

## 定量蛋白质组学分析方法

钱小红

(蛋白质组学国家重点实验室, 北京蛋白质组研究中心, 军事医学科学院放射与辐射医学研究所, 北京 100850)



钱小红: 军事医学科学院放射与辐射医学研究所研究员, 博士生导师。国家蛋白质组学重点实验室副主任, 北京蛋白质组研究中心技术总监。国家“863”计划“蛋白质组技术与分子标志物研发”项目首席科学家, 中国蛋白质组组织(CNHUPO)副理事长, *Proteomics* 编委, 《分析化学》编委, 《色谱》编委, 《质谱学报》编委。发表研究论文 100 余篇, 主编、参编专著和译著 15 部, 培养博士后、博士研究生、硕士研究生 60 余名。

精确测量多个不同生理或病理条件下生物样本中蛋白质表达量的变化是定量蛋白质组学(quantitative proteomics)研究的重要内容。与传统的蛋白质定量方法(见表 1)相比, 组学规模的蛋白质定量可以在一次实验中对成百上千个蛋白质的定量测定和比较分析, 为规模化发现和验证疾病诊断的生物标志物以及发展新的药物靶标提供了重要手段。

表 1 传统的蛋白质定量方法<sup>[1]</sup>

方法	原理	特点
凯氏定氮法	基于蛋白质的元素组成(蛋白质的含氮量在 16%左右)。	测定蛋白质的总量。优点: 适用性好。缺点: 非蛋白氮的干扰, 操作过程繁琐。
Lowry 法(比色法)	双缩脲法(碱性溶液中蛋白质的肽键与 Cu <sup>2+</sup> 反应形成紫色络合物)结合 Folin 酚试剂法(酚试剂被蛋白质中的氨基酸残基还原呈深蓝色)。	蛋白质含量的相对定量分析。优点: 可以对多个样本同时分析。缺点: 需要已知含量的蛋白质标准品, 对样本溶解度要求高, 酚试剂稳定性差, 反应易受多种物质干扰。
考马斯亮蓝法(比色法)	考马斯亮蓝 G-250 与蛋白质的复合物在 595 nm 有很高的光吸收系数, 并且吸光度与溶液中蛋白质浓度成正比。	测定蛋白质总量。优点: 操作简便, 灵敏度较高, 复合物稳定。缺点: 非特性吸附产生干扰, 样品要完全溶解, 样品不能回收使用。
紫外光谱吸收法	蛋白质中芳香族氨基酸在 280 nm 有特异吸收, 肽键在小于 240 nm 波长(通常在 215 nm)有吸收。	测定蛋白质的总量。优点: 操作简便, 灵敏度高, 样本不损失。缺点: 精确度差, 其他吸收物质干扰, 要求溶液完全透明。

近十年来, 随着高精度物质谱技术的发展和基于生物信息学的大量数据处理技术的进步, 规模化地精确测量细胞内蛋白质组的表达变化已经成为现实。定量蛋白质组研究方法也从传统的基于二维凝胶电泳(two dimensional electrophoresis, 2-DE)结合质谱鉴定的策略向着更高准确度、更大动态范围和更高通量的方向发展。

蛋白质组学定量方法根据是否对蛋白质/多肽进行标记可以分为标记和非标记两类(见表 2); 而根据定量目的的不同又可以分为相对定量和绝对定量, 前者是对不同样本中相同蛋白质群表达量的相对变化进行分析, 而后者是对样本中单个蛋白质或蛋白质组的绝对量或浓度进行测定。目前, 基于多重标记的蛋白质组相对定量和基于多反应监测(MRM)针对有限数量的目标蛋白质群的绝对定量方法及其应用发展较快, 而采用非标记技术对蛋白质组进行绝对定量分析是未来的重要发展方向。

