
## （二）、转基因动物：

**转基因动物：**目的基因 + 载体 → 重组DNA → 微量注射法将重组DNA导入受体合子细胞核 → 遗传转化。如：利用转基因羊 → 大量表达人类的**抗胰蛋白酶**。可分泌人类生长**因子IX**的转基因羊“Polly”。



(Wilmut等, Nature 385, 1997)

## 受精卵细胞注射法



Transplantation of nucleus of a fertilized egg into an unfertilized egg using a glass pipette

转基因牛：受精卵注射法

转移有人类生长激素转基因鼠



浙江大学

遗传学第十二章

68



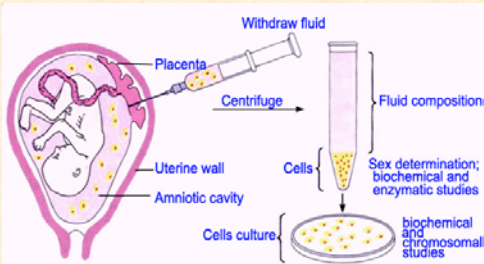
人类生长激素转基因猪

同胞鼠对照

## 四、遗传疾病诊断：

利用重组DNA技术进行遗传疾病诊断，是直接对DNA即基因水平进行诊断，准确度高，速度快。

如进行产前诊断，分析胎儿是否有遗传疾病。



羊膜穿刺术Amniocentesis

浙江

69

## 1. RFLP法：

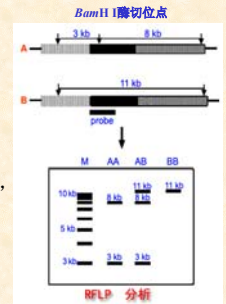
①. 原理：限制性内切酶能识别并切割特异核苷酸序列。

某位点核苷酸序列发生突变，有可能导致缺失或产生新的限制性酶切位点。

若突变只发生在同源染色体的一条上，另一条正常，限制性内切酶酶切DNA片段带型可用于鉴别同源染色体的差异。

酶切后产生的DNA片段长度的差异，称为限制性片段长度多态性 (restriction fragment length polymorphisms, RFLP)。

RFLP差异具有共显性特点。



浙江大学

遗传学第十二章

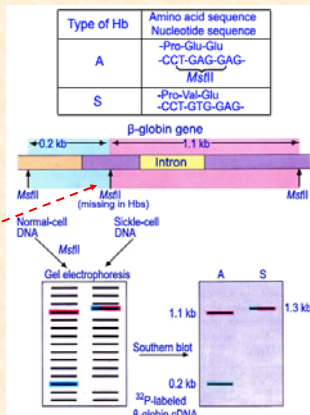
70

## ②. 镰刀性贫血病的诊断：

该病由β-球蛋白基因

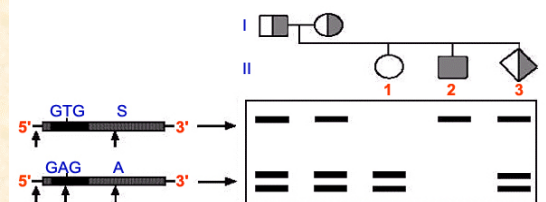
一个核苷酸突变 (由GAG变成GTG) 导致的，这个核苷酸改变也正好导致一个限制性酶切位点改变。

杂合体经酶切产生带型差异，识别该位点。



遗传学第十二章

71



镰刀性贫血病的 RFLP 诊断

一个核苷酸突变 (GAG-GTG) 使限制性酶酶切点发生改变 (A 为正常基因 β<sup>A</sup>, S 为突变基因 β<sup>S</sup>)。箭头处表示酶切位点。从 Southern 杂交结果即可知道这个家系各成员的遗传组成及表型。双亲 (I) 均为贫血病携带者 (杂合)，三个子代 (II) 中一个正常纯合，一个贫血病携带者 (杂合)，一个隐性突变基因纯合 (患病)。粗线表示与 β 球蛋白基因杂交的探针 (probe)。

浙江大学

遗传学第十二章

72

## 2. 等位基因特异寡核苷酸法(allele-specific oligonucleotide, ASO):

当已知突变基因的核苷酸序列时,可用人工合成的寡核苷酸进行PCR反应,将PCR产物用作点杂交分析,鉴定基因型。

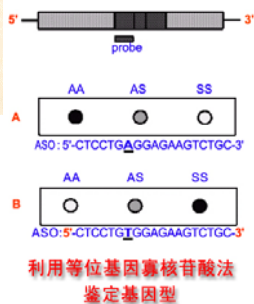
这种用于检测单个核苷酸变异的方法,称为ASO法。

许多遗传疾病可用该法诊断。

从血液细胞中提取DNA,分别用正常及突变的ASO进行PCR,杂交后的两种结果。

A、用正常基因( $\beta^A$ )的ASO进行PCR时,三种基因的信号强度。

B、用突变基因( $\beta^S$ )的ASO进行PCR时,三种基因的信号强度。



## (四)、基因治疗:

特异基因导入并整合到有遗传缺陷患者基因组中→治疗遗传疾病,通常叫做**基因治疗**(gene therapy)。

**单基因遗传病:** 白化病、先天性聋哑、血友病等。

**多基因遗传病:** 唇裂、腭裂、精神分裂症、先天性心脏病。



**方法:** 用减毒的病毒DNA

作载体(retrovirus DNA)

→利用**正常基因**构建重组

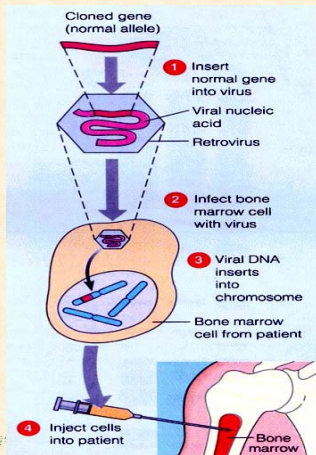
DNA分子→用病毒包装

物包装后形成的减毒病毒

感染患者的细胞(**骨髓细胞**)

