植物科学学报 2018, 36(3): 470~478
Plant Science Journal http://www.plantscience.cn

DOI:10.11913/PSJ.2095-0837.2018.30470
秦爱, 郑晓敏, 王坤. 基于质谱的植物蛋白质组学研究方法[J]. 植物科学学报, 2018, 36(3): 470-478
Qin A, Zheng XM, Wang K. Research methods of plant proteomics based on mass spectrometry[J]. *Plant Science Journal*, 2018, 36(3): 470-478

基于质谱的植物蛋白质组学研究方法

秦 爱, 郑晓敏, 王 坤*

(武汉大学生命科学学院, 武汉 430072)

摘 要: 基于质谱的植物蛋白质组学研究方法, 从定性和定量蛋白质组学两个方向进行了归纳总结, 并对近年来出现的靶向蛋白质组学、DIA/SWATH 技术、化学蛋白质组学, 以及多组学联合分析等蛋白质组学研究的新技术、新方法和新应用进行了综述。

关键词: 植物; 质谱; 蛋白质组学; 研究方法

中图分类号: Q943.2 **文献标识码**: A **文章编号**: 2095-0837(2018)03-0470-09

Research methods of plant proteomics based on mass spectrometry

Qin Ai, Zheng Xiao-Min, Wang Kun*

(*College of Life Sciences*, *Wuhan University*, Wuhan 430072, China)

Abstract: This paper discusses the existing research methods of plant proteomics based on mass spectrometry by summarizing the qualitative and quantitative proteomics in both directions and reviewing the new technologies, methods, and applications of proteomics that have emerged in recent years, including targeted proteomics, DIA/SWATH, chemical proteomics, and multi-omics co-analysis.

Key words: Plants; Mass spectrum; Proteomics; Research methods

蛋白质组学概念最早在 1994 年由澳大利亚的 Marc Wilkins 等提出, 即以蛋白质组为对象, 研究细胞、组织或生物体全套的蛋白质组成及其变化规律的学科, 并将其定义为"一种基因组所表达的全部蛋白质"。蛋白质组学本质上指的是在大规模水平上研究蛋白质的特征, 包括蛋白质的表达水平, 翻译后修饰的种类, 蛋白与蛋白的相互作用等, 从整体水平上研究蛋白质的组成和调控规律。

蛋白质组学最初的研究是通过双向凝胶电泳(2D-PAGE: Two-dimensional polyacrylamide gel electrophoresis)和氨基酸序列测定(Edman 化学降解法)实现, 但因其具有通量和敏感度低、重复性差的缺点, 现在的研究已基本不再使用。随着 2012 年世界上第一台质谱仪的出现, 更高分辨率和灵敏度的质谱设备不断涌现, 与此同时质谱数据分析软件也得到了很好的开发[1,2], 这些都为蛋白质组学研究的进一步发展提供了良好的软硬件基础。因此目前大多数蛋白质组学研究都依靠质谱技术完成, 质谱分析也成为植物系统生物学研究中不可缺少的重要手段[3]。基于质谱的蛋白质组学研究一般包含以下 7 个步骤: (1) 蛋白质样本抽提; (2) 蛋白质初步分离纯化; (3) 蛋白质胰蛋白酶酶解(也可用其他酶代替[4]); (4) 肽段加以标记/富集(可选步骤); (5) 高效液相色谱进行肽段预分

收稿日期: 2018-01-22, 退修日期: 2018-03-19。
基金项目: 国家自然科学基金项目(31770310)。
This work was supported by a grant from the National Natural Science Foundation of China (31770310).
作者简介: 秦爱(1993-), 女, 硕士研究生, 研究方向为植物蛋白质组学(E-mail: 18971445290@163.com)。
* 通讯作者(Author for correspondence. E-mail: wangk05@whu.educn)。

离；(6) LC-MS/MS 质谱检测；(7) 数据库搜索鉴定(开源软件：MaxQuant[1]和 SearchGUI[2]等；商业化软件：Mascot、ProteinPilot、Proteome Discoverer 等；蛋白质数据库可以从官网或 Uniprot[5]数据库下载)。

