

第四节 转录组学

- ◆ 概念
- ◆ 研究内容
- ◆ 研究方法



概 念

□ **转录组**：是某个物种或者特定细胞类型产生的所有RNA的集合。

□ **转录组学**：是对转录水平上发生的事件及其相互关系和意义进行整体研究的一门学科。



为什么研究转录组？

□ 转录组是连接基因组遗传信息与生物功能的蛋白质组的纽带，转录水平的调控是最重要也是目前研究最广泛的生物体调控方式。转录组的研究比基因组的研究能给出更高效的有用信息。与基因组不同，转录组更有时间空间性，能反映特定条件下活跃表达的基因。



转录组学研究内容

- 对特定细胞转录与加工的研究
- 对转录物编制目录
- 绘制动态转录物图形
- 转录物的网络式调节



转录组学的研究方法

■ 基于测序的转录组学方法

- EST
- 全长cDNA文库
- RNA-Seq

■ 基于杂交的转录组学方法

- cDNA芯片
- 寡核苷酸芯片



EST

- 90年代初**Craig Venter** 提出了EST的概念，并测定了609条人脑组织的EST，宣布了cDNA大规模测序的时代的开始 (Adams et al., 1991)。
 - EST (**Expressed Sequence tags, 表达序列标签**) 是从已建好的cDNA文库中随机抽取克隆，从5'末端或3'末端对插入的cDNA片段进行一轮单向自动测序，所获得的约60-500bp的一段cDNA序列。
-



EST相关数据库

储存EST原始数据的一级数据库

- ◆ EMBL
- ◆ GenBank (dbEST)
- ◆ DDBJ

对EST进行聚类拼接的二级数据库

- ◆ UniGene (<http://www.ncbi.nlm.nih.gov/UniGene>)
- ◆ TIGR Gene Indices (<http://www.tigr.org/tdb/tgi/>)
- ◆ STACK (<http://www.sanbi.ac.za/Dbases.html>)



EST的应用 1: EST与基因识别

- 在同一物种中搜寻基因家族的新成员(paralogs)。
- 在不同物种间搜寻功能相同的基因(orthologs)。
- 已知基因的不同剪切模式的搜寻。

EST的应用 2: 大规模分析基因表达水平



EST技术流程:

一、cDNA文库构建

二、序列测定及数据分析

随机挑取克隆进行5'或3'端测序



序列前处理



聚类 and 拼接



基因注释及功能分类



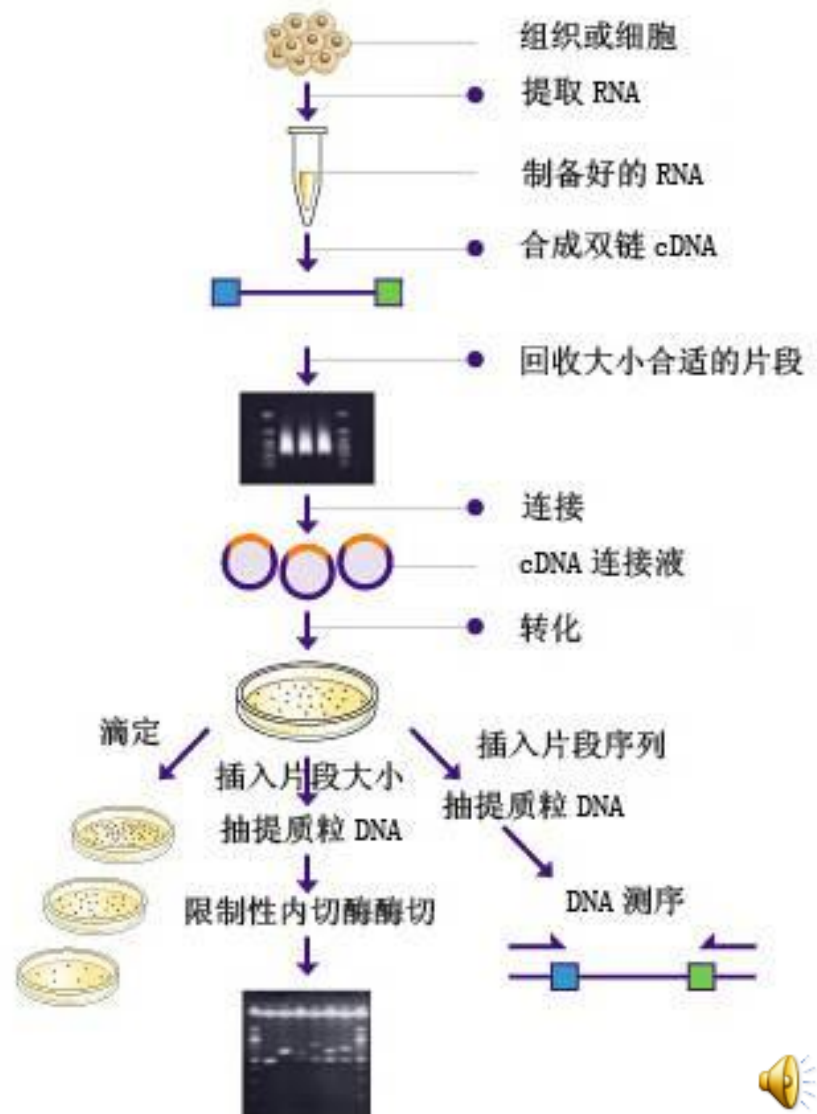
后续分析



cDNA文库

□ **cDNA**:为具有与某RNA链呈互补的碱基序列的单链DNA即**complementary DNA**之缩写。

□ 以mRNA为模板，经反转录酶在体外反转录成cDNA，与适当的载体（常用噬菌体或质粒载体）连接后转化受体菌，则每个细菌含有一段cDNA，并能繁殖扩增，这样包含着细胞全部mRNA信息的cDNA克隆集合称为该组织细胞的**cDNA文库**。



