

版权声明:

本站几乎所有资源均搜集于网络, 仅供学习参考, 不得进行任何商业用途, 否则产生的一切后果将由使用者本人承担! 本站仅提供一个观摩学习与交流的平台, 将不保证所提供资源的完整性, 也不对任何资源负法律责任。所有资源请在下载后 24 小时内删除。如果您觉得满意, 请购买正版, 以便更好支持您所喜欢的软件或书籍!



☆☆☆☆☆生物秀[\[http://www.bbioo.com\]](http://www.bbioo.com)

☆☆☆☆☆中国生物科学论坛[\[http://www.bbioo.com/bbs/\]](http://www.bbioo.com/bbs/)

☆☆☆☆☆生物秀下载频道[\[http://www.bbioo.com/Soft/\]](http://www.bbioo.com/Soft/)

生物秀——倾力打造最大最专业的生物资源下载平台!

■■■■ 选择生物秀, 我秀我精彩!! ■■■■

欢迎到生物秀论坛(中国生物科学论坛)的相关资源、软件版块参与讨论, 共享您的资源, 获取更多资源或帮助。

蛋白质组学(Proteomics) Outline of Proteomics

- 第一章 绪论
- 第二章 蛋白质组学的研究体系
- 第三章 蛋白质组学的应用与发展
- 第四章 蛋白质组学研究的实例

引言

- ❖ 生命科学的主要研究对象是什么？
- ❖ 生命科学的核心是什么？
- ❖ 传统蛋白质研究研究什么？
- ❖ 基因研究 - 基因组研究
- ❖ 蛋白质组研究

生命科学的核心问题—生老病死

- ❖ 生命现象的过程及其变化 - 发生和发展
- ❖ 解释疾病的发生
- ❖ 诊断和治疗疾病

传统蛋白质研究

- ❖ 研究对象：单个蛋白质
- ❖ 核心问题：结构和功能
- ❖ 由于蛋白质结构多变且各异，分离纯化不易，易失活，技术上存在困难，所以长期以来发展步子不大

基因研究到基因组研究

- ❖ 基因是遗传信息的载体，可以视为生命活动的内在因素，推动生命 generation to generation 地稳定且适应环境变化地延续下去。DNA（基因）的发现是生命科学的伟大里程碑，它将生命科学加速地推进了。
- ❖ 生物体是一个协调的整体，单个基因不能反映正确的整体信息

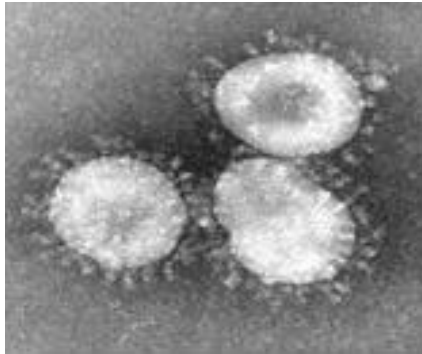
Biological Process(A simplified view)



基因组学 Genomics

Completely Sequenced Genomes (2001/2002)

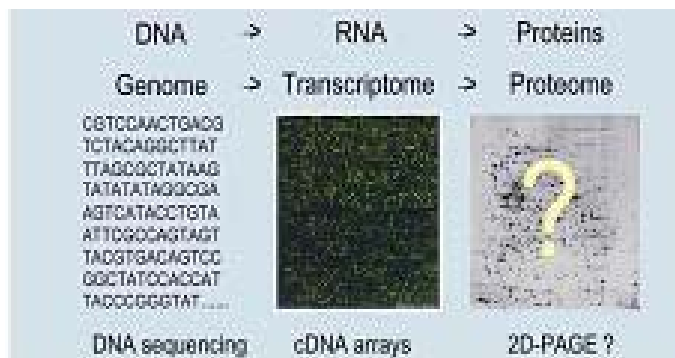
- ❖ 4 /15 eucaryotes + (draft) human
 - ❖ 9 /16 archaea
 - ❖ 32 /89 bacteria
 - ❖ >600 viruses
 - ❖ 正在进行的基因组测序包括271原核生物和152真核生物。
- SARS virus_coronavirus



功能基因组学---基因组的表达

但是，完成了测序，并不一定知道基因组的功能

- ❖ cDNA阵列 (cDNA microarray)
- ❖ DNA 芯片 (DNA chips)
- ❖ 基因表达序列分析 (Serial analysis of gene expression, SAGE)
- ❖ 表达序列标记 (Expressed sequence tag, EST)



蛋白质表达的调控

转录水平	DNA	mRNA
翻译水平	mRNA	Protein
翻译后水平	Protein	Modified protein

蛋白质存在形式和活动规律，必须依赖蛋白质本身的研究

“ If the 1990s were the decade of genomics, the first three years of the new century are set to become the decade of the proteomics. ”

N. Leigh Anderson

Current Opinion in Biotechnology, 2000, 11: 408-412

“ Proteins are central to our understanding of cellular function and disease processes, and without a concerted effort in proteomics, the fruits of genomics will go unrealized. ”

Nature 409, 747 (2001)

第一章 绪论

- 科学发展的历史背景
- 基本概念
- 蛋白质组研究的必要性与必需性
- 蛋白质组学研究的目的是、内容和特色
- 蛋白质组学研究进展

一、科学背景 (Gene and Protein)

- 1860年 Friedrich Miescher 鉴定出细胞核内有酸性和碱性蛋白质当时这些被错误的认为就是遗传物质的携带者
- 1940年 Beadle 和 Tatum 将基因和蛋白产物联系起来，确立了“一个基因，一个蛋白”的概念
- 1953年 Watson 和 Crick 确立了DNA双螺旋结构
- 1956年 Smithies 和 Poulik 结合纸电泳和淀粉凝胶二维电泳法分离蛋白质
- 1961年 在发现信使RNA解释遗传密码，描述蛋白质合成的基因调节理论基础上，提出蛋白表达的概念
- 1967年 Edman 和 Begg 完成蛋白质自动测序
- 1970年 Kenrick 和 Margolis 发明等电聚焦-梯度凝胶电泳的二维电泳技术
- 1972年 Bernstein 等收集10种蛋白质X-射线晶体衍射结构，组建蛋白质数据库 (Protein Data Bank)
- 1975年 O'Farrell 建立现代高分辨形式的蛋白质二维电泳法
- 1981年 Anderson 以二维聚丙烯酰胺凝胶电泳为核心技术进行药理学和毒理学研究
- 1982年 Anderson 提出人类蛋白质组作图谱的概念
- 1986年 Kuska 以 Roderick “基因组学 Genomics” 为名办了杂志，并于1987年开始出版
- 1986年 Swiss Institute Geneva 大学建立第一个蛋白质序列数据库—SWISS-PROT
- 1987年 第一个蛋白质组学公司成立---Large Scale Biology Corporation
- 1995年 Wilkins 定义了蛋白质组
- 1997年 Wilkins 等出版了第一本有关蛋白质组学的书
- 1990年 开始人类基因组计划 (HGP)
- 1994年 人基因组全套遗传连锁图发表
- 1995年 全基因组覆盖率94%的物理图谱问世
- 1995年 开始模式生物与致病微生物等的基因组研究
- 1995年 支原体和流感嗜血杆菌基因组全序列发表
- 1996年 第一个真核生物—酵母的基因组全序列完成
- 1998年 多细胞真核生物---线虫基因组全序列测定
- 目前已完成不少于50种生物基因组全序列测定，97%人类基因组图化，85%精确测序
- 2000年6月 公开人类基因组90%的草图 (DDBJ/EMBL/GenBank)
- 目前涉及10多种生物体，包括人心脏和肝脏在内的10多个蛋白质组数据库初步建立，其中有完整的酵母蛋白质组数据库
- 2001年4月 在美国成立了HUPO (Human Proteomics Organization)，研究重点为人类疾病与健康。参与的人员来自科研机构和各大医药公司两方面
- 2001年5月18日，亿利资源集团公司联合北京师范大学、中国协和医科大学、上海交通大学等八家院校和科研机构对外宣布 成立“高等学校蛋白质组学研究院”及“中国高校亿利蛋白质组学研究中心”

二、基本概念

1. 基因组和基因组学

- ❖ 基因组genome：一个细胞或病毒所包含的**全部基因**。

通常在真核生物中指一个物种的单倍体染色体组所含有的一整套基因。原核生物一般只有一个环状DNA分子，其上含有的基因为一个基因组。

- ❖ 基因组学genomics：从基因组**整体水平**上对基因的活动规律进行阐述。

结构基因组学

- ❖ 通过基因作图（基因连锁图，物理图谱，转录图谱，遗传图谱）确定基因组成，基因定位的科学**比较基因组学**

- ❖ 将已知基因结构进行比较，了解基因的功能，表达机理和物种进化

功能基因组学

- ❖ 利用结构基因组学提供的信息，通过在基因组或蛋白质水平上全面分析基因功能，使生物学研究从单一转向系统。

- ❖ 基因功能发现，基因表达分析，突变检测

营养基因组学

- ❖ 主要用于代谢方面的研究，检测和操纵生物中微量营养代谢途径（代谢基因组学）

2. 蛋白质组和蛋白质组学

蛋白质组 (proteome)

- ❖ 1994年澳大利亚Macquarie大学的Marc Wilkins和Keith Williams首先提出The entire PROTEIn population expressed by a genOME or by a cell or tissue type.

- ❖ 1995年7月发表于Electrophoresis杂志上PROTEOME is the PROTEINS expressed by a genome or a tissue.

- ❖ Swinbanks指出“蛋白质组代表一个完整生物的全套蛋白质。”

- ❖ Kahn认为“蛋白质组反映不同细胞的不同蛋白质组合”

总之，蛋白质组是基因组所表达的全部蛋白质，更为清楚地表述了细胞或组织或机体在特定时间和空间上表达的所有蛋白质。同一概念不同表述。

定义：蛋白质组是由一个细胞，一个组织或一个机体的基因组所表达的**全部**相应的蛋白质。是一个**整体**概念。

蛋白质组学(Proteomics)

- ❖ 是以蛋白质组为研究对象，从蛋白质**整体水平**上来认识生命活动规律的科学。

- ❖ 蛋白质组学研究直接定位于蛋白质水平，从**整体，动态，定量**的角度去研究基因的功能，是后基因组计划的一个重要组成部分。

- ❖ 是对**特定时间或特定环境条件下**，细胞内**完整表达状况**的研究，代表了正在工作的基因组的情况，是一个**动态过程**。

- ❖ 集中于**动态**描述基因调节、对基因表达的蛋白质水平进行定量的测定、鉴定疾病、药物对生命过程的影响，以及解释基因表达调控的机制。

- ❖ 旨在阐明生物体内**全部**蛋白质的**表达模式和功能模式**，其内容包括蛋白质的**表达与存在方式**（修饰方式），**结构与功能**，以及各个蛋白质之间**相互作用**等。

- ❖ 是一门以**全面**的蛋白质性质研究(如表达水平、转录修饰、相互作用等)为基础，在蛋白质水平对疾病机理、细胞模式、功能联系等方面进行探索的科学，包括**表达蛋白质组学**，**细胞谱蛋白质组学**以及**功能蛋白质组学**。

表达蛋白质组学(Expression proteomics)

- ❖ 又称定量调节蛋白质组学 (quantitative regulation proteomics)，是在整体水平上研究生物体蛋白质表达的变化，对细胞或组织中蛋白质**表达量化谱**的反映。

- ❖ 目前蛋白质组学的重点。

- ❖ 通过将细胞、组织中的蛋白建立蛋白表达图谱或扫描EST图，监测一个细胞或组织内大多数蛋白质的表达状况，定量观察在不同环境，如药物存在或病体组织中表达方式的变化
- ❖ 对寻找疾病诊断标志、筛选药物靶点、毒理学研究等有重要作用

细胞谱蛋白质组学(*Cell-map proteomics*)

- ❖ 又称结构蛋白质组学(*structural proteomics*)，旨在研究蛋白质在细胞内行为、运输和相互作用，为研究蛋白质功能发挥作用。
- ❖ 确定蛋白质在亚细胞结构中的位置
- ❖ 系统地鉴定蛋白质复合物和明确的细胞组成情况
- ❖ 通过纯化细胞器或分离蛋白质复合物，来系统研究蛋白质-蛋白质的相互作用 (Blackstock and Weir 1999)
- ❖ 建立各种细胞类型和状态的“物理图谱”
- ❖ 确定蛋白质功能和新的疾病诊疗的靶位点

功能蛋白质组和功能蛋白质组学

功能蛋白质组(*functional proteome*)

- ❖ 由Humphery-Smith 于1998年总结了基因组当时研究结果后提出
- ❖ 细胞内与某个功能有关或在某种条件下的一群蛋白质。
- ❖ 又指特定时间、特定环境和实验条件下基因组活跃表达的蛋白质。
- ❖ 总蛋白质组的一部分

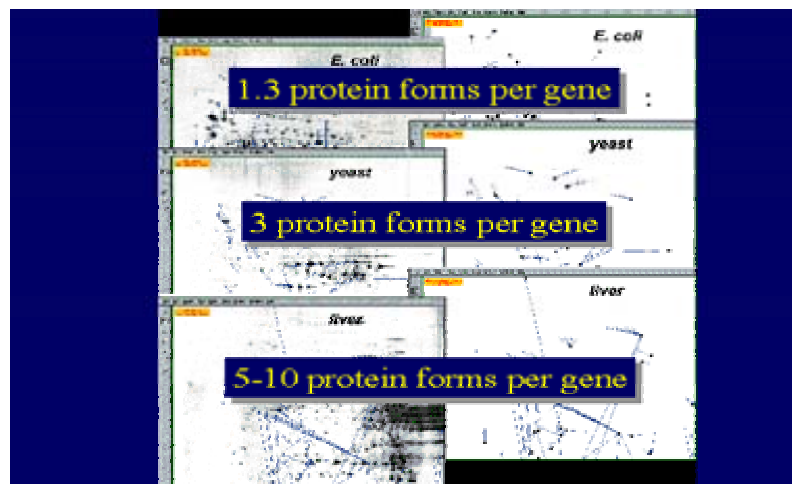
功能蛋白质组学 (*functional proteomics*)

- ❖ 李伯良 1998年提出
- ❖ 是介于传统蛋白质研究（针对个别蛋白质）和蛋白质组研究（以全部蛋白质为对象）之间的层次，是以功能蛋白质组为对象的研究
- ❖ 研究细胞内与某个种条件下的一群蛋白质；功能有关或在某种条件下的一群蛋白质。

3. 基因组与蛋白质组的比较

- ❖ 相同点: 都属于整体概念，蛋白质组是基因组的反映，但不是一个基因组的直接产物。

蛋白质组与基因组基因数(ORF)的关系



❖ **不同点:** 基因组: 1. **均一性和完整性** 2. 非常**稳定**, 不易发生改变

蛋白质组: 1. **特殊性和多样性** 2. **可变性, 高度动态的**

基因组的均一性, 完整性和稳定性

❖ 基因组在同一生物体的所有体细胞中是均一的, 完整的, 稳定的, 成分不随时间变化(终身不变);

❖ 组成该有机体的所有不同细胞都共有同一个基因组

蛋白质组的特殊性和多样性

❖ 蛋白质组具有很高的细胞特异性(不同组织细胞中, 蛋白质的表达有着很大差异); 不同细胞有不同的蛋白质组;

❖ 在生命发育不同阶段, 同一细胞的蛋白质种类也是不一样的;

❖ 正常状态与病理状态下, 同一细胞内的蛋白质组成也不同;

❖ 环境的改变也会造成细胞内的蛋白质组成也不同;

蛋白质组是动态的, 可变的

❖ 一个有机体的不同细胞表达它的全蛋白的不同亚套;

❖ 根据细胞的发育和生理状态, 蛋白表达的亚套一刻不停地变化着; 即使同一细胞在生长活动的不同时期, 在不同条件下, 其蛋白质组也是在不断的改变之中。

三. 蛋白质组研究的必要性与可能性

❖ 生命功能的主要执行者和直接体现者: **蛋白质**

❖ 基因组计划的局限性

❖ 基因组计划的成就与要求

❖ 蛋白质组研究自身发展的需求和研究技术的突破

基因组计划的局限性1

基因组序列本身的局限性, 并不是所有基因同时表达, 而是根据环境和生理状况而定, 即使知道全序列也不能完全清楚表达状况。

❖ 基因产物何时被翻译, 是否被翻译不清楚

❖ 基因产物被翻译的量是多少

❖ 翻译后修饰的程度不清楚

❖ 基因表达的结果如何

❖ 基因相对应的表型

基因组计划的局限性2

❖ 尽管通常情况下mRNA表达丰度决定了对应的蛋白质量, 但是不能仅仅以mRNA的数据为基础预测蛋白质表达水平。例如, 对数生长期的啤酒酵母, 同样丰度的mRNA产生的蛋白质量不同;

❖ 蛋白质动态修饰和加工并非必须取决于基因序列, 许多调节是在蛋白质的结构域中发生。例如蛋白质转录后的糖基化, 磷酸化等作用, 以及许多蛋白质只有与其它分子结合才体现活性, 这些修饰和结合通常是动态和可逆的, 而且不能由基因序列来决定。

基因组计划的成就与要求

❖ 上世纪九十年代初, 美国科学家提出并实施了人类基因组计划, 对人类基因组的全部DNA序列进行测定, 希望在分子水平上破译人类所有的遗传信息, 经过各国科学家的努力, 人类基因组计划已经取得巨大成功, 迄今已测定的EST已涵盖了人类的所有基因, 生命科学进入了后基因组时代

(post-genome era)。在后基因组时代, 研究的重心从揭示人类所有的遗传信息转移到在整体水平上对生物功能的研究。这种转向的首个标志就是功能基因组学, 它采用SAGE, DNA芯片等新技术, 对成千上万的基因表达进行分析和比较, 力图从整体水平上对基因的活动规律进行阐述。但是, 由于生

物功能的主要体现者是蛋白质，DNA和RNA并不能准确反映蛋白质的表达，修饰及相互作用。

- ❖ 在人类基因组计划和功能基因组学的基础上，产生了一门在整体上研究细胞内蛋白质组成和活动规律的学科。

蛋白质组研究自身发展的需求和研究技术的突破

- ❖ 蛋白质组学因为是对细胞内的总蛋白同时进行研究，所以需要高通量鉴定蛋白质点，对自动化程度要求高，且对样品的需求越少越好，纳米技术，电喷雾质谱，蛋白质芯片等新技术的全新登场也促进了蛋白质组学的产生和发展。

四、蛋白质组学研究的目的是、内容和特色

研究目的

- ❖ 分离蛋白，显示差异，将表现型与功能联系起来；
- ❖ 寻找蛋白质之间的相互作用，动态地反映生物体所处的状态。

研究特色

- ❖ 以群体为对象，综合并移植各种业已发展的技术；

研究内容

1. 蛋白质组研究体系的建立与完善

1) 研究技术体系

- ❖ 蛋白质组的样品制备和鉴定
 - ❖ 蛋白质相互作用网络的研究技术等
- 样品 → 样品制备 → 样品分离 → 蛋白质点提取鉴定
- ↓
- 功能确定 ← 生物信息学 ← 图像分析

2) 信息学Bioinformatics

- ❖ 蛋白质组数据库的建立
- ❖ 相关软件的应用与开发
- ❖ 生物信息学等

2. 重要的生物学问题有关的功能蛋白质组研究

- ❖ 功能蛋白质群体的选择
 - ❖ 功能蛋白质群体的研究方法
 - ❖ 功能蛋白质群体的分析
- ##### 1) 确定蛋白质(Protein identification)
- 蛋白质细胞/组织表达谱 Protein identification
- ❖ 一维胶电泳(One-dimensional gels)：如亲和纯化后样品的分析(analysis after affinity purification)
 - ❖ 二维胶电泳(Two-dimensional gels)：如纯化液或生物流体的分析等
 - ❖ 电喷雾质谱(Electrospray ionization mass spectrometry)：如肽的指纹图谱以及肽序列的测定
 - ❖ 蛋白芯片(Protein chips)：如用蛋白或抗体包被的芯片成批鉴定蛋白质 (chips coated with proteins or antibodies)
 - ❖ 非电泳技术鉴定溶液中的蛋白或蛋白复合物(Proteins/protein complexes in solution identified without electrophoresis)

2) 翻译后修饰的研究(Post-translational modifications)

❖ 磷酸化修饰 (Phosphorylation)最被关注

方法：2D电泳分离后，用P32标记或用磷酸化酪氨酸或磷酸化丝氨酸特异性抗体做Western印迹检测磷酸化蛋白质，再用质谱鉴定。

❖ 糖苷化修饰(Glycosylation)

❖ 脂质修饰 (Lipolyzation)

3) 功能的确定(Determining function)

❖ 酶活力分析或底物确定(Assays for enzymatic activity or determining substrates)

❖ 胞浆物质的生物测定及受体/配体结合性分析(Bioassays for cytokines, receptor/ligand-binding assays)

❖ 蛋白在细胞内定位 (Localization within the cells (GFP fusions))

❖ 使用大范围敲除法或RNA干扰法进行蛋白质组分析 (Proteomic analysis using large-scale mouse knockouts or RNA interference.)

用缺陷型菌株进行表型分析 (Phenotypic analysis using deletion strains)

4) 分子药理学(Molecular medicine)

❖ 发现药物分子 (蛋白) 的作用靶点 (Finding molecular (protein) drug targets)

❖ 用药物破坏蛋白质-蛋白质之间的相互作用 (Disrupting protein-protein interactions using drugs)

重组蛋白质、抗体、抑制剂的动物实验 (Large-scale animal assays for recombinant proteins, antibodies and inhibitors)

5) 二维凝胶的差异显示(Differential display by two-dimensional gels)

❖ 二维凝胶的差异显示(蛋白差异显示图谱)

❖ 一种在很多方面超越于以DNA为基础的阵列方法，用于疾病的特异性检测

❖ 局限性：只适用于生物流体，如血浆和尿液
受翻译后修饰的影响大**6) 蛋白质-蛋白质相互作用的研究(Protein-protein interaction)**

是蛋白质组学研究的主要方面，因为细胞的任何活动都需要蛋白质分子之间或和其它分子的相互作用，这也是调节细胞反应适度的必要条件，是蛋白质相互作用的网络。

❖ Direct DNA readout

酵母双杂交 (Yeast two hybrid)

噬菌体显示 (Phage display)

核糖体显示 (Ribosome display)

RNA-多肽融合分析 (RNA-peptide fusions)

❖ 蛋白质鉴定 (Protein identification): 亲和层析, 质谱, 蛋白质芯片, 表面等离子共振技术 (Biosensor)

7) 蛋白质胞内分布与移位(Localization of protein)

❖ 细胞对胞外信号的快速反应往往只是某些蛋白质的区域移位从而在正确的时空发挥作用

胞内移位的研究：分步离心法和顺序抽提法分离亚细胞结构，进行2D电泳和蛋白质鉴定
胞内蛋白荧光标记法

五. 蛋白质组研究进展**1. 发展速度和规模**

❖ 1995年“proteome”一词问世

❖ 从问世至1997年底两年多时间内，相关文献42篇，其中研究论文28篇，各类评述、综述13篇。研究论文中1995年1篇，1996年4篇，1997年23篇。呈现一种研究和实际应用并驾齐驱的态势。

1997年成立专门研究机构和公司5个，参与国家有澳大利亚、美国、丹麦、瑞士、英、法、日、瑞典、

意、德等10国

- ❖ “Center for Proteome Research and Gene Product Mapping, National Innovation Center” Australia
- ❖ “Australia Proteome Analysis Facility” Australia
- ❖ “Center for Proteome Analysis in Life Science” Denmark
- ❖ “Proteome, Inc. Beverly MA” USA
- ❖ “Large Scale Biology Corp. Rockville MD” USA

2. 研究材料

- ❖ 1995年 Wasinger等第一篇关于蛋白质组研究文章的对象为目前已知最小，但能自主复制的原核微生物---阴道支原体(Mycoplasma genitalium)
- ❖ 1996年研究对象扩展到单细胞真核生物---酵母，以及人正常组织及病理标本，同时突破了早期人们普遍认为的“蛋白质组研究只适合于基因组计划已经完成的生物”的界限
- ❖ 1997年研究对象扩展到14种生物，其中绝大多数为原核生物，但也含多细胞真核生物，如线虫。初步建立了包括人心脏、肝脏在内10余个蛋白质组数据库，其中包括一个完整的酵母蛋白质组数据库(YPD, 8021个蛋白质)
- ❖ 目前研究对象无任何限制，无需基因组计划完成(当然完成更好);无原核生物/真核生物;单细胞/多细胞;组织之分。

3. 研究范围

蛋白质

- ❖ 蛋白质作图;
- ❖ 蛋白质成分鉴定;
- ❖ 蛋白质组数据库构建;
- ❖ 新型蛋白质发掘;
- ❖ 蛋白质差异显示;
- ❖ 同功体(isoform)比较;
- ❖ 蛋白质翻译后修饰的分析等

基因

- ❖ 功能基因组计划;
- ❖ 基因产物识别;
- ❖ 基因功能鉴定;
- ❖ 基因调控机制分析

重要生命活动的分子机制

包括细胞周期，细胞发育与分化，肿瘤发生与发展，环境反应与调节，物种进化等

医药靶分子寻找与分析(新药研究领域)

靶分子类型包括：新型药物靶分子，肿瘤恶性标志，人体病理介导分子，病原菌毒性成分等
总之，蛋白质组研究已经涉及到生命科学中一系列热点领域。

4. 主要进展

蛋白质作图是蛋白质组研究的主要领域

- ❖ 1995年已测定10种蛋白质组图
- ❖ 1998年又完成16种生物或组织的蛋白质组图,其中3种生物(线虫、豆科植物根瘤菌属、Ochrobactrum

anthropi)的蛋白质组图超过1600点,3种蛋白质组图(大肠杆菌,盘基网柄菌、人正常组织和病理组织)建立数据库并上互联网

- ❖ 各类生物中10余种蛋白质数据库初步建立(包括人心脏、肝脏)
- ❖ 第一个完成的蛋白质组数据库--酵母蛋白质组数据库(YPD)完成,含6021种蛋白
- ❖ 提出“蛋白差异显示”概念并用于环境应激、基因突变、病理进程等研究,尤其是癌症研究
- ❖ 建立“Proteome contig”方法,进而使蛋白质组图分辨率提高10倍
- ❖ 提出“蛋白质连锁图”的概念,改进了双杂交系统,并用于蛋白质组相互作用网络的分析
- ❖ 改进了二维电泳方法,现在分辨率达10000蛋白点
- ❖ 联合液体自动取样器与LC/ES-MS技术,蛋白质鉴定速度达到20点/天
- ❖ 建立2DE中糖蛋白、膜蛋白、微量鉴定方法
- ❖ 建立“Streamlined 样品处理”胶转膜技术,突破了大规模蛋白测序与氨基酸组成分析这两种严重影响蛋白质组分析中蛋白质鉴定的限度步骤

5. 蛋白质组研究的整体状况

- ❖ 相关技术已基本配套且基本达到实用化水平
- ❖ 已经形成一定规模的专业队伍及专业机构
草创阶段已经结束,发展阶段已经开始
- ❖ 蛋白质组研究已经成为后基因组时代最重大的生命科学命题之一
- ❖ 蛋白质组分析已经作为专门的技术体系广泛用于生命科学众多领域,尤其是热点领域的研究
- ❖ 蛋白质组的基础研究与分析应用正以指数增长方式发展
- ❖ 是基因组计划由结构走向功能的必然和必需,是生命科学由分析走向综合的必经之路,是连接微观分子系统,运行机制与宏观生物系统,生命活动的桥梁
- ❖ 是21世纪生命科学的重要支柱之一,发展无限

思考题

- 为什么说基因研究是二十世纪生命科学的主线?
- 为什么进行蛋白质组研究?
- 蛋白质组内蛋白质的数目与ORF间有何关系?
- 功能蛋白质组学有何意义?
- 蛋白质组研究与传统蛋白质研究的有何差异?
- 简述蛋白质组学与基因组学的异同?
- 基本概念