1　基于质谱的常用研究方法

蛋白质组学研究可分为定性蛋白质组学和定量蛋白质组学。

定性蛋白质组学研究一般有 3 种策略：Top down、Middle down 和 Bottom up[6,7]。Top down 是指从一个完整的蛋白出发，不对蛋白进行酶解操作，直接在质谱中进行碎片化，通过对碎片分子的检测推导蛋白质的序列。Middle down 是采用不同于胰蛋白酶的生物酶对蛋白质进行酶解，得到较长的肽段，再将长片段肽段进行质谱分析。Bottom up 则是蛋白质研究领域常说的“鸟枪法”，蛋白酶将蛋白质在特定位点切断成长度合适的肽段，肽段进入质谱进行检测，最后通过检测到的谱图推测肽段信息，再推测蛋白质的序列信息[8]。“鸟枪法”一次鉴定可获得巨量谱图，即高通量的多肽和蛋白质信息，在蛋白质组学研究中被大多数研究者选用[8,9]，而 Top down 和 Middle down 策略的使用率较低。

定量蛋白质组学研究方法可以分为 2 类：(1) 绝对定量，即在蛋白质酶解成多肽以后加入已知含量的稳定同位素标记肽段，通过对样本中肽段和标记肽段的信号强度的比较，确定样本中肽段的绝对含量，从而定量到蛋白质的绝对含量。因为植物中已知的生物学标志物较少，而且绝对定量的成本很高，所以这项技术在植物蛋白质组学研究中应用不多，只在拟南芥(*Arabidopsis thaliana* L.)转基因植株[10]和玉米(*Zea mays* L.)脂质转运蛋白[11]等研究中有所应用，且只局限在某种蛋白质的绝对定量；(2) 相对定量，即在不同样本酶解后的多肽上加以不同标记，再通过对所加标记进行定量来比较样本间的蛋白质差异。主要的标记方法有 3 种：代谢标记定量(Metabolic labeling)、化学标记定量(Chemical labeling)和非标记定量(Label-free)。因为定量方式的选择较多，而且无样本偏向性，因此在植物中应用极为普遍。

1.1　定量蛋白质组学——代谢标记

稳点同位素体内标记法可以通过动植物代谢过程中对营养物质的吸收实现蛋白质的体内标记，被称为代谢标记定量。在动物细胞研究中首创的 SILAC(Stable isotope labeling with amino acids in cell culture)[12]方法是通过对 2 组动物细胞加以稳点同位素标记的必需氨基酸(赖氨酸或精氨酸等)标记蛋白质，再将一级质谱中形成的一对同位素峰进行比较，从而获得蛋白质的相对定量结果[13]。根据植物营养吸收的特点，植物蛋白质组学研究者开发出 ^{15}N 稳定同位素标记法，该方法分别采用 ^{15}N 和 ^{14}N 作为唯一氮源的培养基来培养植物组织或植株，实现了植物蛋白质的同位素体内标记定量[14,15]。

1.2　定量蛋白质组学——化学标记

化学标记是一种对肽段引入化学基团的标记方式，现在主流的标记技术有：稳定同位素双甲基化标记[16]、iTRAQ(Isobaric tags for relative and absolute quantification)[17]和 TMT(Tandem mass tags)标记[18]。双甲基化标记是将样本中肽段的 N 端和赖氨酸的侧链上标记一个双甲基化基团，并通过它们引入不同的稳定同位素，使得不同样本的相同肽段之间发生一个至少 4 Da 的差异，从而在质谱检测中形成一对质荷比偏移的同位素峰，再通过峰的强度定量出肽段的相对表达量。iTRAQ 和 TMT 两种试剂的定量原理一致，均由 3 部分组成：质量报告基团、质量标准化基团和氨基反应基团。以 TMT 六标标记为例：质量报告基团有 6 种不同的分子量，同时也有 6 种不同分子量的质量标准化基团与质量报告基团匹配，使 2 个基团的分子量之和保持一致，氨基反应基团能与肽段 N 端及赖氨酸侧链氨基发生共价连接使肽段连上标记。一级质谱中，不同样品中的同一肽段表现出相同的质荷比，但经过碰撞室的进一步碎裂之后，质量报告基团和质量标准化基团之间的键断裂释放出 TMT 报告离子，从而在二级质谱低质量区形成 6 个 TMT 报告离子峰，其强度反应了该肽段在不同样品中的表达量，经过数据库检索后得到蛋白质的相对定量信息。目前稳定同位素双甲基标记可以同时标记 2 ~ 3 个样本，商业化试剂盒有 iTRAQ 四标、iTRAQ 八标、TMT 六标和 TMT 十标，因此化学标记技术可以在一次实验中实现较多样本间的相对定