### 1 蛋白质组标记定量方法的研究

标记定量方法是基于在肽段中引入稳定同位素标记(如 <sup>2</sup>H、<sup>13</sup>C、<sup>15</sup>N、<sup>18</sup>O), 使得在同一次质谱扫描中标记“轻”和“重”同位素的多肽具有相同的色谱行为和离子化效率, 此时成对出现的质谱峰信号的相对强度可以精确地反映样品中多肽的丰度比例(蛋白质的比例)。因此该方法具有良好的定量可靠性, 不仅可以用于不同样品间的相对定量分析, 还可以用于蛋白质的绝对定量分析。

根据稳定同位素掺入的方式, 标记方法可以分为代谢标记、酶促标记和化学标记等。经典的体内代谢标

基金项目: 国家自然科学基金项目(31100591); 蛋白质组学国家重点实验室自主研究课题(SKLP-Y201101)。

收稿日期: 2013-07-16

表 2 定量蛋白质组学方法(数据主要来源于文献<sup>[21]</sup>)

方法	代表方法	动态范围	覆盖率	定量精度/通量/应用范围	相关软件(网址)
标记	细胞培养氨基酸稳定同位素标记(SILAC)	1~2	中等	高/低/相对定量 ≤3个样本之间的比较	MSQuant( <a href="http://www.msquant.sourceforge.net/">http://www.msquant.sourceforge.net/</a> ) Maxquant( <a href="http://www.maxquant.org/">http://www.maxquant.org/</a> ) Elucidator( <a href="http://www.rosettabio.com/">http://www.rosettabio.com/</a> ) ASAPRatio( <a href="http://tools.proteomecenter.org/">http://tools.proteomecenter.org/</a> ) Mascot Distiller( <a href="http://Matrixscience.com/distiller.html">http://Matrixscience.com/distiller.html</a> ) PEAKS Q( <a href="http://www.bioinformaticsolutions.com/products/peaks/quantification.php">http://www.bioinformaticsolutions.com/products/peaks/quantification.php</a> )
化学标记	同位素标记: 同位素亲和标签(ICAT), 同位素标记蛋白质技术(ICPL)	1~2	低	高/低/相对定量 2个样本之间的比较	Mascot Distiller( <a href="http://Matrixscience.com/distiller.html">http://Matrixscience.com/distiller.html</a> ) Elucidator( <a href="http://www.rosettabio.com/">http://www.rosettabio.com/</a> ) XPRESS( <a href="http://tools.proteomecenter.org/">http://tools.proteomecenter.org/</a> ) ASAPRatio( <a href="http://tools.proteomecenter.org/">http://tools.proteomecenter.org/</a> ) ZoomQuant ( <a href="http://proteomics.mcw.edu/zoomquant.html">http://proteomics.mcw.edu/zoomquant.html</a> ) PEAKS Q( <a href="http://www.bioinformaticsolutions.com/products/peaks/quantification.php">http://www.bioinformaticsolutions.com/products/peaks/quantification.php</a> ) ProteinPilot( <a href="http://absciex.com/">http://absciex.com/</a> )
	等量异序标签标记: 等重同位素标记相对和绝对定量(iTRAQ), 质量差异同位素标记相对和绝对定量(mTRAQ), 串联质量标签(TMT), 差异凝胶双向电泳(DIGE)	2, 3	中等	中等/低/相对定量或绝对定量 2~8个样本之间的比较	Multi-Q( <a href="http://ms.iis.sinica.edu.tw/Multi-Q-Web/">http://ms.iis.sinica.edu.tw/Multi-Q-Web/</a> ) iTracker( <a href="http://www.cranfield.ac.uk/">http://www.cranfield.ac.uk/</a> ) Scaffold Q+ ( <a href="http://www.proteomesoftware.com/">http://www.proteomesoftware.com/</a> ) DeCyder( <a href="http://www.gelifsciences.com/">http://www.gelifsciences.com/</a> )
酶促标记	<sup>18</sup> O	2	低	高/低/相对定量或绝对定量	MSQuant( <a href="http://msquant.sourceforge.net/">http://msquant.sourceforge.net/</a> ) Mascot Distiller( <a href="http://Matrixscience.com/distiller.html">http://Matrixscience.com/distiller.html</a> ) PEAKS Q( <a href="http://www.bioinformaticsolutions.com/images/stories/pdf/PEAKSposters/peaksq-softwareforms-basedquantification.pdf">http://www.bioinformaticsolutions.com/images/stories/pdf/PEAKSposters/peaksq-softwareforms-basedquantification.pdf</a> )