第二章 蛋白质组学研究体系 Tools of Proteomics

蛋白质组学研究分类

- ❖ 蛋白质表达模式(蛋白质组成)的研究
- ❖ 蛋白质功能模式的研究

蛋白质组学研究体系

- ❖ 蛋白质组学研究策略

❖ 蛋白质组学研究方法与技术

一. 蛋白质组学研究策略

- ❖ 依赖于分离的研究策略
- ❖ 不依赖于分离的研究策略

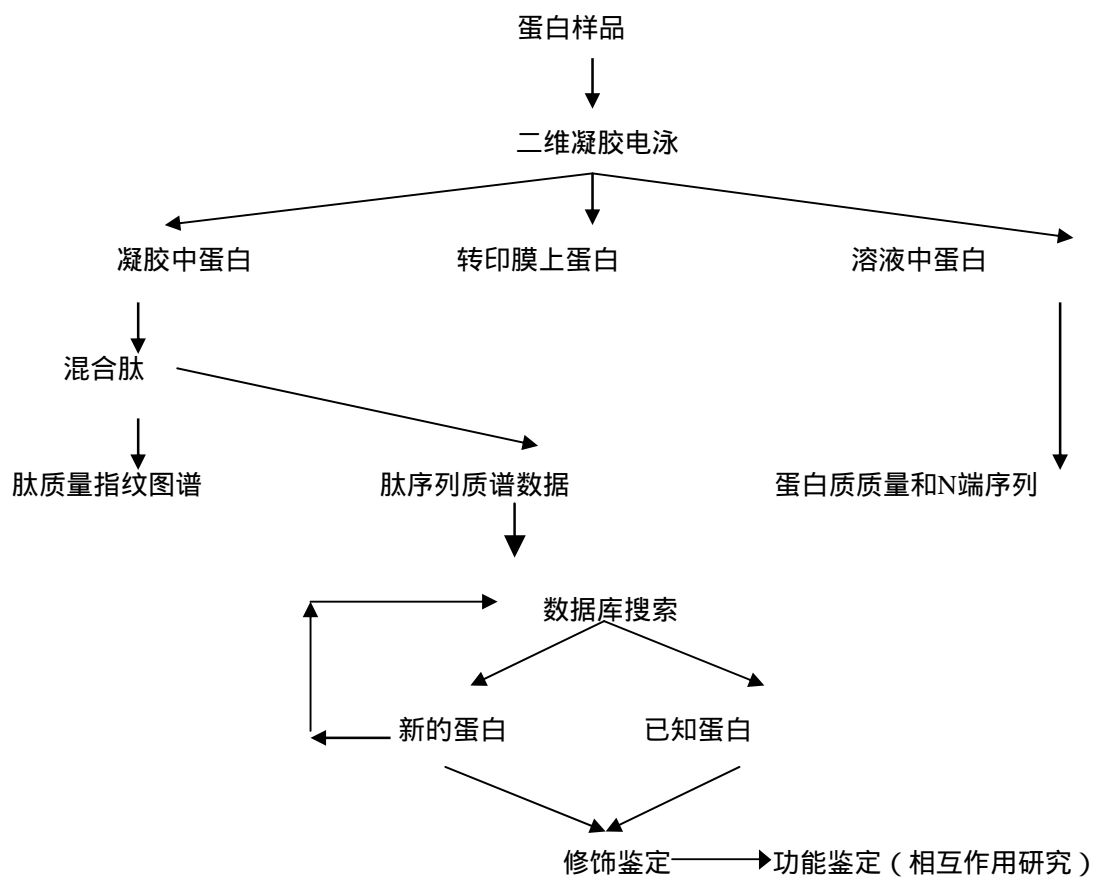
1. 依赖于分离的研究策略

1) 二维凝胶电泳结合质谱：蛋白质组学研究的经典组合

二维凝胶电泳利用蛋白质两种相互独立的物理化学特性进行分离：pI & Mr

利用红外 MALDI-TOF-MS 扫描分析分辨率

基于二维凝胶分离的蛋白质组分析策略



2) 一步分离法：一维凝胶电泳结合电喷雾质谱

- ❖ 利用电喷雾质谱能够分析蛋白质混合物的特性
- ❖ 利用一维凝胶进行初步分离，然后对蛋白质复合物快速分析：1D纯化的蛋白质复合物，酶解相应的蛋白条带（可能含有多种蛋白），再用质谱分析

3) 其它垂直分离法：液相色谱和/或毛细管电泳分离结合质谱

- ❖ 为了克服二维凝胶电泳的局限性
- ❖ 利用两种或多种分离方法达到二维凝胶的分离能力和通量：主要是液相色谱和/或毛细管电泳
- ❖ 液相色谱：高效液相色谱（排阻，反相）
- ❖ 液相色谱和/或毛细管电泳容易实现自动化，且快速，重复性高，灵敏度高；

❖ 缺点：与质谱对接困难（将分离的蛋白移入质谱时易丢失和交叉污染）

2. 不依赖于分离的研究策略

- ❖ 依赖分离的技术局限在于：需要较多样品；需要多步操作
- ❖ 不依赖分离的技术可以减少分析需要样品的量
- ❖ 原理：分子识别
- ❖ 依赖于分子识别的方法：蛋白质芯片
- ❖ 蛋白质识别分子：抗体，分子印迹聚合物
- ❖ 检测方法：基于荧光的光学传感器和生物传感器

1) 蛋白质芯片

❖ 蛋白质芯片关键：空间上固定能严格识别单一蛋白成分分子（蛋白质识别分子）
具备检测单一蛋白成分与其识别分子间相互作用的方法

2) 蛋白质识别分子

- 抗体：最明显，最有效
- 要求：高度特异，甚至可以区分修饰前后的蛋白分子
- 噬菌体显示技术：
- 分子印迹聚合物：人造识别位点模仿通过形状起作用的生化位点，如抗原 - 抗体作用，酶 - 底物作用，受体 - 配体作用

二. 蛋白质组学研究方法与技术

- ◆ 样品准备
- ◆ 样品分离技术
- ◆ 蛋白质鉴定技术
- ◆ 生物信息学
- ◆ 蛋白质功能研究

1. 蛋白质组样品制备

❖ 蛋白质组样品分离成功的关键，直接影响蛋白质组研究结果，样品制备的失败很难通过后续工作的完善或改进获得补偿。

❖ 处理方法必须根据不同样品，不同状态，以及实验目的来合理组合：细胞裂解方法，蛋白质浓缩或稀释，去污剂选择，裂解液配方

1) 蛋白质组样品制备基本原则

- ❖ 尽可能提高样品溶解度，抽提最大量的总蛋白，减少蛋白质的损失
- ❖ 解除蛋白质之间以及与其它生物大分子的相互作用，使蛋白质完全处于变性状态和还原状态
- ❖ 减少对蛋白质以及蛋白质组的人为修饰（假点）
- ❖ 总之：尽可能简单，减少步骤

2) 蛋白质组样品制备的一般过程

对细胞、组织等样品进行裂解、蛋白质进行溶解、变性、解聚，除去非蛋白质部分，提取全部蛋白质。

细胞裂解

- ❖ 机械方法或/和化学方法
- ❖ 保持低温（冰浴，使用预冷的溶液）
- ❖ 直接加入含强变性剂的裂解液

❖ 裂解液中含合适的溶解蛋白成分**温和的裂解方法**

- ❖ 离心收集方法，血细胞，组培细胞
- ❖ 冻融法，细菌，组培细胞
- ❖ 去污剂裂解法，组培细胞
- ❖ 酶解法，植物细胞，细菌，真菌

较激烈的裂解方法

- ❖ 超声波裂解，细胞悬浮液
- ❖ 高压爆破，含细胞壁的微生物（如酵母）
- ❖ 研磨法，固体组织，微生物
- ❖ 匀浆法，固体组织
- ❖ 玻璃珠震荡裂解，微生物，细胞悬浮液

非蛋白物质的去除

- ❖ 核酸，脂，多糖：大分子，可能阻塞凝胶孔径；
- ❖ 核酸：超声（防止产生泡沫）和核酸酶（产生假点）
- ❖ 脂与多糖：超速离心（部分损失）
- ❖ 盐：小分子，浓度过高会降低等电聚焦的电压
- ❖ 盐：透析（时间长），凝胶过滤，沉淀重悬法（部分损失）

避免蛋白水解和降解的措施

细胞裂解的同时，蛋白酶释放

- ❖ 保持低温
- ❖ 变性剂
- ❖ 蛋白酶抑制剂

蛋白质溶解

- ❖ 二维凝胶分离最关键的因素之一：全程保持可溶性
- ❖ 破坏所有非共价结合的蛋白复合物和解聚，形成各个多肽的溶解液 - 一级结构
- ❖ 裂解液主要目的：裂解细胞与溶解蛋白

裂解液中主要试剂

- ❖ 离液剂(Chaotropes)：尿素，硫脲
- ❖ 表面活性剂(Surfactants)：NP-40，CHAPS（两性离子去污剂），SDS
- ❖ 还原剂：DTT，TBP
- ❖ 蛋白酶抑制剂：EDTA，PMSF，或混合抑制剂

裂解液主要成分

- ❖ 变性剂：硫脲比尿素具有更强能力的变性剂，但硫脲水溶性差，只有在浓尿素溶液中可溶性才好，所以尿素 - 硫脲混合物提高样品溶解度
- ❖ 去污剂：SDS：阴离子去污剂，能破坏大多数非共价结合的蛋白，对膜蛋白溶解有效，破坏IEF，所以只用于前期溶解，然后用非离子或两性离子去污剂置换出来。
CHAPS：两性离子去污剂，对膜蛋白溶解有效
- ❖ 还原剂：DTT：含有自由巯基，带电荷，IEF过程中降低溶解度
TBP：不带电荷，增强IEF过程中蛋白溶解度，提高一维向二维转移效率

3) 蛋白质样品准备特别措施

- ❖ 对溶解性差的蛋白质如膜蛋白、与膜相连的蛋白质以及来自具有高抗性组织（头发、皮肤等）蛋白质，需要添加破膜剂、两性离子表面活性剂等以增加蛋白质的溶解度，提高蛋白质的提取率。
- ❖ 有机溶剂提取亲脂性蛋白
- ❖ 不同溶出缓冲液多部提取各种膜蛋白
- ❖ “变性鸡尾酒”：含有14—16个碳的磺基甘氨酸三甲内盐（ASB₁₄₋₁₆）的裂解液，效果最好
- ❖ 使用含离液剂2mol/L硫脲和表面活性剂4% CHAPS的混合液促使疏水蛋白质从等电聚焦胶上转换
- ❖ 三丁基膦（TBP）取代巯基乙醇或 DTT 完全溶解链间或链内二硫键，增加蛋白溶解度使得向第二向转移量增加
- ❖ 固态组织样品的处理：Banks等用激光捕获微解技术降低组织选取面过大，组织异质性样品（组织样品或病理样品往往包含不同类型细胞）带来的干扰，是研究样本定位更准确
- ❖ 亚细胞分级、蛋白质预分级、提高加样量、应用敏感性检测法提高敏感性
- ❖ 处理碱性样品提高电泳效果

激光捕获微解技术(Laser capture microdissection, LCA)

- ❖ 将一张转移膜置于选取部位表面，倒置在显微镜下，用直径30 μ m或60 μ m的激光束聚焦其上，膜活化并与选取部分结合将其从周围其他组织中提取出来。
- ❖ 优点：富集样品使丰度更集中，定位更准确

实例1---疏水蛋白质

- ❖ Molloy用氯仿、甲醇（有机溶剂）、硫脲，去污剂（表面活性剂）、三丁基膦化氢（还原剂）等溶剂提取出许多疏水蛋白质。质谱检测表明13种蛋白质中有9种在从前的双向电泳中从未出现过

实例2---低丰度蛋白

- ❖ 样品分级：降低样品复杂度，富集低拷贝蛋白；亚细胞分级，选择性沉淀，分级提取（蛋白质在一系列溶解能力逐渐增强的裂解液中溶解性能不同而实现）
- ❖ Cibacron Dye色谱技术浓缩
- ❖ 用ProtoClear 技术处理，人血浆中至少95%白蛋白和97%IgG被清除。

实例3---极碱性蛋白

- ❖ 浓缩：TCA/丙酮沉淀法，可以失活蛋白酶，减少蛋白降解，去除干扰物，富集极碱性蛋白，如从全细胞裂解液中富集核糖体蛋白
- ❖ 样品处理提高电泳效果：碱性范围内凝胶基质不稳定，会产生逆向电渗透，除了可以用N,N' 双甲基丙烯酰胺增加基质稳定性外，对等电点超过10的碱性蛋白质，通过1 - 10%的山梨醇梯度和16%的异丙醇减少不稳定性，提高电泳效果

2. 蛋白质组样品分离技术

- ❖ 一维电泳(1D-PAGE)
- ❖ 二维电泳(two dimension polyacrylamide gel electrophoresis, 2D PAGE)：双向电泳
- ❖ 其他方法

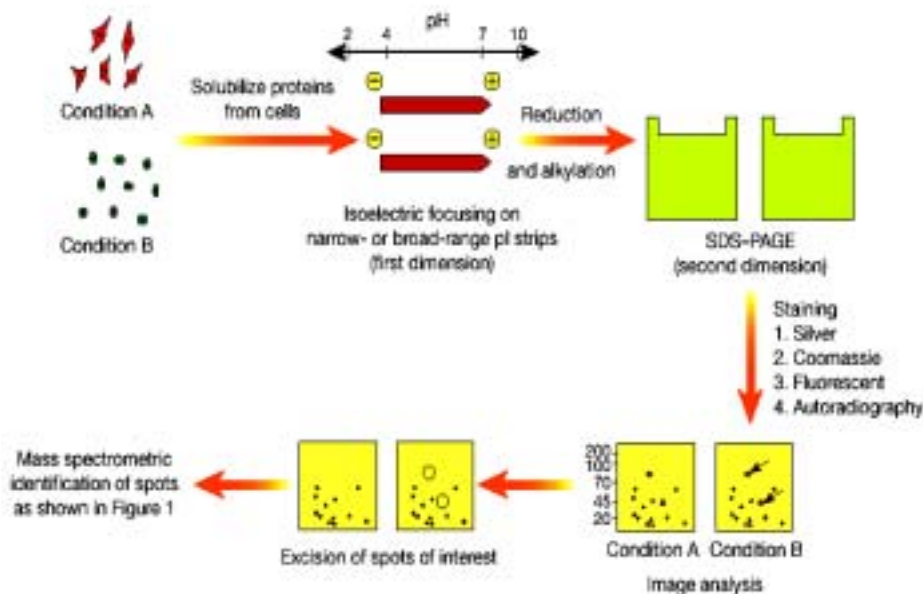
1) 二维凝胶电泳

- ❖ Smithie & Poulik (1956)：纸电泳结合淀粉凝胶电泳
- ❖ 聚丙烯凝胶支持介质以及聚丙烯凝胶浓度梯度
- ❖ 等电聚焦技术(isoelectric focusing, IEF)应用，结合SDS-PAGE的二维凝胶电泳

二维凝胶电泳基本原理

- ❖ 第一向是等电聚焦，根据蛋白质等电点(pI)不同的分离
- ❖ 第二向是SDS-PAGE，根据蛋白质分子量不同的分离。
- ❖ 一般采用垂直电泳或水平电泳，两者差别不大，均适于分子量10-150kd的蛋白质。

二维聚丙烯酰胺凝胶电泳



良好电泳图谱的标准

- ❖ 分辨率高：保证最大程度分离
- ❖ 重现性好：不同次，不同人，不同实验室之间
- ❖ 灵敏度高，低背景，少纹理

等电聚焦技术

- ❖ 蛋白质的等电点：蛋白质是两性分子，在不同的环境下（pH值）可能带正电荷，负电荷或不带电荷，等电点是其净电荷为零时的pH值
- ❖ 等电聚焦技术：在一个pH梯度和外加电场下，蛋白质有移向pH梯度中使其净电荷为零的点的倾向。（带正电荷移向阴极，带负电荷移向阳极）。IEF可以基于极微小的电荷差异而分离蛋白，具有高分辨率。
- ❖ 分辨率的决定因素：pH梯度，电场强度（高电压，低电流），变性条件

等电聚焦技术的发展

- ❖ 根据第一维等电聚焦电泳的条件和方式的不同，二维凝胶电泳分为三种系统：ISO-DALT, Nonequilibrium pH gradients electrophoresis (NEPHGE), IPG-DALT
- ❖ 蛋白质分辨率提高1-2个数量级

ISO-DALT方法

- ❖ 在聚丙烯酰胺管胶中，由载体两性电解质（carrier ampholyte）在外加电场下形成pH梯度，等电聚焦，Anderson称此为ISO-DALT技术
- ❖ 载体两性电解质：可溶性的两性小分子物质，在等电点附近有很高的缓冲能力，在外加电场下，低等电点的载体两性电解质混合物会移向阳极，相反高等电点的载体两性电解质混合物会移向阴极而形成连续的pH梯度

ISO-DALT方法的不足

- ❖ 载体两性电解质是一种高聚物的混合物，批次间的差异影响了IEF的重复性
- ❖ 载体两性电解质形成的pH梯度不稳定，易产生阴极漂移，同时管状胶的玻璃壁上的负电荷也会加剧这种漂移，结果是碱性区蛋白的丢失

❖ 强度差，难操作

非平衡pH梯度电泳

❖ 主要用于分离碱性蛋白质 (pI 7-11.5)

❖ 样品加在酸性端，蛋白质在已经分开但还没有达到各自等电点的时候就停止电泳，电泳展开时间相对比较短

❖ 因为如果聚焦达到平衡状态，碱性蛋白会离开凝胶基质而丢失，故等点区域的迁移需在平衡状态前完成 (pH梯度快速形成)

❖ 很难控制，重复性差。

固相pH梯度 (IPG) 等电聚焦

❖ IPG的产生：通过Immobiline™共价偶联丙烯酰胺，产生固定的pH梯度

❖ 可以精确制作线性、渐进性甚至S型曲线的pH梯度，范围或宽或窄

固相pH介质 - Immobiline™

❖ Immobiline是一类具有弱酸或弱碱性缓冲基团的丙烯酰胺衍生物，与Arc和Bis有相似的聚合行为

❖ Immobiline分子的一个双键在聚合过程中通过共价结合镶嵌到聚丙烯酰胺骨架中，在Immobiline分子的另一端有缓冲基团R (羧基或叔氨基)，在聚合物中形成弱酸或弱碱的缓冲体系，pH梯度的范围和形状通过变化两个相对酸和相对碱的混合溶液配方混合聚合而成，pH值梯度一般分布在3 - 10

❖ Immobiline本身由于水解、自聚合，与pH有关N-氧化物形成等，使得immobiline分子发生变化而不稳定

❖ LKB公司生产出的immobiline II 大大提高了稳定性和重复性

IPG等电聚焦优点

❖ pH梯度稳定，不产生电极漂移 (共价聚合固定梯度)

❖ 可以生成任意的适当的pH梯度用于不同目的分离，甚至做到pH相差1单位的胶条，预制胶条商品

❖ 上样量大，可达几十毫克，可以用于微量制备

❖ 重复性好，使不同人，不同实验室之间的实验结果能够相互比较

❖ 强度好，易操作

IEF实验程序

❖ 胶条选择和准备

❖ 胶条的重溶胀 (rehydration, 再水合作用)

❖ 加样与电泳运行

❖ 二维间的平衡

IPG的选择

❖ 快速筛选或高丰度蛋白质：短胶条

❖ 高分辨率和大上样量：长胶条

❖ 观察全部蛋白质分布：pH3 - 10范围

❖ 使用窄pH范围能更精确分析在这感兴趣的pH范围内的蛋白质 (如放大胶)

❖ 非线性梯度胶条也可以在特定pH范围使蛋白质分布更均匀，通过分辨率

实例

❖ 新的酸性 (pH3-5)、碱性 (pH6-11) 的IPG，联合商品化的pH4-7的IPG，可对蛋白形成蛋白质组重叠群 (proteomic contigs)，从而有效分离 (*Electrophoresis* 1997, 18: 1384-1392)

❖ 使用pH5 - 6的胶条分离的蛋白质斑点间距离明显比pH3 - 10的胶条大

胶条的重溶胀

❖ 总的原则：样品中蛋白质的溶解要求

❖ 典型的重溶胀条件：尿素，去污剂，还原剂，IPG缓冲液，指示剂

加样与电泳运行

- ❖ 加样：利用加样杯或直接加在重溶胀液中 (up to 1mg)
- ❖ 电泳运行：IPGphor

运行与聚焦条件的优化

- ❖ 目标：最好的图谱质量和重复性
- ❖ 考虑因素：蛋白质样品，上样量，所选的pH范围，IPG胶条长度
- ❖ 选择变化：电压和聚焦时间
- ❖ 电压：重涨泡 - 初始化 - 稳定电聚焦
- ❖ 聚焦时间：短 - 水平条纹；长 - 电渗 - 图谱变形

二维间的平衡

- ❖ 目的：蛋白转移（使分离蛋白与SDS更好结合而正常迁移）
- ❖ 措施：在50mM Tris缓冲液（含SDS, 尿素, 30%甘油, DTT/TBP）平衡15min后，再在50mM Tris缓冲液（含SDS, 尿素, 30%甘油, 5%碘乙酰胺）平衡15min
- ❖ 尿素和甘油：减缓电渗，提高转移率
- ❖ 碘乙酰胺：烷基化自由的DTT

二维SDS-PAGE基本原理1

- ❖ 第二向的聚丙烯酰胺凝胶电泳含SDS (sodium dodecyl sulphate)，SDS在样品和胶中的存在使蛋白质所带的电荷不再是第二向蛋白分离的因素
- ❖ 阴离子去污剂SDS以1.4g比1g蛋白质的比率（结合计量常数）包裹在多肽骨架外，掩盖了蛋白质的电荷，形成一个阴离子复合体，每一个复合体具有稳定的净负电荷
- ❖ SDS同时破坏蛋白质间的氢键，阻断蛋白质间的疏水作用，部分打开蛋白分子，减少分子形状对分离的影响
- ❖ 完全打开蛋白分子：DTT的加入破坏胱氨酸间的二硫键，使电荷均匀在分布打开的蛋白分子长度上
- ❖ SDS&DTT：使蛋白质分离仅仅依赖其分子量成为可能，蛋白质分子量和相对迁移距离之间的动力学关系成线性，值得注意的是这种线性关系仅在一定分子量范围下成立，且和一定的胶浓度有关

缓冲液与胶浓度

- ❖ 缓冲液：非连续缓冲系统（Tris-甘氨酸）
- ❖ 胶浓度：单一浓度胶或梯度（线性或非线性）胶，针对不同分子量的蛋白质的分离

单一浓度胶	MWX10-3	梯度胶	MWX10-3
	0.05 36-200		5-15% 14-200
	0.075 24-200		5-20% 10-200
	0.1 14-200		10-20% 10-150
	0.125 14-100		
	0.15 14-60		

二维凝胶电泳的优点

- ❖ 高分辨率：二维凝胶电泳的分离能力与有效分离面积成正比
- ❖ 高重复性：IPG，胶条的预制化，集成化电泳设备的开发
- ❖ 高重复性&高分辨率：二维凝胶电泳可以成功用于蛋白质组学研究的关键

局限性

- ❖ 灵敏度有限——显色反应 - 低丰度蛋白
- ❖ 耗时，耗力
- ❖ 自动化程度低
- ❖ 后续凝胶图谱分析和数据分析是瓶颈

二维凝胶电泳的适用

- ❖ 适于大量未知功能蛋白的分离：依赖于蛋白质生理化学性质分离，与功能无关---适于目前大量未知功能蛋白的分离，可获得一特定代谢状态下，组织或细胞蛋白质现有表达水平的“shooti nggun”式分离
- ❖ 可以比较不同条件下（如不同药物或化合物处理前后，压力诱导等，或生理或疾病状态下），细胞或组织裂解物的定量蛋白表达图，比较不同条件对蛋白表达的影响
- ❖ 适于微量一步分离纯化得到同源纯度的蛋白进行生化或化学分析：直接从胶上洗脱或转移膜上洗脱得到蛋白，用来制备单克隆抗体，分析其受体结合性，酶活性，结构特征

二维凝胶电泳的应用

- ❖ 检测正常和非正常生长分化中，如癌症发生中的蛋白表达的表型依赖的变化
- ❖ 确定诱导剂、carcinogens、激素处理、mitogen刺激、营养变化引起的特定蛋白变化
- ❖ 分析人、动物、植物、微生物组织和细胞的性质

实例

- ❖ 小鼠组织和E. coli 总提取物
O'Farrell用放射自显影法得到1000多个点，并指出理论上 $70 \times 100 = 7000$ 多组分，该差异是混合样品中组分含量差异引起的（掩盖效应）

二维电泳凝胶的蛋白质斑点检测 (spot detection)

- ❖ 通常蛋白质可以通过UV吸收来检测，但是由于聚丙烯酰胺在同一紫外范围也有吸收，所以需要染色
- ❖ 基本要求：染色方法应该对所有蛋白都有效（方便定量合进一步分析）
- ❖ 染色法：有机染料（考马斯亮蓝），银染
- ❖ 荧光标记法：SYPRO
- ❖ 放射标记法：稳定同位素标记

考马斯亮蓝法 (Coomassie blue)

- ❖ 考马斯亮蓝：灵敏度 1 μ g
- ❖ 胶体考马斯亮蓝 (Neuhoff)：1. 灵敏度 200ng；2. 染色过程通常在酸性（TCA）环境中进行，辅以甲醇；由于胶体染料的使用，只有少量自由扩散到溶液中的作用染料分子可以穿透凝胶基质，选择性地对蛋白质进行染色，大大地降低了化学噪音，提高了灵敏度

胶体考染法

- ❖ 优点：考马斯亮蓝定量地与蛋白质结合，可以准确定量蛋白质；
- ❖ 缺点：蛋白质谷氨酸羧基侧链不可逆的酯基化（酸催化下），使得肽谱数据失真

银染法 (silver staining)

- ❖ 银染是最普遍最常见的电泳凝胶染色的技术
- ❖ 灵敏度：~1ng
- ❖ 原理：在碱性条件下利用甲醛将自由银离子还原成银，沉积在蛋白质分子表面而呈色
- ❖ 优点：大大提高了灵敏度；扩展了动态线性范围兼容质谱（去除戊二醛后）；快速，方便
- ❖ 缺点：银离子可与半胱氨酸残基结合；银染试剂戊二醛和甲醛都会对蛋白质的 -和 -氨基烷基化，妨碍后续分析

荧光标记法

- ❖ 包括荧光标记在内的发光染料法较比色法染色更灵敏，动态线性范围更广
- ❖ 共价结合的荧光标记：共价结合需要对蛋白质进行修饰，蛋白质和蛋白质之间荧光基团修饰的功能基团数目不同而造成灵敏度不稳定，标记与非标记蛋白质之间细微的分子量差异可能导致后续分析的偏差
- ❖ 非共价结合荧光标记：与SDS-蛋白质复合物形成非共价结合在疏水环境中显示荧光，结合计量常数稳定，适合定量检测
- ❖ 灵敏度：~2ng
- ❖ 优点：简单，快速（一步法染色）；适合后续定量分析（CCD数码相机或激光扫描仪）；线性范围广；

光稳定（抗淬灭）

- ❖ SYPRO染料
- ❖ 发光金属螯合染料：SYPRO Ruby
- ❖ 灵敏度：~1ng
- ❖ 优点：快速，灵敏，终点式染色易操作；动态线性范围广，染色覆盖范围广；不干扰后续分析

荧光标记法与银染法

- ❖ 鼠纤维原细胞
- ❖ 印染法：1290点，Sypro Ruby：1063点
- ❖ 肽质量图谱和动力学线性范围

稳定同位素标记法

- ❖ 稳定同位素标记提高检测灵敏度： ^{15}N ， ^2H ， ^{13}C
- ❖ 体内标记（稳定同位素代谢标记法）
不同条件（正常培养基或含 ^{15}N 培养基）下平行生长的微生物混合提取蛋白后经二维凝胶电泳分离，选出兴趣蛋白点后酶解，经过质谱分析， ^{15}N 结合的肽比正常肽大一个质量单位，使得不同生长条件下的蛋白样品得以区分，同时这样的肽质谱峰成对出现，可以根据其相对峰面积来定量蛋白质，从而观察到不同蛋白质的表达

同位素编码亲和标签法1 (Isotope-coded affinity tag, ICAT)

- ❖ 胶后标记及不依赖分离的蛋白质组分析
- ❖ ICAT试剂：
 - ❖ 亲和标记（biotin）分离ICAT标记蛋白；
 - ❖ 连接子：整合同位素，重（含8个氘 ^2H 原子），轻（含8个氢 ^1H 原子）；分别标记的ICAT分子量正好相差8
 - ❖ 活性基团，用来特异结合半胱氨酸的巯基，
 - ❖ 不同条件下的细胞提取物经二维凝胶电泳分离，选出兴趣蛋白点进行蛋白酶酶解，用两种含不同同位素试剂分别标记，随后的MALDI-TOF-MS分析可以通过 $^1\text{H}/^2\text{D}$ 比率进行蛋白质的相对定量，甚至可以对凝胶未能完全分离的含多个蛋白质的点进行相对定量
- ❖ 优点：广泛兼容各种条件下的蛋白质
即使在盐，去垢剂，稳定剂（SDS，尿素）等存在也可以进行标记胱氨酸，比较简单
对分离方法兼容
- ❖ 缺点：ICAT是个大修饰物
无法分析不含胱氨酸的蛋白质

2) 其它分离方法

- ❖ 凝胶分离除了受显色方法的影响外，超高灵敏度的质谱技术的出现使凝胶分离的缺点暴露在阳光下，即单块胶呈现蛋白质组全貌，使很多蛋白质得不到充分分离。
- ❖ 毛细管电泳（CE）
- ❖ 高效液相层析（HPLC）
- ❖ 毛细管等电聚焦电泳（CIEF）
- ❖ 微毛细管电泳（MCE）
- ❖ 发展阶段，目前不能实现重复性和复杂蛋白样品的完全分离