量，而且准确性、重复性相对较好，没有样本偏向性，所以在植物蛋白质组学的研究中应用比例最大，广泛应用于样本间差异蛋白质的筛选[9,19]。

1.3 定量蛋白质组学——非标记定量

非标记定量是一种不对蛋白质进行任何标记的定量方式[20]。将实验组和对照组进行平行的样本前处理，无需进行样本的等量混合，而是分开进行各自的质谱检测，得到原始谱图信息后根据谱图计数或质谱离子强度统计的方法进行相对定量，目前非标定量数据解析较好的分析软件是 MaxQuant[1,21]。理论上，非标记定量可用于任何样本的蛋白质定量分析，如果实验条件较为稳定，还可以对不同来源的同一样本的定量数据进行互相比较，数据移植性较高，这是标记定量无法实现的，但是非标记定量在实际应用上还有诸多不足。由于实验组和对照组的样本前处理是分开进行的，虽然是平行实验，但是引入的差异仍然较大，重复性较差，而且二者需要分开进行质谱检测，得到原始数据后需要将2组谱图根据肽段洗脱时间和质荷比进行匹配，因此对实验操作及质谱仪的分辨率、灵敏度和稳定性等均有极高的要求。诸多因素导致非标记定量的准确性和可靠性较标记定量而言较差，但因为非标记定量没有样本限制，且数据具有可移植性，所以在拟南芥[22]，番茄（*Lycopersicon esculentum* Mill.）[23]，玉米[24]和水稻（*Oryza sativa* L.）[25]等植物的蛋白质组学研究中得到一定的应用。

2 新兴研究方法——靶向蛋白质组学

传统基于质谱的植物蛋白质组学研究是从全蛋白质组水平中发现差异蛋白，但当研究目的变为精确定量特定少数目标蛋白的变化时，靶向蛋白质组学[26]应运而生。靶向蛋白质组学最早被广泛应用于医学领域的生物标志物（Bio-marker）开发[27]。其检测通量虽不高，但可在大量的样本中，高度稳定和特异性地检测几十或几百个靶标蛋白质的变化。靶向蛋白质组学技术分为3大类：SRM（Selective reaction monitor）/MRM（Multi reaction monitor）[28]、PRM（Parallel reaction monito）[29,30]和DIA（Data independent acquisition）/SWATH（Sequential windowed acquisition of all theoretical fragment ions）[31]。值得一提的是，因为DIA/SWATH定量蛋白质组学技术可以应用于靶向和非靶向蛋白质组学研究，在鉴定特定复合体或特定蛋白质变化的时候，可以认为是靶向蛋白质组学的应用；在鉴定非目标性群体蛋白质时则可以归类于非靶向蛋白质组学技术，所以分类比较模糊，主要根据研究者的应用方向而定。

2.1 SRM/MRM

SRM和MRM通常在三重四级杆质谱仪中完成，通过对某蛋白质特定肽段的母离子和子离子进行监控达到定量的效果。研究者需要预先在第一个四级杆中设定肽段母离子的大小，只有符合要求的母离子才可以进入第二个四级杆进行碎裂，碎裂形成的碎片进入第三个四级杆后也只有符合特定大小的子离子才可以被质量检测器检测到，然后通过肽段信号的强弱得到蛋白质的相对定量结果。MRM和SRM的区别在于SRM只选择一个肽段离子进行监测，MRM则同时选择多个一起监测。

2.2 PRM

PRM是SRM/MRM的衍生技术，需要在高精度的Orbitrap质谱仪上完成，它与SRM/MRM的区别在于PRM只需要预先设置肽段母离子的大小，因为Orbitrap的高精度和高灵敏性，可以使母离子的所有子离子被准确地测定，所以在拥有仪器支持的前提下，使用PRM做靶向蛋白质组学研究更为便捷和准确。

