

临床蛋白质组学 ——蛋白质组学在临床研究中的应用*

李 峰³⁾ 杨 芳³⁾ 肖志强^{1,2)} 陈主初^{1,2,3)**}

(¹⁾中南大学湘雅医院, 长沙 410008; ²⁾卫生部肿瘤蛋白质组学重点实验室, 长沙 410008;

³⁾中南大学湘雅医学院肿瘤研究所, 长沙 410078)

摘要 临床蛋白质组学是将蛋白质组学技术应用于临床医学研究, 它主要围绕疾病的预防、早期诊断和治疗等方面开展研究, 其中, 恶性肿瘤是临床蛋白质组学研究的一个重点研究对象. 由于肿瘤生物标志物对早期诊断具有重要价值, 所以临床蛋白质组学的主要目标之一是寻找合适的肿瘤生物标志物, 多分子生物标志物已成为寻找肿瘤生物标志物的一个研究趋势. 简要介绍了临床蛋白质组学的基本概念, 实验设计, 临床样本收集与预处理以及蛋白质组学技术在临床研究中的应用与进展.

关键词 蛋白质组学, 临床蛋白质组学, 肿瘤生物标志物

学科分类号 R73

蛋白质组 (Proteome) 与蛋白质组学 (Proteomics) 的提出至今已有 10 年左右的时间, 蛋白质组的经典内涵是指一个细胞、组织或机体的基因组所表达的所有蛋白质, 蛋白质组学则是在蛋白质组整体水平进行研究的一门新兴学科^[1~4]. 从诞生开始, 蛋白质组学研究就有两种互为补充的研究策略. 一种是表达蛋白质组学 (expression profile proteomics) 研究策略, 即尽可能地分析所有蛋白质^[5,6]. 从逻辑意义上来讲, 完整的蛋白质组可能永远无法获得, 因此采用表达蛋白质组学研究策略的科研人员更多地关注于生命科学的基础研究方面; 另一种方法是功能蛋白质组学 (functional proteomics) 研究策略, 即将疾病的整个过程视为一个黑箱而只着眼于与疾病相关的差异表达蛋白质^[7,8]. 临床蛋白质组学 (clinical proteomics) 正是运用这一研究策略的典型代表.

1 概 述

临床蛋白质组学是蛋白质组学新近出现的一个分支学科, 它侧重于蛋白质组学技术在临床医学领域的应用研究, 并围绕着疾病的预防、早期诊断和治疗等方面开展研究, 主要包括以下几个方面: 疾病动物模型的蛋白质组学研究、寻找疾病的生物标

志物和药物治疗靶点开发^[9,10]. 由于临床蛋白质组学具有广阔的应用前景而倍受科研机构、制药公司的关注^[11], 比如美国国家癌症研究所和国家食品药品监督管理局就联合发起临床蛋白质组学研究计划 (FDA-NCI Clinical Proteomics Program) 并投入巨资进行研究.

临床疾病中, 恶性肿瘤由于其发病率不断提高, 对社会危害大而成为临床蛋白质组学研究的一个重点研究对象. 肿瘤通常被认为是遗传性疾病, 是由一系列癌基因的激活与抑癌基因失活而导致的. 但从基因产物的角度考虑, 肿瘤是一种蛋白质组学疾病, 正是由于一组蛋白质发生改变导致整个信号通路出现异常而最终导致肿瘤的发生. 因此, 肿瘤既是临床蛋白质组学研究的良好疾病模型, 同时也是临床蛋白质组学的一个热点研究领域. 其中, 寻找肿瘤早期检测的生物标志物、开发新的肿瘤药

*国家重点基础研究发展计划(973)(2001CB510208), 国家自然科学基金(30240056, 30370642), 教育部跨世纪优秀人才培养计划基金(教技函[2002]48)资助项目, 湖南省重点科研项目(025SY2001-1), 湖南省卫生厅重点科研项目(Z02-4).

** 通讯联系人. Tel: 0731-4805447, Fax: 0731-4805482

E-mail: tcbl@xyism.net

收稿日期: 2005-08-12, 接受日期: 2005-09-30

物治疗靶点, 具有重大的实用价值和经济效益, 成为临床蛋白质组学的一个主攻方向^[12, 13].