RNA-Seq

- 基于新一代**高通量**测序技术的**转录组测序**
- 可以识别比蛋白组高一两个数量级的基因，从而帮助科学家构建完整的基因表达谱以及蛋白质相互作用网络。



第五节 蛋白质组学

基本概念

蛋白质组学研究的特点

研究内容及主要技术

主要研究技术



一、概念

1、蛋白质组和蛋白质组学

□ 蛋白质组（Proteome）

“**proteome**”是由protein一词的前几个字母“**prote**”和genome一词的后几个字母“**ome**”拼接而成

□ 对应于基因组的所有蛋白质构成的整体。

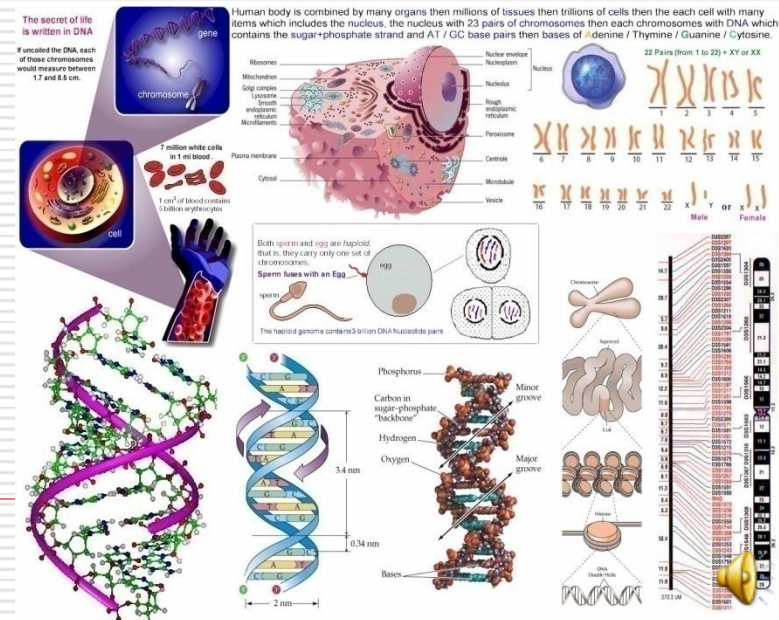
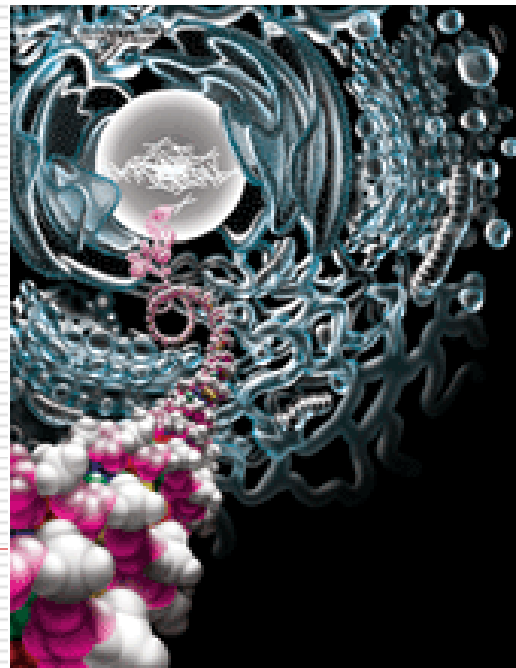
时空 □ 是指特定细胞、组织乃至机体作为一个生命单元中所拥有蛋白质的集合；

动态
变化 □ 是生命体在某一时段所表达的全部蛋白质。



蛋白质组学(Proteomics)

采用大规模、高通量、高速度的技术，通过全局性研究基因组表达的所有蛋白质在不同时间与空间的表达谱和功能谱，全景式揭示生命活动的本质。



蛋白质学与 基因组学关系

Genome --- 总导演

Proteome --- 演 员



基因组与蛋白质组的比较

基因组

蛋白质组

研究对象

DNA

蛋白质

稳定性

相对稳定，静态

变化大，动态

研究内容及
方法

结构基因组

序列分析

多态性作图

结构蛋白质组

晶体结构分析

生物质谱

功能基因组

表达系列分析

基因芯片

功能蛋白质组

双向电泳、生物质谱

蛋白芯片

转基因技术

酵母双杂交技术



Proteome 与 Genome区别

- 稳定性: **Proteome** ---动态 **Genome** ---静态
- 特异性
 - **Proteome** --- 组织和细胞特异性
 - **Genome** --- 无组织和细胞特异性
- 修饰化程度
 - **Proteome** --- 高
 - **Genome** --- 低
- 复杂程度
 - **Proteome** --- 高 (20aa + 高度多样性修饰)
 - **Genome** --- 低 (4nt)
- 孤立行为与相互作用: **Proteome** --- 相互作用
Genome --- 孤立



蛋白质组学的研究意义

- 蛋白质是生理功能的执行者，是生命现象的直接体现者，对蛋白质结构和功能研究**将直接阐明生命在生理或病理条件下的变化机制。**
- 通过蛋白质组学的研究，针对生理或病理条件下的差异，可以**为人类各种疾病的诊断和治疗提供新的途径。**