→将正常基因整合到

染色体上。



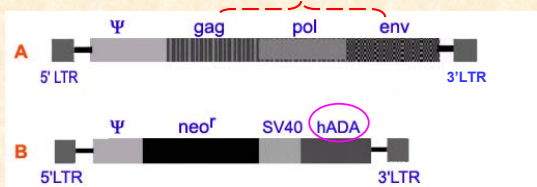
例如,用**莫洛尼氏鼠**白血病病毒(MLV)改造而成的减毒病毒载体,已用于治疗严重**综合免疫缺陷(SCID)**。

SCID是由于**腺苷脱氨酶基因**(adenosine deaminase, ADA)**突变**引起,患者无任何免疫功能。

**2000年法国Fischer**用该方法治愈患有SCID遗传病的两个婴儿(8个月和11个月)→被认为是**20世纪人类基因治疗上的重大突破**。



**方法:** 将ADA基因导入到MLV retro virus 载体→取代该病毒DNA中的**三个基因结构**(gag、pol、env)→将重组后的病毒去感染T细胞,使ADA基因整合到T细胞的染色体中→实现**基因治疗**的目的。



**基因治疗的载体和重组DNA分子**

我国基因治疗研究坚持“**有所为、有所不为,有限目标、重点突破**”的方针,取得了明显的**成绩**。

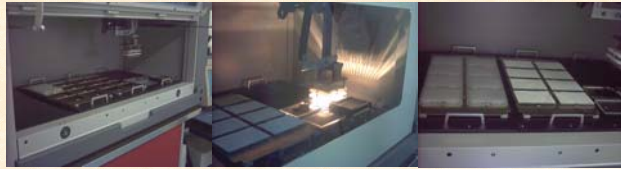
**1991年**,我国在世界上首次对B型血友病进行基因治疗临床研究;

**1995年:**基因治疗列入国家863计划生物领域重大项目。研制成功具有自主知识产权的**基因导入载体系统**;恶性肿瘤、B型血友病、梗塞性外周血管病等**7个基因治疗方案**已进入临床阶段;**20多个**具有自主知识产权的基因治疗方案将于2004~2007年进入临床研究。

**2004年1月:**深圳市**塞百诺基因技术有限公司**研制的**首个重组腺病毒p53基因(抑癌基因)**治疗药物“**今又生**”正式获准上市。

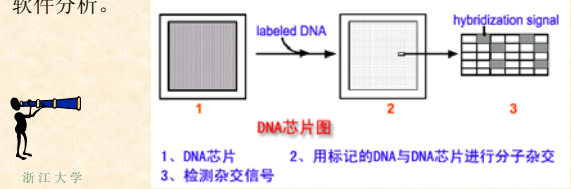
## 六、DNA芯片(DNA chips)：

核酸分子杂交的原理和方法 + 半导体技术结合 → 形成的一门新技术。这一技术可使许多分子杂交反应同时进行。



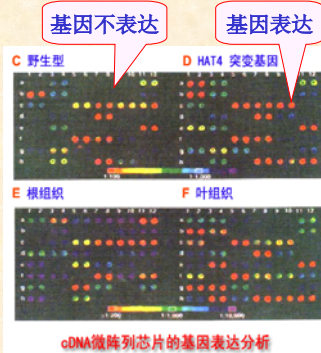
## DNA芯片技术：

在小面积(如2cm<sup>2</sup>)芯片表面分成不同小格 → 有序点阵排列在一定位置、可寻址的核苷酸分子 → 将待分析核苷酸分子标记(如用荧光)，**变性成单链与芯片上序列相同的核苷酸分子杂交** → 与芯片上序列不同的核酸分子被洗掉 → 利用高精度的激光扫描仪记录分子已杂交的荧光信号 → 计算机软件分析。



拟南芥48种cDNA经PCR扩增，微点样制成96孔cDNA芯片 → **野生型**和**HAT4突变型**mRNA探针分别用**荧光素**和**丽丝胺**标记，将两探针以1:1比例混合并杂交于同一芯片 → 分别用该两种荧光素激发波长进行扫描得到C、D图。

桔红色越深，基因表达丰度越高。E、F图也相同。



DNA芯片的**基质**：玻璃、硅片或尼龙等。

DNA芯片从几千个微点(用同位素标记) → 发展出160万点阵的芯片(荧光素标记)。

### 应用领域：

**基因多态性检测**(如**单核苷酸多态性**筛选，single nucleotide polymorphisms, **SNPs**)；

**基因表达分析**(不同细胞和不同组织的RNA群体比较)；

**克隆选择及文库筛选**(如cDNA文库)；

**基因突变检测及遗传病和肿瘤的诊断等**。



## 第二节 基因组学



### 一、概述：

#### 1. 基因组学：

**基因组学(genomics)**遗传学研究进入分子水平后发展起来的一个分支，主要**研究生物体内基因组的分子特征**。

#### 研究对象：

以**整个基因组**为研究单位，而不以**单个基因**为单位作为研究对象。

#### 研究目标：

**认识**基因组的结构、功能和进化；

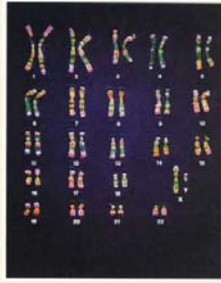
**阐明**整个基因组所包含的遗传信息和相互关系；

**充分利用**有效资源，预防和治疗人类疾病。



## 2.基因组计划

1990年美国启动被誉为“人体阿波罗计划”的“**人类基因组计划**”，投资30亿美元，计划历时15年**测定人类基因组的30亿个核苷酸对**的排列次序**构建**高分辨率的人类基因组**遗传图谱**和**物理图谱****发展**生物信息学，中国承担1%人类基因组计划测序工作。