2.3 DIA

传统的数据采集模式是DDA（Data dependent acquisition），质谱根据肽段母离子的强度依次选择10～20个强度最大的母离子进入二级质谱分析，这就导致了大量母离子的丢失，而且二级谱图在鉴定的过程中也仅有少量谱图得到充分的鉴定，大部分谱图得不到很好的利用[32]。DIA是质谱的另外一种数据采集模式，指在一级质谱中均匀地采集质量范围内每个窗口的肽段信息，不对母离子进行限制性筛选，信息不会有所遗漏，所以DIA具有较高的准确性和可重复性，弥补了DDA检测重复性和覆盖率不高以及低丰度蛋白质丢失较大的弊端。因为DIA不对母离子进行筛选，加上一次谱图采集所开的窗口较大，所以进入二级质谱的肽段较多，会产生极其复杂的碎片离子，对后续的数据分析形成了巨大的挑战。因为DIA采集的谱图数量巨大，如果用传统的数据库鉴定方法，将理论

谱图和采集的谱图进行匹配，将会漏掉很多有用信息，也不方便控制准确率，所以 DIA 的数据处理，必须先通过 DDA 方法收集大批量的高质量谱图，再将这些高可信度的背景谱图和 DIA 采集谱图进行相互匹配，才可以很好地鉴定 DIA 采集到的肽段信息[33]。高质量 DDA 谱图数据库是挖掘 DIA 中隐藏信息的必要前提，但现有的植物谱图信息还不够全面，所以研究者一般需要自行构建一个谱图数据库以供 DIA 数据分析使用。

2.4　SWATH

SWATH 技术是 DIA 方法的一个代表性应用技术。该技术是由瑞士苏黎世联邦理工学院的 Ruedi Aebersold 博士及其团队与 SCIEX 公司合作，在 2012 年将该技术实现了商业化。目前支持 SWATH 分析的质谱仪只有 SCIEX 公司开发的 Triple-TOF 5600+和 Triple-TOF 6600+。其基本原理是：设置母离子的质荷比所在范围(常见扫描范围为 400 ~ 1200 Dalton)，再将扫描范围划分为每 25 Dalton 一个窗口，随后通过超高速的扫描来依次获得每个窗口内全部离子的所有碎片信息[31]。SWATH 技术具有极高的重复性和灵敏度，定量的准确性也非常好，且对样本没有偏向性，但一次能定量到的蛋白质只能达到约 2000 个，因此可以应用于植物样本的蛋白质数目不是特别多，但需要鉴定的信息极其全面的研究中。目前已有研究者使用多组实验数据对 SWATH 技术的重现性和准确性等参数进行了评估[34]，SWATH 技术的发明者也研究出了构建 SWATH 分析所需的高质量 DIA 数据库的方法[33]，同时也有大量 DIA 数据分析软件被研发出来，如 OpenSWATH[35]、Skyline[36]、DIA-Umpire[37] 和 Spectronaut[38,39]等，应用前景十分广阔。

靶向蛋白质组学因其准确性和无样本偏向性，可以很好的应用于植物蛋白质组学研究，例如：(1) 使用某种刺激对植物处理不同的时间，然后分别提取不同时间和不同组织的全蛋白，监控响应蛋白的动态变化，从而研究激素响应蛋白的变化模式[40]；(2) 使用 AP (Affinity purification)-SWATH 进行蛋白质复合体的挖掘，采用亲和纯化技术将目标靶蛋白及其相互作用的蛋白从样本中特异性富集出来，在准确富集到蛋白质复合体之后，结合 SWATH 定量技术实现精准鉴定[41-43]；(3) 因为 SWATH 技术不丢失区间内的任何信息，因此可以进行植物样本中低丰度蛋白质的定性或定量分析，有助于构建更完善的植物蛋白质表达谱[44]。

3　新兴研究方法——化学蛋白质组学

蛋白质组学研究不仅在传统的生物学基础研究中发挥重要作用，还为生物药的研发提供技术支撑[45]。在新药开发过程中，研究者需要在受试样本中进行药物靶标蛋白的筛选。因为药物小分子和靶标蛋白质结合后，会引起靶标蛋白表达量的变化，所以借助蛋白质组学的定量方法对参考组和实验组进行动态比较可以从中筛选出表达量变化较大的蛋白作为候选靶标，随后再从药物与靶标的相互作用、中靶或脱靶反应等方面对药物和靶标的行为进行进一步的验证[19,46]。

