

[文章编号] 1007-385X(2003)01-0060-03

蛋白质组学研究的主要策略及其在肿瘤研究中的应用

夏 晴 综述; 沈倍奋 审阅 (军事医学科学院基础医学研究所分子免疫室, 北京 100850)

[摘要] 肿瘤的发生涉及一系列复杂的分子事件,蛋白质组学研究手段可以大规模地定量分析细胞内的蛋白质表达水平、翻译后修饰等性质以及定义信号网络中的蛋白质间相互作用,从而有望发现控制肿瘤进程的关键分子,为肿瘤的诊断、分型、药物研制带来新的思路和途径。蛋白质组学为肿瘤的研究提供了新的平台。本文综述了目前蛋白质组学研究中的主要策略和技术体系及其应用于肿瘤研究中取得的一些进展。

[关键词] 蛋白质组学; 肿瘤; 信号转导

[中图分类号] Q51; R73 **[文献标识码]** A

* 蛋白质组学的核心在于大规模地对蛋白质进行综合地分析,通过对某种物种、个体、器官、组织或细胞的全部蛋白质性质(包括表达水平、结构、翻译后修饰、细胞内定位、蛋白质相互作用等)的研究,对蛋白质所执行的生理性、病理性生命活动作出最精细、最准确、最本质的阐述^[1]。几年间,蛋白质组学得到了迅猛发展,已广泛地应用于医药卫生、生物学及农业等研究的各个领域,其概念也不断扩展,包含了许多以往被称作功能基因组学的内容。

肿瘤是一种涉及到多种分子事件的疾病,目前研究的热点包括鉴定新的肿瘤标志物以提高早期诊断及肿瘤分型诊断水平;筛选更有效的治疗药靶以及病患的个体化治疗等。蛋白质组学为肿瘤的研究提供了新的平台,可以实现同时大规模地对细胞内的蛋白质表达进行定性、定量分析,从整体网络的水平了解细胞内信号转导通路,从而有望发现控制肿瘤进程的关键分子,为肿瘤的诊断、分型、药物研制带来新的思路和途径^[2]。本文综述了近年来蛋白质组学研究的主要策略和技术体系及其应用于肿瘤研究中取得的一些进展。

1 蛋白质组学的主要研究策略和技术体系

目前,蛋白质组学研究包括2种相互补充的策略。一种策略是表达蛋白质组学(protein expression proteomics),观察某种细胞或组织中大量蛋白质的整体表达,并分析在不同情况下表达图谱的变化;另一种策略是细胞图谱蛋白质组学(cell-mapping proteomics),旨在定义蛋白质间的相互作用,以建立细胞内信号转导通路的复杂网络图^[3]。

1.1 表达蛋白质组学

表达蛋白质组学是现今使用的最为广泛的蛋白质组学研究模式。目前该策略的技术路线主要是通过双向凝胶电泳(two dimensional electrophoresis, 2DE)建立蛋白质组图谱、用专业图像扫描分析软件进行图像分析、通过质谱(mass spectrometry, MS)及蛋白质数据信息处理技术进行蛋白质的鉴定与分析。2DE是1975年由O'Farrell等人发明,其原理是第一向基于蛋白质等电点的不同,通过等电聚焦进行分

离,再根据分子量的不同,在与第一向垂直的方向通过聚丙烯酰胺凝胶电泳进行分离。近年来,商品化的预制固定化梯度胶条的使用(目前已有可低至1个pH单位的窄范围梯度胶条)以及蛋白质提取技术的进步和非荧光及荧光等显示方法的发展,使操作更为方便,并且提高了双向电泳的上样量、重复性、分辨率、低丰度蛋白质的检测能力和定量的准确性,使该分离方法成为目前蛋白质组学研究中最具实用价值的核心技术之一。但目前该系统还面临着一些方法学上的局限,如:低丰度蛋白质(拷贝数小于1 000)的检测;大(>100 kD)或小(<10 kD)分子量蛋白质的检测;极酸或极碱蛋白质的分离;高疏水性蛋白质的检测(其中包括一些重要的膜蛋白)等^[4,5]。为了获得更完整的蛋白质组表达图谱,研究者们也在积极发展别的非二维电泳的方法来获得图谱^[5]。目前主要有2种技术,一种是液相色谱和串联质谱联用(liquid chromatography and tandem mass spectrometry, LC-MS/MS)。该方法将复杂的蛋白质样品酶解消化后,通过一维或多维色谱预分离后,用串联质谱对多肽混合物进行鉴定分析。近年来出现了一种通过色谱联用进行定量蛋白质组学分析的方法,即稳定同位素标记内标的质谱定量分析法,应用同位素编码亲和标记(isotope coded affinity tags, ICAT)定量分析蛋白质混合物,通过质谱同时进行定性和定量的分析,目前已有商品化的ICAT试剂提供^[6]。近几年LC-MS/MS的发展非常迅猛,虽然目前该技术还不能替代2DE-MS在蛋白质组学研究中的核心位置,但由于LC-MS/MS能够弥补2DE-MS的上述缺陷,没有偏向性,能等同分离高丰度和低丰度的蛋白质,并且具有自动化程度高、重复性好的优点,所以展现了极大的应用前景。另一种近年来发展很快的快速筛检复杂蛋白质样品的方法是蛋白质芯片。蛋白质芯片是一种含有多种微量纯化多肽、抗体等蛋白质的微阵列,能够高通量地测定这些蛋白质的生物活性以及蛋白质与生物大分子间的相