多重色谱

- ❖ 多维分离系统：将蛋白或多肽经过两个或多个独立的分离方法，并使一步分离后的两种成分保持分开的方法
- ❖ 二维分离系统的分析能力：理论上等于第一向峰容量和第二向峰容量的乘积
- ❖ 方法

- 二维液相电泳 (2D-LPE)
- 二维高效液相色谱 (2D-HPLC)
- 二维色谱 (液相色谱-毛细管电泳) : 最成功

二维液相电泳 (2D-LPE)

- ❖ 基于分子量和等电点pI
- ❖ 优点: 1 蛋白质保留在液相中完成分析, 分段收集、酶消化后即可质谱鉴定
2 高载样量, 许多低丰度蛋白质也能得到检测和测定
- ❖ 缺点: 1. 目前分辨率不如2D-PAGE
2. 已经用于溶解和鉴定来自于人脑脊髓液和胸膜洗脱液的蛋白质, 发现了少数高丰度蛋白质

二维高效液相色谱 (2D-HPLC)

- ❖ 离子色谱和反向HPLC分析人血浆滤液多肽 (Rai da等), 洗脱液直接流入ESI-MS得到3000余种不同多肽质谱, 但只确定了极少的几种多肽
- ❖ IEF-RPHPLC-MALDI-TOF-MS (Wall 等) 分辨率与2D-PAGE相近, 但只有非常少蛋白质能被鉴定 (38/700)

二维高效液相色谱全自动系统

- ❖ 大肠杆菌裂解物先经阳离子交换柱分离, 再经RP柱分离, ESI-MS在线检测
- ❖ 大肠杆菌裂解物经分子排阻层析后, 用RP-HPLC分离。得到450个蛋白质, 其中14个蛋白质通过MS和Edman测定
- ❖ 优点: 具有较大的蛋白容量和高通量, 很适合自动化
- ❖ 缺点: 缺乏质谱仪的直接界面, 不能完全自动化进行蛋白鉴定 (缺乏鉴定)

二维色谱

- ❖ 大肠杆菌胞质蛋白经强阴离子HPLC分离, 并打断, 每20个片段为一份进行胰酶消化后过RP-HPLC微柱, 串联质谱仪鉴定, 结果80个蛋白质被鉴定, 通过SEQUEST分析鉴定

SEQUEST算法

- ❖ ENG等开发
- ❖ 将观察到的多肽串联质谱与从蛋白质和核酸数据库中得到的理论质谱相匹配, 鉴定混合物中的蛋白质
- ❖ 消化成多肽的混合物上反相微柱 (RP), 借助HPLC泵, 将多肽进行反相梯度洗脱, 进入串联质谱, 结果通过SEQUEST算法鉴定蛋白质 (确定蛋白质的同一性)
- ❖ 分离和鉴定相结合

二维色谱

- ❖ Tong等, 酵母核糖体 (75个蛋白质)
消化多肽—固相RP提取盒—一步法CE柱分离—ESI串联质谱仪鉴定
136个蛋白质进行了测序, 其中66个蛋白质通过SEQUEST算法得到鉴定
- ❖ Link等, 啤酒酵母全细胞
消化混合物—强阳离子交换 (SEX) 和RP两种微毛细管柱---ESI-MS-MS (189种蛋白质被鉴定)

3. 蛋白质鉴定技术

- ❖ 基本分析方法
 1. 图像分析
 2. 质谱技术
 3. 氨基酸组份分析
 4. 微序列测定
- ❖ 修饰蛋白质的分析

1) 图像分析 (Image analysis)

- ❖ 2DE图谱呈满天星式分布, 不能凭直觉靠肉眼分析一个图谱上斑点的深与浅, 有与无以及斑点的位置

置，必须依靠强大的计算机技术进行处理包括定量分析

- ❖ 同一蛋白斑点在一组凝胶图谱上发生的变化对于生物学分析来说可能意义重大
- ❖ 图像分析的内容：图像加工、斑点检测、背景消减、斑点配比、数据库构建等

图像分析的步骤

- ❖ 采集图谱：利用电荷耦合CCD照相机或激光密度扫描仪对二维电泳图像数字化
- ❖ 图像加工：在图像灰度水平上进行图像加工
- ❖ 检测斑点：将有意义的斑点与背景分离，精确测定斑点的强度以及面积等，如控制斑点的最高峰，检测斑点的边缘及邻近
- ❖ 图像分析比较即斑点配比：事实上2DE要达到100%的重复性是极其困难的，因此凝胶间蛋白质斑点的配比是一个挑战
- ❖ 生成参考凝胶图，建立2DE数据库

图像分析的完成手段

- ❖ 斑点配比的最大程度相似性通过斑点配比 **向量算法** 在斑点的长度和平行度方面观测得，另外配合相似性分析，聚类分析以及等级分析方法
- ❖ Malanie II软件
- ❖ Ficker软件(www.lecb.ncifcrf.gov/flicker/)
- ❖ 2DE数据库：储存凝胶图谱中蛋白斑点间的描述性信息，对所分析的一组图像对出明确的注释，如一块综合性标准图谱

比较各种状态下凝胶图的作用

- ❖ 病态与正常态的比较是灵敏的诊断工具
- ❖ 整体水平解析蛋白质功能(如某蛋白质的突变无异常表现，而在2DE蛋白质群水平上却有显现效应，表现为突变影响了相同代谢途径上的其他蛋白质)
- ❖ 不同组织的参考图比较可以揭示管家蛋白和特异蛋白质
- ❖ 跨种族比较提供新的信息

2)氨基酸组份分析

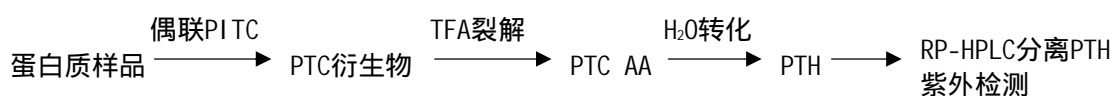
- ❖ 原理：利用不同蛋白质具有特定的氨基酸组份特征来鉴定蛋白质
- ❖ 通过测定蛋白质中各氨基酸摩尔百分数或各氨基酸的摩尔比率，然后与数据库中已知蛋白质的理论值进行比较，给出匹配分数值大小，分数值越小，表示其与真正蛋白质越接近
- ❖ 检索软件：AACompEdent, ASA, FINDER, AAC-PI, PROP-SEARCH等
- ❖ 方法：蛋白质样品经155°C酸水解，使氨基酸自动衍生在经色谱分离，测定
- ❖ 优点：经济、是一种独特的指纹技术
- ❖ 缺点：灵敏度低，约需蛋白样品十几皮摩尔；速度慢；酸水解可能不彻底；酸水解破坏某些氨基酸，需用校正

3)微测序技术(microsequencing)

- ❖ 理论依据：蛋白质的末端氨基酸序列有惊人的专一性，43%-83%的蛋白质可以用N端4个氨基酸残基来确定，74%-97%的蛋白质能用C端4个氨基酸残基确定，提高残基数，鉴定的专一性更高
- ❖ 现在微量测序已经成为蛋白质分析和鉴定的基石(对未知蛋白质的有力鉴定)
- ❖ 目前已经实现自动化，可以用于亚皮摩尔水平蛋白质的鉴定

包括：N端测序，C端测序和Ladder测序

N端测序--Edman降解法的测序原理



Edman降解法特点

C-末端测序法的改进

- ❖ 烷基化TH法，裂解前形成的TH被S-烷基化，使得中间物易于裂解，[NCS]-裂解，一步完成裂解和活化步骤
- ❖ 自动化仪(ABI Model 477或Proclise Instrument Platforms) 50多种重组或天然蛋白(7.6-97kD)最多获得8个C端序列，最少4个
- ❖ AA修饰
- ❖ 其他异硫氰试剂 (DPP-ITC, Ac-ITC)，较少侧链反应，增加测序分析物的初产率

C-末端测序法的不足

- ❖ 目前C端序列仪一般只能测定5个氨基酸残基左右，有些可测到10个左右
- ❖ 分辨率为nmol 水平

Ladder 测序

- ❖ 是一种末端降解与质谱技术相结合的序列测定法
- ❖ 级联由蛋白或多肽的C-或N-末端转化出一系列片段 (每个间相差1个AA的肽序列阶梯)，用质谱分析ladder分子量，由分子量的差别确定AA类型从而测得ladder序列。
- ❖ 优点：高数据获得率、测定时间短、样品需求少、灵敏度高 (皮-飞摩尔)

方法

- ❖ 化学方法序列梯子：基于Edman降解，在Edman降解中按设计好的耦合和裂解步骤用一定量的序列终止剂，获得一系列的N端截断的肽混合物，再经过质谱分析，在质谱图中连续离子间的质量差异与相应的氨基酸质量相对应产生肽序列
- ❖ 酶法序列梯子：通过外肽酶消化肽过程中于不同时间段取出反应物，得到一系列C端或N端序列梯子，混合进行质谱分析，只要包含10个以上连续氨基酸就能达到pmol水平

S-P-A-F-D-S-I-H-A-E-T-L-K (protonated mass 1410.6)			
mass*	b-ions	y-ions	mass*
88.1	S	PAFDSIMAEYLK	1323.6
185.2	SP	AFDSIMAEYLK	1226.4
256.3	SPA	FDSIMAEYLK	1155.4
403.5	SPAF	DSIMAEYLK	1008.2
518.5	SPAFD	SIMAEYLK	893.1
605.6	SPAFDS	IMAEYLK	806.0
718.8	SPAFDSI	MAEYLK	692.3
850.0	SPAFDSIN	AEYLK	561.7
921.1	SPAFDSINA	EYLK	490.6
1050.2	SPAFDSINAE	YLK	361.5
1151.3	SPAFDSINAEI	LK	260.4
1264.4	SPAFDSINAEIL	K	147.2

Ladder 测序示意图

4) 质谱技术 (mass spectrometer, MS)

- ❖ 质谱分析是分析化学中已经广为使用的技术方法，其基本原理是分析样品在特定的条件下转变为高速运动的离子，这些离子根据质量/电荷比的不同在静电场和磁场的作用下得到分离，用特定的检测器可以记录不同质荷比的各种离子的相对强度并形成质谱图
- ❖ 通过质谱分析，可以获得样品分子量、分子式、分子中同位素构成和分子结构等多方面的信息

质谱技术

- ❖ 是蛋白质组学发展具有重要的推动意义的技术，现代质谱仪具有电离蛋白质及其它一些生物大分子的能力，它具有高效、灵敏、准确、自动化、可同时分析多蛋白混合物等特点；已逐步取代了传统的Edman降解测序与氨基酸组份分析法，成为蛋白质鉴定的核心技术
- ❖ 软电离技术为质谱分析进入生物及医学研究领域开辟了道路
- ❖ 计算机技术的发展进一步提高和扩展了质谱技术的应用能力
- ❖ 串联质谱

质谱仪的基本组成

- ❖ 离子源：产生气相离子
- ❖ 质量分析仪
- ❖ 离子检测器

质谱仪是在真空状态下分析离子的质荷比 (m/z)，有不同的离子源与质量分析器 (检测器) 的组合

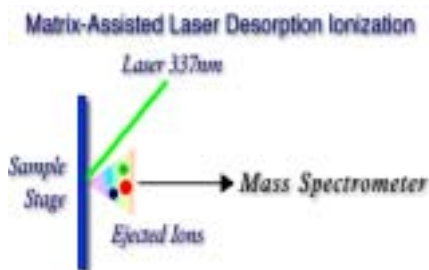
- ❖ 样品准备及进样
- ❖ 数据分析

生物样品的电离方法

- ❖ 离子源—将分子转变为气态的离子化的分子的装置
- ❖ 传统的电子轰击电离 (EI)，化学电离 (CI) 对生物大分子不适用
- ❖ 软电离技术：在离子化过程中不会 (最多只是部分) 破坏分子结构，保留分子完整性，不会形成碎片，能实现分子质量大于10000质量单位的生物大分子的质量分析
- ❖ 现代新型离子源：
 1. 基质辅助激光解析电离(matrix-assisted laser desorption ionization, MALDI)
 2. 电喷雾电离(electrospray ionization, ESI)

基质辅助激光解析电离 (MALDI)

- ❖ 以脉冲方式使样品离子化，从固相标本中产生的离子，并在飞行管中测定分子量
- ❖ 将分析物分散在基质分子 (尼古丁酸及其同系物) 中并形成晶体，用激光照射晶体，由于基质分子经辐射所吸收的能量导致能量蓄积并迅速产热，从而使基质晶体升华，导致基质和分析物膨胀并进入气相
- ❖ 样品被包裹在基质中，大部分激光能量被基质的发色基团首先吸收，从而保护了样品分子的完整性



MALDI 演示图

基质辅助激光解析电离的基本步骤

- ❖ 分析物与基质混合加到金属片上挥发后，形成晶体，送入真空室
- ❖ 激光脉冲照射样品基质混合物，基质晶体吸收激光能量，热释能使基质晶体升华，基质与样品分子气化进入质谱仪气相
- ❖ 在气相中发生一系列质子化和去质子化作用，实现离子化，进入飞行时间质量分析仪，通过飞行管道的离子被检测器检测

MALDI-TOF-MS

- ❖ MALDI离子化过程进入质量检测器的离子具有相同的终动能，因此速度与 m/z 成反比，离子小速度快，所以离子的飞行时间反映了其 m/z ；
- ❖ 通过已知质量和电荷状态的化合物的飞行时间计算
- ❖ MALDI离子化过程产生单电荷离子，其飞行时间与分子质量的平方根成反比

MALDI-MS的特点

- ❖ 产生的基本上是分子离子峰，能提供样品的精确分子量信息
- ❖ 目前MALDI-TOF质谱质量检测数：检测相对分子质量上限为 350×10^3 。
- ❖ 由于MALDI技术中样品是和基质混合上样，因而MALDI技术对样品的纯度要求并不高，可以直接

分析细菌培养液、细菌、菌体、动物组织等非常复杂的生物样品

电喷雾质谱 (ESI-MS)

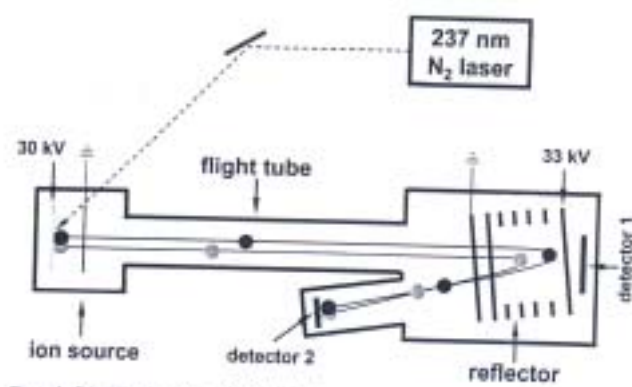
- ❖ 是一种连续离子化的方法，从液相中产生离子，联合四极杆质谱或飞行时间检测器测定分子量
- ❖ 在毛细管出口处施以高电压（几千伏），所产生的高电场使从毛细管喷出的液相样品雾化成细小的带电液滴，随着溶剂蒸发，液滴表面缩小，液滴电荷密度增加，最终导致液滴爆裂，分析物从液相中释放出来，以多电荷离子形式进入气相

质量分析仪

- ❖ 质量分析仪：利用被分析物的物理性质，如电场、磁场或飞行时间等，分离特定质/荷比的离子，并将此离子传送到离子检测器中
- ❖ 生物大分子量通常超过数千甚至数十万，传统的质谱无法达到这种大分子量的测定：
 1. 大轨道半径和薄片叠加的高场强磁铁，提高测定范围，扫描速度及分辨率
 2. 超导技术达到高磁场强度，测定高分子量，达到高分辨率
 3. 采用飞行时间测定，理论上飞行时间测定质量无上限，适合生物大分子质量测定
- ❖ 由于MALDI是脉冲式的，其对应的检测装置也是脉冲式的飞行时间（time of flight, TOF）质谱检测器
- ❖ 与ESI相连的多为四极杆质谱检测器或离子阱质谱检测仪

MALDI-TOF-MS 技术的改进

- ❖ MALDI离子化过程中存在动能分散的问题，稳离子和亚稳离子之间动能差异会削弱质谱的分辨力
- ❖ 加离子镜反射(reflector)：能量聚焦作用，使低动能离子早反射，高动能离子晚折返，所以低速离子赶上高速离子，减少离子之间动能差异，提高分辨率
- ❖ 再加离子镜聚焦降低离子能量分散：延迟取样
- ❖ 源后衰变



MALDI-TOF-MS 的质量测定仪

四极杆质谱与离子阱质谱

- ❖ 在四级杆质谱中，第一和第三级是质量过滤器，而第二级则充当产生碎片离子的碰撞室，从前一级传送来的肽离子在碰撞室内经惰性气体的碰撞诱导，产生正离子，称为碰撞诱导解离(CID)，ESI通常产生带多电荷的肽离子
- ❖ 离子阱质谱中离子的聚集和碰撞在离子阱中发生。通过离子的捕获和选择发射，具一定荷质比的离子从离子阱中分离出来检测。如果这一离子在离子阱中再裂解，得到的碎片也被发射，就可以得到MS/MS串联质谱

纳喷技术

- ❖ 纳喷技术(Nano-electrospray)：需求样品量少，可以在fmol水平上

ESI-MS与MALDI-MS的简单比较

	理论质量检测范围 (实际检测范围)	优点	缺陷	检测样品
ESI-MS	约为 200×10^3 (70×10^3)	能与HPLC、CZE联用;能形成多价离子,可更为准确地测定分子量;质量分辨率为2000以上*;可以直接从水相观察非共价复合物;敏感度为fmol ~ pmol;质量测定精确度为 $\pm 0.01\%$	在分析混合物时,多价离子形成会使结果分析比较困难;仅在低盐($< 1\text{mmol/L}$)条件下给出理想信号;样品组成过于复杂时可能不产生信号	肽和蛋白;多糖;核酸;其余带电的小分子
MALDI-MS	大于 300×10^3 (150×10^3)	能够耐受高盐;具有较高的质量检测上限;能分析混合物;敏感度为fmol;可发展为蛋白及核酸测序手段	质量分辨率 $< 500^*$;质量测定精确度为 $\pm 0.1\% \sim \pm 0.01\%$;不能与液相色谱联用	肽和蛋白;多糖;核酸;组成复杂的生物样品;其余带电的小分子

质谱技术优缺点

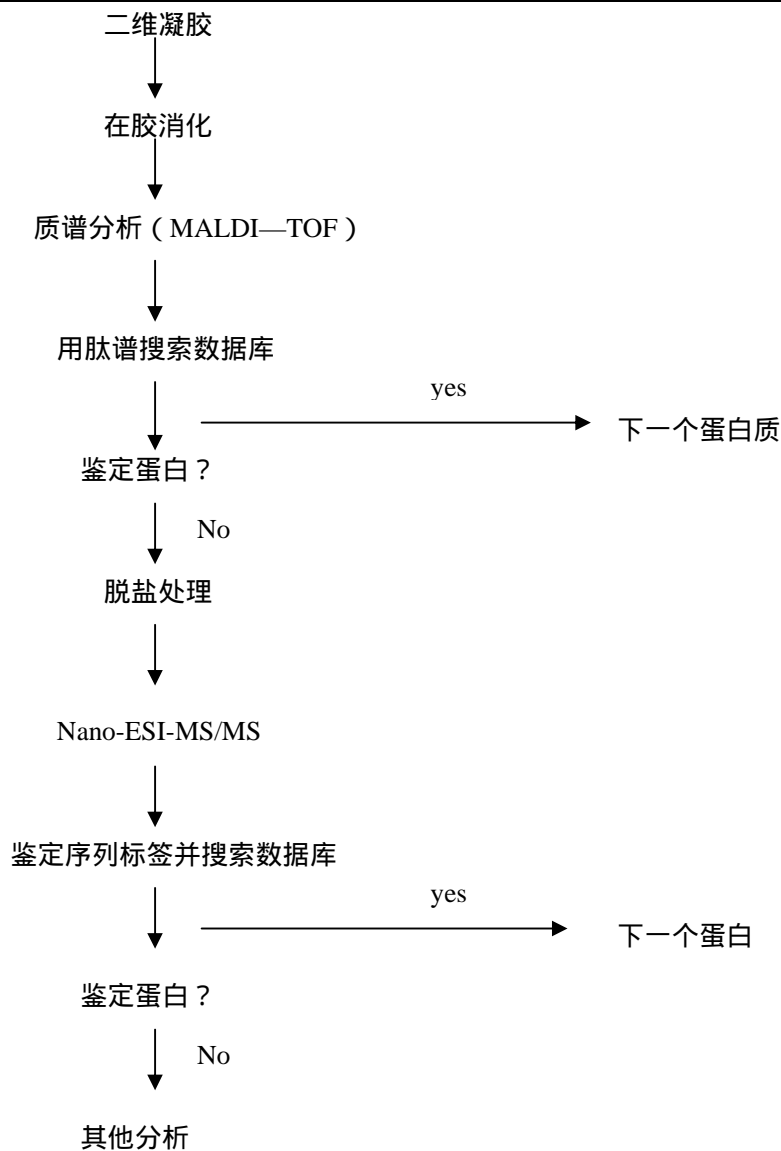
- ❖ 优点：灵敏度高，蛋白质用量较少(fmole or even attomole)
- ❖ 缺点：质谱分析需要蛋白质数据库或EST数据库中已有蛋白的数据，而这两个数据库中蛋白质的数量是有限的

质谱技术的用途

质谱技术可以进行的从基因组到蛋白质功能的研究

问题/任务	有关的质谱技术
在基因组、蛋白质序列库和EST序列库用MALDI或ESI-MS,ESI-MS/MS中筛到的蛋白是已知蛋白吗?	MALDI-PSD做肽谱,或做全蛋白的ESI-MS/MS
如未知,提供足够的序列信息做克隆蛋白鉴定,二级修饰,二硫键,异构体(序列错误)。	ESI-MS/MS, MALDI-PSD分子量测定,再用MALDI或ESI-MS或MS/MS做肽谱
高级结构:折叠,稳定性,单体或多聚体	用ESI-MS监测重氢交换, MALDI-或ESI-MS监测表面标记,非变性条件ESI-MS, MALDI-MS监测交换
蛋白质何时,和什么分子,怎样相互作用	亲和技术与MALDI-或ESI-MS结合, MALDI-或ESI-MS监测表面标记和有限水解

基于质谱的蛋白质组鉴定策略



质谱技术应用

- ❖ 鉴定蛋白质
- ❖ 检查蛋白质完整性,均一性和纯度分析
- ❖ 用稳定同位素标记,进行蛋白质定量分析
- ❖ 鉴定蛋白质复合物组成
- ❖ 研究蛋白质分子的相互作用
- ❖ 鉴定蛋白质的三维结构
- ❖ 确定蛋白质翻译后修饰类型与发生位点等

鉴定蛋白质

- ❖ Ladder测序
- ❖ 肽片段离子谱和片段测序
- ❖ 蛋白全谱分析
- ❖ 肽谱(肽质量谱)

Ladder 测序

- ❖ 是一种末端降解与质谱技术相结合的序列测定法
- ❖ 原理：N-末端常用化学法，C-末端常用蛋白水解消化法
- ❖ 质谱：FAB，PD(plasma desorption), ESI, MALDI-TOF
- ❖ 优点：高数据获得率、时间短、样品损失少、灵敏度高（皮-飞摩尔）

MALDI-TOF-MS的改进

- ❖ MALDI离子化过程的问题：离子化过程中动能有微小分散，即产生亚稳离子，会发生中性小分子（如水，氨）的丢失，甚至肽键断裂
- ❖ 结合源后衰变（post source decay, PSD）技术和串联质谱（tandem MS），也可以获得样品分子的结构信息

源后衰变(PSD)

- ❖ MALDI-MS中的两种不同的片段反应：
 - 源内衰变：发生在离子源内，在激光撞击之后几百纳秒内离子即刻片段化，可以通过线性时间飞行质谱中检测到
 - 源后衰变：发生在离子源后的第一个无场区域，时间跨度为微秒；由于发生在无场区域，所以产生的不同片段离子和母离子保持同样速度，用离子镜反射，将片段离子和母离子分离，按质量大小排列可形成离子谱，称为PSD谱

肽片段图谱

- ❖ 肽片段不仅可以在PSD中被诱导产生，在CID也可以见到
- ❖ 这些在质谱气相中产生的肽碎片离子，断裂方式具有序列特异性，只有沿肽骨架的肽键断裂，如果断裂时肽离子的正电荷保持在碎片离子的N端，称为b系列离子，如果电荷保留在C末端，称为y系列离子
 - ❖ b系列离子：C端缺失，N端完整
 - ❖ y系列离子：N端缺失，C端完整
 - ❖ b系列和y系列离子的混合图谱：碎片离子图谱，肽片段离子图谱

肽片段图谱的特点和用途

- ❖ 特点：包含大量的重复信息，如重叠的b系列和y系列离子，同一母离子的多个中间碎片离子（parent and daughter），而成为极丰富的序列信息
- ❖ 比如：
 - ❖ 在PSD谱中将那些仅有一个氨基酸质量差异的系列片段排列，可以推测肽片段序列
 - ❖ 可以粘贴得到完整肽段（母离子）的质量

蛋白全谱分析

- ❖ 蛋白全谱分析：是组份分析，在完整的组织等或提取液中识别出尽可能多的肽和蛋白质混合物
- ❖ 测得的组份分子量，和已知或预期的蛋白质相应的分子量比较

全谱分析的步骤

- ❖ 提取多肽和蛋白质，并分馏提取液
- ❖ 质谱测定提取液中各个组份的分子量
- ❖ 数据库搜索被测的分子量是否符合特定的肽或蛋白质的质量
- ❖ 进一步确认分析：

全谱分析的特点

- ❖ 优点：不要求蛋白样品的分离，只需要提取或简单分馏
- ❖ 局限：
 - ❖ 对仪器的测定质量范围，分辨率，精确度，灵敏度有要求；
 - ❖ 虽然蛋白质和核酸数据库不断丰富，但是即使经过搜索匹配的分子质量也不能确保分子的确认，需要另外的信息支持

全谱分析的应用

- ❖ 用来定性鉴定给定组织的肽和蛋白，从而检测出肽和蛋白的变异和缺失
- ❖ 定量研究在对疾病，药物，内源调节子和环境压抑的响应中，肽或蛋白的整体响应轮廓。例如：对神经免疫内分泌这个复合系统来说，神经，免疫，内分泌三者共同形成一个复杂网络，通过整个网络，响应可被单一的，局域的或规则的事件触发，因此需要监测所有的肽和蛋白，而不是仅仅个别组份。
- ❖ 定量全谱研究，可以建立未鉴定组份在不同实验条件下的上下限，以确定对那一种未鉴定蛋白深入研究

肽质量谱

- ❖ 通常测定蛋白质的全序使十分困难的，这种困境使我们难以象描述DNA分子那样描述蛋白质的性质和结构，然而利用蛋白裂解后生成的一系列肽片段，可以得到与该蛋白质序列有关的具有指纹特征的肽质量谱，将肽谱与蛋白质数据库信息结合起来成为鉴定蛋白质的有效方法
- ❖ 待测定的蛋白质样品，经特异性酶解为肽段，然后应用单级或串联质谱分析，可得到肽质量指纹(peptide mass fingerprinting, PMF, 即肽谱)与部分氨基酸序列，使用相关数据库比对软件，即可实现已知蛋白质的定性鉴定

全谱分析与指纹分析

- ❖ 全谱分析：组份分析，识别完整组织中的肽和蛋白质混合物；指纹分析：测定单一的肽段质量
- ❖ 全谱分析：未破坏组份分子量与已知或预期蛋白质分子量比配；指纹分析：降解产物分子量与已知或预期蛋白质降解产物分子量比配
- ❖ 都需要数据库支持和搜索

肽谱实验策略

- ❖ 肽谱是目前凝胶分离蛋白质鉴定的最常用手段
- ❖ 肽段由已知切点专一的蛋白酶产生
- ❖ 产生的肽段质量由MALDI或ESI-MS精确测定
- ❖ 计算蛋白质序列库中的每一个蛋白质序列被实验所用的已知切点专一的蛋白酶水解后的理论肽段的质量，并建立数据库
- ❖ 将被测样品的肽质量与数据库中的肽质量理论值相匹配，计算匹配排序值

影响匹配的可能因素

- ❖ 翻译后修饰，翻译后加工，人为修饰而产生的另外质量
- ❖ 非特异的蛋白水解
- ❖ 杂质蛋白的影响
- ❖ 同源蛋白的存在
- ❖ 实验精确度

影响肽谱鉴定的决定性因素

- ❖ 肽质量测量的精确度：可靠性
- ❖ 蛋白酶的专一性：纯度

肽谱的应用

- ❖ 确证蛋白质的序列，进行蛋白质鉴定
- ❖ 鉴定蛋白质修饰（翻译后以及体外化学修饰）
- ❖ 检测单个氨基酸的突变
- ❖ 鉴定不同蛋白构型：巯基和二硫键的测定
- ❖ 对照天然蛋白和重组蛋白的指纹图等等