### (1)人类基因组计划简史

1986年，美国杜伯克在《科学》上撰文，号召大家联合起来，从整体上把人类的基因组搞清。

1990年10月，经5年的辩论以后，美国国会终于批准被誉为生命科学“阿波罗登月计划”的国际人类基因组计划。

1998年5月，在美国罗克威尔组建私人公司：塞莱拉遗传公司，目标是投入3亿美元，到2001年绘制出完善的人体基因图谱，与国际人类基因组计划展开竞争。

1999年12月1日，国际人类基因组计划联合研究小组宣布，完整破译出人体第22对染色体的遗传密码。

2000年5月8日，由德国和日本等国际科研小组宣布，基本完成了人体第21对染色体的测序工作。

2000年6月26日，中、美、日、德、法、英等6国科学家公布人类基因组工作草图，标志着人类在解读自身“生命之书”的路上迈出了重要一步。

2001年2月12日，6国科学家和美国塞莱拉公司联合公布人类基因组图谱及初步分析结果。

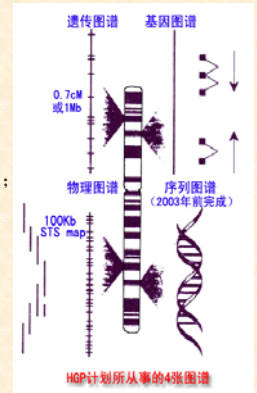
2003年4月14日，6国科学家宣布人类基因组序列图绘制成功，HGP的所有目标全部实现。覆盖人类基因组所含基因区域的99%，精确率达到99.99%，比原计划提前两年多，耗资27亿美元。并公开研究数据和结果，赢得了这场竞赛。

但国际HGP的协作组与私人公司的竞争仍没有停止过，目前在许多领域研究仍进行竞争，如在单个基因型的SNP图谱研究方面。



### (2)主要组成部分

- ①**构建**基因组的**遗传图谱**；
- ②**构建**基因组的**物理图谱**；
- ③**测定**基因组DNA的**全部序列**；
- ④**绘制**基因组的**转录本图谱**；
- ⑤**分析**基因组的**功能**。



### (3)中国参与人类基因组计划

1995年，杨焕明等人呼吁参与国际人类基因组计划。

1998年6月，遗传所人类基因组中心挂牌成立，由4位美国科学院院士及一位诺贝尔奖得主组成国际顾问组。

1999年4月，开始进行人类和嗜热菌基因组测序，大规模基因组测序的技术平台建设。

1999年9月1日，杨焕明在第五次伦敦国际人类基因组战略讨论会上介绍情况。会议正式接受中国加入国际合作，正式承担1%的测序任务。

1999年11月12日，科技部通过这1%的项目成为国家支持的重点项目。

2000年6月26日，包括中国在内的六国科学家共同宣布，HGP“工作框架图”绘制完成，这是人类历史上值得“载入史册的一天”。

2000年6月28日，江泽民就HGP“工作框架图”完成发表重要讲话。

2001年4月1日，运算速度超千亿次的曙光3000超级计算机落户杭州华大基因研究中心，“华大基因”的测序能力名列世界第六。

2001年8月26日，人类基因组计划中国部分测序项目汇报及联合验收会在京召开，标志人类基因组“中国卷”通过国家验收。

### (4)水稻基因组计划

在美国提出人类基因组计划后，**日本和中国**分别启动了“**水稻基因组计划**”。

2002年4月5日《Science》以14页的篇幅刊登和宣布中国科学家独立绘制完成水稻基因组草图序列（**总数：4.6亿**），是继人类基因组工作草图之后完成测定的最大基因组。



**材料:** 籼稻“9311”。

**完成单位:** 华大基因研究中心、中科院遗传与发育生物学研究所等12个单位, 2001年7月启动。

**水平:** 水稻基因组的总基因数约为46022~55615个(接近人类基因数的两倍), 工作框架图序列已覆盖水稻整个基因组92%以上的基因。

**方法:** “鸟枪射击法”, 利用国产曙光2000、曙光3000超级计算机(1000亿次/秒)对随机DNA碎片进行排序和组装, 确定其在基因组的正确位置。

**计划:** 功能基因分析和蛋白质研究。



### 国际水稻基因组测序计划:

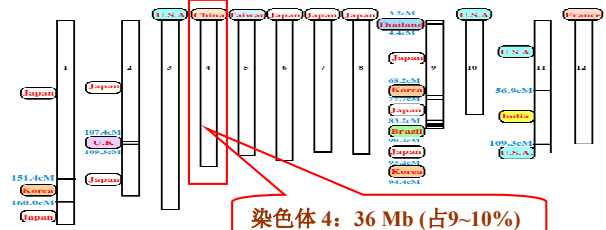
国际水稻(粳稻)基因组计划始于1998年,

日、美、中、法等国家和地区参加。

2004年4月美国Monsanto公司完成粳稻全基因组工作框架图。



chromosome sharing in IRGSP (February, 2002)



2002年12月21日《Nature》宣布中国独立

绘制完成**粳稻基因组第四号染色体精确测序图**, 我国对国际水稻基因组测序计划贡献率为10%。

**材料:** 粳稻“日本晴”。

**完成单位:** 中科院国家基因研究中心等4家单位。

**水平:** 第四号染色体中的**总碱基数目为0.35亿碱基对**, 覆盖了该染色体全长序列**98%**的区域, 只剩下7个小空洞, 碱基序列的精确度达到99.99%。完整测定的**着丝粒序列**在高等生物中属于首次。

**计划:** 对预测鉴定的**4658个**基因作进一步分析, 并对**着丝粒**功能等作研究。



### (5)模式生物基因组研究:

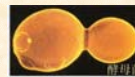
由于以人类为对象的研究实际上受到诸多限制, 也受伦理学的约束, 所以人类基因组计划还将对有关的模式生物, 如酵母、线虫、果蝇、小鼠、拟南芥等进行研究, 为人类的基因组研究提供重要的依据。

1996年, 完成酵母菌基因组测序。

1998年12月, 宣告完成线虫完整基因组序列的测定工作, 这是第一次绘出多细胞动物的基因组图谱。

2000年3月, 已完成了果蝇的基因组测序

2001年12月14日, 美、英等国科学家宣布绘制出拟南芥基因组的完整图谱。



浙江大学



遗传学第十二章



94

2004年3月1日, 科学家宣布绘制完成首幅

禽鸟类物种的**基因组序列草图**。

**材料:** 红原鸡(鸡的远祖)。

**完成单位:** 中国、美国、德国、英国、瑞典、荷兰等。

**水平:** 总碱基数目约为**10亿碱基对**, 该序列草图与人类基因组序列进行了并列比较(由美国华盛顿大学负责)。同时完成一张鸡遗传差异图谱(由中科院负责), 对红原鸡、肉鸡(英)、蛋鸡(瑞典)、乌鸡(中)进行序列差别分析。

**作用:** 鸡是研究**低等脊椎动物**和**人类等哺乳动物**的理想中介, 因此在禽流感防治、改善家禽和人类健康等方面有明显应用价值。



### 一、基因组图谱的构建:

基因组中核苷酸顺序的测定方法:

利用现有DNA测序方法, 每个测序反应通常只能得到800个核苷酸的序列。

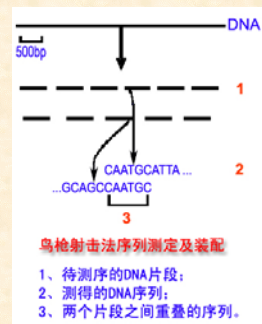
∴小基因组物种常用**鸟枪射击法**。

**大基因组测序存在两个问题:**

◎片段数(n)庞大, 片段间连接和装配非常复杂, 重叠数为 $2n^2-2n$ ;

◎基因组中相同或相似的重复序列在连接和装配时**容易出错**。

**解决方法:** 大规模序列测定前, 构建**基因组图谱**可作为序列测定中制定测序方案的依据, 以便锚定测知的核酸序列在染色体上的位置。





## 测序方法:



### 克隆连续序列法(clone contig):

将基因组DNA切割成长度为0.1~1Mb的大片段 →  
 克隆到YAC或BAC载体上 → 分别测定单个克隆的序列  
 → 再装配连接成连续的DNA分子。

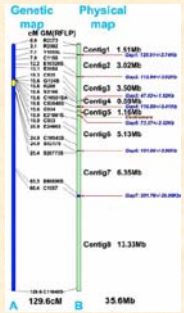
### 定向鸟枪射击法(directed shotgun):

以基因组图谱中标记为依据 → 测序装配和构建  
 不同DNA片段的序列。

## 基因组图谱根据构建的途径,可分为两种:

### 遗传图谱:

根据**遗传性状**(如已知基因位点、  
 功能未知的DNA标记、可鉴别的表型  
 性状)的**分离比例** → 将其定位在基因组  
 中,构建相应的连锁图谱。



### 物理图谱:

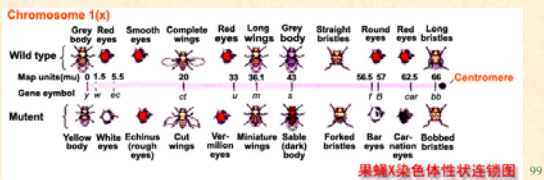
将各种标记**直接定位**在基因库中  
 的某一点上。

## (一)、遗传图谱的构建:

1. **图谱标记:** 可用**基因**或**DNA**作为可鉴别的标记(marker)。

(1). **基因标记:** 基因控制性状的表现,利用可鉴别的形态、  
 生化等表型性状作标记 → 根据**连锁交换原理**来分析基因  
 之间的连锁关系和遗传距离 → 绘制连锁图谱。

**缺点:** 基因数目有限、所构建遗传图谱不详细、标记间  
 的遗传距离较大。



果蝇X染色体性状连锁图

## (2). DNA标记:

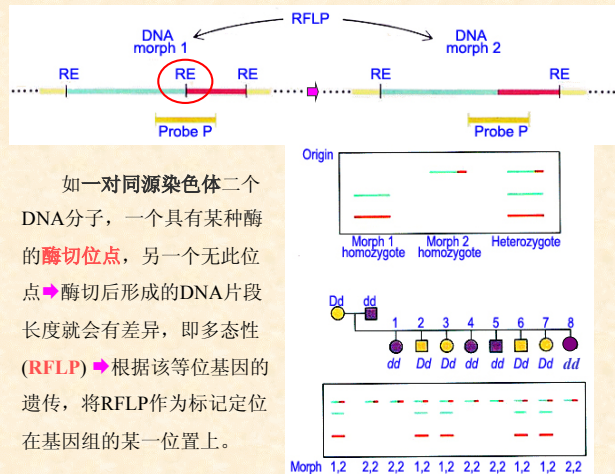


每种生物DNA具有稳定性。

∴ DNA本身可作为构建遗传图谱的标记,主要有以下3种  
 类型:

### ①. 限制性内切酶多态性:

限制性内切酶**能识别和切割特异核苷酸序列**, DNA  
 序列能不能被某一酶酶切,实际上相当于一对等位基因  
 的差异。



如一对同源染色体二个  
 DNA分子,一个具有某种酶  
 的**酶切位点**,另一个无此位  
 点 → 酶切后形成的DNA片  
 段长度就会有差异,即多态性  
**(RFLP)** → 根据该等位基因的  
 遗传,将RFLP作为标记定位  
 在基因组的某一位置上。

Morph 1,2 2,2 2,2 1,2 1,2 2,2 2,2 1,2 1,2 2,2

## ②. 简单序列长度多态性 (simple sequence length polymorphisms, SSLP):

SSLP是一些**长度不等的重复序列**、有**多个等位基因**位点。

- ① **多次重复序列**(variable number of tandem repeats, **VNTR**): 也称**微卫星**(minisatellite), 重复序列长度为几十个核苷酸。
- ② **简单重复序列**(simple tandem repeats, **STRs**): 又称为**小卫星**(microsatellite), 常只有2~4个核苷酸重复单位。

**利用STR作为标记更为普遍:**

∴ **STR**分布在整个基因组中,而**VNTR**主要集中在染色体末端;  
**STR的PCR检测速度更快、准确性更高**,因为STR长度  
 通常在300bp以下,而VNTR相反。



### ③. 单核苷酸多态性

(simple nucleotide polymorphisms, SNPs):

基因中**点突变**数量多, 其中**有些可产生RFLP** (发生在**酶切位点**)。

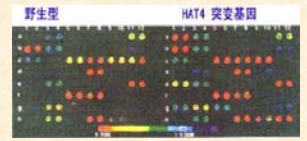
人类基因组中的编码序列中约有20万个SNPs标记, 非编码序列中约有**多10倍以上的标记**。

CTGGTCGT**CAGT**CTTTAGTT  
GACCAGCAGT**CA**GAAATCAA

↑  
SNP

### SNP位点可用寡核苷酸分子杂交来检测:

- ⊙ **DNA芯片技术:** 将不同的寡核苷酸固定在芯片上 → 标记待测DNA → 一次就可检测很多SNPs标记;
- ⊙ **特异等位基因动力学杂交法:** 利用液体杂交法检测, 在酶标偶联反应板的96孔中进行分子杂交 → 用一种只与双链DNA结合的荧光染料检测 → 当两条DNA分子完全互补配对时, 形成双链DNA时才有信号 → 检测SNPs。