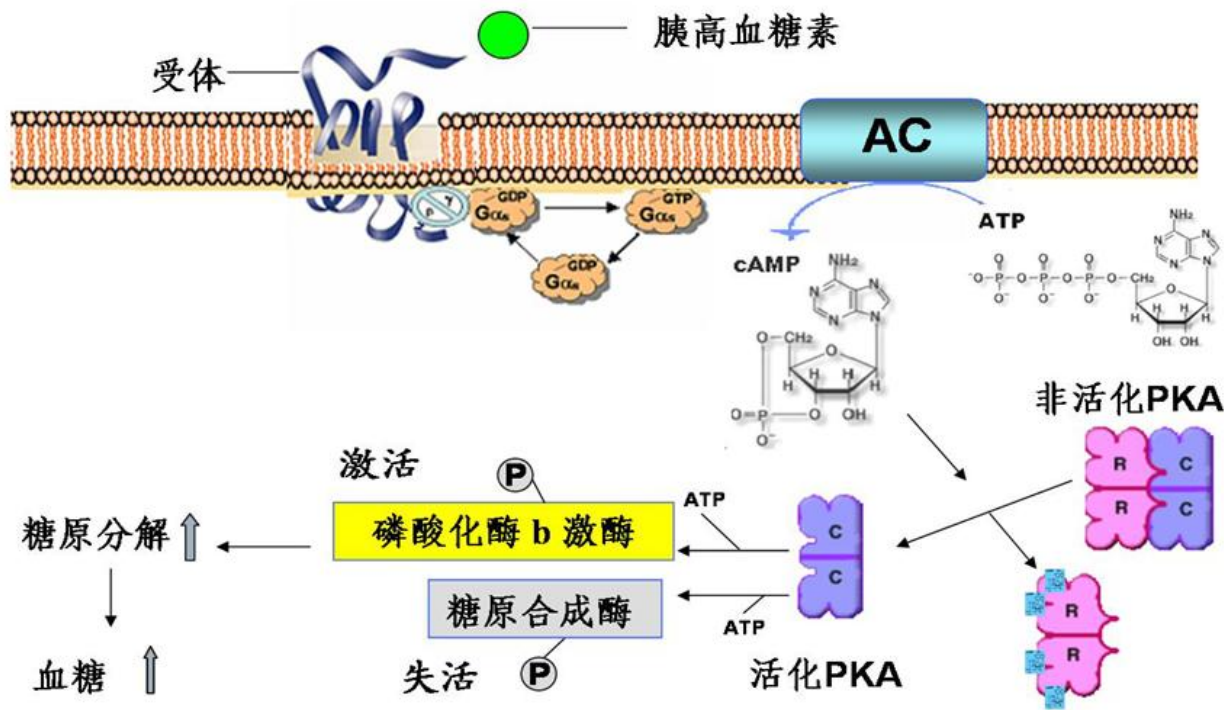


二、蛋白质组学研究的特点

- 整体性
- 揭示生命的动态性过程
- 体现生命现象的复杂性



整体性

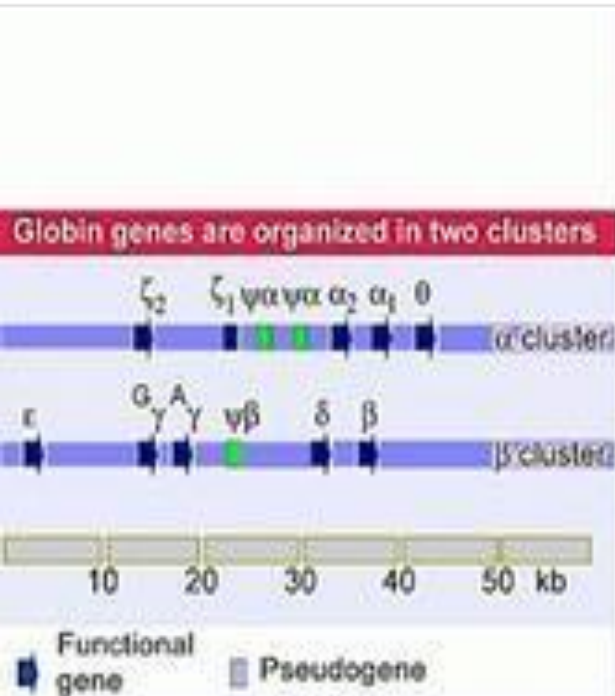


各种蛋白质自身功能的发挥依赖于蛋白质间的相互作用，即**整体性**

胰高血糖素通过cAMP - PKA途径升高血糖



动态性



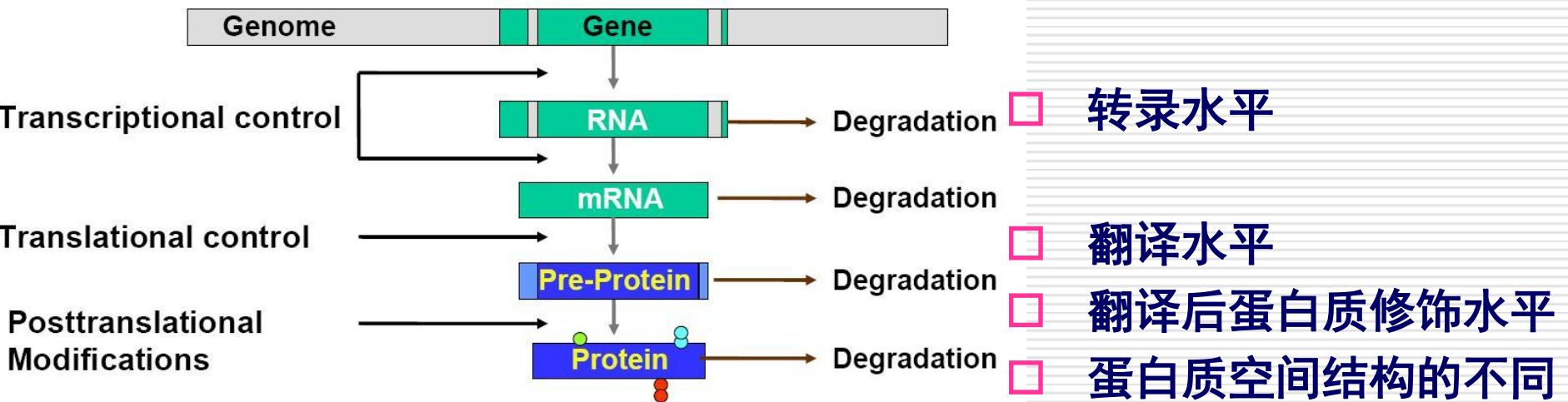
胚胎的血红蛋白	
Hb Gower 1	$\xi_2\epsilon_2$
Hb Gower 2	$\alpha_2\epsilon_2$
Hb Portland	$\xi_2\gamma_2$
胎儿的血红蛋白	
HbF	$\alpha_2\gamma_2$
成年的血红蛋白	
HbA	$\alpha_2\beta_2$
HbA2	$\alpha_2\delta_2$

□ 特定时间
不同发育阶段

□ 特定空间
机体
组织
细胞



复杂性



与传统蛋白质研究的区别

个体 → 整体

分析静态性质 → 比较动态变化

孤立个体 → 相互作用

结构 → 功能

小规模、非连续 → 高通量、自动化



三、蛋白质组学研究的内容

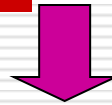
- 蛋白质组及其组成成分的分离
 - 蛋白质组及其组成成分的分析与鉴定
 - 蛋白质结构分析
 - 生理与病理状态或不同发育阶段的蛋白质组表达差异的研究
 - 蛋白质和蛋白质间相互作用的研究
 - 蛋白质与核酸间相互作用的研究
 - 丰富与完善蛋白质信息库
-



蛋白质组学研究流程图

□ 蛋白质样品制备
(sample preparation)

Protein Extraction



Protein Prefractionation



□ 蛋白质的分离

Protein Separation



□ 蛋白质的鉴定

Protein Identification



□ 生物信息学研究

Database utilization



四、蛋白质组学研究采用的主要技术

- 蛋白质分离技术：2-DE
- 蛋白质鉴定技术：质谱...
- 蛋白质-核酸间相互作用研究技术：
凝胶滞后实验...
- 蛋白质与蛋白质间相互作用研究技术：
酵母双杂交...



双向凝胶电泳-分离

(two-dimensional gel electrophoresis, 2-DE)

- 双向凝胶电泳目前仍然是蛋白质组学研究的主要分离方法
- 利用蛋白质的等电点和分子量，结合凝胶化学特性，分离各种蛋白质的方法。



原理

第一向：

等电聚焦凝胶电泳 **IEF**
(**i**so**e**lectro**f**ocusing)