肽谱分析常用手段

- ❖ 二维凝胶电泳
- ❖ 高压液相色谱
- ❖ 毛细管电泳 - - - 混合肽片段图谱

❖ 质谱 - - - 质量肽谱**具体应用**

- ❖ 蛋白质序列的确证
- ❖ 巯基和二硫键测定
- ❖ 蛋白质修饰的测定
- ❖ 在数据库中检索和鉴定蛋白质

蛋白质序列确证

- ❖ 应用DNA重组技术生产重组蛋白的工作中，常常需要确证重组蛋白质的氨基酸序列是否从cDNA推导的氨基酸序列一致
- ❖ 质量肽谱是个理想方法，而且具有快速、准确、蛋白质用量少的特点
- ❖ Gibson 和Biemann在1988年最早应用这一方法确证了从大肠杆菌中重组的谷氨酰基tRNA合成酶和甘氨酰基tRNA合成酶，以及从酵母中重组的甲硫酰基tRNA合成酶
- ❖ 现广泛应用

二硫键测定

- ❖ 二硫键是人们关心的蛋白质结构的一个重要方面
- ❖ 它在稳定蛋白质的三级结构方面有重要作用，并且常常与蛋白的活性有关
- ❖ Glocker等人使用HPLC-FAB-MS的方法研究了重组人M-CSF中所含的二硫键，并确定了蛋白折叠过程。使用溴化氰裂解后再用内切酶进一步裂解，片段用MS分析鉴定二硫键

蛋白质修饰的测定

- ❖ 质量肽谱的方法对蛋白中的很小的修饰可以提供足够的准确度。实测分子量与预测分子量的差异，通常提示着蛋白质发生了修饰。使用ESI-MS测定分子量3万的蛋白时，可以精确到1Da，其误差仅0.01%，这样的准确度能描述化学修饰的特性

蛋白质修饰的测定

- ❖ 翻译后的修饰都能从质量数差异上表现出来
- ❖ N乙酰化后分子量就会增加42Da
- ❖ C端形成酰胺分子量则减少0.98Da
- ❖ 氨基酸缺失的产物引起的质量差在0.2-1%的水平时即可检出

乙酰化修饰

- ❖ Spetzler JC和Bouchon等人分别用MADLI-MS和ESI-MS对蛋白质N端乙酰化的情况进行了分析
- ❖ N端封闭血吸虫属曼纹属抗原蛋白Smp28经胰蛋白酶裂解后测得N端片段分子量为695.8，氨基酸组份分析表明该片段含有等摩尔的EGHAIK，与cDNA推导的序列AGEHIK相比，分子量相差41.7Da，表明N端片段的丙氨酸被乙酰化。
- ❖ 同时，该蛋白分子量比含有N端丙氨酸乙酰化的结果还要大14Da，是由于七个蛋氨酸中有五个被部分氧化

乙酰化修饰

- ❖ 另外，还有两个其它修饰的产物，其中的半胱氨酸与相应的半胱氨酸或谷胱甘肽形成共价二聚体
- ❖ 重组Smp28的研究证实了上述结果
- ❖ Smp28的LC-MS联机质量肽谱分析对一级结构的鉴定比脱机分析时减少了两小时
- ❖ SDS-PAGE分离后使用毛细管柱进行LC-ESI-MS分析，只需要7pmol的样品

在数据库中检索和鉴定蛋白质

- ❖ 把质谱测定的分子量同数据库信息结合起来，鉴定蛋白结构
- ❖ 可用的数据库有：PIR,Swiss-Sport蛋白序列数据库、基因库等

在数据库中检索和鉴定蛋白质

- ❖ John R Yates等应用质量肽谱在PIR数据库中鉴定不同来源的细胞色素C的序列，发现因种属不同而产生的氨基酸序列的差异

- ❖ 先酶解蛋白，然后用FAB-MS测定质量肽谱，再在数据库中检索
- ❖ 结果表明，只需测定总片段肽片段数目一半的肽片段的质量，就可以在数据库中确定蛋白质的序列
- ❖ 同样方法分析了一组GTP结合蛋白
- ❖ 在重组蛋白中，经常可以发现氨基酸的突变或丢失

在数据库中检索和鉴定蛋白质

- ❖ Mann. M等人发现，肽片段分子量的准确度达到0.1-0.01%时，就可以在数据库中找到相应的序列
- ❖ 该方法可以快速准确地利用数据库在分子水平上定位蛋白质或基因的序列
- ❖ 随着数据库的不断扩充，该应用会越来越广，成为一种常规的研究蛋白质性质的方法

研究蛋白质修饰

- ❖ 糖蛋白
- ❖ 脂蛋白
- ❖ 磷酸化蛋白：蛋白质修饰研究的重点在蛋白质的磷酸化

糖蛋白的分析

- ❖ 联机的LC-MS方法为分析蛋白的糖基化提供了有效的方法

例1-- rtPA的糖基化

- ❖ John V.O.等
- ❖ 将rtPA经膜蛋白酶裂解后，进行了肽谱分析
- ❖ 糖基化发生在裂解片段T45上
- ❖ 在分离过程中可以看到二级、三级和四级糖肽

例2-- tPA的性质

- ❖ Guzzetta等
- ❖ 用LC-ESI-MS在短时间内对tPA的性质进行了鉴定
- ❖ 酶解---液相色谱分离---ESI-MS检测
- ❖ 得到在不同保留时间相应的质量肽谱(m/z 对保留时间作图)
- ❖ 进一步用糖酶酶切，进行LC-MS分析，从而确定了糖基化位点；同时还鉴定T11片段上二种杂化糖的连接
- ❖ 使用LC-MS选择性离子数据文件，形象地表示了相近糖肽的洗脱顺序
- ❖ 检出tPA突变体， ^{103}Thr 被Asp取代。
- ❖ 同时，LC-MS质量肽谱研究表明，突变物中含有新的糖链。
- ❖ 质量肽谱法为快速分析微量糖蛋白，获取结构信息提供了有效的方法。

MALDI-TOF-MS与糖蛋白的研究

- ❖ MALDI-TOF-MS的质量肽谱方法研究重组糖蛋白比FAB-MS研究非衍生化糖蛋白提高10-100倍的灵敏度
- ❖ 使用正离子或负离子MALDI-TOF-MS分析酶解混合物，表明唾液酸化的糖蛋白易于在阴离子模式下检出
- ❖ 分析糖降解前后的蛋白，通常不需要移去缓冲液就可以测出糖化的信息
- ❖ 结合已知糖的分子量，可以确定糖的结构

脂蛋白的分析

- ❖ 脂质修饰蛋白的方式：
 - 氨基的烷基化
 - 胱氨酸通过二硫键的脂酰化
 - 通过磷酸酯键与C端的羧基结合
- ❖ 脂蛋白的分析方法：
 - 利用有机溶剂（如氯仿/甲醇）将脂蛋白从富含脂肪的组织中抽提出来，然后利用Lipidex-5000的反

相层析初步纯化

- 利用C18反相HPLC分离
- 在蛋白序列已知情况下，ESI-MS可以清楚地判断肽链的共价修饰，有机溶剂液适合电喷雾

磷酸化蛋白的分析

- ❖ 磷酸化是翻译后最重要的修饰之一，它常与代谢调节途径或是信号传递有关，可逆的蛋白质磷酸化/去磷酸化在生命调控中非常重要
- ❖ 磷酸化的存在及位点一直是研究蛋白质结构与功能的一个重要方面
- ❖ 最常见的磷酸化类型是：丝氨酸、苏氨酸、和酪氨酸的磷酸酯化（真核生物中）

磷酸化分析的策略

- ❖ 定位：确定蛋白质磷酸化氨基酸残基位点
- ❖ 鉴定磷酸化过程的有关的激酶
- ❖ 观察磷酸化对功能的影响
- ❖ 细胞内磷酸化蛋白的含量非常少，有时，某一蛋白质中仅有一小部分被磷酸化，同时同一个蛋白质中也会存在多种不同的磷酸化类型

磷蛋白的检测和分离

- ❖ 产生磷酸肽样品的方法：凝胶分离蛋白，水解蛋白条带或斑点，提取磷酸肽混合物样品
- ❖ 确定磷酸化蛋白质的方法：
- ❖ ³²P标记后放射自显影法：放射标记的ATP被蛋白激酶用于磷酸化底物
- ❖ Western印迹法：利用酪氨酸-磷酸盐特异抗体

磷酸肽的分离

- ❖ 磷酸肽通常很少，容易淹没在背景中，肽分离技术浓缩分析物，以提高信噪比
- ❖ 肽分离技术也可以除去非肽类杂质，更有利于检测和分析低丰度的磷酸肽
- ❖ 分离技术：2D-PP,RP-HPLC,凝胶电泳，固相金属亲和色谱

确定磷酸化氨基酸类型

- ❖ 可以限定可能的磷酸化位点
- ❖ 磷酸氨基酸分析：水解³²P标记磷酸蛋白质与磷酸氨基酸标准品混合后TLC分离，蛋白质中存在的磷酸氨基酸可以提高染色标准品和放射标记种类的相关分析得到
- ❖ 特殊磷酸氨基酸抗体分析

确定磷酸化位点

- ❖ 磷酸化蛋白质首先纯化，然后特异裂解产生肽混合物 - 磷酸肽
- ❖ 磷酸肽质谱分析：依靠测定磷酸酯基团使肽段增加的质量数，通常磷酸化研究阶段，蛋白质的序列是已知的，从理论上可以通过对磷酸基团加合在丝氨酸、苏氨酸、和酪氨酸残基上引起的80质量单位数差分析而鉴定磷酸肽
- ❖ 磷酸化位点的确定通过串联质谱的产物离子扫描技术

例1--- MAP Kinase Kinase

- ❖ Resing等人对MAP Kinase Kinase(MAP-KK)的磷酸化情况进行了分析
- ❖ 用ESI-MS的质量肽谱法确定了MAPKK经V-Mos催化的磷酸化位点
- ❖ S218和S222，与cAMP依赖蛋白的T环结构相吻合；并且S222位的磷酸化优于S218
- ❖ 自动磷酸化过程发生于几个残基位置上，如S298，Y300，极少数的自动磷酸化也发生于T23，S299，S218或S24，S25。

例2---视紫质的磷酸化

- ❖ 与视觉有关的视紫质的磷酸化是由视紫质激酶催化的，发生在多重Ser和Thr位点
- ❖ 以往的研究表明，修饰多发生在C端区域
- ❖ 用ESI-MS和MS-MS的质量肽谱法，确认磷酸化是在视紫质激酶作用下发生在三个位点上

- ❖ Asp-N内切酶裂解后C端片段含磷酸化的位点。
- ❖ 分离单磷酸化、双磷酸化和三磷酸化的片段，确定了每个磷酸化的位点

例3--牛磺蛋白的磷酸化

- ❖ Hasegaiva等人应用ESI-MS质量肽谱技术确定了从Alzheimer病的病人脑组织中分离出来的牛磺蛋白的磷酸化位点
- ❖ 患者蛋白常比正常人富含牛磺酸
- ❖ 比较去磷酸化前后两种蛋白的肽谱
- ❖ 病人体中的蛋白在Thr231和Thr235明显地发生了磷酸化
- ❖ 191位至225位，386位至435位也存在着多余的磷酸化，Ser262至少存在部分磷酸化，显示了病变者的蛋白与正常人相比的不同之处

研究蛋白质分子的相互作用

- ❖ 实例：Mark等，利用ESI-MS和固定化酶技术筛选抑制剂文库；对比固定化酶作用前后的抑制剂文库的质量谱，如果哪个峰消失，说明该峰所代表的化合物是与此酶结合紧密的配体

4. 生物信息学(Bioinformatics)

- ❖ 蛋白质组学也是一门信息科学，因为蛋白质组学涉及到一系列的信息，如序列信息、序列加工信息、表达信息、结构信息、功能信息，以至于蛋白质之间的相互作用网络更象现代意义上信息网络
- ❖ 蛋白质组具有“整体性”和“动态性”的特点，其研究对象的数量大、种类多，而且随着时间和空间变化而不同，所积累的研究数据复杂多样；数据库和分析软件，可用于这些数据的存储、管理与分析，是蛋白质组的研究不可缺少的信息学手段

生物信息学

- ❖ 蛋白质组较基因组更复杂
- 1. 基因组是固定的，蛋白质组则随发育阶段、驻留组织甚至所处的环境的变迁而变化
- 2. 98%的人类疾病是多基因的(polygenic)和/或后成的(epigenetic)，疾病的诊断和治疗的选择可能完全依赖于后成的调节和患者所处的环境
- 3. 组织中mRNA的丰度与蛋白质丰度的统计相关不显著
- 4. 几乎所有的蛋白质都会经过翻译后的不同类型和不同程度的修饰

生物信息学的组成

- 数据库
- 计算机网络
- 应用软件

蛋白质组与互联网

- ❖ 蛋白质组在互联网上
蛋白质组计划是国际性的，为了方便快捷和有效地交换及共享信息资源，应用互联网WWW规程，用超文本标志语言(html)，将超文本文件按超文本格式描述，可以不依赖与计算机的类型和操作系统
- ❖ 第一个关于蛋白质组学的服务器Expasy(蛋白质分析专家系统)：<http://www.expasy.ch>，建于Geneva大学，由瑞士生物信息学研究所(SIB)与欧洲生物信息学研究所(EBI)共同协作维护

蛋白质组数据库

- ❖ 基因组相对比较好定义，即全部基因的碱基序列，以及由此推导出的开放阅读框架、外显子、内含子、启动子、增强子以及染色体的结构区域
- ❖ 蛋白质组由于涉及蛋白质的动态变化，所以完整的蛋白质组资源不能简单地看作氨基酸序列的组合
- ❖ 目前没有一个数据库可涵盖蛋白质组学研究的各个领域，所以出现了蛋白质组学相关领域的各种数据库
- ❖ 各种不同类型的蛋白质数据库，是蛋白质组研究数据存储、管理的主要形式，是实现已知蛋白质的分析鉴定与未知蛋白的发现前提，也是总结蛋白质结构、性质与功能的规律，实现模拟与预测的基础。

蛋白质组相关数据库分类

- ❖ 蛋白质氨基酸序列(sequence)数据库,常用的有SWISS-PROT、PIR、TrEMBL、GenPept等
其中PIR(Protein Information Resource)和SWISS-PROT是目前最常用的两大序列数据库,除序列之外,还提供已知蛋白质的功能、结构域、翻译后修饰、突变体等详尽的注释以及和其他重要相关数据库的交叉检索,是最重要的蛋白质初级数据库
- ❖ 蛋白质结构域与家族数据库,常用的有ProSite、BLOCKS、DOMO、ProDom、Profam等
- ❖ 蛋白质高级结构及分类数据库,常用的有PDB、SWISS-MODEL、NRL-3D、BioMagResBank、MMDB、SCOP、CATH等
其中PDB(Protein Data Bank)最为著名,收录有蛋白质原子坐标、晶体结构和核磁共振的数据、一级和二级结构信息及相关文献引用
- ❖ 蛋白质二维凝胶图谱数据库,包括SWISS-PROT的2DPAGE、SEI NA的2DEPAGE等综合性2DE数据库和不同生物、不同器官、组织、细胞的专一性2DE数据库
- ❖ 蛋白质相互作用数据库,常用的有ProNet、DIP、INTERACT等
- ❖ 除以上数据库之外,还有各种类型的专一性蛋白质数据库
- ❖ 上述数据库资源,均可通过访问Expasy北京大学镜像cn.expasy.org获得链接

SWISS-PROT蛋白质序列数据库

- ❖ SWISS-PROT蛋白质序列数据库,是目前最广泛应用的蛋白质组数据库,是Expasy服务器的一个组成部分
- ❖ SWISS-PROT蛋白质序列数据库除了包含蛋白序列信息外,还包含大量的注释
- ❖ SWISS-PROT蛋白质序列数据库信息来源是文献,引用了一千多种期刊上发表的蛋白质序列及相关信息

SWISS-PROT的特点

- ❖ 可提供蛋白质序列的详尽注释信息,包括蛋白质功能、蛋白质翻译后的修饰、结构域、结合位点、二级结构、四级结构、蛋白质缺陷相关疾病
- ❖ 避免过多的重复,即对同种蛋白质的多个记录进行仔细比较后归结到一个记录中,分别检索
- ❖ 与其它专业数据库可以超级连接,交叉参考,以求一次检索可获得蛋白质的各方面信息
数据库中的记录显示格式规范化

TrEMBL蛋白质序列数据库

- ❖ TrEMBL是计算机注释的蛋白质序列数据库,是SWISS-PROT的辅助数据库
- ❖ 包含尚未结合到SWISS-PROT的EMBL核酸序列数据库中所有编码序列的翻译内容
- ❖ 可以看作SWISS-PROT的初级部分,所以TrEMBL的数据分离总是标明“Preliminary”,而非“Standard”
- ❖ TrEMBL与SWISS-PROT的记录结构以及检索方法是一致的

SWISS-PROT/ TrEMBL数据库的检索

- ❖ 可以通过Expasy和EBI
- ❖ 常规检索方法:通过首页上的检索窗口输入检索词检索,包括:登记号、标志号、序列描述内容、基因名称、物种名称等
- ❖ 检索途径:全文检索、登记号或标志号检索、序列描述或标识检索、著者检索、文献源检索
- ❖ 检索结果显示
- ❖ 序列数据上传:研究者可以将直接测序的肽序列数据上传至SWISS-PROT数据库

其它蛋白质组相关数据库

- ❖ 二维凝胶图谱数据库SWISS-2DPAGE

应用软件

- ❖ 蛋白质二维凝胶电泳图谱分析软件
- ❖ 蛋白质鉴定工具软件

❖ 蛋白质的结构分析与功能预测工具软件**蛋白质二维凝胶电泳图谱分析软件**

比较知名的有

- ❖ Geneva Bioinformatics提供的Melanie 3
- ❖ 美国Proteome Works公司的商用软件包PDQuest6.1
- ❖ 英国公司Nonlinear Dynamics Ltd.的Progenesis
- ❖ 德国Decodon GmbH的Delta 2D等

二维凝胶电泳图谱分析软件

应用此类软件可完成:

- ❖ 电泳图谱背景消减与条纹消除
- ❖ 斑点探测
- ❖ 图形匹配和定量测定
- ❖ 数据存储与分析等
- ❖ 个别功能强大的软件还具有数据统计、不同凝胶的比较、数据库联接检索、斑点注释等特殊功能

Melanie III

- ❖ Melanie III兼容性好,可以在SUN工作站、Machintosh和PC上运行,并提供友好的菜单界面,高度自动化,同时还可以超级链接各种蛋白质数据库,互参数据

蛋白质鉴定工具

- ❖ 蛋白质组分离工作完成后,要低费用、高效率和大通量地识别和鉴定数量众多的蛋白质必须采取不同于基因组测序方案直接测序,而是测定一些能够刻画蛋白质属性的各种属性参数,同时把已有数据库中的序列也转换成相应的属性参数,形成属性化数据库,通过搜索和匹配鉴定蛋白质,如果搜索不到,在进行生化鉴定

蛋白质的属性参数

- ❖ 分成:

1. 一级参数:完整蛋白质分子的属性,包括蛋白质质量、等电点、N端和C端的序列标记、蛋白质的氨基酸组份构成
2. 二级参数:蛋白质的片段属性,包括肽质量标记、肽序列标记

- ❖ 对鉴定蛋白质来说,若独立的属性参数少,则蛋白质鉴定出错率高,所以通常组合各种属性参数,以提高鉴定的正确率

蛋白质鉴定工具(见下页表)

- ❖ 蛋白质的鉴定,需要利用实验得到的相关数据,通过相关的算法与程序,进行已知蛋白质数据库的搜索比对来完成
- ❖ 目前已有的鉴定工具:氨基酸组成比对、肽片段质量比对和部分肽段序列比对三类
- ❖ 常用的有AAComplident、PeptIdent、SEQUEST、MultiIdent等

蛋白质的结构分析与功能预测工具

- ❖ 蛋白质翻译后修饰的预测
- ❖ 蛋白质序列分析与二级结构预测
- ❖ 蛋白质结构域预测
- ❖ 功能域预测
- ❖ 蛋白质跨膜区预测等
- ❖ 这些工具的应用可以对蛋白质相关结构、性质与功能的实验研究,起到一定的指导参考作用

蛋白质结构

- ❖ 蛋白质功能与其结构密切相关,蛋白质通过其结构发挥作用,了解结构是研究其功能的先决条件
- ❖ 蛋白质分子结构的特征是:结构的层次性
- ❖ 一级结构:氨基酸序列

- ❖ 二级结构：主链之间借助氢键形成的构象
- ❖ 三级结构：二级结构基础上氨基酸残基侧链之间相互作用形成的结构
- ❖ 四级结构：亚基之间空间关系

蛋白质鉴定工具

属性参数	工具名称
蛋白质质量 + 序列标记	PeptideSearch
	TagIdent
肽质量标记	MassSearch
	PeptideSearch
	MS-Fit
	ProFound
肽质量标记 + 序列标记	MS-Edman
肽序列标记	MS-Tag
	PepFrag
	PeptideSearch
氨基酸组份	Seguest
	AACompIdent
	ProSearch
氨基酸组份 + 序列标记 + 肽质量标记	MultiIdent
模拟酶解	MS-Digest
解释翻译后修饰	

蛋白质结构预测

- ❖ 分为两类：

一、采用分子力学、分子动力学方法，根据物理化学原理，理论上计算蛋白质分子的空间结构。理论依据：一个蛋白质分子的溶液中的天然构象应是热力学上最稳定的、自由能最低的构象

二、通过对已知的空间结构的蛋白质进行研究和分析，找出蛋白质一级结构和空间结构之间的关系，总结出一定的经验规律；如：根据一级结构预测二级结构；根据二级结构建立蛋白质三维模型；再根据经验规律和能量最低原理排除不合理模型以及修正模型

5. 蛋白质功能研究

1) 蛋白质相互作用

蛋白质组学

- ❖ 了解细胞内蛋白的功能
- ❖ 研究细胞内各种蛋白质相互作用及相关物理化学、生物化学性质及功能

细胞功能

- ❖ 具体细胞过程（代谢途径、调节途径、信号级联传导等）中一个蛋白质是呈蛋白质复合组团的形式起作用的
- ❖ 可以通过蛋白质与其它分子间物理和功能的相互作用确定
- ❖ 一个蛋白发挥正常功能首先需要定位正确

❖ 蛋白活力的调节（如调节特定组织、细胞或时间基因表达）在细胞功能中非常重要

分子功能

- ❖ 特定结合位点（底物、小分子效应物、核酸、其他蛋白质）及动力学特性
- ❖ 通过一系列化学结构（如氨基酸序列）和动力学性质（如构象状态）的分析得到
- ❖ 三维结构对于体现功能了解功能尤其重要
- ❖ 翻译后修饰（糖基化、磷酸化、原肽切割、蛋白剪接等）也影响/决定分子功能

蛋白质相互作用研究

- ❖ 蛋白质的功能往往体现在它与其他分子（蛋白质/核酸等）的相互作用中
- ❖ 细胞内各种重要生理过程，如信号传导、细胞对外界环境及内部环境变化的反应等，都是以蛋白质相互作用为纽带，并形成网络而实现
- ❖ 蛋白质相互作用的结构基础：介导蛋白质 - 蛋白质之间相互作用的结构域(domain)

蛋白质的相互作用

- ❖ 蛋白质：功能大分子
- ❖ 蛋白质功能：催化、运输、营养和贮存、运动、结构物质、防御蛋白质依靠什么行使这一系列功能？
蛋白质分子与其他物质分子之间的相互作用

蛋白质分子与蛋白质分子之间的相互作用对于一个生命体来说，其细胞所包含的全部蛋白质都不是孤立存在的，只要蛋白质分子存在，它们之间或者它们与其它物质之间必定有着或长或短、或强或弱、或简单或复杂、或动态或静态的相互作用，这种相互作用对于蛋白质执行各种生理功能是必须的。在蛋白质组学的研究中，尤其关注蛋白质之间的相互作用。

单个蛋白质分子的结构

- ❖ 一级结构 (Primary Structure): 多肽分子中氨基酸的顺序
- ❖ 二级结构 (Secondary Structure): 多肽分子内氢键和其它弱键相互作用产生的规则重复的局部构象
有序构象：-helix, -sheet, turn
无序构象：coil
- ❖ 超二级结构 (Supersecondary Structure): 由若干相邻的二级结构单元 (Secondary Structure Element) 组合在一起，彼此作用，形成空间上能辨认的二级结构组合体，也称基序 (Motif)
- ❖ 三级结构 (Tertiary Structure): 指多肽链的整体三维结构。

蛋白质分子的结构域

对于较大的蛋白质分子来说，多肽链三级结构往往由两个以上相对独立的三维实体缔合而成，每个三维实体称为一个结构域 (Domain)。而对于小分子蛋白质来说可能只包含单个结构域，因而对小分子蛋白质来讲结构域等同于三级结构。由此也可以说结构域是“准”独立的功能单位。事实上，在含有多个结构域的蛋白质大分子中，每个结构域不仅可以联合在一起作为一个整体发挥功能，也能独立执行某个特定的功能。

蛋白质的四级结构(quaternary structure)

- ❖ 概念：蛋白质分子之间通过相互作用聚合在一起形成的空间结构。
- ❖ 它反映了蛋白质分子之间的空间组织和它们之间形成的化学键。
- ❖ 所以，研究蛋白质分子之间的相互作用，很大程度上是研究蛋白质四级结构的组织形式。

蛋白质分类(按外形分)：纤维状蛋白质(fibrous protein)(轴比>10) 球状蛋白质(globular protein)(轴比<10) (分子轴比=分子长轴/分子短轴)

两者的结构有所不同，在讨论它们的相互作用时，有必要分开研究。

纤维状蛋白质(fibrous protein)

- ❖ 广泛存在于动物体基本支架和外保护成分。可分为两类：不溶性纤维状蛋白(硬蛋白)：角蛋白、胶原蛋白、弹性蛋白 可溶性纤维状蛋白：肌球蛋白、纤维蛋白纤维状蛋白的典型代表——**角蛋白**
- 结构特点：二级结构为右手-螺旋构象，超二级结构为三股螺旋向左缠绕，形成原纤维，原纤维以

9+2 的形式排列为微纤维，微纤维包埋在含硫的基质中，数百微纤维组成大纤维，很多大纤维轴向排列构成毛发基质。和 α -角蛋白结构相类似的典型纤维状蛋白还有胶原蛋白和肌球蛋白。其中肌球蛋白由两条 α -螺旋肽（重链）螺旋而成，头部为四条轻链，有酶活性，催化 ATP \rightarrow ADP 的转化，很多肌球蛋白聚集成粗肌丝（thick filament）。

纤维状蛋白质分子间的相互作用力

❖ 氢键：存在于蛋白质分子链的链间，由于纤维状蛋白分子氨基酸序列（一级结构的）规则性，蛋白质分子之间存在致密的结合氢键。交联键：主要是二硫键，存在于蛋白质分子链间，由于氨基酸序列（一级结构的）规则性，蛋白质分子之间每隔开一定螺旋数就出现二硫键的交联。此种作用力很大。范德华力：即分子间力，由蛋白质分子量决定，存在于蛋白质分子链间或片层结构之间（如 α -角蛋白）。纤维状蛋白的聚集方式（四级结构）比较特殊，单个蛋白质分子在到达超二级结构后直接通过分子间的相互作用聚集，这和此类蛋白质一级结构——氨基酸顺序的规则性是密切相关的。例如胶原蛋白的氨基酸序列中存在明显的“三联体”（Gly-x-y），即每第三个位置必定是甘氨酸，这种结构从空间上来说十分适合链内氢键的形成。

❖ 氨基酸顺序的规则性（一级结构）决定 规则的线性二级结构 决定 蛋白质分子的规则的线性结构（超二级结构） 直接相互聚集 四级结构。

❖ 纤维状蛋白质分子间的作用力十分类似于球状蛋白质分子维持自身三级结构的分子内作用力。

❖ 注意：与骨骼肌粗丝紧密缔合在一起的细丝（thin filament）由肌动蛋白组成，肌动蛋白在体内以两种形式存（globular-肌动蛋白，fibrous-肌动蛋白），F-肌动蛋白其实是以球状蛋白单体（G-肌动蛋白）聚合而成的双股绳索结构的形式。所以形态上类似纤维，但其本质属于球状蛋白。同样的情况也存在于细胞中的微管（microtubule）和鞭毛（flagella），它们同属球状蛋白长向的聚集体。**球状蛋白质（globular protein）**

一般情况下，在谈及蛋白质的结构时，都是指球状蛋白质的情况，蛋白质组学的研究中，重心更是主要放在对球状蛋白的认识上。原因是纤维状蛋白在整个蛋白质分子种类中比重很小，而且它们结构单一，有很多共同点，更主要的是它们很多都是“静态”的，变化速度不快，所以研究难度相对较小。在球状蛋白质中，多肽链被折叠成致密的球形构象，比纤维蛋白要复杂地多，而且生物功能多样（绝大多数酶、运输蛋白、抗体、部分激素、结构物质），它们的活性是动态(dynamic)的而不是静态(static)的。所以球状蛋白分子之间的相互作用也更加复杂。**球状蛋白在机体内的状态** 单条多肽链自我折叠成一定构象，不与其它蛋白结合

