

## 第七章 基因组学及相关组学

### 第一节 基因组学

#### Section 1 Genomics

基因组学 (genomics) 是指对整个基因组的结构、结构与功能的关系以及基因之间相互作用进行探索的一门科学。

#### 一、人类基因组计划的主要研究内容

- 人类基因组计划 (human genome project, HGP) 最初由美国科学家于20世纪80年代中期提出, 其目标是用15年时间, 测定出人类基因组DNA的全部碱基序列。
- 该计划从1990年开始实施, 先后有六个国家的数千位科学家直接或间接参与了HGP的工作。
- 实际上, HGP最终需要完成四张图谱: 遗传图、物理图、转录图和序列图。

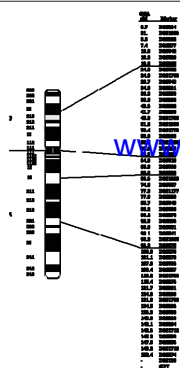
人类染色体图谱



#### (一) 遗传图谱

- 遗传图谱 (genetic map) 是以具有遗传多态性的遗传标记作为“位标”, 以遗传学距离作为“图距”所作出的基因组图谱。
- 常用的遗传标记包括限制性片段长度多态性 (RFLP) 位点、短串联重复序列 (STR) 和单核苷酸多态性 (SNP) 位点等。
- 遗传学距离通常用“分摩尔根 (cM)”表示。

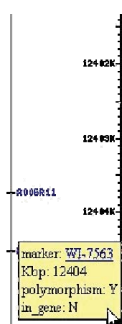
遗传图谱示意图



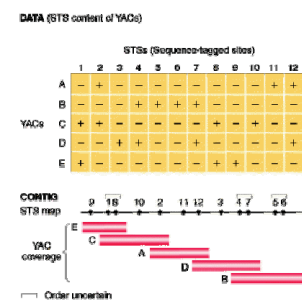
#### (二) 物理图谱

- 物理图谱 (physical map) 是以一段已知碱基序列的DNA片段 (碱基序列具惟一性) 作为“位标”, 以DNA分子的实际长度作为距离所作出的基因组图谱。
- 这些已知碱基序列的DNA片段被称为序列标记位点 (sequence tagged sites, STS)。

物理图谱示意图

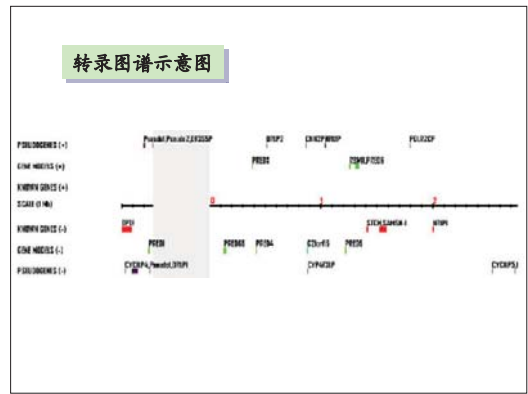


使用STS排列克隆的DNA序列片段



### (三) 转录图谱

- 转录图谱 (transcriptional map) 是以已知的基因表达序列作为“位标”，以DNA分子的实际长度作为距离所作出的基因组图，即基因表达序列图。
- 这些已知的基因表达序列被称为表达序列标记 (expressed sequence tags, EST)。



### (四) 序列图谱

- 序列图谱 (sequence map) 即指完整的人类基因组DNA碱基序列图谱。
- 人类基因组全部DNA碱基序列的测序工作于2003年4月基本结束；2006年4月，最后一个染色体的测序工作完成。

## 二、人类基因组计划的医学中的意义

- 人类基因组计划在医学中的意义主要体现在：
  - 通过对基因结构与功能的研究，推动了在分子水平上对各种生理功能的认识。
  - 推动了基因组信息与疾病易感性的研究。
  - 促进了对癌症的研究。
  - 促进了对疾病遗传学背景的研究。
  - 促进了药物基因组学的研究。