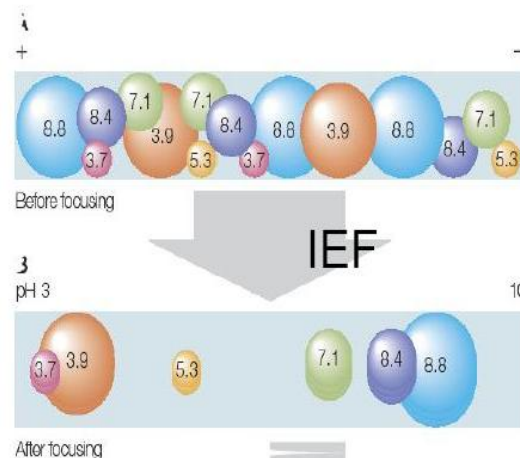
以蛋白质等电点差异为基础

第二向：

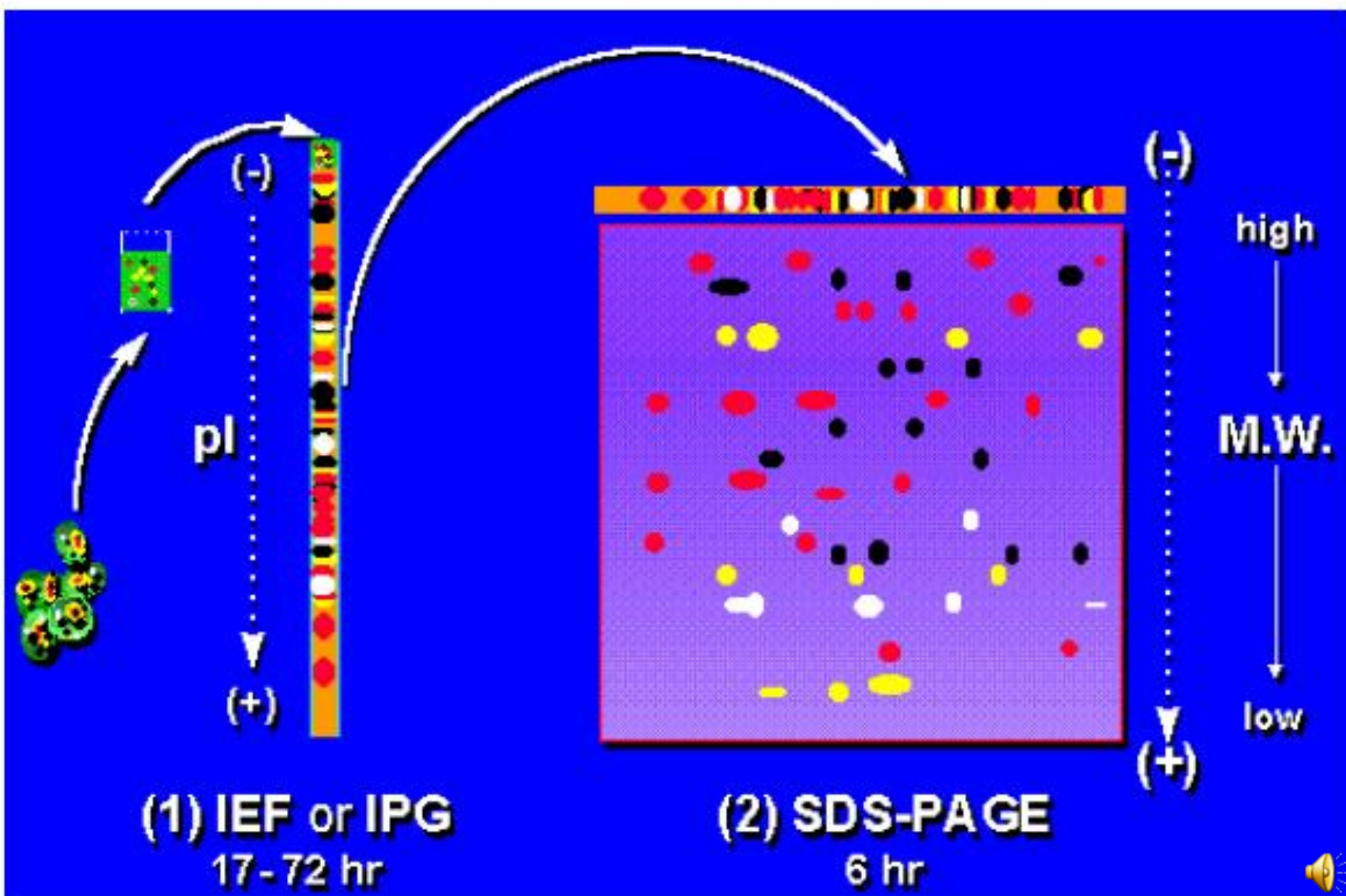
SDS-PAGE (**SDS**-聚丙烯酰胺凝胶电泳)