[基金项目] 国家高技术研究发展计划资助项目(863计划)
(No. 2001AA233041)

互作用。由于蛋白质芯片具有高通量、微型化及自动化等优点,在蛋白组学研究中的应用越来越广泛^[7]。

1.2 细胞图谱蛋白质组学

细胞图谱蛋白质组学,也称作相互作用蛋白质组学(interaction proteomics)。目前研究蛋白质的相互作用有两种不同的方案。一种即“经典的”蛋白质组学手段,主要依赖于蛋白质复合物的亲和纯化,后续通过质谱鉴定它们的组成。具体来看,相互作用蛋白质的获得主要有以下几种途径:(1)采用抗某种感兴趣蛋白质的抗体进行经典的免疫沉淀实验;(2)将带有亲和纯化标签的某种感兴趣蛋白质的编码基因转入细胞内,将与该表达产物相互作用的蛋白质亲和纯化后加以鉴定。(3)大分子复合物的分离。另一种方案则通过表达某种感兴趣的基因组中的基因产物,建立一整套蛋白质,后者可用于蛋白质相互作用的体外检测(蛋白质芯片)或蛋白质相互作用的体内检测(酵母双杂交系统)。这些方法所基于的原理互不相同,只有同时联用这些方法才能最终发现某一特定细胞或组织中尽可能完全的蛋白质相互作用。

2 蛋白质组学策略在肿瘤研究中的应用

一个正常的细胞转化为恶性细胞的过程中,在蛋白质水平上发生着一系列变化,包括表达量、翻译后修饰(如磷酸化、糖基化、剪接)及细胞内易位、活性等,这些变化都可能影响细胞的功能。蛋白质组学研究为确定和理解这些变化提供了新的平台,从而可能发现可用于早期诊断的新的肿瘤标志物、肿瘤分子分型方案、治疗靶标以及个体化的治疗方案。

2.1 样品的制备

对于肿瘤蛋白质组学的研究而言,样品的制备是实验成功与否的关键。从活体内分离切割的组织或体外培养的细胞都是重要的样品来源。两者各有优点,互为补充,前者不受体外培养对实验结果的影响,而细胞系可分析特定的细胞亚组分、分泌蛋白和膜蛋白等。诸多因素,如组织经手术切除后至蛋白提取间隔时间长短、固定剂类型、包被介质、固定时间长短、样品处理温度等都会影响蛋白质组的完整性。因此,从组织的切取、靶细胞群的分离、培养细胞的收集到蛋白质的溶解,每一步都应小心谨慎。为确保实验结果只与肿瘤

本身的变化有关,应避免周围结缔组织、间质、血管组织、血细胞、坏死组织的污染。Page 等^[8]采用钙饥饿处理后免疫磁珠分离的方法,用于乳腺癌原代培养中与正常肌内皮细胞的分离。这种方法所得到的细胞纯度与荧光激活细胞分选法(fluorescent activated cell sorting, FACS)相当。目前最有前景的肿瘤细胞分选方法是激光捕获微切割技术(laser capture microdissection, LCM),该技术可在显微镜直视下精确分离组织切片中特定的细胞或细胞群体,回收的细胞可适用于各种检测分析^[9]。