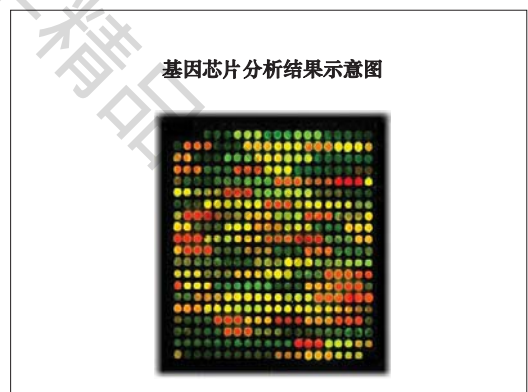
## 第二节 后基因组学

### Section 2 Functional Genomics

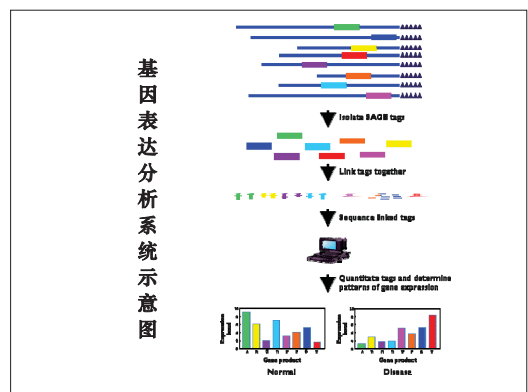
- 功能基因组学研究基因组中所有基因的功能，它从整体水平上研究某种组织或细胞在同一时间或同一条件下所表达的基因的种类、数量及功能，或同一细胞在不同状态下基因表达的差异。

### (一) 基因表达谱的研究

- 基因表达谱是指某种细胞在特定条件下所表达的全部基因的种类和数量。研究基因表达谱的主要技术方法有：
  - 基因芯片技术 (DNA chip) :
    - 即将大量已知序列的cDNA探针固定于支持物上，然后与样品所表达的全部cDNA进行杂交。



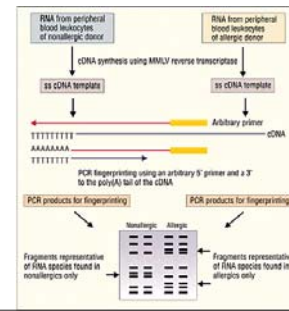
- 基因表达分析系统 (SAGE) :
  - 将细胞内所表达的mRNA逆转录为双链cDNA后，获取不同cDNA的序列标签并进行连接和扩增，再分析其表达模式图。
  - 此法可快速详细地分析成千上万的基因。



### 3. 差异显示PCR (DDRT-PCR) :

- 将若干种不同mRNA的5'-端引物与3'-锚定引物随机组合后进行RT-PCR扩增, 然后分析其产物的种类和表达量。

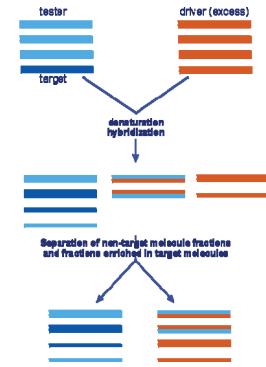
差异显示PCR示意图



### 4. 消减杂交法 (SH) :

- 即将对照组的mRNA或cDNA与实验组的cDNA或mRNA进行杂交, 去除杂交体中所含对照组的成分即为实验组中独特的表达成分。

消减杂交示意图



### (二) 基因产物-蛋白质功能的研究

- 通过基因重组并进行体外表达其蛋白质以确定基因的功能。

### (三) 基因组多样性的研究

- 人类基因组的多样性决定了个体之间生物性状的差异以及对疾病的易感性的不同, 因此开展基因组多态性的研究有助于了解人类的起源、进化过程以及与疾病的相关性。

### (四) 疾病相关基因的再测序

- 疾病相关基因的再测序有助于了解基因及其变异与疾病发生的相关性。

### (五) 寻找人类基因共同变异型

- 通过对单核苷酸多态性的研究并进行基因型, 有助于了解疾病易感群体和遗传疾病群体。

### (六) 基因组功能的注释

- 即利用生物信息学手段, 高通量地注释基因组所有编码产物的生物学功能。主要包括:

1. 基因组组成元素的识别。
2. ORF产物功能的预测。
3. 基因之间相互作用及比较基因组学研究。

## 第三节 蛋白质组学

### Section 3 Proteomics

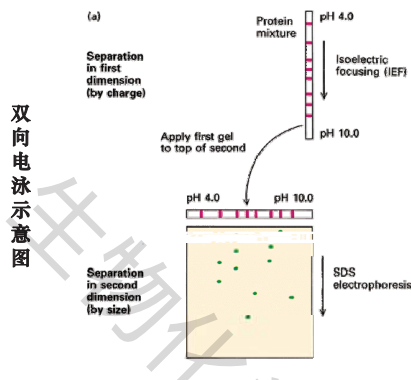
- 蛋白质组学 (proteomics) 是指在亚细胞、细胞、组织或器官的整体水平上, 研究其蛋白质组成及其活动规律的一门科学。
- 这些研究结果促进了21世纪一门新的学科——系统生物学 (system biology) 的产生和发展。

- 蛋白质组学主要研究某种生物体、组织器官、细胞或亚细胞结构在特定条件下所表达的全部蛋白质的种类和数量。
- 蛋白质组学的研究远比基因组学的研究更为复杂、难度更大。

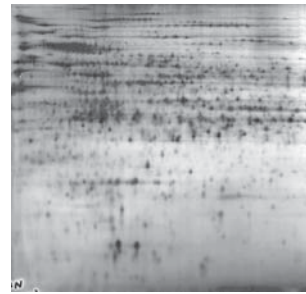
双向电泳分离技术、生物信息学处理技术以及蛋白质鉴定技术被称为蛋白质组研究的三大技术。

### (一) 双向电泳分离技术

- 利用蛋白质分子所带电荷及分子量的不同，将混合的蛋白质样品首先置于IPG胶上进行等电聚焦（IEF）分离，然后再置于变性聚丙烯酰胺凝胶（SDS-PAGE）上进行分离。
- 此法通常可分离到1000~10000种不同的蛋白质。



双向电泳分离结果图



### (二) 生物质谱与蛋白质鉴定

目前，对蛋白质的鉴定主要根据其属性参数，如分子量、等电点、氨基酸组成、肽序列标签及肽质量指纹图谱等，在生物信息数据库中进行检索进行鉴定。

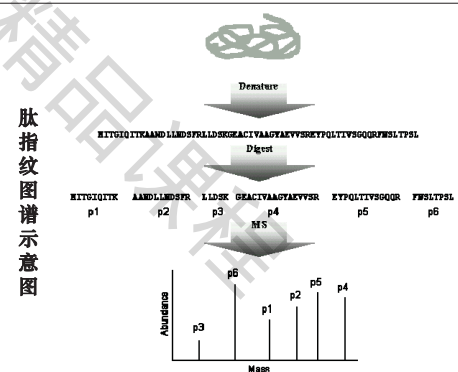
#### 1. 肽质量指纹谱鉴定技术：

- 将蛋白质用特异的蛋白酶水解成片段后，再使用质谱仪测定其肽片段质量所获得的指纹图谱。
- 不同的蛋白质被酶水解后产生的肽片段不同，所获得的指纹图谱也不同，故可用于蛋白质鉴定。

### 基质辅助激光解吸电离(MALDI)质谱仪



[www.med126.com](http://www.med126.com)



#### 2. 肽序列标签串联质谱鉴定技术：

- 采用串联质谱仪测定蛋白质多肽链中一定长度的肽片段的序列，并以此作为标签鉴定蛋白质。

### 串联质谱仪



## 二、蛋白质组数据库

- ✦ 进行蛋白质组学研究广泛使用的生物信息数据库主要是Swiss-Prot/TrEMBL建立的蛋白质分析专家系统（ExPasy）。
- ✦ 该数据库已收集了大量已知的蛋白质序列、肽指纹图谱及序列标签，对于蛋白质的鉴定非常有用。

## 三、蛋白质组学的应用

- ✦ 随着蛋白质组学研究的发展，目前已衍生出其他相关的蛋白质组学研究，如比较蛋白质组学、功能蛋白质组学、疾病蛋白质组学等，因此对蛋白质组学的研究可深入到医学研究领域的各个环节中。

生物化学与分子生物学精品课程