以蛋白质分子量差异为基础。

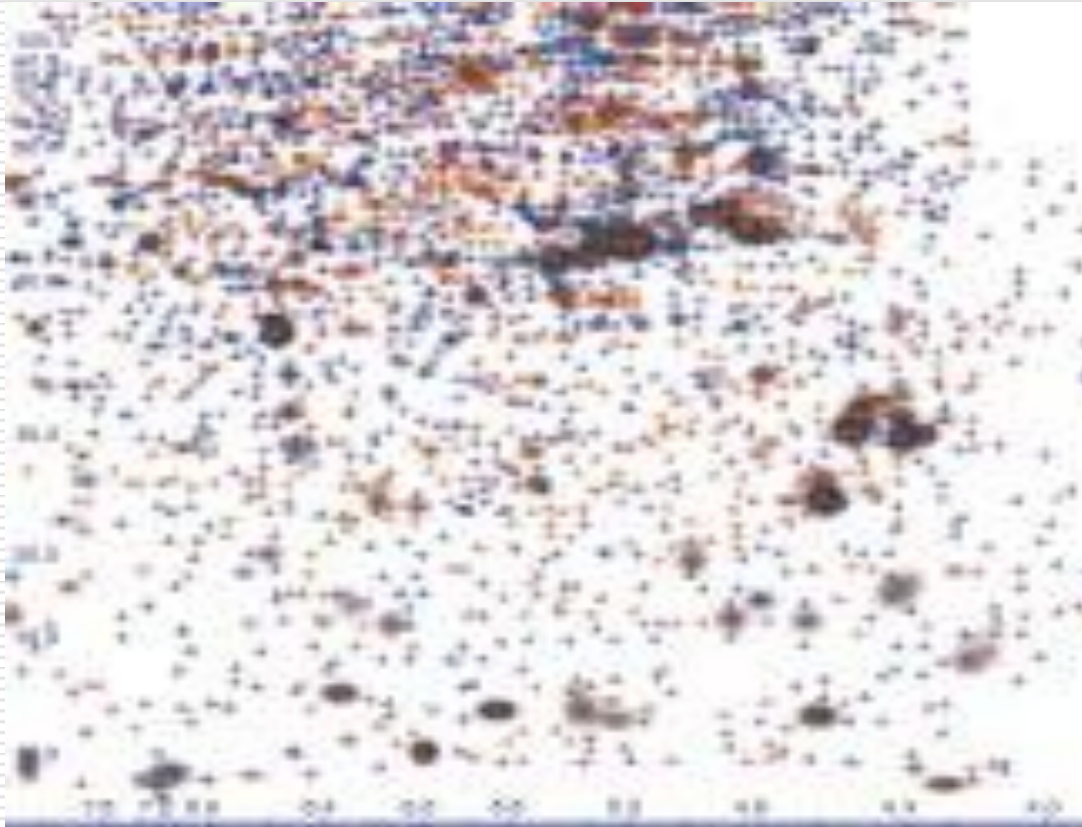
By isoelectric point (pI): IEF



Proteomics – 2D Electrophoresis



二维凝胶蛋白电泳图



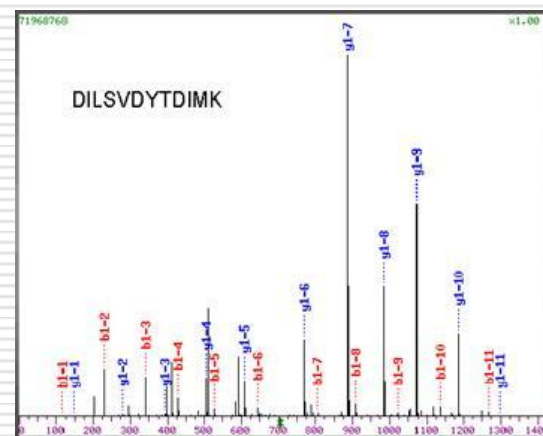
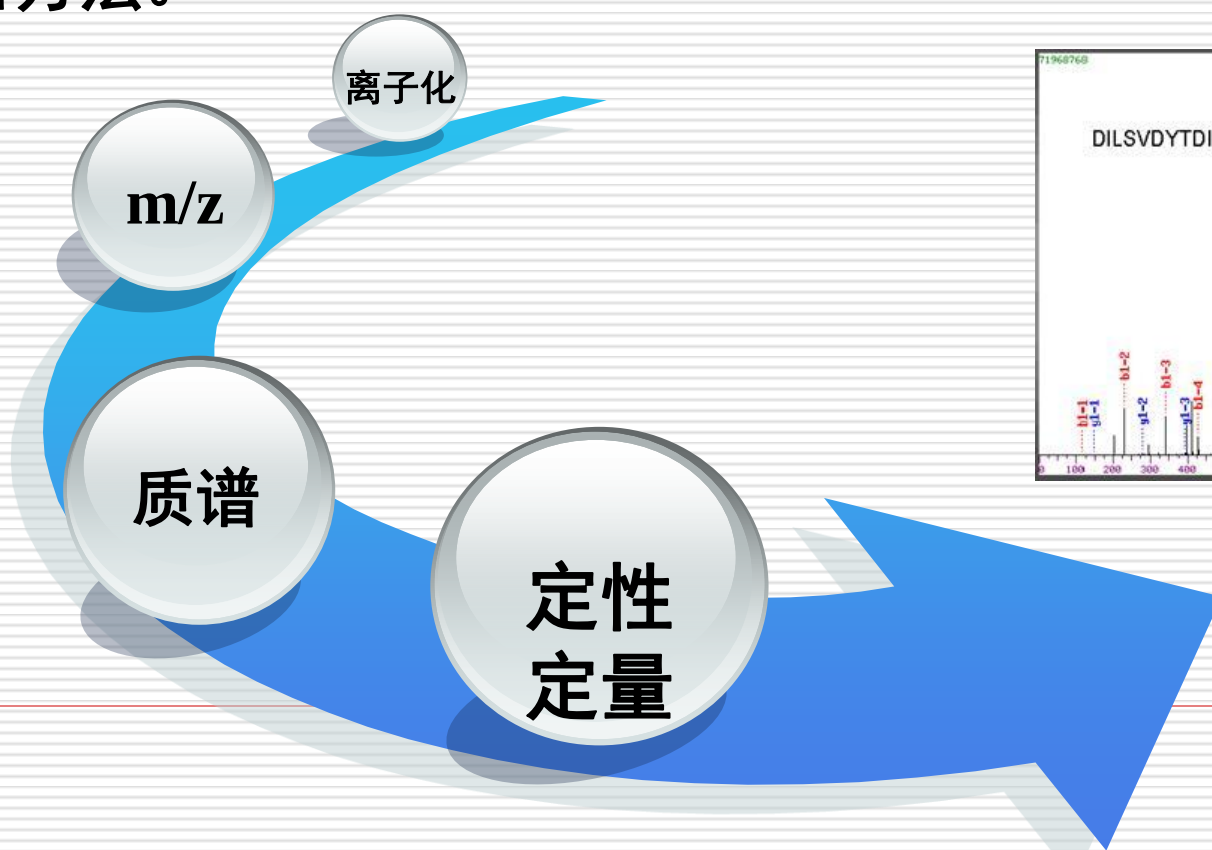
有颜色的点为每一不同的蛋白质，2-DE结果呈现满天星状图案





质谱技术(mass spectrometry, MS)-鉴定

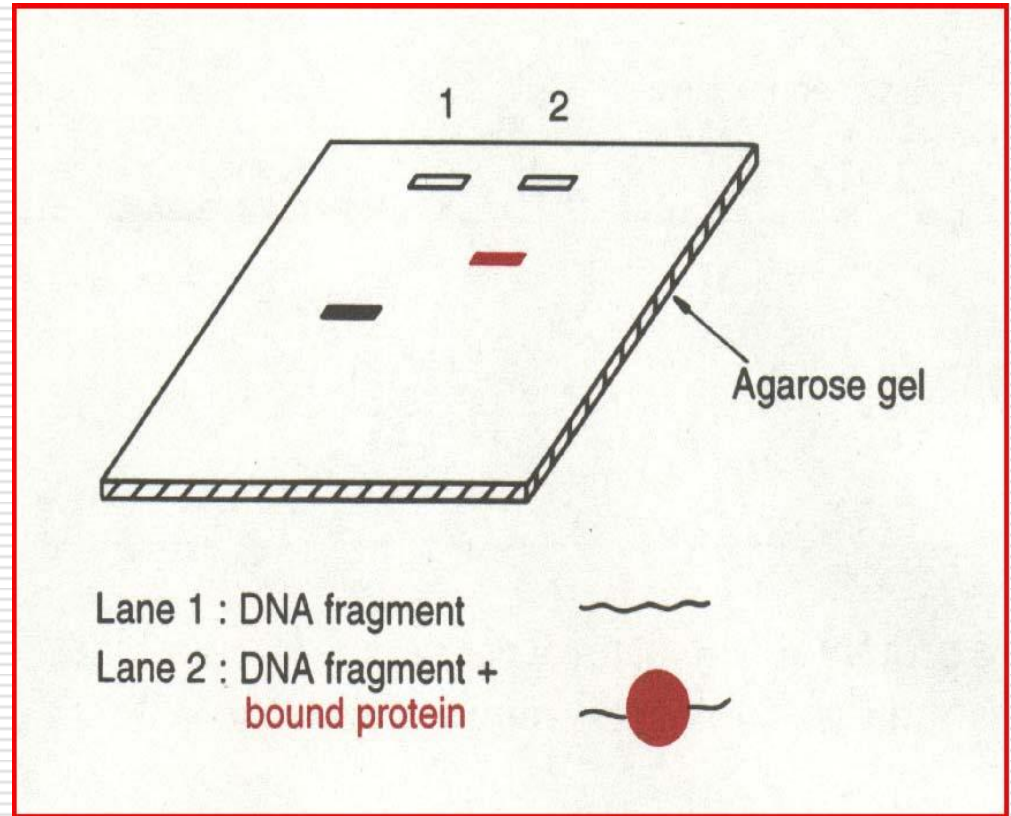
质谱分析是先将物质离子化，按离子的**质荷比**分离，然后测量各种离子谱峰的强度而实现分析目的的一种分析方法。