❖ 相同或不同的蛋白质分子（本身也可以是不同更小蛋白质分子的聚合体，称为原聚体(protomers)）聚集成寡聚或多聚复合物

❖ 多个蛋白分子相互作用形成的聚集体的空间结构就被称为蛋白质的四级结构。形成四级结构的蛋白质中每个球状蛋白成员称为亚基或亚单位。

❖ 亚基构成四级结构的缔合作用力包括很多类型的化学键（共价和非共价）。

蛋白质分子（亚基）的聚合

❖ 根据蛋白质组成亚基的异同，蛋白复合物可分为

同聚多亚基蛋白（homomultimers）：复合物包含同一多肽的多个拷贝。异聚多亚基蛋白（heteromultimers）：复合物包含不同多肽。

❖ 根据蛋白复合物行使功能的方式，蛋白复合物又可分为

亚基独立行使功能，不依赖聚集体 以聚集体（复合物）为功能单位行使功能由以上分类标准，可以将所有蛋白复合物进行分类

- 1、亚基独立行使功能的同聚多亚基蛋白
- 2、复合物整体行使功能的同聚多亚基蛋白
- 3、亚基独立行使功能的异聚多亚基蛋白
- 4、复合物整体行使功能的异聚多亚基蛋白

亚基独立行使功能的同聚多亚基蛋白

胰岛素，在单体蛋白质分子浓度高时，形成二、四、六、乃至多聚体，这对防止单体分子非水解有很大作用。

复合物整体行使功能的同聚多亚基蛋白

- ❖ 肌动蛋白，其分子聚合与肌肉收缩联系密切血纤维蛋白，分子的聚合与凝血
- ❖ 病毒外壳蛋白

亚基独立行使功能的异聚多亚基蛋白

❖ 丙酮酸脱氢酶复合物，一种多酶复合物，包含丙酮酸脱氢酶、二氢硫辛酸乙酰转移酶、和二氢硫辛酸脱氢酶色氨酸合成酶，一种双功能寡聚酶（2 分子蛋白 A+一分子蛋白 B）固氮酶，包含 2 分子铁蛋白和 1 分子铁钼蛋白，各有功能，相互偶联

复合物整体行使功能的异聚多亚基蛋白

- ❖ 组蛋白，结合顺序 H2A-H2B-H4-H3-H3-H4-H2B-H2A
- ❖ 血红蛋白，珠蛋白 (globin) 家族，2 分子 亚基和 2 分子 亚基的四聚体
- ❖ E. coli 的 DNA 聚合酶，三个亚基的聚集体，分管三种功能

亚基的聚集方式同种蛋白质亚单位或分子之间的的聚合有各种方式，产生不同形状的聚合体，如环状、螺旋状、线状、球状等。

- ❖ 对于异种蛋白质亚单位来说结合方式比较复杂，很多目前还未知，比如丙酮酸脱氢酶复合物。
- ❖ 环状聚合体，有二聚体、三聚体、四聚体、五聚体等，聚合体呈环状，有对称性。
- ❖ 螺旋聚合体，聚合体呈螺旋状，有各种形式。大肠杆菌菌毛是单股非整数螺旋(B)，每圈含 3.125 个亚单位；TMV 病毒壳体蛋白也是单股非整数螺旋(E)，每圈含 16.3 个亚单位；细菌鞭毛(D)是五股整数螺旋，每圈含 11 和亚单位，F 型肌动蛋白是双股螺旋(C)，每圈含 13 个单体。
- ❖ 线状聚合体，棒状的血纤维蛋白分子按照末端对末端的方式直接聚合，生成细长聚合体。
- ❖ 球状聚合体，聚合体是空心的球体。如小球病毒的蛋白壳体，电镜下观察，发现它由许多相同的结构单位(壳微体)聚合而成。

蛋白质分子之间的相互作用归根到底来源于分子之间的特异性识别，这种识别可以发生在不同的蛋白质分子之间，也可以发生在相同分子的不同部位之间。

❖ 两个蛋白质分子能互相识别(即有相互作用)，要具备两个条件

1. 在两种蛋白质分子的结合部位之间，其结合部位的微区构象要能够相互嵌补。
2. 两个部位含有化学基团，能使两种蛋白质分子结合起来。分子或亚基之间的聚合不是任意的，结合部位必须互相嵌补，有时，为了满足互相嵌补，还要引起亚基某种程度的构象变化，虽然亚基能量暂时有所升高，但结合的结果总是比之前的总能量更低，从而使得结合产物更加稳定，是符合物理化学的要求的。当然，有的时候，由于亚基的形变，造成催化功能方面的一些改变，如催化效率降低等，甚至会改变酶的催专一性。例如乳糖合成酶由蛋白 A 和蛋白 B(一种 β -乳清蛋白，本身无催化活性)结合而成，蛋白 A 在独立状态下不能催化乳糖的生成，但与蛋白 B 结合后蛋白 A 改变了底物的专一性催化。对于这种结合，很多作用力起了作用：

1. 非共价键力，包括疏水键(如胰岛素)、范德华力、离子键力(互相结合的两个部位表面电荷差异)、氢键(存在于不同蛋白分子接触部分链间)
2. 共价键力，包括和二硫键(如组蛋白中的两个 H3 之间)

❖ 以上描述了蛋白质相互作用的一大类，即蛋白亚基之间的聚合和解离，同时蛋白质的相互作用的形式还包括：蛋白分子识别(抗体抗原补体的识别)

自装配

多酶复合物(多个酶复合完成一个代谢途径)

自装配的实质

其实，任何生物体内有活性的大分子都有相互识别的特性，例如有名的 TMV 的重构实验就说明了这一点。这种自装配的实质就是生物大分子之间的相互作用。

影响蛋白质分子聚合和解离的因素

- ❖ 蛋白质分子单体的浓度（化学平衡的影响）周围液体环境的 PH 值（影响蛋白质表面电荷分布、影响蛋白质分子单体的构象）周围液体的离子强度（影响蛋白质分子周围双电层，进而影响蛋白结合的活性）**蛋白质结构与功能**
- ❖ 蛋白质功能与其结构密切相关，蛋白质通过其结构发挥作用，了解结构是研究其功能的先决条件
- ❖ 蛋白质分子结构的特征是：结构的层次性
- ❖ 一级结构：氨基酸序列
- ❖ 二级结构：主链之间借助氢键形成的构象
- ❖ 三级结构：二级结构基础上氨基酸残基侧链之间相互作用形成的结构
- ❖ 四级结构：亚基之间空间关系

介导蛋白质相互作用的结构域

- ❖ BRCT结构域
- ❖ LIM结构域
- ❖ POU结构域
- ❖ Bromo结构域
- ❖ 亮氨酸拉链结构域(Leucine zipper)
- ❖ POZ结构域
- ❖ 环指结构域（锌指结构域）

BRCT结构域

- ❖ BRCT结构域最先发现于乳腺癌抑癌蛋白BRCA1的羧基端，进化上并不是保守
- ❖ 由大约95个氨基酸组成
- ❖ 由三个螺旋围绕四条平行的折叠构成
- ❖ 分布在与细胞周期相关的蛋白质、与DNA损伤识别及修复相关的蛋白质中
- ❖ 含有BRCT结构域的BRCA1蛋白质可以通过自身的BRCT结构域和Rb结合蛋白(RbAp46, RbAp48)产生作用；同时组蛋白去乙酰化酶(HDAC1)也可与BRCA1的BRCT结构域结合
- ❖ 与DNA损伤修复以及凋亡相关的p53蛋白也是BRCT结构域的作用对象，发生作用后可激活相应基因
- ❖ DNA修复蛋白XRCC1含有两个BRCT结构域，它利用BRCT结构域形成同源二聚体，同时XRCC1利用BRCT结构域分别独立地与DNA连接酶III及多聚(ADP-核糖)多聚酶结合，参与DNA单链断裂的修复

LIM结构域

- ❖ 由60个富含胱氨酸的氨基酸残基组成，进化上相对保守，保守区域为7个胱氨酸残基和1个组氨酸残基；LIM结构域可结合两个锌离子，且可随机定位
- ❖ LIM结构域不能结合DNA分子
- ❖ LIM结构域含一个疏水核心以维持其立体折叠
- ❖ LIM结构域每一个锌离子结合区的氢键对维持锌指的几何构象是必需的

LIM结构域蛋白分类

- ❖ 含LIM结构域的蛋白质分为三种：
 1. 有成对的LIM结构域，分为A型和B型，靠近N末端常常含有一个同源结构域
 2. 含有一或两个单一序列型的LIM结构域，称为C型
- ❖ 一些不同来源的LIM结构域的集合，大多靠近C末端，和一些附加结构域相连，具有和细胞骨架成分结合的能力
- ❖ 来自第二和第三组的LIM结构域可以单独和不同配体作用
- ❖ 在立体构型层面看，LIM结构域可能有不止一个的结合界面

LIM结构域和转录因子家族成员

- ❖ 含LIM结构域的转录因子家族成员大多在N端含两个LIM结构域，在C端含一个能与DNA结合的同源结构域

- ❖ 两个LIM结构域转录因子家族成员要建立联系，必须通过各自的LIM结构域分别和NLI单体的LIM结构域相互作用，形成四聚体结构，从而促进远距离的增强子和启动子的相互作用
- ❖ NLI (Nuclear LIM domain interactor, 胞核LIM结构域作用子，是一种核蛋白)

LIM结构域结合蛋白

- ❖ LIM结构域结合蛋白 (LIM domain binding protein) 通过其C端和含有LIM结构域的蛋白质相互作用，形成复合体作为传递激活造血信号的桥梁

POU结构域

- ❖ 由70到75个氨基酸构成，位于许多真核转录因子的同源盒结构域的上游
- ❖ 有两个保守区域，一个为POU特异结构区域，一个是POU同源结构域，两个亚结构域由一个“螺旋 - 转角 - 螺旋”模块作为DNA结合功能区，同时也参与介导蛋白质 - 蛋白质之间的作用
- ❖ 许多转录因子都可以和POU结构域发生作用

POU结构域与八聚体结合转录因子

- ❖ 八聚体结合转录因子 - 1 (octamer-binding transcription factor, Oct-1) 含有POU结构域，在细胞内广泛表达，参与激活含有八聚体模块的启动子，这些启动子与细胞周期调控基因H2B以及核小RNA基因的表达有关
- ❖ Oct - 1或Oct - 2蛋白可以通过其POU结构域与维甲类受体 (RXR) 的DBD (DNA binding domain, DNA结合结构域) 结合，影响甲状腺素受体元件 (RXR/TR) 异源二聚体和甲状腺素受体元件 (TRE) 的作用，从而调节包含TRE的启动子的转录活性
- ❖ 同样，含POU结构域的转录因子Pit-1也可以利用其POU结构域和TR作用从而激活生长激素基因

Bromo结构域

- ❖ 大约由70个氨基酸组成
- ❖ 立体结构由四个螺旋束(Z, A, B, C)，一个左手扭曲，一个介于 Z和 A之间的长环 (ZA环) 构成
- ❖ 四个螺旋束紧密捆绑，双双成反平行，包裹在 Z和 A之间ZA环构成一个疏水口袋，成为蛋白质间相互作用的场所
- ❖ 几乎所有与组蛋白去乙酰酶(HAT)相关的共激活因子都含有该结构域
- ❖ 核受体共激活因子复合物的一个组分p300/CBP，含有Bromo结构域，能识别乙酰化的Lys，从而介导组蛋白乙酰化酶(HAT)到染色质的特异部位，调控DNA与组蛋白的作用，继而调节基因表达

亮氨酸拉链结构域

- ❖ 亮氨酸拉链 (Leucine zipper) 通常为一系列周期重复亮氨酸残基组成，每个周期含4 - 7个亮氨酸，大约跨越8个螺旋转弯
- ❖ 有时典型的亮氨酸残基会被异亮氨酸所取代
- ❖ 亮氨酸拉链由平行的卷曲型 α 螺旋构成，同时通过亮氨酸侧链的疏水作用相互缠绕，这样有利于形成同源二聚体
- ❖ AP-1，一个典型的具有亮氨酸拉链结构域的蛋白质家族
- ❖ Jun&Fos是AP-1家族成员，它们可以形成同源或异源二聚体
- ❖ Jun可以利用C端的亮氨酸拉链核Rb蛋白发生相互作用，从而影响细胞的生长和分化
- ❖ Myb蛋白是一种转录因子，它的N端具有DNA结合域，中央是一个反式的激活结构域，C端是一个亮氨酸拉链结构域
- ❖ 现在已知Myb的亮氨酸拉链可以和很多蛋白质如P26，P28，P67，P160作用

POZ结构域

- ❖ 由120个氨基酸组成，进化上保守
- ❖ 含有C₂H₂锌指结构的蛋白质的N端，大约5 - 10%含有POZ结构域
- ❖ 含有POZ结构域的蛋白质多数为转录因子，它们通过POZ结构域形成高度紧密扭曲结合的同源二聚体，每个单体由1/4的表面用于形成二聚体的分子间作用

- ❖ 整个二聚体由一个较大的疏水表面，二聚体表面由保守氨基酸残基构成的沟用于和其它蛋白质分子作用

POZ结构域蛋白质

- ❖ POZ结构域可以和组蛋白去乙酰化酶复合体作用，从而发挥转录抑制活性
- ❖ PLZF蛋白中，蛋白质可以通过POZ结构域形成二聚体，中心构架为一组 α 螺旋，头尾由 α 折叠封住
- ❖ 在APL（急性早幼粒细胞白血病）中由t(11;17)染色体易位引起的PLZF-RAR α 蛋白，不但形成同源二聚体，还可以通过POZ结构域与PL蛋白形成异源二聚体，影响PL蛋白正常的亚细胞定位

环指结构域

- ❖ 也称为C₃H₂C₄锌指结构域
- ❖ 富含半胱氨酸，由40 - 60个氨基酸残基组成
- ❖ 结合两个锌离子，锌离子被4(3)个半胱氨酸及1个组氨酸分子螯合
- ❖ 三维结构有 - - - 或 - -环 - 构象
- ❖ 相对保守区域包括保持分子基本构象的 α 链，锌离子连接系统，保守残基的包裹形式（疏水核心）

环指结构域蛋白质

- ❖ c-Cbl 蛋白，信号传导途径中一种接头蛋白，其环指结构可以和人类泛素结合酶7(UbcH7)结合，对细胞中配体诱导的表皮生长因子受体的泛素化起作用；同时，c-Cbl 蛋白还可以通过环指结构域与泛素结合酶(E2)结合，使磷酸化的底物泛素化，进而从蛋白体途径降解
- ❖ MDM2蛋白，C端含有环指结构域，可以通过环指结构域与MDMX蛋白形成异源二聚体，从而阻止MDM2经过蛋白体途径被降解，同时，MDM2蛋白也可以通过这个C端环指结构域与p53蛋白结合，从而通过蛋白体途径降解p53

可能参与蛋白质相互作用的结构域

- ❖ Set结构域：与一种类似双重特异磷酸酶蛋白家族作用，代表蛋白质：Rb结合锌指蛋白
- ❖ Sam结构域：进化上保守，参与调节真核细胞的发育过程，可凭借Sam结构域形成同或异二聚体，代表蛋白质：丝氨酸/酪氨酸蛋白激酶
- ❖ Tpr结构域：34个氨基酸残基组成，分布广泛，成脚手架样结构，且常形成较大的复合体，代表蛋白质：分裂后期启动复合体亚单位(cdc16, cdc23, cdc27)

可能参与蛋白质相互作用的结构域

- ❖ WD40结构域：真核细胞G蛋白 α 亚基中有8个随机重复序列，每个由40个氨基酸残基组成，每一个都含有中央的Trp-Asp模块；代表蛋白质：凋亡蛋白酶活化因子1(Apaf1)
- ❖ SRCR结构域：见于细胞表面分子和一些分泌蛋白，参与配体介导结合；代表蛋白质：Mac2结合蛋白
- ❖ SH2结构域：由大约100个氨基酸构成，可以与靶肽上磷酸化的酪氨酸残基高特异结合，调节细胞信号传导链，这种结合使序列特异的，且依赖磷酸化；代表蛋白质：STAT蛋白

2) 蛋白质相互作用分析

- 生物细胞的生命活动是由相互作用的蛋白质控制的，这些蛋白质参与代谢和信号传递通路以及组成复合体诸如合成和使用腺苷三磷酸(ATP)、复制和翻译基因或构成细胞骨架基础结构等等的分子机器
- 分析蛋白质之间的相互作用及其方式，认识与特定生理活动相关的蛋白质网络，绘制出蛋白质相互作用的图谱(interaction mapping)，是功能蛋白质组学的重要内容。

蛋白质相互作用研究传统方法

- ❖ 酶联免疫吸附法
- ❖ 蛋白质亲和层析法
- ❖ 亲和印迹
- ❖ 免疫沉淀
- ❖ 蛋白交联
- ❖ 蛋白探测

蛋白质功能研究

杂交系统

❖ 酵母杂交系统

1. 酵母双杂交系统
2. 酵母单杂交系统
3. 酵母三杂交系统
4. 反向杂交系统
5. 核外双杂交系统

❖ 哺乳动物双杂交系统

❖ 细菌双杂交系统

酵母双杂交系统(Yeast two-hybrid system)

- ❖ 1989年，Song和Field首先建立基于酵母的细胞内检测蛋白质间相互作用的遗传系统（Nature 1989, 340: 245-246）
- ❖ 理论依据---真核转录因子的结构特征
- ❖ 实验原理
- ❖ 应用
- ❖ 存在的问题和改进

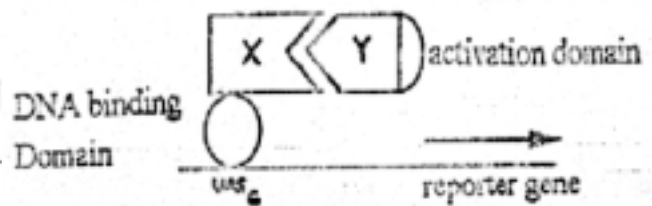
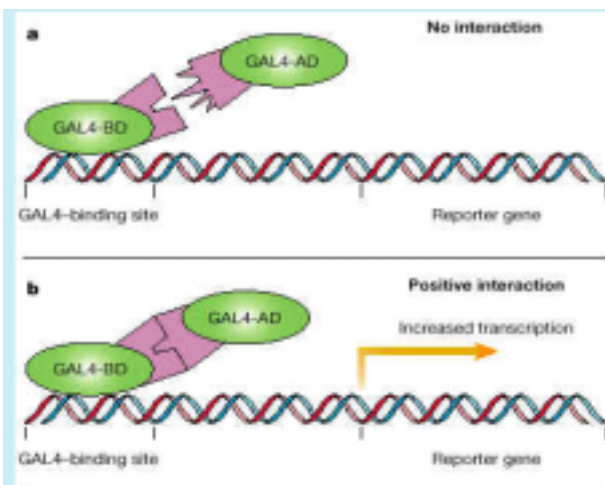
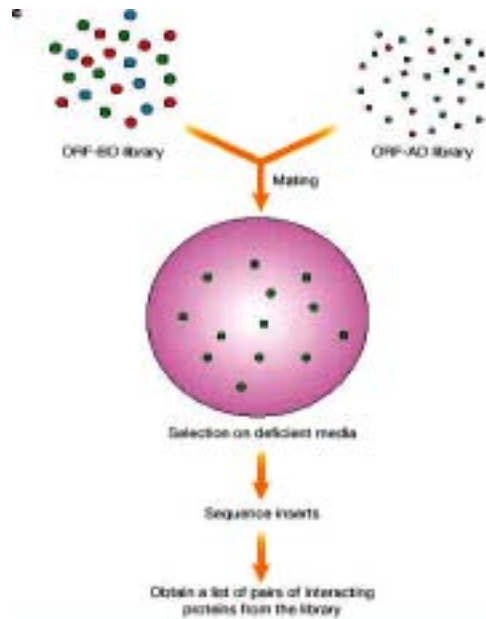


图1 酵母双杂交系统

酵母双杂交系统三个组成部分

- 1) 与BD融合蛋白表达载体，被表达的蛋白成诱饵蛋白 (bait)；
 - 2) 与AD融合蛋白表达载体，被其表达的蛋白称为猎物或靶蛋白 (prey)；
 - 3) 带有一个或多个报告基因的宿主菌株
- 常用的报告基因有HIS3, URA3, LacZ和ADE2等；有GAL4系统和LexA系统



酵母双杂交系统应用

- 1) 检测功能已知蛋白间的相互作用
- 2) 研究蛋白间发生相互作用所需的结构域或重要活性部位。通常须对待测蛋白作定点突变或缺失突变处理
- 3) 寻找与靶蛋白相互作用的新蛋白
- 4) 设计寻找能与细菌或病毒致病性蛋白相互作用而具有一定治疗价值的蛋白多肽
- 5) 寻找调控蛋白相互作用的化合物
- 6) 蛋白质相互作用图谱

问题与改进

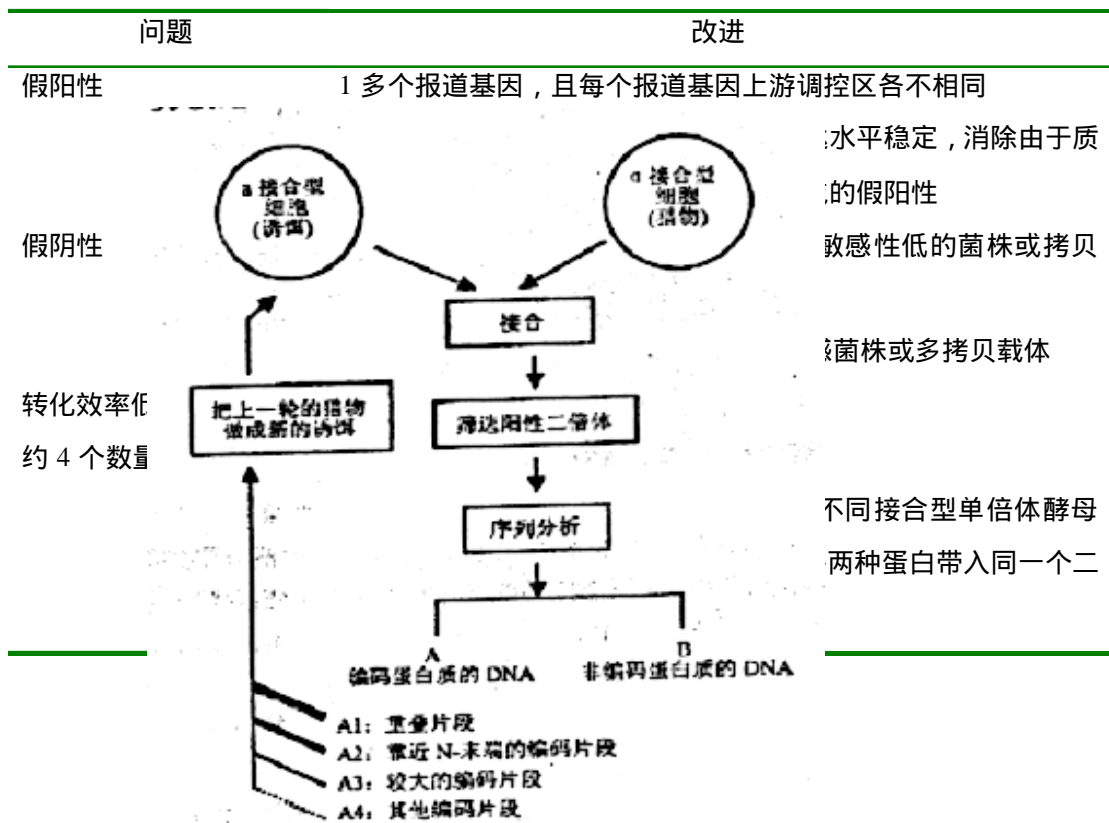


图 1 Fromont-Racine 等人的多轮双杂交筛选策略

酵母单杂交系统(Yeast One-hybrid system)

❖ 原理：DNA结合蛋白（即转录因子）与DNA顺式作用元件结合，调控报道基因的表达

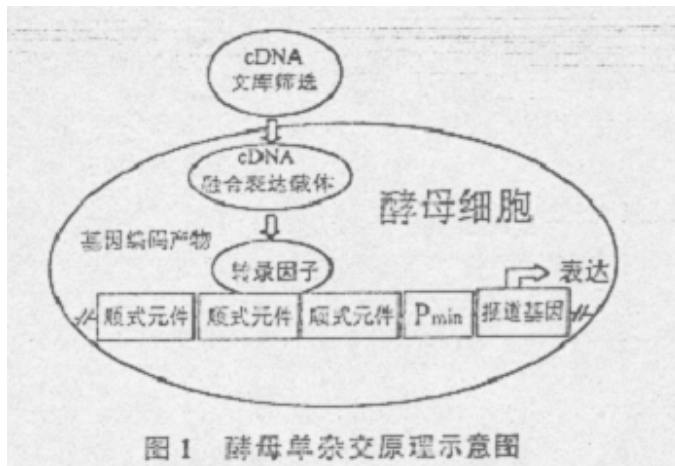


图1 酵母单杂交原理示意图

应用

- 1) 分离鉴定与特异DNA序列结合作用的蛋白质并获得其编码基因
 - 2) 在细胞内分析鉴定转录因子与顺式元件结合的有效性
- ❖ 确定已知DNA-蛋白质之间是否存在相互作用
 - ❖ 分离编码结合于目的顺式结合元件或其它短DNA结合位点上蛋白的新基因
 - ❖ 定位已经证实的有相互作用的DNA结合蛋白的DNA结合区结构域以及准确定位与之结合的核苷酸序列

酵母三杂交系统(Yeast Three-hybrid system)

- ❖ SenGupta在酵母双杂交思想上提出的
 - ❖ 主要特征：蛋白X与Y的相互作用通过第三个分子蛋白Z的参与实现
- 目的：研究两个蛋白质与第三个成分间的相互作用，按第三个成分的不同，可以是蛋白质、RNA 或小分子药物

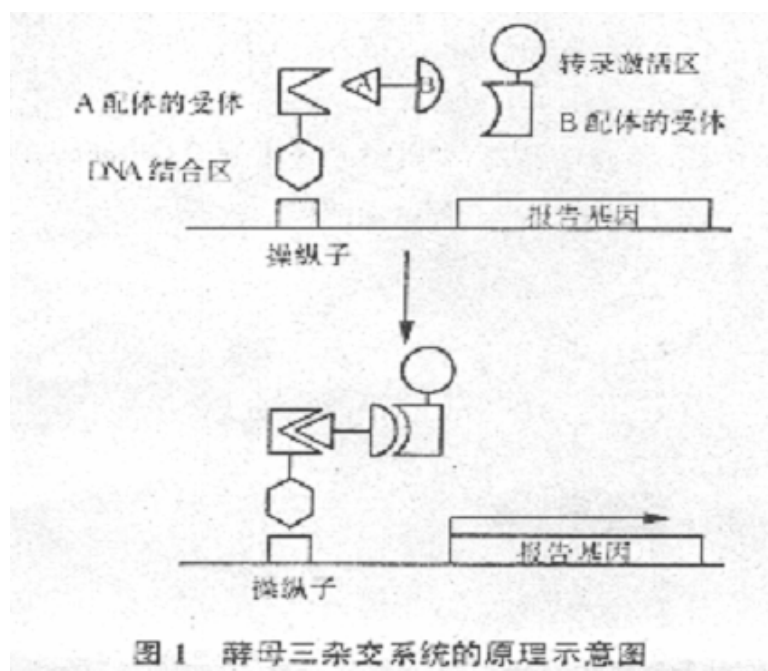


图1 酵母三杂交系统的原理示意图

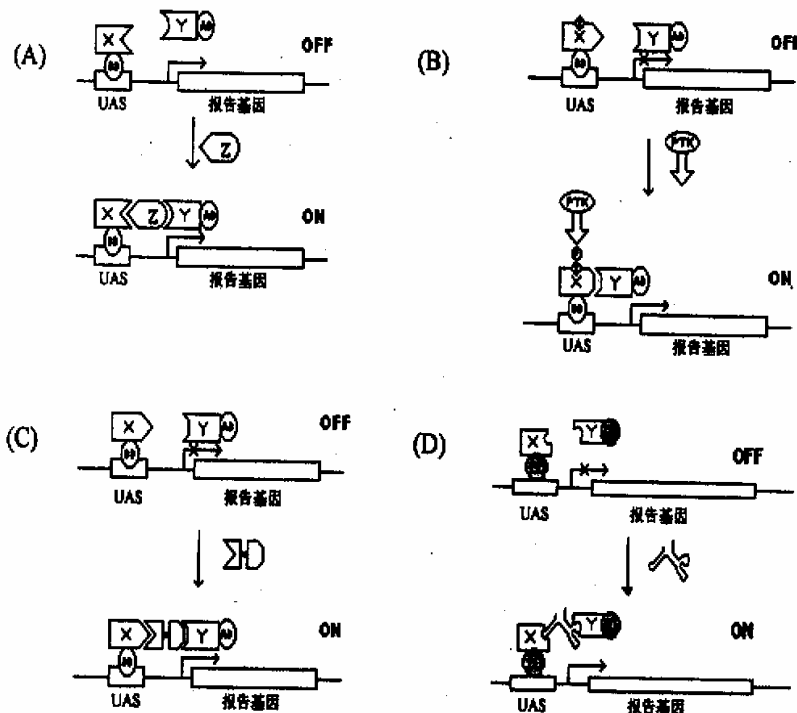
三杂交系统需要满足的条件

- ❖ 第1、2个融合蛋白必须通过与成分3的结合，间接起到激活报道基因转录的作用，两者间没有直接作用，不能直接激活报告基因的转录；
- ❖ 成分3具有分别与第1、2个融合蛋白结合的区域。