肿瘤生物标志物 (biomarker) 对早期诊断具有重要价值, 一直以来是临床研究的重点领域. 受以往少数几种肿瘤中较为成功的标志物影响, 如肝癌中的甲胎蛋白 (Alpha-fetoprotein, AFP)、前列腺癌中的前列腺特异抗原 (prostate specific antigen, PSA), 人们的兴趣总是集中在单一肿瘤标志物上, 试图寻找单一的过表达、低表达蛋白质或疾病过程中释放入血液的某种蛋白质. 由于肿瘤是一种蛋白质组学疾病, 单一肿瘤标志物的改变几乎不可能特征性地反映肿瘤的整体变化特征. 事实上, 多年来寻找单一肿瘤标志物进展缓慢也说明单一肿瘤标志物对于绝大多数肿瘤无法作为有效的判断依据. 因此, 多分子标志物 (multi-marker) 成为寻找肿瘤生物标志物的一个研究趋势, 其实质是采用多个肿瘤标志物联用并据此来反映肿瘤的整体特征, 显然这一方法优于单一肿瘤标志物. 有研究表明, 采用新发现的肿瘤标志物与已有的单一肿瘤标志物联用所形成的多分子标志物对于肿瘤的判别能力明显提高. 比如 Rai 等^[14]发现了 3 种卵巢癌的候选标志物, 当它们分别作为单一 biomarker 时, 对卵巢癌鉴别能力不如 CA125, 但当它们与 CA125 联用则可以显著提高鉴别能力. 其中两种与 CA125 联用时敏感性可以达到 94%, 而单独使用 CA125 仅为 81%. Zhang 等^[15]发现了 3 个区分卵巢癌和非癌的标志物, 这 3 个 biomarker 与 CA125 在联合检测卵巢癌时的敏感性 (74%) 高于单一的 CA125 (65%). 所以, 多分子标志物的筛选成为肿瘤标志物寻找过程中的一个新方向.

2 实验设计与样本收集

临床蛋白质组学研究的对象即实验的样本材料直接来自于临床的医学实践, 包括组织标本与体液 (其中主要是血液). 由于存在个体之间的差异以及随疾病变化引起的个体自身差异, 同时临床样本难以像实验室研究材料 (比如培养细胞) 那样严格控制实验条件, 所以临床蛋白质组学必须进行精心的实验设计并仔细考虑样本收集和处理等诸多过程, 尽可能减少由干扰因素所引起的差异表达蛋白质. 尤其是血液样本, 由于血液容易受外界因素影响而引发蛋白质的波动, 因而特别需要注意血液的样本存储、处理和实验操作等过程. 从差异表达蛋白质的干扰来源进行分析, 可分为以下 3 种: 分析前的

差异、分析中的差异和生物学的差异. 其中, 生物学的差异包括个体间差异和个体内差异, 个体间差异来源于年龄、性别和人种等所造成的差异, 而个体内差异随个体在不同时间、状态 (如禁食) 或者不同年龄所出现的差异. 分析中的差异是指在蛋白质组学研究产生的大量数据中, 必然包含着仪器设备本身的变异而带入的实验误差. 分析前的差异则是由于样本在实验之前的处理过程所产生的差异, 比如样本在收集、储存等过程中由于条件不同而引入的差异^[16]. 尽管这些差异可能与疾病无关, 但确实增加了寻找疾病相关差异蛋白的复杂性, 并可能影响蛋白质组分析的结果. 为此, 2002 年人类蛋白质组组织 (HUPO) 专门成立了一个标本委员会 (Specimen Committee) 来研究和解决这些问题^[17].

要区分甄别蛋白质的变化是否为疾病发生发展过程中产生的差异蛋白质, 除了需要良好的实验设计, 还需仔细考虑不同蛋白质组研究技术的优势和缺陷, 利用正确的统计学方法, 采用自动化机械装置减少实验误差等措施. 对于样本而言, 必须有足够多的样本收集来源, 并且含有尽量多的各种疾病亚型, 这样才能最好地代表疾病的整体情况. 考虑到原始数据中因样本来源不同而可能存在疾病亚型之间的差异, 最好对样本群 (sample cohort) 进行合理分配, 以多中心方式进行工作并相互验证^[18]. 例如, 由一样本群数据中发现的候选肿瘤生物标志物可以用另一样本群数据进行横向比较, 并利用这一样本群进行验证. 这种工作模式模仿了多中心验证程序 (multicenter validation process), 可以更为准确地考察候选肿瘤生物标志物, 同时这也是所有肿瘤生物标志物在进入临床应用前所必需遵守的循证原则.