## 2. 遗传图谱的构建:

### (1). 人类基因组遗传图谱的构建:

**家系分析法:** 分析8个家系134个成员 (186个减数分裂)

→ 根据5264个**STR(简单重复序列)**标记绘制而成 (X染色体分析是利用12家系170个成员 (105个减数分裂)) → 将5264个标记定位在2335个位点 → 构建的人类基因组遗传图谱密度为**每个标记599 kb**。



## (2). 植物基因组遗传图谱的构建:

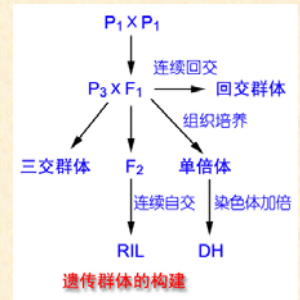
### ①. 选择亲本:

亲缘关系远, 遗传差异性大, 亲本间分子标记多态性强。

### ②. 产生构图群体:

配制组合 → **建立分离群体:**

- ⊙ 单交组合产生的F<sub>2</sub>;
- ⊙ 衍生的F<sub>3</sub>、F<sub>4</sub>家系;
- ⊙ 由连续多代自交或姐妹交产生的重组近交系(recombinant inbred lines, RIL);
- ⊙ 通过单倍体加倍而成的双单倍体(doubled haploids, DH);
- ⊙ **利用回交或三交(复交)**产生的后代群体。



### ③. 遗传标记的染色体定位:

有单体、三体、代换系与附加系分析等方法, **依据染色体剂量的差异** → 将遗传标记定位在特定染色体上。

⚡ 当供体材料总DNA等量时, **DNA杂交带的信号强弱**与该标记位于的染色体剂量成正比。

### ④. 标记间的连锁分析:

通过分析分离群体内双亲间有**多态性的遗传标记的连锁交换情况和趋于协同分离**的程度 → 即可确定标记间的连锁关系和遗传距离。



## (二)、物理图谱的构建:

### 1. 构建物理图谱的原因:

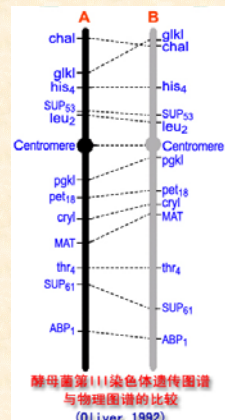
#### (1). 遗传图谱的分辨率有限:

**低等生物**可重复进行多次杂交、测交 → 在短期内可产生大量群体、得到大量重组交换的子代群体 → 构建高密度的遗传图谱。

**人类或其它高等生物**难以获得大量的子代群体, 减数分裂的后代有限 → 限制连锁分析。

#### (2). 遗传图谱的精确性不高:

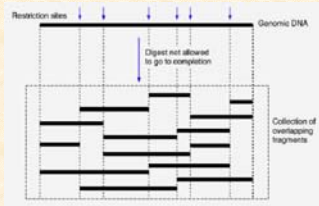
染色体上有一些**重组热点**, 其发生重组的频率高于其它位点 → 影响重组热点邻近区段的遗传图谱**准确性**。



## 2. 构建物理图谱的三条途径:

### (1). 限制性酶图谱:

利用限制性内切酶绘制图谱, 对于DNA分子长度 **小于50kb** 的片段, 一般没有困难。而对于 **>50kb** 的DNA分子, 可选用稀有切点内切酶酶切DNA。



### ◎ 用识别位点较多核苷酸的内切酶:

如 *Not I*, 为8个核苷酸出现的频率为  $1/4^8=1/65536\text{bp}$ ; 而识别位点为6个核苷酸的出现频率为  $1/4^6=1/4096\text{bp}$ 。

### ◎ 选用其酶切位点在基因组中出现频率低的内切酶:

人类基因组中,  $5'-\text{CG}-3'$  出现的频率很低:

*Sma I* ( $5'-\text{CCCGGG}-3'$ ) 酶切DNA, 每78kb有一个切点;

*BssH II* ( $5'-\text{GCGCGC}-3'$ ) 酶切DNA, 每390 kb有一个切点;

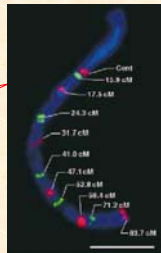
*Not I* ( $5'-\text{GCGGCCGC}-3'$ ) 酶切DNA, 每10 Mb只有一个酶切点。

### (2). 荧光原位杂交 (fluorescent in situ hybridization, FISH):

另一种 **物理图谱构建方法**: 荧光标记的探针与DNA分子杂交  $\rightarrow$  染色体上 **杂交信号** (即探针DNA在染色体上的图谱位点) 在显微镜下可直接观察。

**步骤**: 取有丝分裂中期的细胞制片  $\rightarrow$  将染色体变性成单链  $\rightarrow$  将标记DNA探针变性后杂交到染色体上  $\rightarrow$  观察分析。

应用多色荧光原位杂交技术将10个BAC探针定位于水稻的第10号染色体的长臂上 (引自Cheng等, *Genetics*, 2001)



### (3). 序列标签位点:

**已知序列为标签的位点** (sequence tagged sites, **STS**)

作探针, 与基因组DNA杂交  $\rightarrow$  绘制物理图谱。

**STS的要求:**

- ◎ 为已知序列, 便于PCR检测;
- ◎ 基因组中仅此一个位点, 无重复。



## 常用的两大类型STS:

### ①. 特异DNA序列:

- ◎ **表达的序列标签** (expressed sequence tag, **EST**), 表达的序列来自cDNA。
- ◎ **简单序列长度多态性 (SSLP)** 也可用于物理图谱的构建, 比较遗传图谱和物理图谱。
- ◎ **DNA随机片段**, 测序后经分析比较, 非重复序列可作为STS。

### ②. DNA片段:

分为两种不同类型:

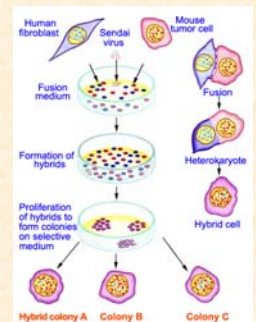
- ◎ **辐射杂交系 (radiation hybrid)**: 指啮齿动物细胞含有另一个生物染色体片段的细胞株。



## 细胞遗传学研究表明:

用3000~8000radX射线处理后  $\rightarrow$  **人类染色体会出现随机断裂**  $\rightarrow$

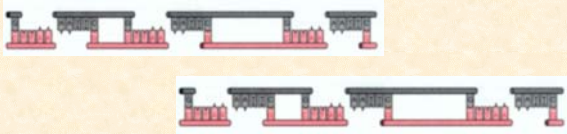
将处理细胞与正常仓鼠或其它啮齿动物细胞融合  $\rightarrow$  断裂的人类染色体可与仓鼠染色体融合  $\rightarrow$  复制  $\rightarrow$  利用100个人类基因组辐射后与缺失仓鼠细胞融合形成的辐射杂交系 (能在HAT培养基上生长的就是含有人类染色体片断的细胞  $\rightarrow$  进行人类基因组物理图谱的构建。



### ◎克隆的DNA片段序列:

利用YAC或BAC克隆的DNA, 测序后可作为分子标记。

不同克隆测定后, 进一步分析克隆之间的重叠部分, 将各克隆连接进来, 可作为STS使用。



### 3. 人类基因组物理图谱:

人类基因组序列开始测定时, 已有45万个EST, 其中有一些重复序列→经计算机分析筛选后得到49625个, 各代表一个基因→再从中筛选出3万个EST、二个辐射杂交系库(分别有83和93个细胞株)、一个有32000个克隆的YAC文库用于构建图谱。

构建的物理图谱密度为每个标记183kb, EST分布结果表明基因在染色体上的排列是不均匀的。

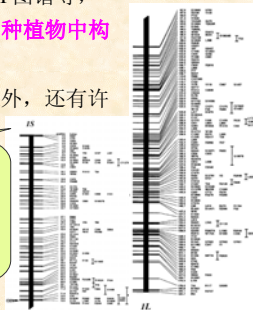
将上述遗传图谱及其它物理图谱整合→构成更加完整的人类基因组图谱, 作为基因组序列测定的框架和分析依据。

### 二、基因组图谱的应用:

基因组图谱中除了构建遗传图谱和物理图谱外, 还可进一步根据标记类型细分为不同的称谓。如RFLP图谱、RAPD图谱、AFLP图谱等, 目前已在拟南芥、番茄、水稻等30多种植物中构建了基因组图谱。

基因组图谱的构建除了进行测序之外, 还有许多用途。

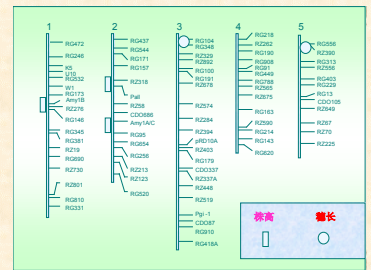
高密度水稻遗传连锁图(Harushima等, Genetics 1998): 2275遗传标记, 覆盖水稻基因组1521.6cM, 标记间距离为: 0.67cM。



### 1. 基因定位:

借助基因组图谱, 可使基因定位在精度、深度、广度等方面有极大的提高。已陆续在一些农作物上定位了许多重要农艺性状和经济性状的基因,

在复杂数量性状位点(QTL)定位分析方面, 也取得了很大进展。



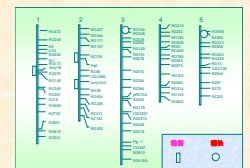
### 2. 基因组比较分析:

已经在禾本科玉米、水稻、高粱、大麦、小麦和黑麦, 茄科番茄、胡椒、马铃薯, 十字花科拟南芥、甘蓝、花椰菜、油菜等进行遗传图谱比较分析, 并从分子水平上了解物种间同源性, 研究基因组的进化和染色体的演变。



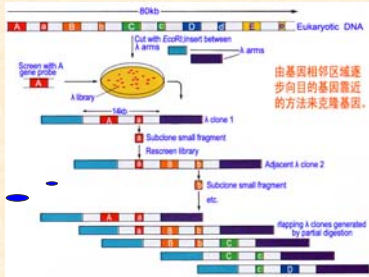
### 3. 标记辅助选择(marker-assisted selection MAS):

饱和基因组图谱可用来确定与任何一个目的基因紧密连锁的分子标记→根据图谱间接选择目的基因, 可降低连锁累赘, 加速目的基因的转移与利用, 提高回交育种的效率。



#### 4. 基因的克隆与分离:

根据饱和基因组图谱,可找到一个与目的基因紧密连锁的分子标记,作为染色体步行(chromosome walking)起始点→进行基因的克隆和分离,此法亦称图位克隆法(map-based cloning),为基因产物未知的基因克隆提供了捷径。



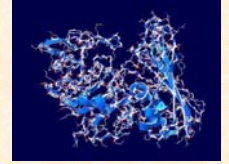
染色体步行法

### 三、后基因组学(post-genomics):

#### (一)概念及问题

##### 1.概念

在完成基因组图谱构建以及全部序列测定的基础上→研究全基因组的基因功能、基因之间相互关系和调控机制为主要内容的学科。



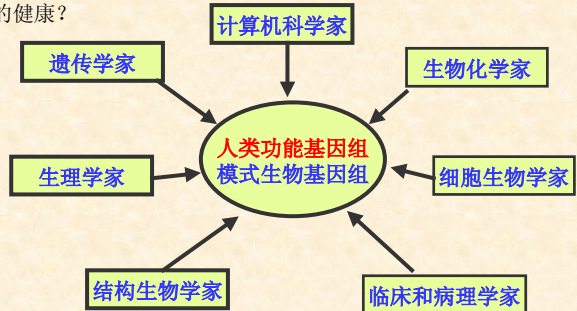
#### 2.面临的最重要问题

2003年4月14日,宣布人类基因组序列图绘制成功。目前,虽然完成了99%人类基因的序列分析,但90%以上的人类基因尚不了解功能。

自从1996年酵母菌基因组全序列测定以来,全世界已有1000多个实验室、5000多名科学家从事酵母菌后基因组学的研究,共发表论文7000多篇,鉴定1060个新基因的功能,但仍然还有约1600个阅读框架(基因)的功能不清楚。



如何将人类和模式生物体基因序列资料更好地用于生命科学?如何对这些基因加以利用,使之能够造福于人类的健康?