凝胶滞后实验

□ 基本原理

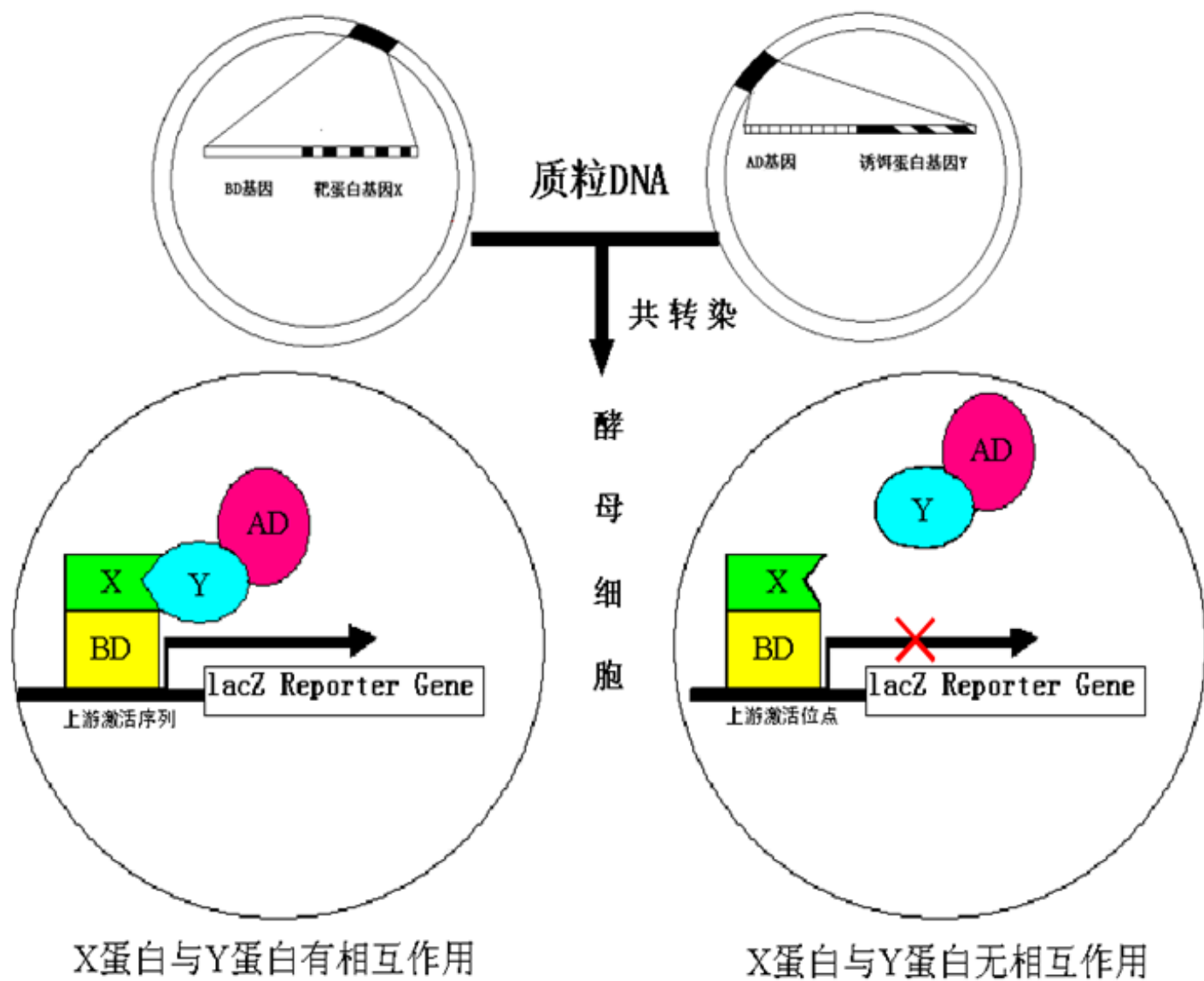
蛋白质可以与末端标记的核酸探针特异性结合，电泳时这种复合物较无蛋白结合的探针在凝胶中的泳动速度要慢——即表现为相对滞后。



酵母双杂交系统 (yeast two-hybrid system)

- 1989年Field 和Song等人在酵母细胞中设计的分析蛋白质相互作用的方法。
- 利用酿酒酵母的GAL4/Lex蛋白质N-端和C-端分别融合两种蛋白质相互作用可激活具有UAS (upstream activating sequence, 上游激活序列) 的报告基因表达, 从而判断两种蛋白分子间是否发生了相互作用的方法。





酵母双杂交技术原理



酵母双杂交技术的应用

- **1、** 利用酵母双杂交发现新的蛋白质和蛋白质的新功能
- **2、** 利用酵母双杂交在细胞体内研究抗原和抗体的相互作用
- **3、** 利用酵母双杂交筛选药物的作用位点以及药物对蛋白质之间相互作用的影响
- **4、** 利用酵母双杂交建立基因组蛋白连锁图



四、蛋白质组数据库

蛋白序列数据库 (SWISS-PROT, TrEMBL)

基因序列数据库 (Genbank, EMBL)

蛋白模式数据库 (Prosite)

蛋白二维凝胶电泳数据库

蛋白三维结构数据库 (PDB, FSSP)

蛋白翻译后修饰数据库 (O-GLYCBASE)

基因组数据库 (GDB, OMIM)

代谢数据库 (ENZYME)



五、蛋白质组学的应用及进展

- ❖ 用于寻找疾病相关的蛋白质，如疾病蛋白质标志物
- ❖ 用于微生物蛋白组研究，彻底阐明病原微生物的致病机制并寻找全新的药物作用靶点
- ❖ 蛋白质组数据库将成为药物设计的路标
- ❖ 对不同生物的蛋白质组进行比较性研究，可以对生物进化途径提供参考，对多细胞生物的起源提供线索
- ❖ 追踪胞内信号分子的移位，阐明目标蛋白质在信号转导途径中的位置



第六节 代谢组学

- 概念
- 研究内容
- 研究方法



概 念

代谢组 (metabolome)

是指基因组的所有下游产物也即最终产物的组合，这些产物是一些参与生物新陈代谢、维持生物体正常功能和生长发育的小分子化合物，主要是相对分子量小于**1000Da**的内源性小分子。

代谢组学 (Metabonomics/Metabolomics)