2013 年，瑞典卡罗林斯卡医学院的研究人员开发出了一种可以直接检测药物与靶标蛋白结合及结合强度的方法——CETSA (Cellular thermal shift assay)[47]。CETSA 方法利用了当药物分子与目标蛋白质结合时，靶标蛋白质会变得具有更好的热稳定性的特点。其主要原理是：(1) 通过对不同温度下蛋白质折叠程度的测量得到蛋白质的热稳定曲线，从而了解药物和靶标蛋白相互作用的强弱[48]；(2) 保持温度恒定，测定不同药物浓度下的蛋白质热稳定曲线，了解不同剂量药物对靶标蛋白所带来的影响；(3) 保持温度恒定，对不同处理时间的靶标蛋白质绘制热稳定曲线，从而获得药物进入细胞和靶标蛋白进行相互作用的时间。借助上述 3 种方法，CETSA 可以对靶标蛋白的药物靶点进行验证，监测药物转运、活化、脱靶效应、癌细胞株耐药性等，还可以了解药物在组织中的分布情况[47]，为药物设计与优化提供了高价值的研究方法。2014 年 Savitski 等[49]巧妙地将 CETSA 与蛋白质组学技术结合开发了高通量的药物靶标鉴定方法，实现了细胞全蛋白质组范围内的药物靶标和下游影响因子的同时监测，克服了传统 CETSA 监控蛋白数目单一的缺点，极大提高了实验的通量。

和动物细胞中的药物靶标筛选原理一致，同样的研究思路也可以应用到植物研究中[50]。细菌、病毒、小分子和激素等对植物的刺激影响分析，研究者也可以从蛋白质组学的角度出发，应用类似化学药物靶标鉴定的质谱检测技术，从植物全蛋白水

平挖掘靶标蛋白质分子，并借助 CETSA 对其相互作用进行进一步的验证。该方法在植物蛋白质组学研究中具有很好的应用前景。

4 蛋白质组学与其他组学相结合的复合型研究方法

仅仅依靠蛋白质组学对生物系统进行探究不够全面，所以与其他组学研究进行良好的结合可以达到更系统的分析效果。

4.1 蛋白质组学与翻译后修饰组学结合

许多生命进程不仅需要蛋白质表达强度的控制，还需要翻译后修饰的精准调控，所以将蛋白质组学和翻译后修饰蛋白质组学进行整合分析将有助于理解蛋白质调控的多样性。由于自然状态下，发生翻译后修饰的蛋白质在样本中含量较低而且分布范围较广，但质谱检测对上样量有一定的要求，所以在质谱检测前需要进行亲和富集、多维分离等操作。翻译后修饰因为需要良好的富集技术和更高灵敏度、更高精度的质谱仪，因此研究受到一定程度的限制，目前只集中在修饰类型和修饰位点变化的鉴定上，而不是更精准的鉴定其修饰程度变化。目前研究方法相对成熟的有磷酸化修饰，糖基化修饰和泛素化修饰等，其中应用最多的是蛋白质组学和磷酸化蛋白质组学整合分析的方法，已广泛应用于水稻[51,52]、小麦（*Triticum aestivum* L.）[53]和玉米[54]等植物。研究者从各个修饰水平和常规的蛋白质表达水平分别解释蛋白质的动态变化，从而对植物信号通路调控进行更深入的研究[55,56]。

4.2 蛋白质组学与基因组学、转录组学、翻译后修饰组学结合

2014 年 Nature 杂志连续发表的 3 篇关于人类蛋白质组学的文章在蛋白质组学领域引起了很大轰动，该研究将蛋白质组学与基因组学、转录组学、翻译后修饰组学相结合形成了一种复合型研究方法[57-59]。因为传统的蛋白质组学鉴定过程中所使用的数据库一般只包含基因组组装注释的蛋白质信息，这样便限制了从谱图中挖掘新信息的可能。因此研究学者们开始尝试将各个组学进行有机结合，以期得到更全面完整的分析结果。一项关于人类蛋白质组学草图的研究中，研究者除了使用传统的蛋白质序列数据库外，还构建了其余的 6 种数据库：参考基因组六框翻译数据库、转录组三框翻译数据库、假基因三框翻译数据库、非编码 RNA 三框翻译数据库、蛋白质 N 末端序列数据库和蛋白质信号肽序列数据库[57]。多蛋白数据库的使用和多组学的结合，使研究者可以从质谱结果中利用更多谱图信息来挖掘出更多隐藏的肽段信息。研究结果除了鉴定到大量编码基因的蛋白质产物以外，还鉴定到了一些新的信号肽剪切位点、ORF、uORF、蛋白质 N 端序列等[57]。同样是人类蛋白质组学的相关研究，癌症虽然被认为是基因组疾病，但是行使破坏功能的主体是蛋白质，因此 Zhang 等[60]通过整合蛋白质组学与基因组学、转录组学、磷酸化蛋白质组学的方法，从多个角度尝试解析高级别浆液性卵巢癌中可能的关键性蛋白和潜在的作用机制。