酵母三杂交系统应用

三个蛋白质间相互作用；

RNA-蛋白质间相互作用--- RNA三杂交系统 (RNA Three-Hybrid System) ----一个RNA分子和两个都能结合该分子的蛋白。检测由RNA介导的两个蛋白质之间的相互作用， 筛选鉴定新的蛋白结合RNA；
小分子药物与蛋白质间相互作用---小配体三杂交系统 (Small Ligand Three-Hybrid System)



酵母三杂交系统的前景

- ❖ 为以前局限于方法得不到解决的问题带来希望。
- ❖ 例如，细胞因子融合蛋白除提高原有的生物活性外，同时具有新增加的生物活性，但其机理不明。有的认为是细胞因子融合蛋白分别与各自的受体结合；有的认为是细胞因子融合蛋白与各自受体外的新型受体结合激活该受体引起级联反应所致。酵母三杂交系统的出现为探索这一机理提供了可能。

酵母三杂交系统不足

1. 局限于三种成分间的相互作用，其中两种为蛋白质，第3种成分可以是RNA或小分子药物，对于超过三种成分更为复杂的相互作用无能为力；
2. 酵母双杂交系统可以检测解离常数 (Kd) 小于5 nmol /L的两种为蛋白质间的相互作用，而对于酵母三杂交系统的检测灵敏度没有明确的研究结果；
3. 在对蛋白质-药物小分子的相互作用研究中，药物小分子要能透入酵母细胞膜，对于不能透入酵母细胞膜的药小分子，该方法不适用；
4. 与酵母双杂交系统一样，只适用于可溶蛋白，不适用于膜结合型蛋白的研究。

反向双杂交系统 (Reverse two hybrid system)

❖ 特点：反选择筛选策略

❖ 原理：URA3基因的启动子内引入Gal 4的结合位点。改造的酵母菌株在缺乏尿嘧啶的选择性培养基上只有当“诱饵”和“猎物”相互作用激活URA3基因的表达才能生长。在含有5-F0A的完全培养基上“诱饵”和“猎物”的相互作用则抑制细胞的生长。然而如果目的蛋白，即与DB 或AD融合的蛋白质发生了突变或者由于外加药物的干扰不再相互作用，URA3基因不表达，则细胞能在含有5-F0A的完全培养基上生长。

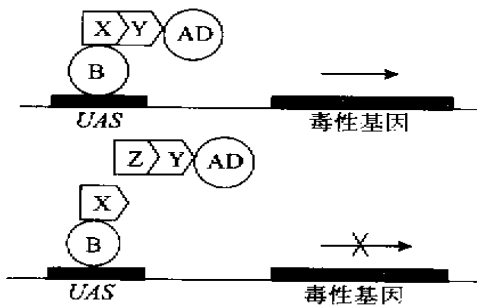


图1 逆向双杂交系统基本原理示意图 1

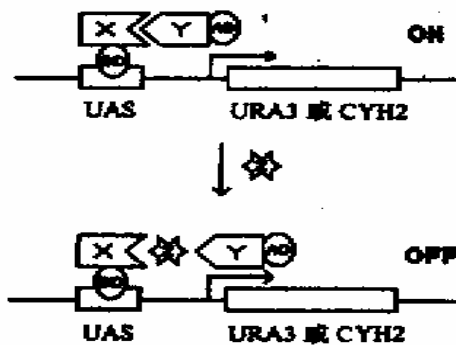
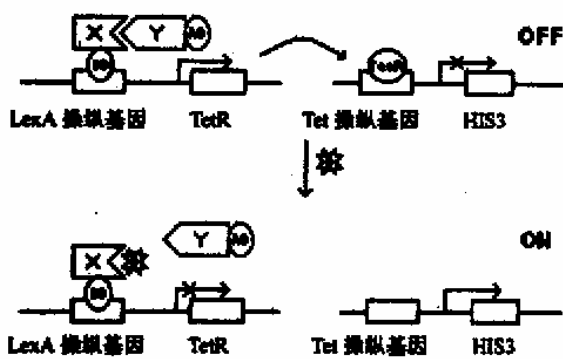


图2 逆向双杂交系统原理示意图 2

分裂双杂交系统(Split hybrid system)



核外双杂交系统

- ❖ 优点：可以检测某些在核外蛋白之间的相互作用
- ❖ 两种：SRS和USPS
- ❖ 问题：假阳性（ras族，类cdc25蛋白）

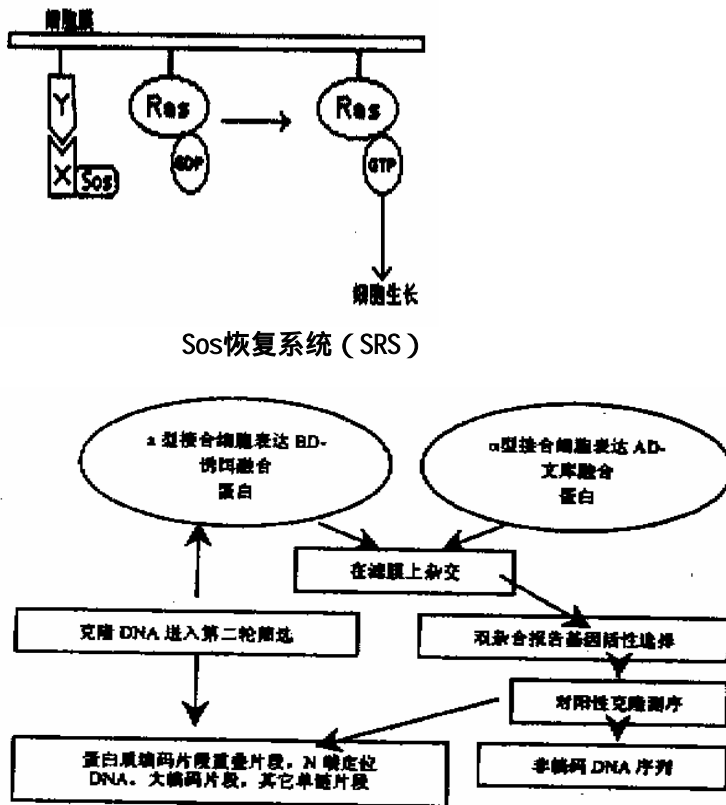


图 7 Fromont-Racine 等人采用迭代的方式从基因组文库中不断筛选出相互关联的一系列诱饵蛋白质。

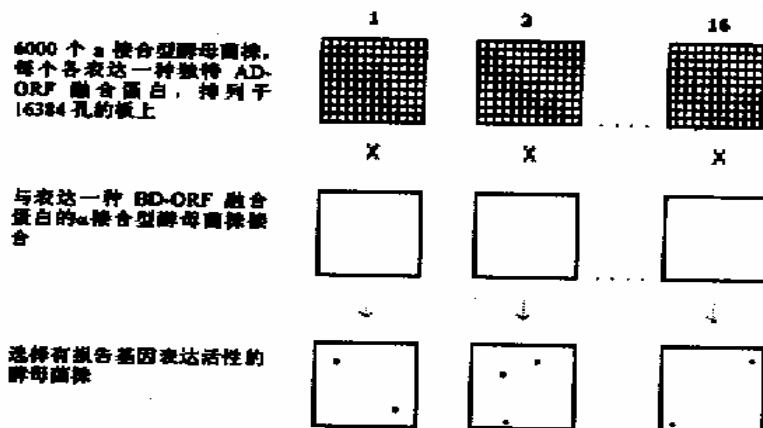
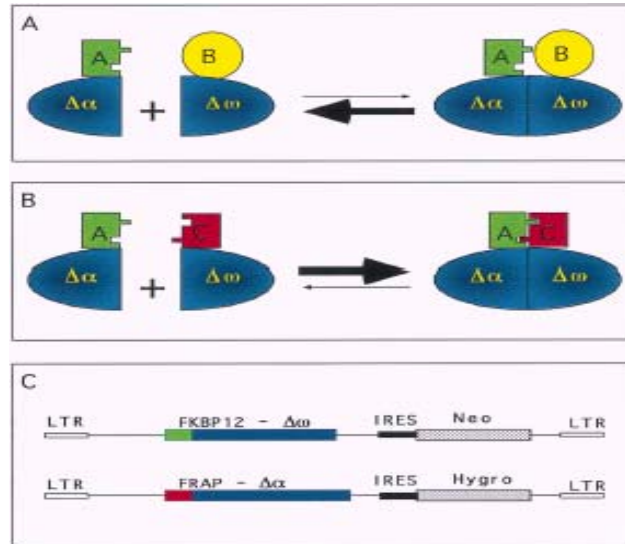


图 8 Fields 等人采用的方法。用表达 6000 种 ORF BD 融合蛋白的每一种酵母菌株重复整个操作, 使得 36,000,000 种组合全部得以检验。

哺乳动物双杂交系统

❖ Rossi F 和 Charlton C A 于 1997 年提出的

- ❖ Proc Natl Acad U S A 1997, 94: 8405-8410
- ❖ 可以分析翻译后加工的蛋白质的相互作用
- ❖ 不能大量筛选转化体



细菌双杂交系统

原理：百日咳博代杆菌 *cya* 基因编码的钙调蛋白 (CaM) 依赖性腺苷酸环化酶 (cAMPase)，其激活区分为 T18 和 T25 两个相对独立的功能片段。此两片段引入 *E. coli* DHP1 (*cya*) 作为两独立实体表达，它们不能彼此识别与作用，不能互补 *E. coli* DHP1 的 *cya* 缺陷型性能，表现的 - 半乳糖苷酶活性低，在 MacConkey 麦芽糖平板上为白色菌落。若这两个功能区分别与一对有相互作用的蛋白质融合表达，形成有功能的 cAMPase，互补 *E. coli* DHP1 的缺陷型，菌株 - 半乳糖苷酶活性上升，平板上为红色菌落。

细菌双杂交系统优点

1. *E. coli* 生长速度快，大大缩短了试验周期
2. 通过是否有活性 cAMPase 判断两蛋白质是否有相互作用，可将它们同 T18 或 T25 片段形成融合蛋白，检测混合液的 cAMPase 活性，体外检测蛋白质的相互作用，并直接研究环境分子或其他分子等对蛋白质相互作用可能产生的影响。
3. 如果被研究蛋白质之一自身有非 CaM 依赖性 cAMPase 活性，用原细菌双杂交系统会有干扰（假阳性），需改进方法以避免。
4. 该系统与原核生物蛋白质相似，用于研究原核生物蛋白质相互作用时比酵母杂交系统优越。

细胞内的蛋白作用网络

- ❖ 细胞内复杂的蛋白作用网络对我们一直是难以研究的，现在应用高通量双杂交筛选技术，这个秘密已经逐步的显现在我们面前了
- ❖ Drees and colleagues 在 Journal of Cell Biology 上报导了用杂交方法筛选与细胞极性有关蛋白的研究结果
- ❖ Drees, B. L. et al. A protein interaction map for cell polarity development. J. Cell Biol. 154, 549-571 (2001)
- ❖ 用一个已知与细胞极性有关的蛋白 68 作为鱼饵，筛选到的蛋白既有 Rho-GTPase 又有分泌性的蛋白。共发现 128 个新的蛋白质相互作用，其中 44 个包含未知功能的蛋白。
- ❖ 最大的发现是在细胞信号通路之间新的联系，例如 Rho1 和 cdc42 细胞途径，这些信号途径将肌动蛋白和形态建成检查点结合了起来。为了确证它们之间的相互作用，研究者将这些因子用黄色荧光蛋白融合在细胞内表达，观察它们的亚细胞定位
- ❖ 一步就是用遗传学和生化的方法将这些因子在细胞内的作用进行研究，看看细胞内的复杂网络究竟是什么样子的了

噬菌体展示技术

- ❖ 原理
- ❖ 分类：
 1. 丝状噬菌体系统
 2. T₄噬菌体系统
 3. 噬菌体系统

主要应用

- ❖ 抗体基因文库的构建 - - 筛选基因工程抗体
- ❖ 随机多肽基因文库的构建 - - 筛选随机多肽药物

特点

- ❖ 优点：直接得到基因；高选择性筛选复杂混合物；直接评价蛋白质结合特异性
- ❖ 缺点：插入片段大小有限；细菌为宿主有些蛋白不能正确折叠或修饰；编码的为融合蛋白，可能改变天然蛋白的结构与功能；体外检测

酵母表面展示技术

- ❖ 原理：

酵母是单细胞真核生物，根据交配型的差异分为两种单倍体：MAT α 和MATa，在其细胞表面分别表达。 α -凝集素和a-凝集素，可以介导细胞之间的交配融合。

酵母表面展示系统应用的是酿酒酵母的 α -凝集素，外源蛋白借助它在细胞表面得到展示。
- ❖ α -凝集素由核心亚单位(Aga1)和结合亚单位(Aga2)两部分组成，Aga1共有725个氨基酸，它与酵母细胞壁的 β -葡聚糖共价连接。Aga2共有69个氨基酸，它通过两个二硫键与Aga1结合，表达于细胞表面
- ❖ Aga2的C端部分参与二硫键的形成，外源蛋白通过与Aga2的C端融合可展示于酵母细胞表面

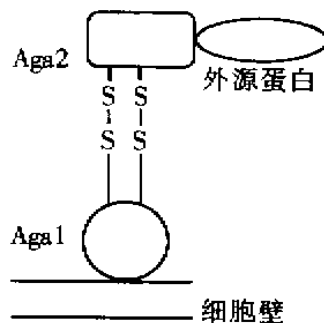


图1 酵母表面展示系统示意图

杆状病毒表面展示技术

- ❖ Autographa californica核型多角体病毒(AcNPV)编码的表面糖蛋白gp64介导病毒与昆虫细胞的融合及侵染过程gp64是I型跨膜糖蛋白，它的N端带有一个内质网加工的信号肽序列，近C端是跨膜结构域。
- ❖ gp64通过二硫键相连形成寡聚体，位于AcNPV颗粒表面，介导病毒的细胞侵染。
- ❖ 外源蛋白在信号肽的C端和gp64 N端之间融合，加工后信号肽被切除形成的N端融合蛋白能稳定的表达于病毒表面
- ❖ 外源蛋白与杆状病毒膜结合蛋白如流感病毒的红细胞凝集素相融合也能在病毒表面展示

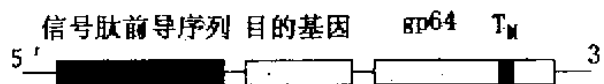


图2 杆状病毒 AcNPV gp64 融合载体示意

- ❖ 使用硅的光刻印刷或者金单层，结合使用微球传感器的微孔或喷墨到聚苯乙烯膜上。
- ❖ 这样的模式牵涉到通过定型单个蛋白质（比如BSA、亲和素和单克隆抗体）制作微型免疫分析模式。

2) 不同标志物

a. 抗体芯片

- ❖ 免疫传感芯片已经被研制，可以同步检测临床分析物。
- ❖ 捕获抗体和分析物被排列到有十字交叉样流室的显微镜片上。检测是通过荧光标记和CCD为基础的光学读数装置。
- ❖ 是以低密度模式（6×6模式）完成的，但是此步骤有高通量潜能，因为涉及到自动化成像分析和微流体力学，它已经成为酶活力和其它分析的一个未来模式

b. 寡聚核苷酸芯片

- ❖ 一个凝胶光学或过硫酸盐诱导的共聚合技术已经被用于制作聚丙烯酰胺凝胶垫上的寡聚核苷酸、DNA和蛋白质微芯片三维阵列（从10×1 μm到100×100 μm），其为一个憎水玻璃界面所分隔。
- ❖ 三维聚丙烯酰胺凝胶提供了比二维玻璃支持物超过100倍固定的能力，因此显著地增强了测量的敏感性。

c. 组织芯片

- ❖ 组织微阵列已经被研制用于制作肿瘤标本的高通量分子分布图
- ❖ 它使用自动穿孔柱（0.6mm宽，3 - 4mm高），采取1000个包埋在石蜡中的个体肿瘤活检组织，将它们排列到45×20mm石蜡块中
- ❖ 连续切片后，肿瘤标本进行免疫组化、荧光原位杂交（FISH）和RNA/RNA原位杂交的平行分析。
- ❖ 能在2小时内显微扫描含有645个标本的免疫组化阵列玻片。它对于许多处于不同疾病阶段的不同患者的同步分析，快速建立新的候选标记基因中有应用

3. 蛋白芯片的制备

4. 蛋白芯片的检测

5. 蛋白芯片的应用

- ❖ 蛋白质芯片能够同时分析上千种蛋白质的变化情况，使得在全基因组水平研究蛋白质的功能(如酶活性、抗体的特异性、配体-受体交互作用以及蛋白质与蛋白质或核酸或小分子的结合)

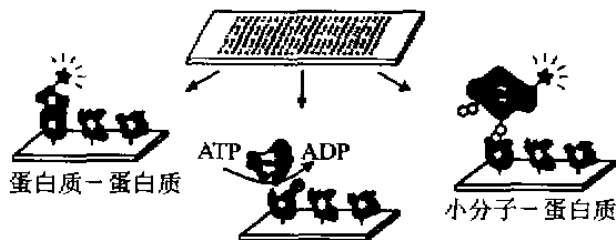


图2 蛋白质芯片的应用

应用1 用于疾病诊断和疗效判定，即生物学标志的检测

- ❖ 蛋白质芯片能够同时检测生物样品中与某种疾病或环境因素损伤可能相关的全部蛋白质的含量变化情况，即表型指纹(phenomic fingerprint)。对于疾病的诊断或筛查来讲，表型指纹要比单一标志物准确可靠得多

加快生物标志物的识别速度

- ❖ Ci phergen Bi osystem公司利用蛋白质芯片检测了来自于健康人和前列腺癌患者的血清样品，在短短的三天之内发现了6种潜在的前列腺癌的生物标志物。
- ❖ 如果利用过去的方法，也许要花费数月或数年的时间

例：多肿瘤标志物蛋白芯片检测系统

- ❖ 2001年12月17日获得国家药品监督管理局颁发的生物制品一类新药证书及生产批文

- ❖ 目前已在国内部分大医院投入使用。
- ❖ 基于生物芯片技术和免疫学技术，利用抗原抗体特异性结合的免疫学原理，只需抽取人体的一滴静脉血，通过检测血液中是否有12种肿瘤标志物，达到对原发性肝癌、肺癌、前列腺癌、胰腺癌、胃癌、食道癌、卵巢癌、子宫内膜癌、结/直肠癌、乳腺癌等多种癌症的辅助诊断目的，特异性达80%以上。

应用2 研究蛋白质间的相互作用

- ❖ 2001年7月 *Science* 网上发表一篇论文，制造了一种蛋白芯片可以同时分析5800个酵母蛋白质。这种芯片通过揭示与数千个蛋白质相互作用的“伙伴分子”而有助于科学研究以及临床诊断工作。
- ❖ 蛋白固定在玻片上，在该玻片上喷洒上被检蛋白溶液。然后增加一个荧光标签突出了被检蛋白质结合的点。
- ❖ Snyder等用该技术迅速鉴定到与数千个排列的酵母蛋白质相互作用的蛋白质。
- ❖ 新发现了33个结合钙调蛋白的蛋白质，52个结合磷脂酰肌醇肽的蛋白质。钙调蛋白是参与钙信号传到的广泛分布的一种蛋白质，磷脂酰肌醇肽是一种细胞膜蛋白，参与生长，分化和细胞骨架重排

应用3 发现药物或毒物新靶点或作用机制

- ❖ 疾病的发生发展与某些蛋白质的变化有关，如果以这些蛋白质构筑芯片，对众多候选化学物进行筛选，直接筛选出与靶蛋白作用的化学物，将大大推进药物的开发。
- ❖ 蛋白质芯片有助于了解药物或毒物与其效应相关蛋白质的相互作用。
- ❖ 蛋白质芯片技术允许对化学物(药物或毒物)作用机制细节不够清楚的情况下，直接研究蛋白质谱，这可能将化学物的作用和疾病联系起来，并进一步建立和发展外源化学物与蛋白质表达谱的数据库，促进药理学和毒理学的研究。

表面等离子共振技术

- ❖ 定性和定量检测蛋白质之间相互作用的简便而快捷的方法
- ❖ 原理：感兴趣的蛋白固定在葡聚糖聚合物上，构建称生物传感器，将待测蛋白质（肽）溶液流过该固相载体，在聚合物表面固定有相互作用的蛋白质后改变冲击光的共振角（变化值与蛋白浓度正相关）

例 - - 淋巴细胞的激活

- ❖ 由T细胞受体（TCR）与主要组织相容性分子（MHC）和抗原呈递细胞递呈的抗原(anti gen)组成的复合固相互作用启动
- ❖ 研究MHC与anti gen、TCR与MHC / anti gen相互作用的动力学常数
- ❖ 前者高，半衰期长；后者低，半衰期短
- ❖ 说明短暂的和低亲合力的相互作用就可激活T淋巴细胞

优点

- ❖ 蛋白用量少(1-10ug)
- ❖ 传感器可重复使用50次
- ❖ 快速：一般一个实验10mi n
- ❖ 待测蛋白不需修饰
- ❖ 复杂环境中也可测定相互作用
- ❖ 易于获得结合与解离常数

计算机分析

- ❖ 蛋白种系发生模式
- ❖ Rosetta stone序列分析
- ❖ 相关mRNA表达方式

第三章 蛋白质组学研究介绍

第一部分 蛋白质组研究的应用与发展

- 第一节 蛋白质组学在基础研究中的应用
- 第二节 蛋白质组在人类疾病研究中的作用
- 第三节 蛋白质组学在药物开发中的应用

第一节 蛋白质组学在基础研究中的应用

- ❖ 深入探索基因和蛋白质关系的本质
- ❖ 研究某个细胞的蛋白质组
- ❖ 结构生物学
- ❖ 蛋白质的折叠与修饰
- ❖ 生物过程分析等

一. 结构生物学---亚细胞结构

- ❖ 过去在细胞生物学领域还没有得到一个主要亚细胞结构的完整的分子图。
- ❖ 运用蛋白质组学方法，该问题得到初步解决
- ❖ 目前已研究了细胞膜、细胞器（如线粒体、溶酶体、内质网、高尔基体等）、胞质复合物（如核糖体）、细胞核（如核质、剪接体、纺锤体、核孔复合体等）的蛋白质组

1. 线粒体

- ❖ 线粒体是细胞内数量最多的细胞器，为细胞产生能量的部位，在细胞代谢中起重要作用
- ❖ 线粒体功能失调与一些疾病相关，包括恶性肿瘤、糖尿病、衰老等。
- ❖ 认识线粒体蛋白质以及其不同表达将有助于阐明这些疾病的发生过程。
- ❖ 复旦大学生命科学院杨芑原等教授负责的“细胞病理过程相关的功能蛋白质组学研究”课题中包括“线粒体（亚细胞）相关的功能蛋白质组谱和关键蛋白的结构与功能的研究”
- ❖ “与线粒体分裂有关的蛋白质研究进展”已被生命的化学杂志接受
- ❖ MI TOP，线粒体蛋白质组资料库，包括与线粒体有关的基因、蛋白质和疾病资讯
http://www.mips.biochem.mpg.de/proj/medgen/mi_top/
- ❖ SWISS-PROT数据库中含有269种人类线粒体蛋白质，而特殊的线粒体蛋白质组数据库含有225种线粒体蛋白质，它们均于人类线粒体疾病相关。
- ❖ Rabi Iloud等建立人胎盘来源线粒体蛋白质2-DE参照图谱，发现了53种蛋白质
- ❖ 线粒体膜通透状态是细胞死亡的限速条件，此过程部分由通透转运孔复合物(PTPC)的开放介导。
- ❖ 通过蛋白质组方法分析PTPC开放后线粒体释放的蛋白质组成，发现79种已知蛋白多肽，21种有相配的人表达序列标签（EST）。

2. 核孔复合体

- ❖ 核孔复合体是一个巨大的跨核膜的八角形结构，是控制大分子在胞质和核质间双相交流、运输的通道。

酵母核孔复合体
蛋白质组分析
40种相关蛋白

结合免疫电镜定位、定量

- ❖ 29种核孔蛋白11种转运因子及核孔相关蛋白质

二．研究蛋白质折叠

Houry等的工作

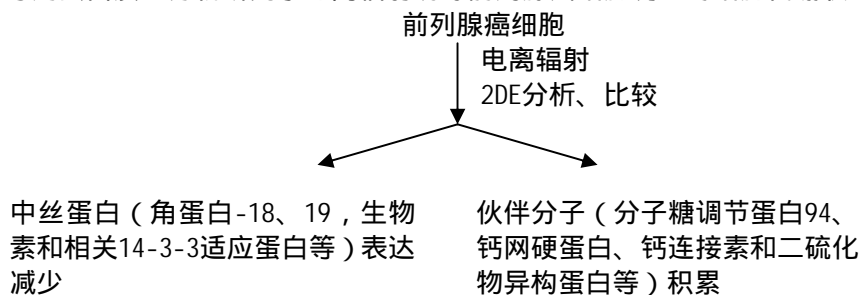
大肠杆菌胞质新生多肽链 (2500种)
 免疫共沉淀
 与GroEL结合的肽链
 2DE+数据库比较
 全面分析新生多肽链与分子伴侣GroEL的关系
 300多条有相互作用
 鉴定其中50种
 揭示蛋白质与GroEL相互作用的关键结构特征

三．转录后修饰研究

- ❖ 抗磷酸酪氨酸和抗磷酸色氨酸抗体的免疫印迹常被用来分析磷酸化蛋白
 小鼠成纤维细胞
 2DE分析
 血小板衍生的生长因子(PDGF)刺激前后细胞蛋白质谱
 Western blot
 显示磷酸化蛋白 (500多个)
 蛋白质组图比较分析
 刺激后磷酸化程度发生变化的蛋白 (至少100个)
 质谱分析
 获得 PDGF信号传导和其他信号途径相关蛋白
- ❖ 利用体外肽耦合(dA)30寡核苷酸方法分离了内皮素B受体和缓激肽B受体，发现它们有广泛的磷酸化、软酯酰化
- ❖ 揭示PDGF受体下游推测信号的新传导途径，并使用到其它信号传导途径

四．凋亡过程分析

- ❖ Prasad 等用蛋白质组分析研究了电离辐射诱导前列腺癌细胞凋亡时细胞代谢状态的蛋白变化模式



第二节 蛋白质组在人类疾病研究中的作用

- ❖ 基因只是遗传信息的载体，蛋白质才是生理功能的执行者。
- ❖ 疾病的发生，往往与蛋白质数量、结构和性质的异常密切相关；而且，相当多的药物都是以蛋白质作为作用靶点，或者药物本身即是蛋白质。
- ❖ 因此，从基因水平向蛋白质水平的深化，也成为医学研究发展的必然趋势。

- ❖ 通过蛋白质组的研究,比较研究体系内不同状态(如不同生长发育时期,给药前后,生理与病理状态)的蛋白质表达图谱,定量分析各蛋白质的变化,从而可以对体系(菌体,细胞,体液等)内复杂代谢,调控等生物功能进行动态监测。
- ❖ 蛋白质组学方法能对机体渐进发展的调控与失控加以掌握,综合环境因素的影响,正成为临床诊断与治疗的新发展方向。
- ❖ 对在分子水平进行疾病的早期诊断、治疗研究,寻找疾病的特异分子标记物,探讨某些重大疾病产生的生理、病理过程和机制是具有实用价值的全新方法。

一. 临床诊断

- ❖ 早在1982年临床化学杂志就有了有关双向电泳应用于临床研究上的文章。
- ❖ 现双向电泳数据库包括许多疾病相关蛋白,如皮肤病,膀胱癌,心脏病等
- ❖ 可以利用数据库发现新的或特异性蛋白质作为诊断标志。
- ❖ 前列腺癌患者与良性前列腺瘤(血浆过多症)个体内前列腺特异性抗原PSA(prostate specific antigen)有差异表现
- ❖ 哮喘个体的鼻腔灌洗液和支气管肺泡灌洗液中Lipocalin-1的表达与在健康个体中的表达状况不同。
- ❖ Bohring等联合双向电泳与Western Blotting技术鉴定分离了高纯度精子膜蛋白
- ❖ 发现14种膜抗原。
- ❖ 能与男性不育病或输精管切除患者体内精子自抗体SPAb(sperm autoantibodies)结合,影响精子运动及顶体反应,
- ❖ 为深入了解精子不育机理、确定SPAb的存在找到一个可行的方法