#### 3.目前已衍生出了许多新兴学科:

##### ①功能基因组学研究

对基因及其编码蛋白的功能进行研究。

##### ②疾病基因组学研究

发现疾病相关基因和致病基因,从疾病诊断到疾病易感性研究。

##### ③药物基因组学(Pharmacogenomics)

研究不同个体对药物敏感性的基因基础,特别是SNPs。

##### ④环境基因组学(Enviromental Genomics)

鉴定机体暴露在特定环境下的那些显示易感性或抵抗性基因的DNA多态性。如DNA修复基因、细胞周期相关基因、激素代谢基因、受体基因、参与免疫和感染反应的基因和信号转导基因等。

##### ⑤蛋白质组学(Proteomics)

利用双向电泳技术研究细胞或组织的基因组表达的全部蛋白质。

##### ⑥生物信息学(Bioinformatics)

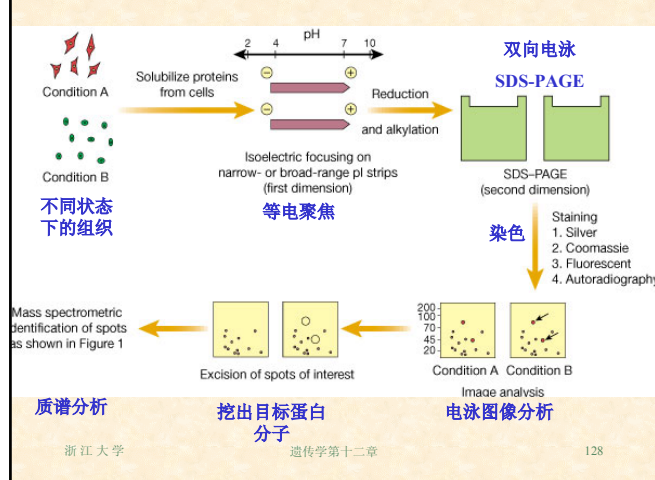
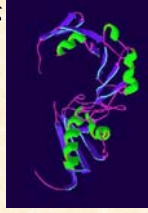
现代生物技术与计算机科学的结合,收集、加工和分析生物资料和信息。

## (二)蛋白质组学 (Proteomics)

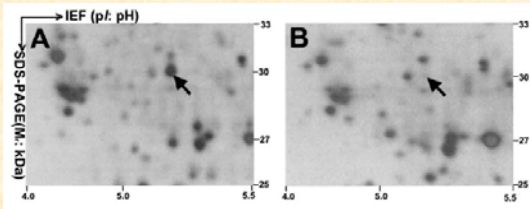
在蛋白质水平研究基因组的基因表达→分析基因组的蛋白质类型、数量、空间结构变异以及相互作用的机制。

蛋白质组学比基因组学更为复杂:

∴DNA线状结构与二级结构的功能差异不大→但多肽链需折叠成一定的三维空间结构才形成有功能的蛋白质;同一种蛋白质经不同的加工修饰可形成不同的功能→因此蛋白质的多样性远复杂于基因本身。



### 蛋白质2D电泳分析

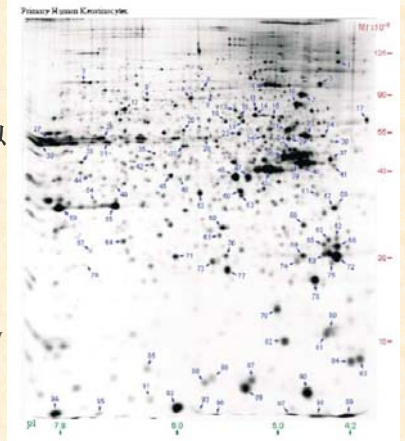


A对照; B为人的垂体瘤蛋白质2D电泳  
箭头示有明显差异的蛋白质。

引自 Zhan et al. (2003) Pituitary 6: 189-202.

人的角化细胞的2D电泳蛋白质图谱,经S35放射自显影显示,可以分辨出100种以上的蛋白质。

(引自 Gromov et al. Progress in Biophysics & Molecular Biology 2002)



## (三)、生物信息学(bioinformatics)

生物信息学是现代生物技术与计算机科学的结合,收集、加工和分析生物资料和信息学科。

应用生物信息学可以来自不同的基因组理论和应用综合并标准化,利用大量的生物信息资料了解遗传网络系统、信号传递及相互关系,计算机还可进行一些生物模拟研究。

∴利用生物信息学能够分析从微生物、动物、植物以及人类基因组序列测定产生的大量资料→阐明遗传信息。

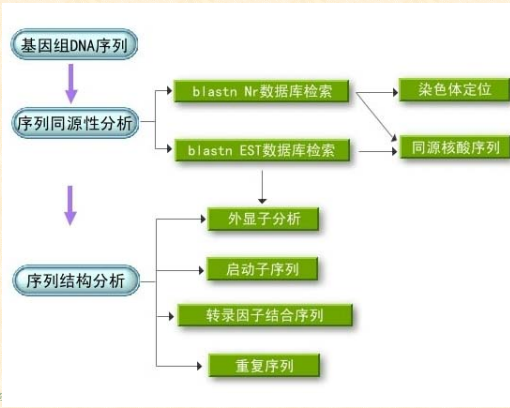
研究内容两大类: DNA 数据分析;  
蛋白质数据分析。



### 1.DNA数据分析

- ①基因结构域分析,包括启动子、转录因子结合序列、内含子、外显子、重复序列、开放读码框架等。
- ②同源分析和检索,包括Nr数据库、EST数据库、STS数据库、Unigene数据库、Swissprot数据库等。
- ③计算机基因克隆化:利用EST数据库的重叠序列克隆新基因。