是通过考察生物体系（细胞、组织或生物体）受到刺激或扰动后（如将某个特定的基因变异或者环境变化后），其代谢产物的变化或其随时间的变化，来研究生物体系的一门科学。

为什么要研究代谢组学呢？

□ 转录组学或蛋白组学上的变化不总能反映基因变更时在生化层面的**表现型的改变**。

- 基因组的变化不一定能够得到表达，从而并不一定对系统产生影响；
- 某些蛋白的浓度会由于外部条件的变化而升高，但由于这个蛋白可能不具备活性，从而也不对系统产生影响；
- 由于基因或蛋白的**功能补偿**作用，某个基因或蛋白的缺失会由于其他基因或蛋白的存在而得到补偿，最后反应的净结果为零；
- 小分子的产生和代谢才是这一系列事件的最终结果，它能够**更准确的反映生物体系的状态**。

-
- ✦ 与转录组学和蛋白质组学比较，代谢组学专门研究生物体系受外部刺激后所产生的所有代谢产物的变化，能够更准确地反映生物体系的状态，且位于系统生物学的最下游，是生物体系整体功能或状态最终结果的表现。
 - ✦ 因此有人认为，“基因组学和蛋白质组学告诉你什么可能会发生，而代谢组学则告诉你什么确实发生了。”（**Bill Lasley, UC Davis**）
-

代谢组学研究的目

◆代谢组学研究的目的是定量分析一个生物系统内所有代谢物的含量。代谢组学分析可以指示细胞、组织或器官的生化状态，协助阐释新基因或未知功能基因的功能，并可以揭示生物各代谢网络间的关联性，帮助人们更系统地认识生物体。



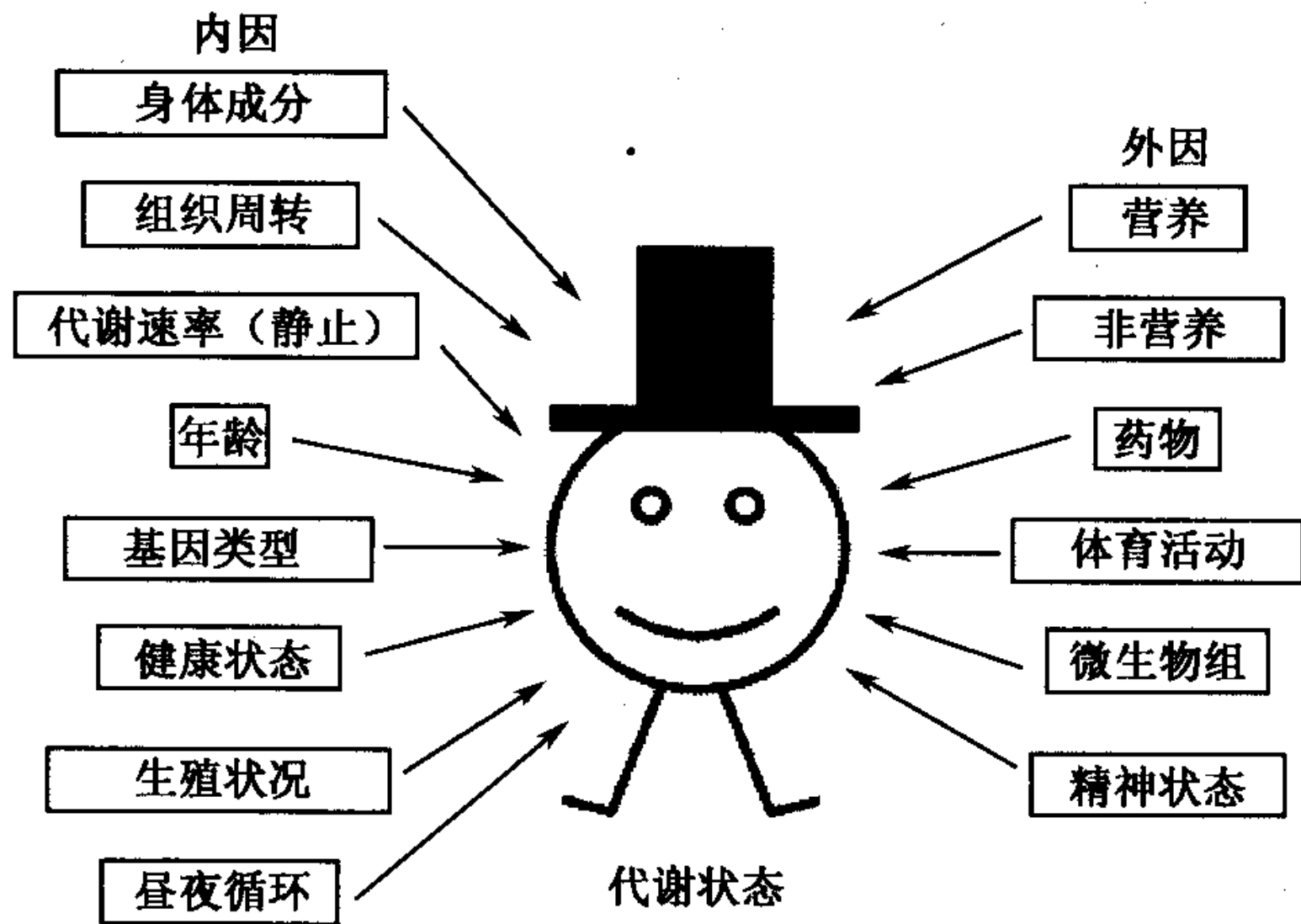


图 1-9 影响人代谢状态的内部和外部因素

这些状态可用代谢组学测量^[192]