2.2 表达蛋白质组学在肿瘤研究中的应用

表达蛋白质组学策略应用于肿瘤的研究,包括建立某种肿瘤组织或细胞(特定分期、分化程度、转移能力、药物敏感性等)及其相应的正常组织或细胞的蛋白质图谱、某种药物处理前后的比较图谱、特定基因的敲除或过量表达前后的比较图谱等^[3,10]。例如造成癌症患者高死亡率的主要原因之一是肿瘤的转移,利用蛋白质组学技术,通过对高低转移细胞株蛋白质图谱的比较来寻找与肿瘤转移相关的蛋白质,同样也可以以高转移株中特异表达的蛋白质为靶点,开发抑制肿瘤转移的新药。再如,现今大多数抗癌药物都伴有严重的毒副作用,同时对癌症患者进行单一化疗或联合放疗、过热疗法时,经常面临癌细胞对细胞抑制剂耐药性的问题,如果能发现与细胞毒性密切相关的蛋白质、在耐药细胞系中表达异常的蛋白质,就可能以此为靶点设计用于联合用药的新方案^[11]。目前在肿瘤蛋白质组学研究中常用的鉴定肿瘤标志物的方案有以下几种^[2,12]:一种方案是比较分析肿瘤组与对照组的组织或细胞全组分图谱;第二种方案是提取组织或细胞系的亚细胞组分、膜蛋白或分泌性蛋白质,建立差异图谱比较分析;另有一种方案是基于体液免疫反应,分析蛋白抗原及其对应血清或体液中的自身抗体,从中寻找新的肿瘤标志物。目前各国科学家们已对肝癌、肺癌、乳腺癌、白血病等多种肿瘤进行了蛋白质组水平的研究,建立了多种组织或细胞的双向电泳图谱、不同分期、不同特性、药物处理前后等等的差异图谱,一些实验室还将他们的研究结果建立了数据库通过互联网对外公布^[3,10,13-14]。(表 1)列出了部分肿瘤蛋白质组学研究的双向电泳数据库。

表 1 肿瘤蛋白质组双向电泳图谱数据库

肿瘤类型	组织或细胞系	数据库网址
膀胱癌	组织	http://biobase.dk/cgi-bin/celis
直肠癌	人结肠息肉、鼠结肠息肉、LIM1215、LIM1863	http://www.ludwig.edu.au/jpsl/jpslhome.html
	新鲜直肠癌组织	http://www.pmma.pmfhk.cz/2d/2d.html
乳腺癌	乳腺导管癌及正常组织	http://www.bio-mol.unisi.it/2d/2d.html
	MDA-MB231	http://www.ludwig.edu.au/jpsl/jpslhome.html
白血病	Jurkat 细胞系	http://www.mpiib-berlin.mpg.de/2D-PAGE/
	淋巴瘤	http://au.expasy.ch/ch2d/

一些学者致力于将色质联用的方法应用于肿瘤的研究。Simpson 等^[3]采用非多孔的反相高效液相色谱法(non-porous reversed-phase high-performance liquid chromatography, RP-HPLC)分离,在线连接电离离子喷雾质谱直接进行鉴定的方法,检测不同肿瘤进程中蛋白质表达的变化,鉴定了一些与乳腺癌或直肠癌相关的癌蛋白。另有学者采用两维液相色谱分离(即强离子交换和非多孔的反相高效液相色谱法分离),基质辅助激光解析电离飞行时间质谱鉴定的方法分析了人红白血病细胞内蛋白质。(表2)列出了近年来肿瘤蛋白质组学研究中,新鉴定的一些具有标志物潜力的肿瘤相关性蛋白质分子。

表2 肿瘤蛋白质组学研究中发现的具有肿瘤标志物意义的相关性蛋白质

肿瘤类型	肿瘤相关性蛋白质
肺癌	TAO1,TAO2
乳腺癌	P50. 9
膀胱癌	Psoriasin, Galectin7, Stratifin
直结肠癌	Calgranulin B
肾细胞癌	Ubiquinol, Cytochrome C reductase
卵巢癌	OP18, PHSP60, HSP90, Calreticulin, Tropomyosin-1, 2
前列腺癌	Oncoprotein18, Elongation factor-2, Glutathion S, Transferase P, Superoxide dismutase, Triosephosphatase isomerase, Cytokeratin 18
神经母细胞瘤	p19/nm23-H1