表 2 (续)

方法	代表方法	动态范围	覆盖率	定量精度/通量/应用范围	相关软件(网址)
非标记	二维凝胶电泳	1 ~ 4	低	中等/低/相对定量 质谱鉴定	PDQuest ( <a href="http://www.bio-rad.com/">http://www.bio-rad.com/</a> ) Progenesis SameSpots ( <a href="http://www.nlinear.com/">http://www.nlinear.com/</a> ) Melanie ( <a href="http://www.genebio.com/">http://www.genebio.com/</a> ) Phoretix ( <a href="http://www.perkinelmer.com/">http://www.perkinelmer.com/</a> ) msInspect ( <a href="http://proteomics.fhcr.org/">http://proteomics.fhcr.org/</a> ) MSight ( <a href="http://web.expsy.org/">http://web.expsy.org/</a> ) TOPP ( <a href="http://open.ms.sourceforge.net/">http://open.ms.sourceforge.net/</a> ) PEPPER ( <a href="http://www.broadinstitute.org/">http://www.broadinstitute.org/</a> ) SuperHirn ( <a href="http://www.waters.com/">http://www.waters.com/</a> ) DeCyder MS ( <a href="http://www.gelifesciences.com/">http://www.gelifesciences.com/</a> ) SIEVE ( <a href="http://thermo.com/">http://thermo.com/</a> ) ProteinLynx ( <a href="http://www.waters.com/">http://www.waters.com/</a> ) Scaffold ( <a href="http://www.proteomesoftware.com/">http://www.proteomesoftware.com/</a> ) Elucidator ( <a href="http://www.rosettatabio.com/">http://www.rosettatabio.com/</a> ) ProteofQ ( <a href="http://bioinquire.com">http://bioinquire.com</a> ) APEX ( <a href="http://pfigr.jcvi.org/index.php/bioinformatics/apex.html">http://pfigr.jcvi.org/index.php/bioinformatics/apex.html</a> ) Mascot ( <a href="http://www.matrixscience.com/">http://www.matrixscience.com/</a> ) ( <a href="http://www.nature.com/nature/journal/v473/n7347/abs/nature10098.html">http://www.nature.com/nature/journal/v473/n7347/abs/nature10098.html</a> ) ( <a href="http://mcpomline.org/content/11/3/M111.013987.abstract?sid=e88dc482-515d-4499-a9e2-b1b1bf1370e2">http://mcpomline.org/content/11/3/M111.013987.abstract?sid=e88dc482-515d-4499-a9e2-b1b1bf1370e2</a> ) Skyline ( <a href="http://brendanx-uw1.gs.washington.edu">http://brendanx-uw1.gs.washington.edu</a> ) MaxQuant ( <a href="http://maxquant.org/">http://maxquant.org/</a> ) ATAQS ( <a href="http://tools.proteomecenter.org/ATAQS/ATAQS.html">http://tools.proteomecenter.org/ATAQS/ATAQS.html</a> ) MRMer ( <a href="http://proteomics.fhcr.org/CPL/MRMer.html">http://proteomics.fhcr.org/CPL/MRMer.html</a> )
	质谱离子强度 (ion intensities)	3	高	中等/中 ~ 高/相对定量 依赖于 LC	
	谱图计数 (spectrum count)	3	高	差/中 ~ 高/相对定量 依赖于 LC	
	APEX (absolute protein expression profiling), 指数修饰蛋白丰度指数 (empAI)	3 ~ 4	高	差/高/相对定量	
	基于信号强度的绝对定量 (IBAQ)	3 ~ 4	高	中等/高/相对定量或绝对定量	
	每个蛋白质的 3 信号最强肽段信号强度定量 (TOP3)	3 ~ 4	高	中等/高/相对定量或绝对定量	
目标蛋白质 组定量	与 mTRAQ, TMT, <sup>18</sup> O 等标记方法结合	5	低	精确/高/相对定量或绝对定量 需要系统的方法学发展	