## 基因组序列分析过程



浙江大学

## cDNA序列分析

cDNA: Complementary DNA, 与RNA序列互补的DNA序列, 一般代表基因的外显子序列, 即蛋白质编码序列。



浙江大学

遗传学第十二章

134

## 2.蛋白质的数据分析

### ①蛋白质一级结构分析:

结构特点分析, 包括等电点、信号肽、穿膜区、DNA结合序列等

同源分析和检索, 包括Nr数据库、Swissprot 数据库等功能区分析, 包括Prosite、Emotif、Identify分析等。

### ②蛋白质空间结构分析:

蛋白质晶体结构数据库检索, 如PDB数据库。

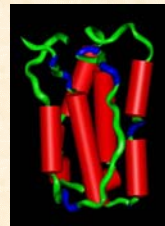
蛋白质空间结构预测, 如Homology等软件分析。

浙江大学

遗传学第十二章

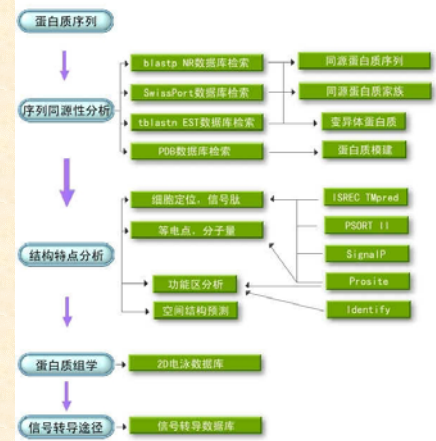
135

## 蛋白质序列分析过程



干扰素α1b分子建模图

浙江大学



## (四)、功能基因组学 (Functional Genomics)

功能基因组学研究将是21世纪生命科学研究的新热点;

功能基因组学研究涉及众多的新技术, 包括生物信息学技术、**生物芯片技术**、**单核苷酸多态性**、**转基因**和**基因敲除技术**、**酵母双杂交技术**、**基因表达谱系分析**、**蛋白质组学技术**、**高通量细胞筛选技术**等。

浙江大学

遗传学第十二章

137

## 1.DNA芯片技术(DNA chip)

又称**微阵列(DNA microarrays)**, 利用DNA芯片技术同时**进行大量分子杂交**, 分析比较不同组织或器官的**基因表达水平**, 筛选**突变基因**, 从**核酸水平**分析**基因表达模式**。

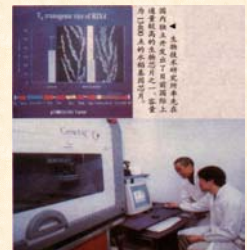
可用于**疾病诊断**、**药物筛选**等领域研究。

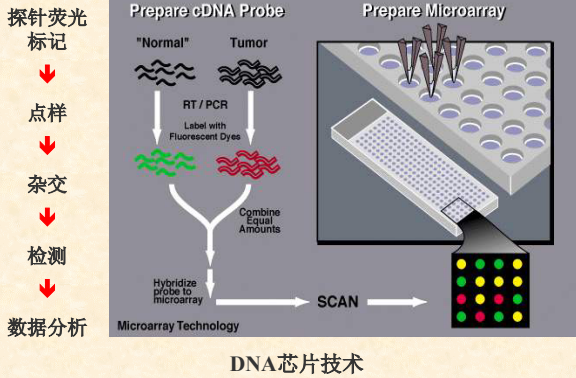


浙江大学

遗传学第十二章

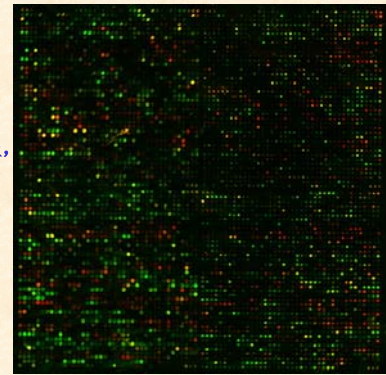
138





## 基因芯片展示酵母菌在发酵状态转向呼吸状态下的基因表达

(在 18 x 18 mm 面积上, 6400 基因表达情况, 引自 Science, 278: 680, 1997)



## 2. 酵母菌双杂交系统(two hybrid-systems) :

利用酵母菌同一个细胞中共同表达不同蛋白质, 以**鉴定蛋白质之间互作**的一种分析方法。

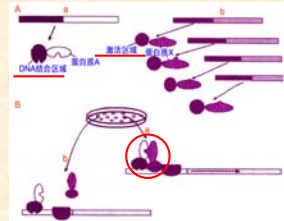
真核生物的基因表达受转录因子控制。

**转录因子**具有**DNA结合**及**激活转录**两功能, 分别由不同区域完成, 而且将这两部分**分割成两个片段**后分别克隆→仍能装配成一个完全有功能的转录因子。



## 酵母双杂交系统设计思路如下:

- (1) 当**报告基因**(如半乳糖苷酶基因)缺少转录因子时→处于关闭状态;
- (2) 将酵母菌**转录因子 DNA结合区域**和**RNA酶激活区域**分开→分别克隆;
- (3) 将**DNA结合区域**与待测目的基因(**蛋白质A**)和报告基因共同串连起来→形成融合基因。同时将**激活区域序列**与另一类DNA序列(**蛋白质X**)也串连起来;
- (4) 将两类融合分子**共转化酵母细胞**, 如**蛋白质X**能与**蛋白质A**互作结合、RNA聚合酶激活因子与启动子区域DNA结合→使报告基因表达, 易于鉴定。



## 本章小结

### 1. 遗传工程:

- 转基因技术的现状和发展;
- 载体条件与限制性内切酶;
- 目的基因获取与体外重组;
- 基因工程在动植物育种、人类医疗中的应用。



### 2. 基因组学:

- 基因图谱构建:** 遗传图谱和物理图谱;
- 基因图谱应用:** 基因定位、比较基因组学、MAS、基因克隆;
- 后基因组学:** 蛋白质组学、功能基因组学、生物信息学等。

- 第一章 绪言
- 第二章 遗传的细胞学基础
- 第三章 孟德尔遗传
- 第四章 连锁遗传和性连锁
- 第五章 数量性状遗传
- 第六章 基因突变
- 第七章 染色体变异
- 第八章 细胞质遗传
- 第九章 细菌和病毒的遗传
- 第十章 遗传物质的分子基础
- 第十一章 基因表达与调控
- 第十二章 基因工程与基因组学
- 第十三章 遗传与发育
- 第十四章 群体遗传与进化
- 返回首页