3 样本预处理

由于临床标本的细胞类型、组织结构或蛋白质组成成分十分复杂, 在进行实验之前采用合理的样本预处理技术能有效地降低样本的复杂性, 便于后续的实验分析工作. 针对组织和体液两大类样本类型, 分别出现了不同的样本预处理技术. 对于组织样本, 激光捕获显微切割技术 (laser capture microdissection, LCM) 是一种在显微镜下用激光对特定类型的细胞 (比如癌巢) 进行切割分离的新技术, 它能从含有多种细胞类型的组织中有效分离纯化出单一的细胞类型, 从而解决了临床蛋白质组学组织样本的纯度问题^[19]. 因此, LCM 技术预处理后

的组织样本用于蛋白质组学研究，能获得更为可靠的结果. 对于血液样本，其信息含量十分丰富、蛋白质组成成分也异常复杂并且高、低丰度蛋白质含量的动态范围非常宽(可达到 10^{10} 数量级). 在如此复杂的血液蛋白质组中，为了发现低丰度蛋白质中的候选 biomarker，必须采用合理步骤尽量去除白蛋白和免疫球蛋白等高丰度蛋白质，这样可以实现富集低丰度蛋白质的目的. 但同时也要尽量避免低丰度蛋白质的丢失，必须注意的是部分短肽作为候选 biomarker 有可能与高丰度蛋白质如白蛋白结合，因而需要避免因白蛋白等的去除而抛弃这些短肽^[20].

4 研究技术与应用

从蛋白质组学研究的技术流程来看，可分为蛋白质分离技术、蛋白质鉴定技术和蛋白质组信息学三大部分. 就蛋白质的分离而言，又可分为基于凝胶系统的分离技术和基于非胶系统的分离技术. 至于先用凝胶粗分再辅以非胶技术细分的“杂合”分离技术，因其主要分离部分与瓶颈在于凝胶系统，因此大多将其归于基于凝胶系统的分离技术. 在蛋白质鉴定技术方面，自 20 世纪 90 年代逐渐发展起来的生物质谱 (mass spectrometry, MS) 技术已扮演起主要的角色，其优点在于 MS 技术的高通量、高灵敏度和高准确性. 蛋白质组信息学是蛋白质组学研究的又一个重要支撑，其研究与应用已深入到蛋白质组学的各个领域，包括获取各种蛋白质组学分离、鉴定技术得到的实验结果并将之输入到计算机中，再对其进行存储、加工、分析、处理，最终得到其中含有生物学意义的信息.

基于凝胶系统的蛋白质组学技术中，传统的双向凝胶电泳 (two-dimensional electrophoresis, 2-DE) 技术依然是临床蛋白质组学研究中寻找 biomarker 的一种重要工具. Celis 等用 2-DE/MS 技术发现膀胱鳞状细胞癌与正常对照组织之间存在差异表达蛋白质，如 Keratin10, 14, 16、MRP8, 14 和 Psoriasin (银屑素)，其中银屑素还存在于膀胱鳞状细胞癌病人的尿液中而无法在正常人的尿液中检出. 从而将银屑素作为膀胱鳞状细胞癌的生物标志物，用于膀胱鳞状细胞癌的早期诊断、疾病鉴别和病情监控^[21]. Zoom 胶、极窄胶以及差异凝胶电泳技术 (differential in gel electrophoresis, DIGE) 等一系列新技术有效提高了 2-DE 技术的分辨率与灵敏度，使得低丰度的差异蛋白质更容易被发现，并在多种

肿瘤中得到应用. 如 Zhou 等^[22]用 LCM/DIGE/MS 技术研究食管癌组织的差异表达蛋白质，发现食管癌中上调大于 3 倍的蛋白质点 58 个、下调大于 3 倍的蛋白质点 107 个，其中包含有已确定的下调蛋白质 annexin I 和上调蛋白质 gp96.

血清蛋白质组学分析方法 (serum proteomics analysis, SERPA) 是 2-DE 和免疫印迹技术相结合的一种蛋白质组分析方法. 从肿瘤免疫学的观点看，肿瘤细胞是“非己细胞”，它或多或少表达区别于正常细胞的肿瘤抗原，因而在患者血清中极有可能存在其相应的抗体，所以可分别采用肿瘤病人与正常人的血清作为一抗筛选肿瘤组织中的差异蛋白质即肿瘤抗原^[23]. 由于 SERPA 技术可结合蛋白质组学能分离肿瘤细胞内成千上万个蛋白质的优势和抗原抗体反应高度特异性的特点，从而能够快速筛选出与肿瘤密切相关的蛋白质，进而识别出肿瘤的分子标志物. 目前，国外已有不少研究采用 SERPA 技术对包括肺癌、乳腺癌、神经母细胞瘤、肾细胞癌、肝癌等多种肿瘤进行了研究，并寻找到了一些潜在的 biomarker (表 1)^[24].