记包括  $^{15}\text{N}$  体内代谢标记<sup>[3]</sup>、SILAC<sup>[4]</sup>等。在此基础之上对 SILAC 技术进行扩展,发展了一系列新的方法,如 culture-derived isotope tags (CDITs)<sup>[5]</sup>和 SILAC 双重(使用两种不同的“重”氨基酸)标记<sup>[6]</sup>等。近年来,随着 MRM 的广泛应用,对于大批量内标的需求催生了基于 SILAC 策略进行重组表达串联体蛋白质(QconCAT, contaminants of Q peptides)技术,以规模化制备内标肽段<sup>[7]</sup>。利用该技术可以将来源于不同蛋白质的特异肽段整合在一个 QconCAT 蛋白质中。通过在蛋白质重组表达时加入重标记氨基酸的方法,得到大量蛋白质的内标肽段,避免了传统的化学合成肽段费时、费力的问题,有利于对目标蛋白质进行相对或绝对的定量分析<sup>[8]</sup>。

2001 年, Fenselau 实验室<sup>[9]</sup>首先将  $^{18}\text{O}$  酶切标记的方法用于蛋白质组相对定量的研究。随后,人们不断对此方法进行改进,重点解决其标记效率、标记稳定性和回交等问题。Zhao 等<sup>[10]</sup>使用酶促进剂 Rapigest<sup>TM</sup> SF 改善肽段的分散度,使用微波加热提高反应效率,使标准肽段的  $^{18}\text{O}$  标记效率高达 100%;再使用高浓度还原剂和烷基化试剂对胰酶彻底灭活,抑制了  $^{18}\text{O}$ - $^{16}\text{O}$  回交反应,标记肽段 6 天内稳定。现在,固定化酶技术的发展大大提高了  $^{18}\text{O}$  酶切标记的效率和稳定性,利于该标记定量方法的进一步广泛应用。

在标记方法中,化学标记的种类最为繁多,应用也最为广泛。依据标记标签的不同,化学标记可以分为同位素标记和等量异序标签标记。代表性的化学标记方法包括标记多肽(蛋白质)N 端氨基的 iTRAQ<sup>[11]</sup>、标记肽段 C 端羧基(-COOH)的酯化标记、针对半胱氨酸巯基(-SH)的 ICAT<sup>[12]</sup>、以及与质谱多反应监测(MRM-MS)技术联用的 mTRAQ 标记技术<sup>[13]</sup>等。化学标记定量技术发展的一个重要方向是多重定量标记,如 iTRAQ 最大的特点就是可以进行八重定量标记,从而大大提高了分析效率。最新的串联质谱标签(tandem mass tag, TMT)最多可以实现十八重标记,再加上高分辨质谱仪的使用<sup>[14]</sup>,规模化精确定量多种细胞状态下相同蛋白质的表达变化已经变为现实。

金属元素化学标记蛋白质组定量也是近年来发展的一类新方法,也拓展了蛋白质组定量的新思路。与稳定同位素标记方法相比,金属元素标记方法具有试剂价廉易得<sup>[15]</sup>、可发展成多重标记技术、可与生物质谱和元素有机质谱联用、可用于蛋白质组相对和绝对定量的双重优势<sup>[16]</sup>等特点。