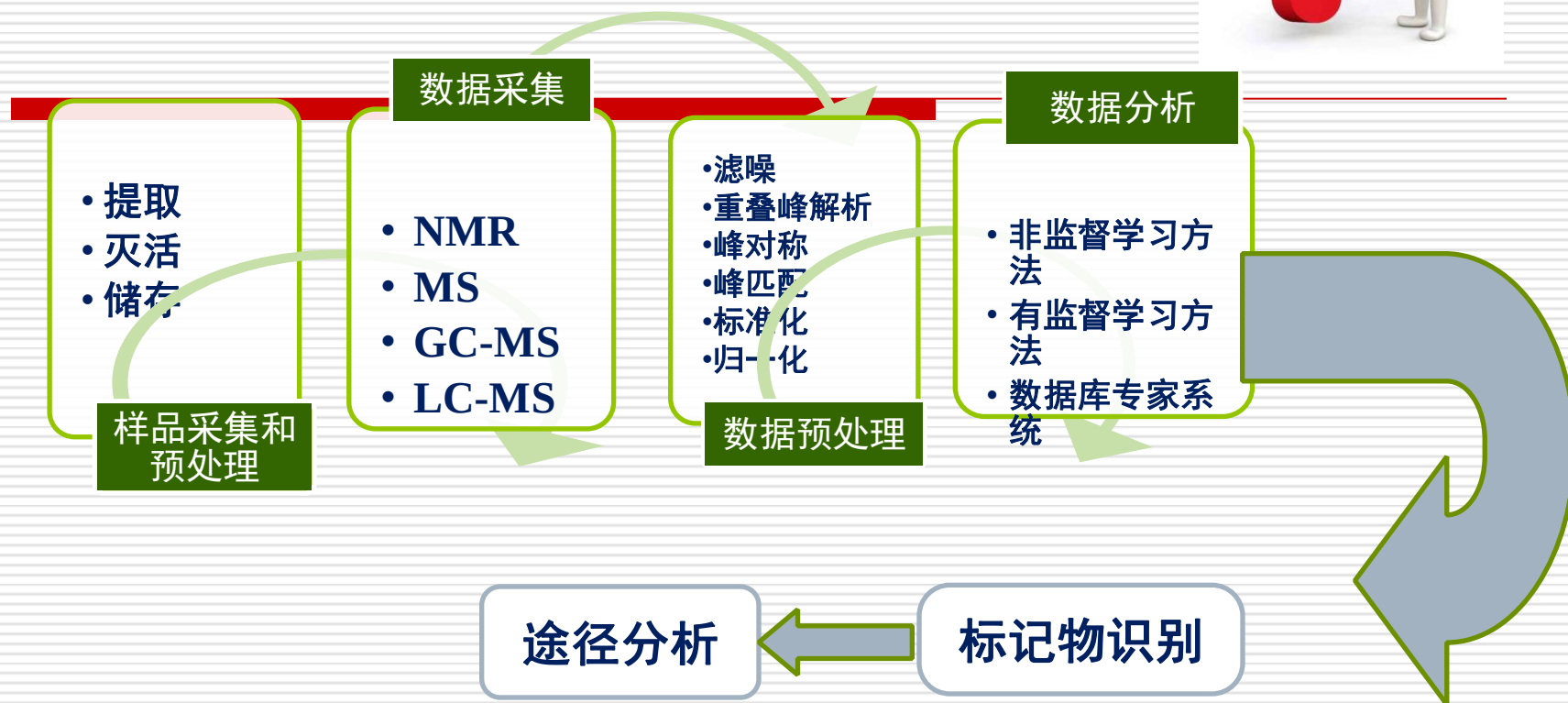
研究方法

- 运用核磁共振(NMR)、质谱(MS)、气质联用技术(GCMS)、高效液相色谱(HPLC)等高通量、高灵敏度与高精确度的现代分析技术对细胞提取物、组织提取物和生物体液随时间变化的代谢物浓度进行检测，结合有效的模式识别方法进行定性、定量和分类，并将这些代谢信息与病理生理过程中的生物学事件关联起来，从而了解机体生命活动的代谢过程。
-

代谢组学的技术平台

- 红外线光谱技术（**IR**），
 - 核磁共振技术（**NMR**），
 - 稀薄气液色谱技术（**TLC**），
 - 高效液相色谱技术（**HPLC**），
 - 高效毛细管电泳技术（**HPCE**），
 - 毛细管电泳与紫外线吸光率检测连用技术（**CE/UV**），
 - 毛细管电泳与激光诱导荧光检测连用技术（**CE/LIF**），
 - 毛细管电泳与质谱连用技术（**CE/MS**），
 - 气相色谱与质谱共用技术（**GC/MS**），
 - 液相色谱与质谱共用技术（**LC/MS**），
 - 液相色谱与质谱先后使用技术（**LC/MS/MS**），
 - 高效液相色谱与质谱和核磁共振技术功用（**LC/NMR/MS**）
-

代谢组学研究步骤



代谢组学的研究应用

◆ 药物研发

- 药物筛选
- 药效及毒性评价
- 作用机制和临床评价



◆ 疾病研究

- 病变标记物的发现
- 疾病的诊断
- 治疗和预后的判断



◆ 微生物代谢组学

- 微生物表型分类
- 突变体筛选
- 代谢途径及微生物代谢工程
- 发酵工艺的监控和优化
- 微生物降解环境污染物



◆ 植物代谢组学

- 特定种类植物的代谢物
- 不同基因型植物的代谢物组学表型研究
- 某些生态型植物的代谢物组学
- 受外界刺激后植物自身免疫应答



系统生物学

基因组学

- 研究生物系统的基因结构组成, 即DNA的序列及表达

转录组学

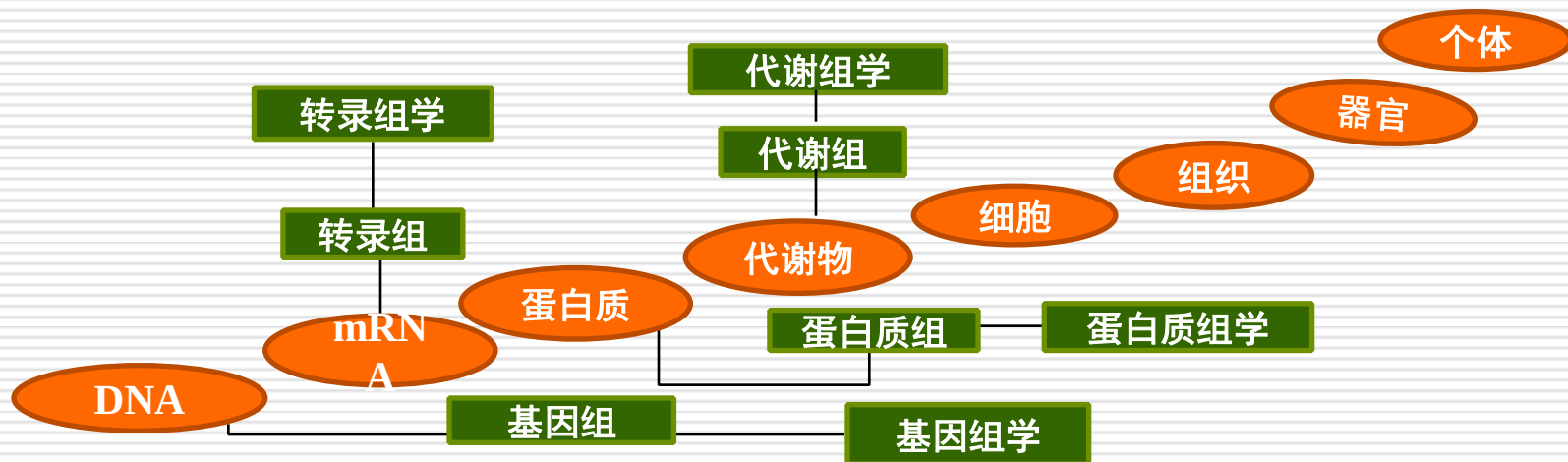
- 研究核糖核酸转录过程

蛋白质组学

- 研究生物系统表达的蛋白质及由外部刺激引起的差异

代谢组学

- 研究生物体系(细胞, 组织或生物体)受外部刺激所产生的所有代谢产物的变化



小 结

- 名词：蛋白质组、蛋白质组学、基因组、人类基因组计划
 - 蛋白质组学的研究内容有哪些？
 - 目前在蛋白质组学的研究中主要采用了哪些技术？
-

Thanks