Table 1 Tumor specific antigen identified by SERPA
表 1 SERPA 技术发现的特异肿瘤抗原

肿瘤	特异肿瘤抗原
神经母细胞瘤	β-tubulin-1, 3 异构体
肺癌	PGP 9.5, UCHL1, annexins
乳腺癌	3种 RNA 结合蛋白调节亚单位异构体
肝癌	Calreticulin异构体等 8 种蛋白质
肾细胞癌	Hsp27, Hsp60, Hsp96, gp96, vimentin 等

基于非胶系统的蛋白质组学技术，常见的有液相分离技术、蛋白质芯片技术、亲和质谱技术及同位素亲和标签技术等，这些技术能克服与弥补 2-DE/MS 技术的一些缺陷与不足. 表面增强激光解析飞行时间质谱 (surface-enhanced laser desorption/ionization time of flight-MS, SELDI-TOF-MS) 是亲和质谱技术的一种，同时也是一种广义上的芯片^[25]. SELDI 芯片表面为诸如离子交换树脂酯之类的色谱介质，具有某种特殊的亲和力，能捕获具有相关理化性状的特定蛋白质，然后用 TOF-MS 对滞留蛋白质进行分析. 与 2-DE 技术相比，SELDI-MS 具有快速、高通量、高灵敏度、适合于低分子量蛋白质 (2~20 ku) 检测、样本用量少以及容易用于临床等优点. Petricoin 等^[26]采用

SELDI-TOF-MS技术对卵巢癌与正常人的血清样本进行分析,高通量获取血清蛋白质组指纹后通过计算机辅助分析,发现由5个特征性质谱峰所组成的卵巢癌血清诊断模式对卵巢癌诊断的阳性率达到95%,准确率达到100%。

蛋白质微阵列 (protein microarray) 技术是一种新型蛋白质分析技术,包括(正相)蛋白质微阵列和反相蛋白质微阵列两种,具有集成化、高通量化、微量化和自动化等分析优势.它可以直接分析微量的生理或生物样品,如血清、细胞裂解液等,同时检测、识别和纯化不同的生物分子和研究分子间相互作用,是研究信号转导通路中分子靶的有力工具. Paweletz 等^[27]用反相蛋白质微阵列研究了LCM获得的前列腺正常上皮细胞、癌前病变细胞和侵袭性前列腺癌细胞,结果发现AKT磷酸化是肿瘤发生过程中的一个早期关键事件,所以抑制AKT分子靶的活性对于前列腺癌的预防和治疗具有重要作用.另外,蛋白质微阵列技术在肿瘤标志物抗体芯片、高通量药物筛选、疫苗靶分子的鉴定、实时疗效评估等方面也具有有良好的应用前景。

5 结语与展望

临床蛋白质组学是一个非常年轻而活跃的领域,因其直接面向临床应用而有着良好的开发应用前景.可以预见,随着肿瘤标志物研究的深入,今后不再是依靠单一的蛋白质分子作为肿瘤标志物而是通过一组或一群蛋白质协同作用共同作为肿瘤的生物标志物.血清中低分子质量蛋白质因其能反映血清中蛋白质的全貌而必将成为挖掘肿瘤分子标志物的一座金矿.对蛋白质-蛋白质相互作用网络理解的深入,更多的分子靶将会被发现,这将为个体化治疗提供坚实的理论基础,并且针对肿瘤多个分子靶的药物联合运用将在提高疗效的同时显著减低药物的毒副作用.随着多中心工作模式的展开,将会建立起一批标准化的临床蛋白质组样本库和临床蛋白质组研究网络。