二. 疾病的蛋白质组研究

- ❖ 运用蛋白质组的研究手段,通过比较正常状态下和病理状态下的细胞或组织中,蛋白质在表达数量、表达位置和修饰状态上的差异,可以发现与病理改变有关的蛋白质和疾病特异性蛋白质。
- ❖ 这些蛋白质,既可以为认识疾病发病机理提供线索,也可以作为疾病诊断的分子标记,还可以作为治疗和药物开发的靶点。

1. 肿瘤

- ❖ 肿瘤是蛋白质组的研究应用最为广泛的领域。通过比较对肿瘤细胞和正常细胞的蛋白质表达情况,可以鉴定出肿瘤特异性标记(marker)或特异性抗原,为诊断与治疗提供线索。
- ❖ Celis等结合免疫组织化学的技术对于膀胱鳞状细胞癌(squamous cell carcinoma, SCC)进行比较蛋白质组研究
- ❖ 鉴定了一系列区分SCC与移行细胞癌(TCC)的特异性表达的蛋白质分子标记
- ❖ 对SCC发生的细胞分化过程进行了系统研究。
- ❖ 获得的蛋白质,可用于SCC的早期诊断与监控

2. 鳞状细胞癌

- ❖ 日本学者Shugo Nawata等对鳞状细胞癌的重组抗原1、2(rSCC antigen-1, rSCC antigen-2)进行2D电泳分析
- ❖ SCCA1基因编码的rSCC antigen-1可分为pI为6.5、6.4、6.3、6.0,分子量为44500的四个蛋白点
- ❖ SCCA2基因编码的rSCC antigen-2为 pI 5.95的一个点。
- ❖ 肿瘤组织提取物中有3个pI为5.7-5.5、分子量为44500的新蛋白点;正常组织没有。
- ❖ 根据新蛋白点不被羟甲基-番木瓜酶水解的特点,认为这3个新蛋白点属于rSCC antigen-2,由SCCA2基因编码。
- ❖ 该新蛋白点的出现与肿瘤分化与肿瘤恶性行为有关。

3. 结肠息肉

- ❖ Mells等用2DE、Melanie II软件分析15例正常结肠和13例多发性结肠息肉上皮细胞蛋白质丰度变化,

共分辨出700个蛋白点，其中59个蛋白点的染色强度有差异。

- ❖ 在多发性结肠息肉中numatrin(nucleophosphine/B23)、热休克蛋白70、60表达强度增加，脂肪酸结合蛋白(L-FABP)、14-3-3、citic角蛋白20、细胞色素c氧化酶多肽Va、RhoGDP-解离抑制因子(RhogDI)、-actin等降低，这个结果相对恒定。

4. 心血管疾病

- ❖ 心血管疾病是疾病蛋白质组研究的重要领域。
- ❖ 目前，已经建立了人和大鼠、狗等实验动物的心房、心室与内皮细胞的蛋白质2DE图谱的一系列数据库(www.expasy.ch/ch2d/2d-index.html)，鉴定了数百个心血管系统特有的蛋白质。心血管疾病
- ❖ 英国Harefield医院实验室对扩张型心肌病(DCM)的进行了系统的蛋白质组研究，发现了大约100个表达异常的蛋白质，并进行了分类。
- ❖ 将2DE与免疫标记相结合，鉴定了一些能和自身抗体发生反应的肝脏特异性抗原。这些抗原与心脏移植中的急性或慢性排斥反应有关。
- ❖ 通过对一些心脏病动物模型，如狗的起搏诱导性心衰和牛的糖尿病心肌病(DCM)的心脏蛋白质组分析，他们发现了与人DCM相似的蛋白质表达异常，并探讨了某些蛋白质的异常修饰作用与疾病发生的关系。

5. 神经系统疾病

- ❖ Rohlf等对脑脊液的蛋白质组分析，发现了一系列与神经细胞信号传递有关的蛋白和多肽，这对于神经精神病的诊断具有重要价值。
- ❖ 老年痴呆症(Alzheimer症)是一种常见的神经系统疾病，其神经系统的特征性病理损伤主要表现为老年斑和胞质神经原纤维缠结。
- ❖ 对老年斑的蛋白质组分析，发现了一个在神经毒性的蛋白质沉积过程中，起关键作用的4000的多肽碎片——amyloid”。神经系统疾病
- ❖ 通过对神经原缠结的2DE及质谱分析发现，其主要蛋白质成份-微管相关蛋白(MAP—)在疾病状态下发生了磷酸化、糖基化等多种修饰作用。
- ❖ 上述工作对阐明Alzheimer症的发病机理有十分重要的意义。
- ❖ P. Edgar等分别将七例阿尔茨海默症(AD)患者、七例精神分裂症患者的海马组织的双向凝胶电泳蛋白质组图谱与正常对照组比较。发现海马蛋白质组谱中约有 549 ± 35 个蛋白质组分。
- ❖ 其中，阿尔茨海默症患者有35个蛋白含量下降，73个蛋白含量升高；而精神分裂症患者有8个蛋白含量下降，8个蛋白含量升高。
- ❖ 表明海马组织蛋白质水平改变可能是导致阿尔茨海默症与精神分裂症的分子事件
- ❖ 克—雅氏病、牛海绵状脑病等感染性蛋白质疾病，也是神经系统蛋白质组研究的热点

6. 感染性疾病

- ❖ 病原微生物蛋白质组研究对于阐明感染性疾病的发病机制，制备和设计新的疫苗与药物，实现有效的预防与治疗，有着重要的应用价值。
- ❖ 目前主要集中在以下几方面：确定致病菌毒力；寻找新的诊断标记物；为疫苗的研制寻找新的候选抗原；阐明药物抗菌作用机制与病菌耐药性发生机理，为新的抗生素的研制寻找靶点。

三. 病原微生物的蛋白质组研究

1. 结核分支杆菌

- ❖ 结核病是一种古老的疾病，目前在全球呈回升之势，对人类健康构成严重威胁，其致病源是结核分支杆菌
- ❖ 结核分支杆菌的基因组测序工作已经完成，共有3924个基因，该工作促进了结合分支杆菌的蛋白质组研究
- ❖ Rosenkrands等用双向聚丙烯酰胺凝胶电泳对结合分支杆菌的培养滤液、细胞壁和胞质溶胶的蛋白质组进行了研究，分别分离出376、413、395个蛋白质，经测序和免疫检测分别鉴定出12、9、10个新

蛋白质，其中一个特殊的蛋白质令人感兴趣，它在结核分支杆菌的基因组中未能找到对应的基因序列

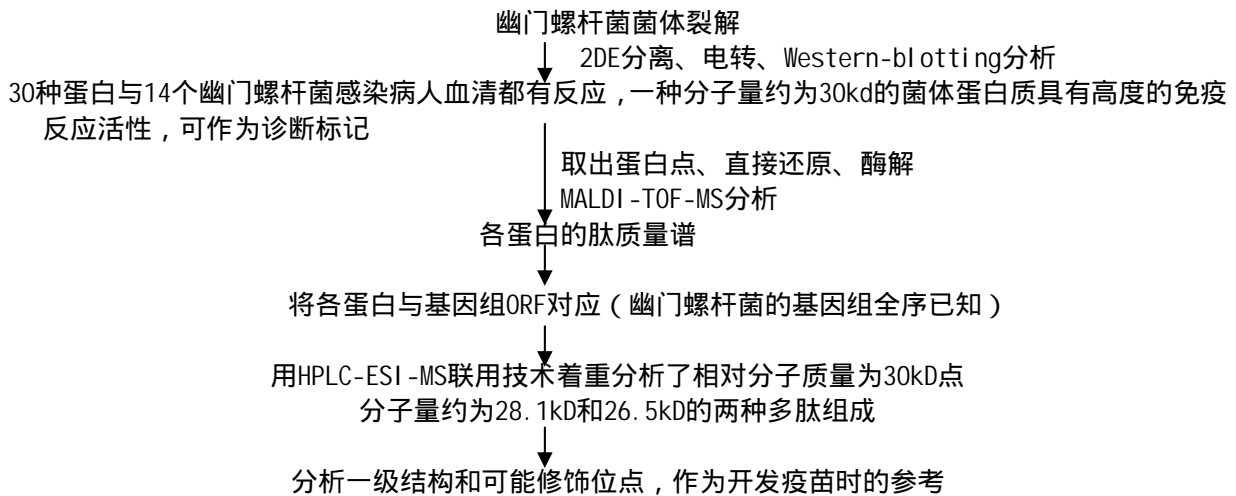
- ❖ Hendrickson等用蛋白质组学的方法对结合分支杆菌进行了研究，对分离的蛋白质结合串联质谱分析，鉴定出结核分支杆菌中的一种新的蛋白质，命名为Mtb81，用它作抗原建立ELISA法检测了27例结核病人的血清，其中25例呈阳性反应，阳性率为92%，而45例健康输血员的血清学全部呈阴性反应。
- ❖ Wel di ngh等通过结核分支杆菌蛋白质组分析，筛选到了一些有潜在的疫苗开发价值的新抗原。
- ❖ Mahai ras等对结核分支杆菌、牛结核分支杆菌和减毒中分支杆菌(M. bovis)的基因组和蛋白质组比较分析，发现了与病菌毒力决定有关的基因区域，为改善M. bovis BCG疫苗的免疫效力提供了线索。
- ❖ 这些工作，为实现结核病更有效的免疫、诊断和治疗提供了可能。

2. 包柔氏螺旋体

- ❖ 莱姆病(Lyme disease)是一种由包柔氏螺旋体引起的传染病
- ❖ Jungblut等对包柔氏螺旋体的蛋白质组进行了研究，他们对2DE-PAGE分离得到的蛋白质进行了质谱分析，同时结合免疫印迹检测
- ❖ 检测出65种包柔氏螺旋体抗原与莱姆病患者血清发生反应，对其中20种抗原进行鉴定后又发现了包柔氏螺旋体的两种新的蛋白质，即甘油醛3-磷酸脱氢酶(glyceraldehydes-3-phosphate dehydrogenase, GAPDH)和ABC转运寡肽渗透酶(ABC transporter oligopeptide permease)

3. 幽门螺杆菌

- ❖ 幽门螺杆菌(Helicobacter pylori)是引起胃炎、消化性溃疡等上消化道疾病的主要致病源之一
- ❖ McAtee等结合血清学对幽门螺旋菌Z-170株蛋白质组分进行了分析



- ❖ 该研究可以迅速分析抗原成分，无疑为微生物疫苗的开发、血清学试剂的优化等提供了新的手段和思路

4. 伤寒沙门氏菌

- ❖ O'Conner等对伤寒沙门氏菌的蛋白质组进行了研究，用2D-PAGE分离生长于不同环境条件下的、毒力不同的伤寒沙门氏菌包膜蛋白
- ❖ 随机选取53种蛋白质进行测序分析，发现有20%的蛋白质序列与当前数据库中的蛋白质不同
- ❖ 可见用蛋白质组学进行伤寒沙门氏菌毒力分析是十分有益的

5. 肺炎链球菌

- ❖ Cash等对肺炎链球菌的蛋白质组进行了研究。
- ❖ 发现蛋白质双向电泳图谱中有一种特殊的蛋白质(pI 6.27, 38.5kD)，而对红霉素敏感的肺炎链球菌则不存在该蛋白质。
- ❖ 经鉴定，该蛋白质是GAPDH

- ❖ 根据等电点的不同，GAPDH又有3中不同的存在形式，从抗红霉素的肺炎链球菌内鉴定出来的是碱性最强的GAPDH，它可降低细胞内红霉素的水平，其作用机制目前仍未能明了。

6. 弓形虫

- ❖ 原生动物弓形体 (Toxoplasma gondii) 寄生感染引起弓形虫病
- ❖ 全世界约有30%的人携带此种寄生虫。在欧洲，弓形虫病是发生频率最高的传染病之一
- ❖ 对免疫功能缺陷或受抑制的患者以及孕妇的健康构成严重的威胁，如引起孕妇流产、胎儿畸形等。

症状

- ❖ 健康人群中，寄生虫的感染通常是无症状的或症状极其轻微的。
- ❖ 怀孕期感染，寄生虫会通过胎盘，并造成胎儿死亡。随着怀孕时间的增加，寄生虫穿透的可能性会增加。
- ❖ 因此确定感染的时间非常重要。

后果

- ❖ 怀孕不同时期的感染后果不同
- ❖ 怀孕早期，器官形成过程时的感染危害可能是致命的
- ❖ 怀孕后期，胎儿的感染经常会导致一些并发症，如视网膜色素异常等
- ❖ 若怀孕期间感染的妇女得到了充分治疗，胎儿感染的可能和后果严重性都会大大降低
- ❖ 因此及时诊断和准确判断感染时间对于该病的治疗非常重要

目前状况

- ❖ 90%以上的怀孕妇女的初期感染不能被及时发现
- ❖ 现诊断手段：血清学方法和PCR法
- ❖ 血清学法检测抗体，对于一些无免疫应答的和怀孕病人有缺陷，而潜伏性感染致病恰恰是经常发生在无免疫应答的人中，如艾滋病患者中，T. gondii 就是导致脑内病变并致死的主要原因

操作

用2D-PAGE分离弓形体的蛋白质组分

与不同感染时期弓形体感染者血清进行免疫印记

寻找和感染相关的抗原

检测的血清

- ❖ 急性感染弓形虫病的怀孕妇女血清
- ❖ 急性弓形虫病的非怀孕病人血清
- ❖ 潜伏性感染弓形虫的尚未发病者血清

研究结果

- ❖ 9个点可以和感染者血清中任何一类型的免疫球蛋白反应，且这种反应和感染的状态和发病与否无关。（这9个点可用来作为T. gondii 感染的标记）
- ❖ 7个点与抗体类型或发病情况有关，可用来区分不同疾病状况，如潜伏期和急性期等（可以作为进一步判断感染状态的诊断标记）

第三节 蛋白质组学在药物开发中的应用

- ❖ 新药研究是高风险、高投入、高回报的产业，分药物发现和药物开发两个阶段。据统计平均每个新药上市需耗时7~12年，耗资近300万美元。而I期临床试验通过率仅为8.3%。因而当今新药研究所面临的重点和难点是如何加快新药发现的速度，如何增加可进入新药开发阶段的新化学实体(New Chemical Entities, NCE)的数量。

1. 前面已提到过的运用蛋白质组分析发现的, 在病理状态下表达异常或者特异性表达的蛋白质, 以及细胞信号传递通路中的关键性蛋白, 都可能作为药物设计与发现的靶分子。此外, 病原微生物蛋白质组研究, 也有助于了解其致病的机理和发现对药物敏感的蛋白质, 为新的抗生素筛选提供更合理的靶点。
2. 通过对已知药物治疗前后病理组织的蛋白质组进行比较分析, 不仅可以评价和确定前述“可能的靶分子”在药物设计中的价值, 加快药物靶点的认定, 减少后续工作的盲目性, 还有助于阐明药物的分子药理, 构建更为合理的筛选模型。此外, 还可以将蛋白质组分析与组合化学的方法相结合, 对结构类似物的构效关系作出比较, 加速先导化合物的筛选与优化。
3. 通过发生损伤的组织器官与正常组织器官之间的蛋白质组比较, 有助于阐明药物毒副作用的发生机制。

- ❖ 随着人类基因组测序工作大规模的展开, 获得全序列的基因已近万余, 其中约有90%的基因已测得了其表达序列标志 (EST), 这就提供了更多的可供筛选的靶标
- ❖ 但核酸作为药物的靶分子, 存在着几点致命的缺陷

一. 核酸作为药物的靶分子的缺陷

- 1) 核酸结构很多情况下是同源性的, 作用于DNA的药物大多选择性差, 毒性大, 特别是形成共价结合的药物都有严重的细胞毒作用。
 - 2) 大多数疾病都是表征在蛋白质水平上, 而不是基因水平上。
 - 3) 如何将核酸结合的药物释放到相应组织仍是一个困扰。
 - 4) 很多疾病如癌症、心血管疾病往往是多基因共同作用的结果或具有遗传上的多态性, 因而很难找到关键基因作为靶标
- ❖ 蛋白质组的研究, 不仅能为药物开发提供治疗机理上的启示与线索, 还可以提供筛选、评价方面的高效率的技术手段。
 - ❖ 其应用主要有以下几方面: 药物靶点的发现。 靶点的评价、认定与筛选的优化 药物毒理学分析与临床前安全性评价。

二. 药物蛋白质组学的重要研究内容

- ❖ 临床前: 1) 新药和靶的发现, 2) 药物作用模式研究, 3) 毒理学研究,
- ❖ 临床研究: 1) 疾病特异性蛋白作为有效患者选择的依据和临床试验的标志, 2) 应用类似于药物遗传学的方法, 按照蛋白质谱来分类患者, 并预测药物作用疗效。
- ❖ 蛋白质水平的分析不仅为生物分子体系提供了最有效的实时分析模式, 而且也能获得在DNA和RNA水平上不易获得的信息
- ❖ 例: 对肝癌患者的最近研究表明, 甲胎球蛋白 (AFP) 的糖基化变化是比甲胎球蛋白更为灵敏的指标。
- ❖ 正常和异常的糖基化甲胎球蛋白在氨基酸序列上则没有差异, 应用遗传学水平的分析则不可能发现这种差异。
- ❖ 蛋白质组研究已经鉴定出与肝癌有关的许多特殊蛋白, 其中包括血清转铁蛋白、纤维蛋白原b链、结合珠蛋白1。
- ❖ 二维凝胶电泳和质谱的结合与精致的生物信息学建立了一套强有力的、通用的方法, 用于蛋白质的分布作图和确定蛋白质的结构特征, 使得药物蛋白质组学至少具有与遗传学水平上相同的作用。
- ❖ 随着蛋白质组学、药物蛋白质组学研究的兴起, 人们将在蛋白质水平上重新认识诸如生长、发育和代谢调控等生命活动的规律, 为研究重大疾病的机制、疾病诊断、防治和新药开发提供重要的理论基础, 并正在成为生物技术药物发展的根本动力, 明显地加快开发新的治疗和诊断方法。
- ❖ 美国Human Genome Science公司首先将系统的基因组和蛋白质组知识作为临床实验的潜在产品, 他们从中发现了骨髓祖细胞抑制因子 - 1 (MPLF - 1) 和角质细胞生长因子 - 2 (KGF - 2), 目前它们已进入 期临床, 血管内皮生长因子 - 2 (WEGF - 2) 也处于 期临床。
- ❖ 该公司的功能基因组计划, 分离全部人类基因95%以上的mRNA, 鉴定了12000个新基因、编码分泌蛋

白,已完成序列分析大约9000个基因和11000个蛋白。其中包括35个新白介素,40个新生长因子和100个新7-跨膜受体,已取得3000个以上的专利及新发现基因的潜在医学用途

三. 蛋白质组表达图谱用于基因组功能提示的可行性

- ❖ 谢涛等以EC02DBASE (Edition6) 为研究材料,研究了利用蛋白质组表达图谱提供的生命动态活动信息提高基因组功能提示效果的可行性。
- ❖ 在设计出一套较为完整的细胞功能簇(CRC)聚类方案的基础上,考察发现:79个蛋白质聚成4个不同的CRC。
- ❖ 显示:功能相关的蛋白质趋向于聚集在相同的CRC中,如9种氨酰tRNA合成酶和4种热休克蛋白分别准确的聚合到CRC2和CRC3中。
- ❖ 提示:在蛋白质组研究中比较充分的前提下,通过有效的算法,蛋白质表达图谱可以为基因组功能提示提供非常重要的序列相似性之外的功能信息。

四. 分子药理筛选模型

- ❖ 采用分子水平的筛选模型的大规模药物筛选已被普遍用于第一部初筛,其特点是特异性强、灵敏度高、微量快速。
- ❖ 蛋白质组学技术将来应用于筛选模型构建中的一个最大的优势是更清楚、更详尽的阐明分子药理机制,提供更为有效、合理的药理模型。
- ❖ 根据现在蛋白质组学的研究结果,表明蛋白质基因表达的调控构成了真正的药物作用机制

1. 抗生素作用机理研究

- ❖ 最近10年,用于临床的抗生素几乎都是原有抗生素的衍生物。
- ❖ 目前这种状况有所改变,寻找作用于细菌内新的靶子的研究工作已经展开,找到了许多有效的抗菌化合物。
- ❖ 目前遇到的困难:难于揭示新的化合物的作用靶子及其作用机理。
- ❖ 有时虽在体外发现新化合物能够使某种蛋白质失活,但在体内是否有这种现象和这是否是抗菌的主要机制任然未知,双向电泳分析提供了一个有效的手段。

例:抑制核糖体类抗生素对细菌作用机制的研究

- ❖ 对12种作用于翻译过程的不同阶段和核糖体内不同分子的抗生素加以考察
- ❖ 发现其中8种诱导冷休克反应(cold-shock response)的一系列蛋白质,另外有4种诱导热休克反应的一系列蛋白质,
- ❖ 预计新的作用于核糖体的化合物也是诱导这两种反应,并且籍此可推测大肠杆菌对冷热的反应发生于核糖体水平。

2. 分子药理筛选模型

- ❖ 例1:在黄曲霉素诱发的肝癌细胞中加入dithioethiones(一种治疗黄曲霉素诱发肝癌的药物),通过与对照组细胞的2DE图进行比较,可以清楚地看到,在它的作用下,一种名为黄曲霉素B1醛还原酶的蛋白质得以表达,表明这种蛋白质具有使黄曲霉素失去毒性的作用。因而基因表达水平的改变是用来反映疾病和药物作用机制的分子印迹(molecular imprints)。
- ❖ 例2:Anderson等组建了两个蛋白质组数据库。一个是分子解剖学和病理学数据库(MAP™),另一个是药物的分子效应数据库(MED™)。前者主要用来界定蛋白质在不同组织中的表达模式和在各种疾病过程中的表达改变,后者着眼于药物作用机理的阐明。这些数据库将成为新药发现的指南针,同时亦将指引人们跃进到分子药理学又一新的纪元

五. 毒理学

- ❖ 双向电泳是筛选毒素,显示其毒性机理的高灵敏的手段。
- ❖ 比较给药前后的蛋白表达状况,观察可能与药物效果或毒性有关的系列蛋白的变化,识别生化途径

的改变过程。

- ❖ 若建立了大量已知毒性化合物的蛋白质组库，就能用来预测新化合物的毒性。
- ❖ 但目前几乎没有大范围的毒理学蛋白质组研究公开化。
- ❖ 进行已知药物的毒理学蛋白质组分析，可以鉴定和积累特定组织损伤的蛋白质标志物，为临床前安全性评价提供指标，实现新药毒性预测，减小药物临床实验的风险，避免安全事故和不必要的资金浪费。
- ❖ 用2DE和NMR研究puramycin aminonucleoside 处理小鼠的肾小球中毒性肾损伤，通过监测尿中蛋白，对与肾小球中毒性肾损伤有关的蛋白尿的性质和过程有了更详细了解。
- ❖ 在导致毒性的兔模型中检测到许多随暴露而变化的蛋白，其中一些是谷胱甘肽转移酶的突变体，它们可以作为导致人毒性的可变标志物。
- ❖ Aicher等运用比较蛋白质组的方法对环孢菌素A(CsA)的肾脏毒性加以分析
- ❖ 发现在药物作用下负责钙排泄的钙结合蛋白D(calbindin—D)的表达明显降低。
- ❖ 结合病理组织分析，说明钙在肾脏中的沉积是CsA造成肾脏损伤的主要原因。
- ❖ 过氧化物体扩增因子还原剂处理后肝脏蛋白改变的研究。

1. 肾脏功能

- ❖ 肾脏是抗生素攻击的主要靶器官，而肾脏皮质和髓质组织结构的差异，使得药物毒性在肾脏内有区域易感性
- ❖ 用2DE、基质辅助激光吸收/离子-质谱等技术分析了大鼠肾脏皮、髓质细胞，皮质中分辨出727个蛋白点，髓质中716个，其中127个蛋白点的丰度有差异（86个在皮质、41个在髓质），26个蛋白只存在于皮质中，4个只存在于髓质中。这些蛋白质图谱的变化有助于对肾脏功能的了解。毒理学--肾脏功能
- ❖ Aicher等用2DE研究了肾脏种28kD的钙结合素（与钙离子结合，减少肾脏的钙化）的作用。
- ❖ 狗、猴的肾脏由于缺少环孢菌素-A介导的肾脏毒性，而人接受移植肾脏的肾内血管和肾小管却有肾毒性表现。毒理学--肾脏功能
- ❖ 钙结合素-D再经过环孢菌素-A处理的肾脏中有明显下降，且有剂量依赖性，说明钙结合素-D是环孢菌素-A肾脏毒性的标志物。
- ❖ 该标志物对临床器官移植的排斥反应有重要的指导意义。

六. 生物技术药物的开发

- ❖ 生物技术药物包括基因工程药物、寡核苷酸药物、基因工程抗体药物、DNA药物等。
- ❖ 生物技术药物开发的主要动力来源是基因组和蛋白质组的研究，其中最令人瞩目的进展是抗体药物。
- ❖ 抗体是机体抵抗病原体的主要防御系统，它们由免疫系统自主产生的蛋白质。
- ❖ 每个抗体都可以特异性识别作为抗原靶分子的特征，哪怕是极其微小的差别。目前大多数抗体药物都是鼠人嵌合体抗体和人源化抗体。生物技术药物的开发
- ❖ CAT生物技术公司正在利用蛋白质组学独特的技术体系，及组合噬菌体抗体文库，加上后基因组学的一些技术，开发能够成为人类今后治疗药物的人抗体

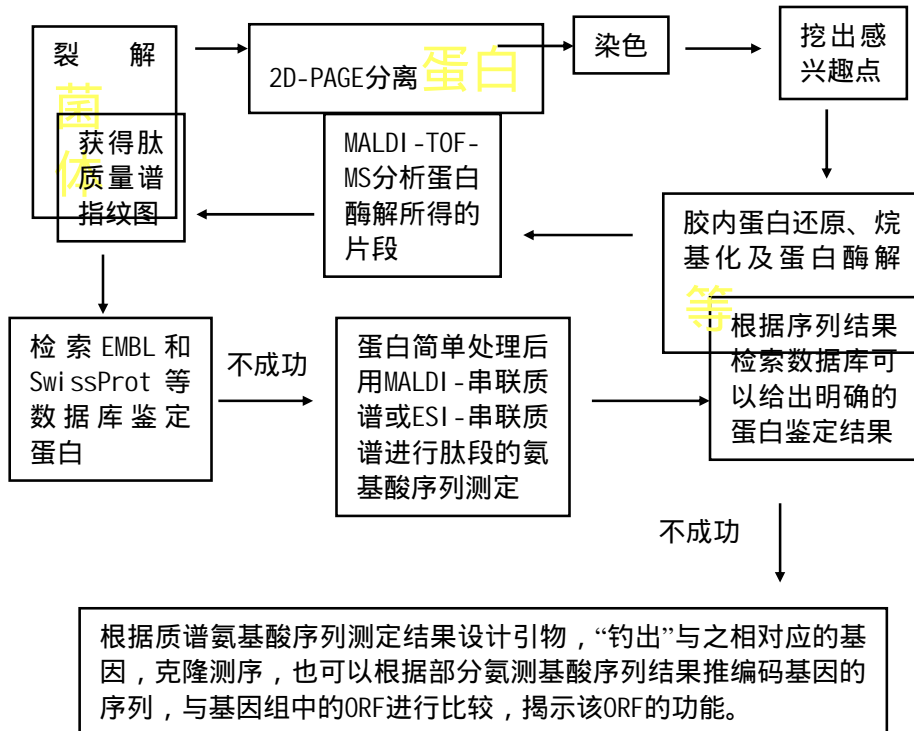
七. 中医药与蛋白质组学

- ❖ 真核细胞生物基因复杂，基因不是生物功能的执行体，它总要表达为相应的蛋白质
- ❖ 一般情况下，基因组只表达一部分基因，其表达类型与表达程度均受到生存环境及内在状态变化的影响而有较大差异。
- ❖ 采用对一个基因组的全套蛋白质、一个细胞乃至一个组织的全套蛋白质的蛋白质组学研究方法，进行中医药学的理论与药物研究同样具有重要的意义。中医药与蛋白质组学
- ❖ 目前，高分辨率的双向电泳分离、图象数字化及质谱分析等蛋白质研究技术上的突破，为中医药学与蛋白质组学的结合研究奠定了基础。

第二部分 各类生物蛋白质组研究现状

- ❖ 目前，蛋白质组研究还主要集中在原核生物和一些简单的真核生物，尤其是一些基因序列被完全搞清或大部分已知的生物如支原体、细菌、酵母等。

一般的蛋白质组研究技术路线



- ❖ 通过蛋白质组研究，对比细菌在某种处理条件下的二维-SDS-PAGE图与正常情况下的区别，将特征点取出，即可获得与这种处理条件应激有关的蛋白信息
- ❖ 选择基因序列已知的生物，将双向电泳得到的蛋白质谱和基因组分析相结合----建立蛋白质-基因联合数据库----可得到更多有用信息。

建立蛋白质-基因联合数据库的目的

- 1) 所有编码基因的产物在双向图谱上的定位
- 2) 确立各种蛋白质的丰度
- 3) 分析不同条件下各种蛋白质表达水平及合成速率的变化
- 4) 各种蛋白质在细胞内的定位