## 2 蛋白质组非标记(label free)定量方法的研究

相对于传统的标记定量方法,非标记定量方法不需要在样本分析前对蛋白质/多肽进行标记,避免了样本标记处理环节造成的可能的样品损失,在检测肽段的数量、蛋白质的覆盖率和分析通量方面具有较大优势,且不受样品来源和数量的限制,因而近年来得以迅速发展<sup>[17-20]</sup>。

传统的非标记定量方法包括准确质量结合色谱保留时间标签法(accurate mass and retention time tag method, AMT)<sup>[21,22]</sup>、Decyder-MS<sup>[23,24]</sup>和基于质谱归一化的非标记定量方法。基于质谱的归一化定量方法的理论基础是多肽的丰度与其在质谱中被检测到的次数或者质谱峰强度具有一定的相关性。如蛋白质丰度指数方法(protein abundance index, PAI)<sup>[25]</sup>、多肽匹配数(peptide hits)<sup>[26]</sup>、谱图计数(spectral count)<sup>[27,28]</sup>、多肽离子强度计数(peptide ion intensity counting)<sup>[29,30]</sup>等。在此基础之上又发展了综合多肽计数(peptide counting)、谱图计数(spectral count)和碎片离子峰强度(fragment-ion intensity)等 3 种信息的归一化非标记定量方法 SIN(normalized spectral index)<sup>[31]</sup>以及将谱图计数的非标记定量方法与 SILAC 定量方法整合的 SILAC 多肽计数比例分析(SILAC peptide count ratio analysis, SPeCtRA)定量方法<sup>[32]</sup>等。这些方法均具有较好的准确性和灵敏度,能应用于复杂生物样品的高通量蛋白质组定量分析。

近年来, Aebersold 等<sup>[33,34]</sup>发展了将 MRM 绝对定量和非标记定量相结合进行高通量蛋白质组定量分析的策略,采用了基于每个蛋白质中信号最强肽段(“best-flyer” peptides)信号强度的定量策略(Top3),取得了很好的定量结果。Schwanhäusser 等<sup>[35]</sup>将基于 iBAQ(intensity based absolute quantification)的定量方法与多肽计数、PAI、intensities、Top3 等定量方法进行了详细的比较研究,发现 iBAQ 定量方法具有更好的准确性。若结合样本的上样量,可以实现对大批量蛋白质的准确绝对定量分析<sup>[36]</sup>,其定量效果甚至可以与 MRM 绝对定量结果相媲美,突破了传统非标记定量方法一般用于相对定量分析的局限,扩展了非标记定量方法的应用范围。

鉴于非标记定量技术对 LC-MS/MS 分析平台的稳定性、样本处理的重现性、数据处理软件以及海量数据处理方法的可靠性等都有极高的要求,因此需要对上述条件进行严格系统的质控,才能保证定量结果的准确性。

### 3 前景与展望

总之,蛋白质组学研究已经从单纯提高覆盖率的定性研究向更加真实地描述生物体本质的定量研究发展。尽管蛋白质组学定量方法已经取得了显著的进展,但距离真正准确和方便地描述细胞和生物体内蛋白质组的丰度及其表达量的变化还有距离,在定量标记的分析效率、分析通量和准确性上仍存在较大的提升空间,非标记定量方法的准确性特别是数据处理软件等均成为今后的重点发展方向。