参 考 文 献

- Kahn P. From genome to proteome: looking at a cell's proteins. *Science*, 1995, **270** (5235): 369~370
- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*, 2003, **422** (6928): 198~207
- Pandey A, Mann M. Proteomics to study genes and genomes. *Nature*, 2000, **405** (6788): 837~846
- Hochstrasser D F. Proteome in perspective. *Clin Chem Lab Med*, 1998, **36** (11): 825~836
- Washburn M P, Wolters D, Yates J R, 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol*, 2001, **19** (3): 242~247
- Wu C C, MacCoss M J, Howell K E, *et al.* A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol*, 2003, **21** (5): 532~538
- Hanash S. Disease proteomics. *Nature*, 2003, **422** (6928): 226~232
- Kraemer K H. From proteomics to disease. *Nat Genet*, 2004, **36** (7): 677~678
- Petricoin E F, Zoon K C, Kohn E C, *et al.* Clinical proteomics: translating benchside promise into bedside reality. *Nat Rev Drug Discov*, 2002, **1** (9): 683~695
- Petricoin E F, Liotta L A. Proteomic analysis at the bedside: early detection of cancer. *Trends Biotechnol*, 2002, **20** (12 Suppl): S30~34
- Hanash S. HUPO initiatives relevant to clinical proteomics. *Mol Cell Proteomics*, 2004, **3** (4): 298~301
- Johann D J, Jr., McGuigan M D, Patel A R, *et al.* Clinical proteomics and biomarker discovery. *Ann N Y Acad Sci*, 2004, **1022**: 295~305
- Wulfkuehl J D, Liotta L A, Petricoin E F. Proteomic applications for the early detection of cancer. *Nat Rev Cancer*, 2003, **3** (4): 267~275
- Rai A J, Zhang Z, Rosenzweig J, *et al.* Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med*, 2002, **126** (12): 1518~1526
- Zhang Z, Bast R C, Jr., Yu Y, *et al.* Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res*, 2004, **64** (16): 5882~5890
- Colantonio D A, Chan D W. The clinical application of proteomics. *Clin Chim Acta*, 2005, **357** (2): 151~158
- Rai A J, Gelfand C A, Haywood B C, *et al.* HUPO Plasma Proteome Project specimen collection and handling: Towards the standardization of parameters for plasma proteome samples. *Proteomics*, 2005, **5** (13): 3262~3277
- White C N, Chan D W, Zhang Z. Bioinformatics strategies for proteomic profiling. *Clin Biochem*, 2004, **37** (7): 636~641
- Emmert-Buck M R, Bonner R F, Smith P D, *et al.* Laser capture microdissection. *Science*, 1996, **274** (5289): 998~1001
- Liotta L A, Ferrari M, Petricoin E. Clinical proteomics: written in blood. *Nature*, 2003, **425** (6961): 905
- Ostergaard M, Rasmussen H H, Nielsen H V, *et al.* Proteome profiling of bladder squamous cell carcinomas: identification of markers that define their degree of differentiation. *Cancer Res*, 1997, **57** (18): 4111~4117
- Zhou G, Li H M, DeCamp D, *et al.* 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Molecular Cellular Proteomics*, 2002, **1** (2): 117~124
- Le Naour F. Contribution of proteomics to tumor immunology. *Proteomics*, 2001, **1** (10): 1295~1302
- Seliger B, Kellner R. Design of proteome-based studies in

- combination with serology for the identification of biomarkers and novel targets. *Proteomics*, 2002, **2** (12): 1641~1651
- 25 Merchant M, Weinberger S R. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis*, 2000, **21** (6): 1164~1177
- 26 Petricoin E F, Ardekani A M, Hitt B A, *et al.* Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 2002, **359** (9306): 572~577
- 27 Paweletz C P, Charboneau L, Bichsel V E, *et al.* Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene*, 2001, **20** (16): 1981~1989

Clinical Proteomics: The Application of Proteomics in The Research of Clinical Medicine*

LI Feng³, Yang Fang³, XIAO Zhi-Qiang^{1,2}, CHEN Zhu-Chu^{1,2,3)**}

⁽¹⁾Xiangya Hospital, Central South University, Changsha 410008, China;

⁽²⁾Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Changsha 410008, China;

⁽³⁾Cancer Research Institute, Xiangya School of Medicine, Central South University, Changsha 410078, China)

Abstract Clinical proteomics is the application of proteomics in clinical medicine, which focuses on the prevention, diagnosis and therapy of disease. Cancer has been the primary research object in clinical proteomics. Because biomarker is invaluable in early diagnosis of cancer, one of the central goals in clinical proteomics is to discover appropriate biomarker and the interest shifts to looking for multi-biomarkers combined. Study design, sample collection and pre-process in clinical proteomics and the application and progress of clinical proteomics technology have been introduced.

Key words proteomics, clinical proteomics, cancer biomarker

*This work was supported by grants from National Basic Research Program of China (2001CB510208), The National Natural Science Foundation of China (30370642, 30240056), Ministry of Education of China for Outstanding Scholars of New Era (2002-48), Key Research Program from Science and Technology Committee of Hunan (02SSY2001-1) and Key Research Program from Public Health Bureau of Hunan Province (Z02-04).

**Corresponding author. Tel: 86-731-4805447, Fax: 86-731-4805482, E-mail: tcbl@xysm.net

Received: August 12, 2005 Accepted: September 30, 2005