2.3 细胞图谱蛋白质组学在肿瘤研究中的应用

研究表明许多与肿瘤进程相关的基因突变所影响的都是编码信号通路相关蛋白质的基因,定义肿瘤形成中的信号网络具有重要的意义。细胞图谱蛋白质组学对蛋白质间相互作用的研究,可以阐明肿瘤细胞中信号转导、功能性调节的一些基本问题。目前在定义肿瘤进程中信号转导网络方面也已取得了一些进展。如 Pandey 等^[15]采用表皮生长因子、血小板衍生因子处理 Hela 细胞,用抗磷酸酪氨酸抗体进行免疫沉淀,同时分析信号转导过程中的一系列磷酸化的蛋白质,揭示了 Vav-2 及其它一些蛋白质在生长因子信号转导中的作用。Lewis 等^[16]选择性激活或抑制活化丝裂原激活蛋白激酶(mitogen-activated protein kinases, MAPK)通路,鉴定了 MAPK 通路中的 20 种新的转导蛋白。而以往对上述通路中关键信号转导分子的鉴定常常需要很多研究小组历经十余年的时间才能完成。

3 展望

生命科学已进入后基因组时代,这是一个整体细胞生物

学的时代,芯片技术、质谱技术、生物信息处理等技术的进步,基因水平、蛋白质水平研究平台的不断完善,可以实现对成千上万基因或蛋白质的同时分析。通过 DNA(基因组)、RNA(转录子组)和蛋白质(蛋白质组)3 个不同水平信息的整合,一定可以更深刻地揭示肿瘤发生的本质,从而发现临床诊断和治疗肿瘤的新的武器。

[参 考 文 献]

[1] Blackstock W, Mann M. A boundless future for proteomics[J]. Trends Biotechnol, 2001, 19(suppl): s1-s2.

[2] Hanash SH, Bobek MP, Rickman DS, *et al.* Integrating cancer genomics and proteomics in the post-genome era[J]. Proteomics, 2002, 2: 69-75.

[3] Simpson RJ, Dorow DS. Cancer proteomics: From signaling networks to tumor markers[J]. Trends Biotechnol, 2001, 19(suppl): s40-s47.

[4] Anderson NL, Anderson NG. Proteome and proteomics: New technologies, new concepts and new words[J]. Electrophoresis, 1998, 19(11): 1853-1861.

[5] Herbert BR, Harry JL, Packer NH, *et al.* What place for polyacrylamide in proteomics[J]. Trends Biotechnol, 2001, 19(suppl): s3-s9.

[6] Gygi SP, Rist B, Gerber SA, *et al.* Quantitative analysis of complex protein mixtures using isotope coded affinity tags[J]. Nat Biotechnol, 1999, 17: 994-999.

[7] Macbeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination[J]. Science, 2000, 278: 1760-1763.

[8] Page MJ. Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammaplasties[J]. Proc Natl Acad Sci USA, 1999, 96: 12589-12594.

[9] Bonner RF, Emmert-Buck M, Cole K, *et al.* Laser capture microdissection: Molecular analysis of tissue[J]. Science, 1997, 278: 1481-1483.

[10] Srinivas PR, Srivastava S, Hanash S, *et al.* Proteomics in early detection of cancer[J]. Clin Chem, 2001, 10(47): 1901-1911.

[11] Friedrich MJ. Genomics and proteomics may help clinicians individualize cancer treatment[J]. JAMA, 2002, 22(287): 2931-2932.

[12] Banks RE, Dunn MJ, Hochstrasser DF, *et al.* Proteomics: New perspectives, new biomedical opportunities[J]. Lancet 2000, 356: 1749-1756.

[13] Ostergaard M, Wolf H, Orntoft TF, *et al.* A purative urinary marker for the follow up patients with bladder squamous cell carcinomas [J]. Electrophoresis, 1999, 20: 2100-2110.

[14] Chuman Y, Bergman A, Ueno T, *et al.* Napsin A, a member of the aspartic protease family, is abundantly expressed in normal lung and kidney tissue and is expressed in lung adenocarcinomas[J]. FEBS Lett, 1999, 462: 129-134.

[15] Pandey A. Analysis of receptor signaling pathways by mass spectrometry: Identification of Vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors[J]. Proc Natl Acad Sci USA, 2000, 97: 179-184.

[16] Lewis TS. Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry[J]. Mol Cell, 2000, 6: 1343-1354