致谢 感谢卫军营博士在文献收集和文章撰写中提供的帮助。

#### 参考文献:

- [1] Yao L B. Medical Molecular Biology Laboratory Technology. 2nd ed. Beijing: People's Medical Publishing House (药立波. 医学分子生物学实验技术. 2 版. 北京: 人民卫生出版社), 2011
- [2] Wasinger V C, Zeng M, Yau Y. International Journal of Proteomics, DOI: org/10.1155/2013/180605
- [3] Oda Y, Huang K, Cross F R, et al. Proc Nat Acad Sci USA, 1999, 96(12): 6591
- [4] Ong S E, Blagoev B, Kratchmarova I, et al. Mol Cell Proteomics, 2002, 1(5): 376
- [5] Ishihama Y, Sato T, Tabata T, et al. Nature Biotechnol, 2005, 23(5): 617
- [6] Imami K, Sugiyama N, Tomita M, et al. Mol Biosyst, 2010, 6(3): 594
- [7] Beynon R J, Doherty M K, Pratt J M, et al. Nat Methods, 2005, 2: 587
- [8] Ding C, Li Y H, Kim B, et al. J Proteome Res, 2011, 10(8): 3652
- [9] Yao X, Freas A, Ramirez J, et al. Anal Chem, 2001, 73(13): 2836
- [10] Zhao Y, Jia W, Sun W, et al. J Proteome Res, 2010, 9(6): 3319
- [11] Ross P L, Huang Y N, Marchese J N, et al. Mol Cell Proteomics, 2004, 3(12): 1154
- [12] Gygi S P, Rist B, Gerber S A, et al. Nat Biotechnol, 1999, 17(10): 994
- [13] Kang U B, Yeom J, Kim H, et al. J Proteome Res, 2010, 9(7): 3750
- [14] Vincent C E, Rensvold J W, Westphall M S, et al. Anal Chem, 2013, 85(4): 2079
- [15] Liu H, Zhang Y, Wang J, et al. Anal Chem, 2006, 78(18): 6614
- [16] Wang X, Zhang Y J, Wang X Y, et al. Anal Method, 2012, 4(6): 1629
- [17] Li Q, Singh C R, Ma S, et al. J Proteome Res, 2011, 10(5): 2425
- [18] Skarra D V, Goudreault M, Choi H, et al. Proteomics, 2011, 11(8): 1508
- [19] Soderblom E J, Philipp M, Thompson J W, et al. Anal Chem, 2011, 83(10): 3758
- [20] Schilling B, Rardin M J, Maclean B X, et al. Mol Cell Proteomics, 2012, 11(5): 202
- [21] Andreev V P, Petyuk V, Brewer H, et al. J Proteome Res, DOI: 10.1021/pr3001546
- [22] Liu K, Zhang J, Wang J, et al. Anal Chem, 2009, 81(4): 1307
- [23] Johansson C, Samskog J, Sundstrom L, et al. Proteomics, 2006, 6(16): 4475
- [24] Liu S, Bai S, Qin Z, et al. J Cell Mol Med, 2009, 13(8A): 1586
- [25] Thingholm T E, Bak S, Beck-Nielsen H, et al. Mol Cell Proteomics, 2011, 10(9): M110.006650
- [26] Li Q. International Journal of Proteomics, 2010, DOI: 10.1155/2010/731582
- [27] Trudgian D C, Ridlova G, Fischer R, et al. Proteomics, 2011, 11(14): 2790
- [28] Sun A, Zhang J, Wang C, et al. J Proteome Res, 2009, 8(11): 4934
- [29] Gao B B, Stuart L, Feener E P. Mol Cell Proteomics, 2008, 7(12): 2399
- [30] Jantos-Siwj J, Schiffer E, Brand K, et al. J Proteome Res, 2009, 8(1): 268
- [31] Griffin N M, Yu J, Long F, et al. Nat Biotechnol, 2010, 28(1): 83
- [32] Parker S J, Halligan B D, Greene A S. Proteomics, 2010, 10(7): 1408
- [33] Malmström J, Beck M, Schmidt A, et al. Nature, 2009, 460(7256): 762
- [34] Ludwig C, Claassen M, Schmidt A, et al. Mol Cell Proteomics, 2011, 11(3): M111.013987
- [35] Schwanhäusser B. [PhD Dissertation]. Berlin: Humboldt Universität, 2010
- [36] Ding C, Jiang J, Wei J Y, et al. Mol Cell Proteomics, 2013, DOI: 10.1074/mcp.O112.025023