- ❖ 某些蛋白质翻译后修饰的方式和水平。

一. 原核生物的蛋白质组

1. 流感嗜血杆菌的蛋白质组研究

- ❖ 流感嗜血杆菌 (*Haemophilus influenzae*) 是一种感染人呼吸道的致病微生物。
- ❖ 是第一个获得基因组全序列的生物。其基因组长约1.8Mb，比大肠杆菌基因组的一半还小，是一环状DNA分子。已鉴别有6个拷贝的rRNA基因，54种不同tRNA基因和1743个有能力编码蛋白质的区域。
- ❖ 在基因组信息的基础上进行蛋白质组研究已成为研究这一细菌的热点。
- ❖ Langen通过2D-SDS-PAGE和MALDI-TOF-MS分析，将所获得的肽质量谱指纹图对库检索，一次性鉴定了流感嗜血杆菌Rd株的119种蛋白质
- ❖ Fountoulakis通过均质Tris-甘氨酸凝胶(5~25KD小分子量蛋白质，从流感嗜血杆菌Rd株分离了80种小分子量蛋白，全部用MALDI-TOF-MS获得了肽质量谱指纹图，其中40种是首次报道。

瑞士一小组的研究实例

- ❖ 通过2D-PAGE研究了流感嗜血杆菌Rd株的蛋白质组分，在pH3 ~ 10的范围内鉴定了约300个蛋白质组分，
- ❖ 用有两个尿素浓度的麦黄酮（tricine）胶分离了很难分离到的5 ~ 20ku范围内的蛋白质，并鉴定了其中的80种。
- ❖ 用肝素亲和层析将碱性蛋白质富集后，在pH6 ~ 11的范围内又鉴定了102个蛋白质，其中许多是核酸结合蛋白尤其是核糖体蛋白。

华盛顿大学的研究实例

- ❖ Link及其同事研究了流感嗜血杆菌模式株 NCTC 8143蛋白质组主要成分，
- ❖ 将双向电泳中等电点在4 ~ 7之间，相对分子质量在10 ~ 100kD之间的303个蛋白质斑点通过MALDI-TOF-MS以及串联质谱获得肽质量谱指纹图，进行肽段氨基酸序列测定
- ❖ 鉴定了各个蛋白质，并将各点与基因组ORF对应起来，发现其中最为丰富的蛋白质为外膜蛋白P2，其次是能量代谢和大分子合成相关蛋白。
- ❖ 发现了NCTC 8143株中，包括色氨酸酶在内的几种蛋白的编码序列是Rd株基因组中所没有的
- ❖ 所研究蛋白大多数等电点和分子量与基因组ORF推测值相吻合，大约22%蛋白的等电点和分子量与推测值相距甚远，显示它们可能经过了翻译后加工，也可能是根据基因组直接推断蛋白质特性时产生了偏差。
- ❖ 提醒人们，仅仅依靠基因组研究是不够的。

预解决的问题

- a. 所有编码基因的产物在双向图谱上的位置；
- b. 各种蛋白质的丰度；
- c. 不同条件下各种蛋白质表达水平及合成速率的变化；
- d. 各种蛋白质在细胞内的定位；
- e. 某些蛋白质翻译后修饰的方式和水平。

2. 大肠杆菌的蛋白质组研究

- ❖ 大肠杆菌由4.7Mb组成的双链环状DNA分子，若按1000bp编码一个基因算，约有4000个基因
- ❖ 目前大肠杆菌全部4288个基因的序列已被测定并定位
- ❖ 大肠杆菌的联合数据库已更新到第7版

大肠杆菌的蛋白质组研究现状

- ❖ 包括1600个以上蛋白质斑点的数据库，其中大约400个已与大约350个基因相对应。
- ❖ 对于这些蛋白质，此数据库可以提供基因名称、蛋白质名称、EC编号、功能范畴、Swiss-Prot数据库编号、Genbank序列号、基因图谱的位置、染色体上的转录方向以及一些生物信息（如在不同生长条件下该蛋白质在细胞内的丰度）等。
- ❖ 双向电泳技术的最大优势：可以对细胞内复杂的蛋白组分进行整体性的定量分析，从而可以分析当条件变化时，一系列蛋白质而不仅仅是某一蛋白质的表达变化。
- ❖ 在基因水平，“调节子”被用来描述受控于一个调节蛋白的一系列基因；在蛋白质和基因水平，刺激子（stimulon）用来描述一系列蛋白质对同一刺激物的反应。大肠杆菌的蛋白质组研究现状2
- ❖ 研究了碳、氮、磷及硫元素限制导致的细胞内蛋白质谱的变化
- ❖ 在对营养成分磷限制的研究中，发现当磷饥饿时有137个蛋白质的合成速率明显改变，其中大部分（118个）的表现诱导合成，其他则被抑制。
- ❖ Dainese等人将蛋白质组技术与基因敲除（gene knockout）技术结合起来，成功地研究了大肠杆菌在硫饥饿下蛋白质表达的情况
- ❖ 对比2D-SDS-PAGE图，分离到8种与硫饥饿相关的上调蛋白
- ❖ 通过酶解和MALDI-TOF-MS及ESI-MS分析得到各蛋白的肽质量谱
- ❖ 检索EMBL和SwissProt蛋白数据库，将各个蛋白与大肠杆菌基因组中的ORF对应，并由基因敲除实验

加以证实。

3. 致病微生物的蛋白质组研究

- ❖ 近年来，关于传染病的研究变得比原来更重要。一些新的传染原，如Borrelia burgdorferi, HIV, Ebola病毒等的出现，使得一些原来认为已经被控制的疾病如结核，多抗药性的链球菌属感染等又有所增加。
- ❖ 32种微生物的基因组测序已经完成，另有60多种微生物的基因组测序正在进行中，这些基因序列的信息和相对真核组织来说少得多的基因组数量为他们的蛋白质组研究提供了良好的基础。

4. 微生物蛋白质组研究的应用及进展

1) 系统细菌学研究

- ❖ 系统细菌学包括细菌分类、鉴定、系统进化等多方面的内容，以色谱技术等化学分析手段为基础的化学分类方法已经成为系统细菌学中的重要分支
- ❖ 由于包括MALDI-TOF-MS和ESI-MS（与HPLC等联用）在内的软电离质谱技术能为研究者提供复杂样品中各组分的分子量信息，它们正在成为系统细菌学研究者新的工具
- ❖ Hagg用MALDI-TOF-MS技术直接分析菌体的水-乙腈可溶组分，获得了嗜血杆菌属的4个种，包括流感嗜血杆菌、副流感嗜血杆菌、杜克氏嗜血杆菌和嗜沫嗜血杆菌，以及放线杆菌属和奈瑟氏菌属某些细菌的蛋白质量指纹图。
- ❖ 发现水-乙腈可以溶解细菌体表面松散结合的蛋白，在质谱图上，质量数5~15kD区域最具鉴别特性
- ❖ 所有菌株可以通过肉眼分开是因为这一区域，它们没有共同的峰。
- ❖ 分析发现：杜克氏嗜血杆菌的不同临床分离株间也存在着质谱图的差异
- ❖ 对炭疽芽孢杆菌，鼠疫耶尔森氏菌和马尔他布鲁氏菌进行了质谱分析，
- ❖ 病原菌产生的质谱图与之紧密相关的非致病菌的质谱图有着肉眼可以识别的显著差别
- ❖ 病原菌毒力不同的菌株之间也有差异
- ❖ 提示MALDI-TOF-MS用于细菌鉴定和菌株分型的前景
- ❖ Nilsson对用不同基质对幽门螺杆菌的菌株进行分型研究，获得了一系列稳定可靠，用于幽门螺杆菌鉴定和分型的质谱峰。
- ❖ 实验表明，将细菌的裂解悬液用HPLC-ESI-MS分析也可以进行细菌的鉴定。
- ❖ Claydon系统研究了多种革兰氏阳性菌和革兰氏阴性菌的全细胞MALDI-TOF-MS质量指纹图，研究重点放在质量数为550~2000区域。

革兰氏阴性菌	革兰氏阳性菌
分子量550~2000区域，有一系列质量数相差129的质谱峰	无
有一些共有峰，但不同种，乃至同一种不同菌株都可以直接用肉眼观察加以鉴别	大部分的质谱峰是相同的，只有少数几个不同的峰可以用作菌种的鉴定

- ❖ 在该区域，革兰氏阴性菌都有一系列质量数相差129的质谱峰（可能是肽聚糖逐步失去戊糖或谷氨酸的产物）；而革兰氏阳性菌中没有观察到这些离子，这可以作为MALDI-TOF-MS鉴定革兰氏阳性菌和革兰氏阴性菌的指标。
- ❖ 所研究的革兰氏阴性菌都有一些共有峰，但不同种，乃至同一种不同菌株都可以直接用肉眼观察加以鉴别；
- ❖ 而葡萄球菌属的3个种在这一质量区域内，大部分的质谱峰是相同的，只有少数几个不同的峰可以用作菌种的鉴定。
- ❖ 分枝杆菌（M. smegmatis）质谱图中1150~1130的范围内有强度很大的一群峰

- ❖ 峰间质量数的差异为14，即甲基的分子量
- ❖ 这些峰是由分枝菌酸的同系物产生的
- ❖ 质谱图中的这一区域的峰可以作为分枝杆菌及其相关菌属的生物标记物 (biomarker)

2) 病毒学研究

- ❖ 病毒由核酸及蛋白质组装而成，部分病毒有脂蛋白或糖蛋白构成的包膜，其结构与细菌比较而言要简单得多，因此从理论上说，病毒更适合直接用软电离质谱技术进行分析
 - ❖ 病毒的衣壳蛋白单体的相对分子质量通常在 10×10^3 左右，将病毒用一定基质溶解，使衣壳解体即可直接用于MALDI-TOF-MS分析，研究者所要做的只是优化质谱图的信号。
 - ❖ 病毒衣壳蛋白MALDI-TOF-MS分析所用的基质一般是3,5-二甲基-4-羟基肉桂酸或是芥子酸，病毒在基质中的浓度为1mg/ml，可以直接准确地测定衣壳蛋白的相对分子质量
 - ❖ Bother等人用MALDI-TOF-MS分析了一种动物病毒FHV (flock house virus) 的悬液，成功地测定了蛋白前体 及其自身酶解产物 和 亚单位的精确分子量
 - ❖ Marmey等人首次用质谱技术精确测定了水稻病毒RTBV (rice tungro bacilli form virus) 衣壳蛋白的相对分子质量
 - ❖ MALDI-TOF-MS技术可以成为病毒衣壳蛋白分子量测定的常规手段
 - ❖ 质谱也可用于研究病毒衣壳蛋白的动力学特征
 - ❖ FHV的晶体衍射照片显示 亚单位和 亚单位的N-端和C-端处于病毒的内部，用MALDI-TOF-MS分析病毒经蛋白酶作用不同时间后所产生的肽段可以找到衣壳蛋白最容易受蛋白酶作用的部分，也就是暴露在病毒表面的部分
 - ❖ 发现这些肽段包括了 亚单位和 亚单位的N-端和C-端的产物，这说明它们在溶液中处于一种动态，经常转移到表面
 - ❖ 值得注意的是 亚单位和 亚单位的N-端和C-端与FHV核酸的释放密切相关。
 - ❖ 使用类似的手段对人鼻病毒 (HRV14) 进行研究
 - ❖ 也证实了衣壳蛋白的这种动态“呼吸”结构，同时发现病毒抑制剂WI N52084可以阻断这种动态结构，从而抑制病毒的生长，
 - ❖ 时序性的酶解结合质谱技术可以用于抗病毒药物的筛选和评价
 - ❖ 质谱记录分子量的变化来研究病毒蛋白质的转录后修饰以及变异情况。
 - ❖ 人鼻病毒HRV14的衣壳蛋白VP4是一个高度肉豆蔻酰化的蛋白，Lewis用MALDI-TOF-MS测定了它的相对分子质量，发现比从蛋白数据库中检索到的蛋白质相对分子质量要大212，而这正是一个肉豆蔻酰残基的相对分子质量，这说明HRV14 VP4蛋白还有一个未被发现的肉豆蔻酰化位点。
 - ❖ 对VP4蛋白进行酶解和串联质谱分析，确定该位点是在N-端。
 - ❖ Jonscher等用MALDI-四极杆离子阱质谱成功地发现了仙台病毒P蛋白的两个磷酸化位点]。
 - ❖ 有人将人呼吸道合胞病毒的黏附蛋白经HPLC纯化后多重酶解，酶解肽段用MALDI-TOF-MS分析后，确定了两个二硫键的位置。
 - ❖ Lewis用ESI-MS结合酶解技术分析肽段的质量变化，准确鉴定了HRV衣壳蛋白1199位由半胱氨酸到酪氨酸的自然变异和烟草花叶病毒 (TMV) 中由基因工程介导的人工变异。
 - ❖ 质谱检测和鉴定病毒变异不需进行测序，是病毒学家研究变异的新选择。
- ## 3) 微生物蛋白质组研究进展
- ❖ 微生物蛋白质组研究在技术上的新进展是使用二维液相色谱取代二维-SDS-PAGE电泳进行蛋白质的分离，随后的研究仍需要MALDI-TOF-MS和ESI-MS
 - ❖ Jensen用毛细管等电聚焦技术结合傅立叶转换离子回旋加速共振质谱进行蛋白质组研究，只需300 ng的大肠杆菌蛋白即可完成分析。
 - ❖ 随着蛋白质组研究逐渐深入，EMBL、EBI、Swiss-Prot等数据库中蛋白质的MALDI-TOF-MS酶解肽质量谱指纹图的数据以及肽段氨基酸序列信息越来越多，这极大地方便了后续的研究工作。MALDI-TOF-MS图谱检索蛋白质数据库<http://www.mann.embl-hedelberg.de/MassSpec/software>

- ❖ Fuerstenau用自制的ESI-MS系统成功地分析了相对分子质量高达 10^6 的DNA分子,这种高质量检测上限的质谱如果商品化,将成为微生物质谱研究最理想的手段。
- ❖ 有人利用外源凝集素(Lectin)结合在MALDI-TOF-MS电离点的表面,可以从成分非常复杂的模拟样品中分离到可与Lectin特异结合的糖以及一种细菌,随后直接进行质谱分析。
- ❖ 这是一种极有前途的技术,将不同的分子如单抗,单链DNA等固定在质谱仪电离点上,可以利用分子间的相互作用直接从混合样品中分析目标分子,这使得不经培养直接检测微生物成为可能,在临床微生物、环境微生物和军事医学方面具有重大的意义。

二. 酿酒酵母的蛋白质组研究

- ❖ 利用双向电泳技术分析酵母蛋白质谱的工作早于蛋白质组概念提出
- ❖ 1996年其基因组的完成及其蛋白质数据库在网上的建立大大加速了这一工作。
- ❖ 最新版的数据库包括6000页,每页代表一个已知或推测的酵母蛋白
- ❖ 借助数据库的有力推动,在双向电泳图谱上最明显的蛋白质斑点的大部分得到鉴定。

1. 数据库提供的信息:

1. 序列为基础的已知或推测酵母蛋白质的特征信息,如分子量、等电点、氨基酸组成、多肽片段大小等
2. 一些研究得到的关于各种蛋白质在翻译后加工、亚细胞定位、功能分类方面的信息。
3. 从以往的超过5000篇关于酵母蛋白质的研究文章中获得的有关各种蛋白质功能、相互作用、突变表型的信息

2. 对单一基因缺失的研究

- ❖ 丹麦一项研究计划分别作出野生型和将基因系统缺失的缺陷型酵母的双向蛋白质图谱。
- ❖ 比较野生型和有单一基因缺失的缺陷型酵母的双向蛋白质图谱发现:单一基因缺失的结果并不是仅缺乏此基因编码的蛋白质,还能导致其它一系列蛋白质的改变,即引起蛋白质组全局性的变化
- ❖ 这种的改变有的是量的减少,有的是量的增加,还有一些被加工修饰而导致性状的变化。
- ❖ 大约20%的酵母基因缺失是致死性的,另外80%则不然,所以机体能通过调节蛋白质的表达来对抗细胞的这种内源性变化。
- ❖ 基因缺失的研究使人们了解细胞内蛋白质分子间相互“对话”和“作用”的重要信息,如蛋白质间的物理联系(有些蛋白质可能会组成蛋白质复合物)、信号传递途径,从而了解细胞是如何构成的以及如何协同工作的。

三. 多细胞真核生物的蛋白质组研究

1. 线虫的蛋白质组研究

- ❖ 利用双向电泳技术,Bini等对线虫(*C. elegans*)进行了蛋白质组分析。
- ❖ 在等电点3.5~9和分子量10~200ku的范围内可分辨2000个以上的蛋白质斑点。线虫的蛋白质组研究
- ❖ 利用Edman微量测序技术对其中24个斑点进行分析,得到其中12个蛋白质的N端序列。
- ❖ 已测出的12个中有一个未能找到与其匹配的基因。另外11个与能量代谢、酸性核蛋白、G蛋白等对应。
- ❖ *C. elegans*的19,000个基因已于1998年12月全部测出,预计会对蛋白质组的研究提供帮助。

2. 果蝇的蛋白质组研究

- ❖ 不同性别果蝇(*Drosophila melanogaster*)成虫的头、胸、腹部的蛋白质组图谱已分别被作出
- ❖ 共有约1,200个蛋白质被检出
- ❖ 大多数检出蛋白质在头、胸、腹中是相同的,
- ❖ 也发现了一些部位、性别特异的蛋白质。

四. 人类蛋白质组的研究

- ❖ 真核生物基因组的一个普遍特性是存在大量非编码序列，多数真核生物基因组的复杂性是由于几种不同类型的非编码序列丰富造成的。
- ❖ 人的基因组比大肠杆菌基因组大700倍，而人基因组理论上估计包含约 10^5 个基因—仅比大肠杆菌基因数目多25倍。
- ❖ 目前人类基因组草图已经创立，需进一步进行功能研究。人类蛋白质组的研究
- ❖ 人类的蛋白质组研究吸引了最多的注意力
- ❖ 由于人有大量的组织、细胞类型和发育阶段，对人蛋白质组的研究主要聚焦在特异组织、细胞和疾病上。如心脏组织、肝脏组织；肿瘤等
- ❖ 已有的证据表明，尽管人的不同组织有着很大的差别，但其中许多蛋白质是看家蛋白，因此他们的双向图谱可能也是近似的。这点与简单生物不同。
- ❖ 对人类来说，一个高质量的基本的双向图谱是非常有用的，它可以作为其他组织、细胞的参照图谱。
- ❖ 人的各类组织、器官、细胞乃至各种细胞器已被广泛研究。
- ❖ 人的各种体液（血液、淋巴、脊髓、乳汁和尿等）都被用于研究与某些疾病的关系。
- ❖ 澳大利亚科学家利用双向电泳技术研究眼泪中的蛋白质与生理状态的关系。
- ❖ 发现一种新的蛋白质，它非常相似于在乳腺癌细胞里高表达的另一蛋白质。
- ❖ 这一发现可能会提供疾病诊治的新手段。
- ❖ 利用蛋白质组研究技术进行的酒精对人体毒性的研究中发现，乙醇会改变血清蛋白糖基化作用，导致许多糖蛋白的糖基缺乏，如转铁蛋白。
- ❖ 人类蛋白质组学的其他研究还有：单个疾病相关蛋白的比较、疾病相关蛋白的整体研究、肿瘤发生过程的研究等

1. 单个疾病相关蛋白质的寻找

- ❖ 在疾病发生过程中，由于和疾病相关的遗传信息的变化常常会导致蛋白的种类和数量发生变化，这些变化可通过高解析度的凝胶电泳检测到，这就是利用蛋白质组学寻找和鉴定疾病相关蛋白的**依据**
- ❖ 目前已用此方法得到一些肿瘤的标志蛋白

2. 肿瘤

- ❖ 肿瘤至今仍是人类的一个顽敌，癌细胞不仅有许多特异蛋白质产物，而且这些产物还会影响其他蛋白质的翻译后加工及许多蛋白质的表达水平。
- ❖ 通过各种肿瘤组织与正常组织间的蛋白质组差异的研究，已找到了一些肿瘤标志物，可能对揭示肿瘤的发生机制有帮助
- ❖ 肿瘤的蛋白质组研究能提供一些以往其他技术所不能提供的早期诊断依据
- ❖ 利用不同细胞组织的特征蛋白质谱，可以判别一些难以判定来源的肿瘤细胞的来源。

例1--膀胱鳞状细胞癌

- ❖ 丹麦的Celis等利用蛋白质组研究技术结合组织冰冻切片技术分析了150例膀胱癌病人的组织
- ❖ 利用数据库识别出几个膀胱鳞状细胞癌特异的蛋白质，并被用来制作抗体识别转移损伤
- ❖ 发现在所有的鳞状细胞癌的尿液中均能发现一种化学引诱剂 - 银屑素 (psoriasin)
- ❖ 推测其极可能作为这种肿瘤的早期标志物，提供有用的、非创伤性的方法进行疾病监控。

例2—结肠癌

- ❖ 结肠癌的产生是一个包括了多个基因突变的多步过程，其中包括抑癌基因的功能、癌基因的活化等。但肿瘤发生的具体机制仍不清楚。
- ❖ 对于这样一种涉及多种蛋白的疾病，人们已经开始利用蛋白质组学来分析结肠粘膜发生恶性转化后的多态变化
- ❖ 比较15例结肠癌病人和13例正常人的接长表皮的双向电泳发现：两者分别含有882个和861个点，其中1个分子质量为13ku，等电点5.6的蛋白只在肿瘤组织中专一表达。
- ❖ 15例病人中13例的此蛋白表达上调占到87%

- ❖ 不同程度癌症引起的发育异常中，该蛋白质也有明显的表达差异
- ❖ 该蛋白质是什么？
- ❖ 从双向电泳凝胶上获得该蛋白，进行胰蛋白酶水解，得到的肽片段用 μ -HPLC分离后测序。
- ❖ 测序得到两个序列：LGHPDTLNQ和VI EHMEDLDTANADK。与钙粒蛋白B吻合
- ❖ 进一步用MALDI-MS分析证实该蛋白就是钙粒蛋白B。
- ❖ 结合以前的发现，即由钙粒蛋白B和A组成的异二聚体蛋白钙防卫蛋白在胃肠肿瘤病人的粪便样品中含量有很大提高，钙粒蛋白B在肿瘤性转化的组织中的高转移性存在显示出它在结肠癌的产生中具有重要作用

例3---肝细胞癌

- ❖ 双向电泳已被成功地用于发现化学诱导的鼠的肝癌相关蛋白
- ❖ 双向电泳和蛋白质化学方法的联合应用，深化了对癌症相关蛋白的具体特征的认识。
- ❖ N-甲基-N-亚硝基脲诱导产生鼠的肝癌，
- ❖ 双向电泳分析表达有变化的蛋白
- ❖ 对变化蛋白基因氨基酸序列分析
- ❖ 得到一个蛋白是来源于肝癌的醛糖还原酶样蛋白(hepatoma-derived aldose reductase-like protein)。
- ❖ 该蛋白分子质量35ku, pI=7.4。是一种在肝癌和胚胎的肝中特异性表达的蛋白。
- ❖ 用直接针对来源于肝癌的醛糖还原酶样蛋白的抗体进行免疫组化分析，发现该蛋白在化学诱导的肝癌小鼠的发生肿瘤转化的前期和转化早期就已经有很强的表达了，而在正常组织中无表达。
- ❖ 说明该蛋白涉及肿瘤发生过程

例4—醛糖还原酶

- ❖ 属于还原酶超家族
- ❖ 在山梨糖醇途径中催化葡萄糖向山梨糖醇的转化
- ❖ 在一些糖尿病的并发症的发生中也有作用。
- ❖ 作为一种酶，可以水解一些生物异源物质等，故参与了一些解毒过程。
- ❖ 肝癌发生中，一些解毒酶表达水平或活力增高是一公认的事实
- ❖ 对于醛糖还原酶这一类有解毒功能的蛋白来说，只有由双向电泳发现的肝癌来源的醛糖还原酶样蛋白是与肝癌相关的。
- ❖ 它首先在胚胎肝中表达，但在成年的肝中就不表达了，肝癌发生时又重新表达。
- ❖ 人肝癌中也发现鼠的醛糖还原酶样蛋白的同源蛋白，同样在人的不同组织中选择性表达

例5 - 乳腺癌

- ❖ 过去几年有一些研究组已经描述了乳腺癌的蛋白质组研究
- ❖ 在正常腔和肌上皮乳腺癌细胞中分离超过1700个蛋白点，他们在两种类型中差异表达
- ❖ 该两细胞类型的分离是此研究的一个重点，因为约95%的乳腺肿瘤是腔细胞起源的
- ❖ 本研究下一目标：确定肿瘤细胞的蛋白质组以及鉴别其与健康腔细胞蛋白质组的区别。

一般步骤

- ❖ 利用双向电泳得到一种可能与癌症相关的蛋白质
- ❖ 利用蛋白质化学的方法研究该蛋白和疾病的相关性

3. 疾病相关蛋白的整体研究

- ❖ 大多数疾病造成的往往不只一个或几个蛋白的变化，参与疾病过程的蛋白的数目也是很大的，因此出了通过双向电泳寻找与疾病相关的单个蛋白外，通过蛋白质组对表达情况有变化的蛋白在整体水平上的研究同样是非常重要的。

例1—扩张性心肌病(Dilated cardiomyopathy, DCM)

- ❖ 扩张性的心肌病是一种严重的心脏疾病，该病的致病机理和涉及的分子都还不清楚，且该复杂疾病

也不可能仅有一种致病机理造成。----（整体蛋白质组水平研究的必要性）

- ❖ 相对其他组织而言，主要是由心肌细胞组成的心脏是一种相对均一的组织，为双向电泳研究提供了良好的基础---（可行性）
- ❖ 对DCM的蛋白质组研究起始于20世纪90年代初
- ❖ 目前已经建立了心肌的双向凝胶电泳数据库
- ❖ 虽然国际上不同实验室间存在样品制备、不同等电聚焦条件、不同凝胶大小等差异，但数据的标角证明，大多数条件下，不同蛋白的点的位置相对稳定，可以进行大规模比较研究。
- ❖ Knecht等研究得到一高解析度的具有约3300个心肌蛋白点的双向电泳结果，
- ❖ 对其中150个蛋白点进行了氨基酸分析、N端和中间的Edman降解以及MALDI-MS等一系列鉴定。
- ❖ 对几百个正常和扩张性心肌病的病人的2DE结果比较发现：两者的条带具有可比性。
- ❖ 除一些可能由不同的疾病有关参数如患病程度、用药情况、病人年龄等因素造成的无重复性的点的多少和强度的变化外，患者和正常人有25个蛋白在统计上具有显著差异，这些即DCM相关蛋白。
- ❖ 对这几十种疾病相关蛋白，可以用一些其他方法，如免疫组化、酶活测定等作进一步鉴定，确认它们与疾病的相关性以及它们在疾病中的作用等

例2-精神分裂症

- ❖ 新西兰的奥克兰大学构建了人海马蛋白质组图谱
- ❖ 识别出精神分裂症的18个异常表达的蛋白质，其中一些定位在6号染色体的同一区域
- ❖ 这些蛋白在此疾病的发病原理中的作用目前正被研究中

4. 肿瘤发生过程的研究

- ❖ 许多科学家致力于各种肿瘤组织与正常组织之间蛋白质谱差异的研究，这些组织包括脑、胸腺、乳腺、肺、直肠、肾、膀胱、卵巢、骨髓等
- ❖ 这些研究已经找到了一些肿瘤特异的标志物，可能会对揭示肿瘤发生的机制有所帮助。

例1 - 肾癌

- ❖ 对肾癌的研究发现4种蛋白质存在于正常肾组织，而肾癌细胞中缺失。
- ❖ 其中两种分别是辅酶Q蛋白色素还原酶和线粒体泛醌氧化-还原复合物I。
- ❖ 提示线粒体功能低下可能在肿瘤发生过程中起重要作用。

版权声明:

本站几乎所有资源均搜集于网络, 仅供学习参考, 不得进行任何商业用途, 否则产生的一切后果将由使用者本人承担! 本站仅提供一个观摩学习与交流的平台, 将不保证所提供资源的完整性, 也不对任何资源负法律责任。所有资源请在下载后 24 小时内删除。如果您觉得满意, 请购买正版, 以便更好支持您所喜欢的软件或书籍!



☆☆☆☆☆生物秀[\[http://www.bbioo.com\]](http://www.bbioo.com)

☆☆☆☆☆中国生物科学论坛[\[http://www.bbioo.com/bbs/\]](http://www.bbioo.com/bbs/)

☆☆☆☆☆生物秀下载频道[\[http://www.bbioo.com/Soft/\]](http://www.bbioo.com/Soft/)

生物秀——倾力打造最大最专业的生物资源下载平台!

■■■■ 选择生物秀, 我秀我精彩!! ■■■■

欢迎到生物秀论坛(中国生物科学论坛)的相关资源、软件版块参与讨论, 共享您的资源, 获取更多资源或帮助。