植物和动物或其他模式生物相比，可用的数据库较少，可能无法挖掘蛋白质 N 端序列和信号肽序列等信息，但是结合已有的基因组学、转录组学、翻译后修饰组学相关数据进行分析，也可以给我们提供巨大的数据基础。2008 年，Baerenfaller 等[61]使用基因组、转录组和蛋白质组结合的分析方法对拟南芥进行了深度挖掘，该研究的主要工作包括：（1）构建高密度的拟南芥蛋白质组谱图和定量蛋白质组谱图；（2）分析基因在转录组水平和蛋白质水平上的表达相关性，再根据各个组织的蛋白质表达情况分析组织之间的相似性；（3）将基因组六框翻译后进行谱图匹配来识别基因组上未被注释的蛋白质编码区，从而帮助修正基因组注释结果。该研究以复合型的组学研究方法，提高了对拟南芥特定组织或器官中基因表达控制和基因组活性的认识。Marx 等[62]在苜蓿（*Medicago sativa* L.）的研究中，通过定量蛋白质组学和转录组学、翻译后修饰蛋白质组学的结合分析，找到了苜蓿多部位共表达的核心蛋白质组和参与根瘤生长调控的关键蛋白质组，深入揭示了豆科（Leguminosae）植物与根瘤菌（*Rhizobium*）之间的共生机制。Walley 等[63]在玉米的研究中，同样从定量蛋白质组学、转录组学、翻译后修饰蛋白质组学 3 个角度进行了信息挖掘，整合 3 个表达水平的数据完成了表达网络构建和基因调控网络构建，该网络的构建是预测全系统范围内的单个基因功能的重要工具。

4.3 蛋白质组学与代谢组学结合

在组学研究领域，代谢组学是最为复杂的研究领域之一，代谢物是细胞调控过程的最终产物，它

们的种类和数量的变化代表了生物系统对基因或者环境变化做出的响应。研究者需要对大量小分子、初级代谢物和次级代谢物等进行定量，得到大量的数据集后借助较为完善的代谢组学数据库进行相应的鉴定工作。蛋白质组学与代谢组学的结合，可以很好地对植物体内更细微的反应进行探究，已应用于植物药的研究[64]，以及缺氧条件下水稻和小麦胚芽鞘的差异反应等研究中[65]。

在所有的整合分析方法当中，将蛋白质组学与基因组学、转录组学、翻译后修饰组学进行整合分析所占比例较大。蛋白质组学与基因组学结合可以帮助鉴定新的基因编码区和对已有的基因编码区进行修正；蛋白质组学与转录组学结合不仅可以鉴定样本间的差异蛋白质和构建植物的表达网络，还可以借助转录本的多样性和蛋白质证据的直接性挖掘更多的蛋白质异构体，mRNA 可变剪切和 uORF 等事件；蛋白质组学和翻译后修饰组学结合则可以更充分的解释信号通路上蛋白质的精准调控。近年来植物多组学研究越来越多，多组学数据分析也有了专业分析平台的支撑[66]，从多个组学的角度对植物特征进行解析，可以帮助我们更系统地理解植物生物学。

5　小结与展望

综上所述，基于质谱的植物蛋白质组学研究内容包含以下几个方面，构建植物蛋白质表达谱；鉴定差异表达蛋白质(差异来源：野生型植株和突变体植株、正常植株和非生物胁迫植株、抗性植株和敏感植株等)；靶向蛋白质组学对目标蛋白进行精准定量和蛋白质复合体鉴定；与其他组学研究结合深度挖掘物种特征和信号调控网络。

目前基于质谱的植物蛋白质组学研究方法有：(1) 定性分析；(2) 相对定量或绝对定量分析；(3) 靶向蛋白质组学分析；(4) 化学蛋白质组学分析；(5) 多组学复合型分析。各个研究方法之间不存在明确界限，可以根据研究需要选择一个或多个研究方法进行使用，组合方式较多，如多组学复合分析中蛋白质组学数据可以进行相对或绝对定量；靶向蛋白质组学可以对目的蛋白进行定性或定量分析；非标记定量和 SWATH 检测结合分析[67]；传统 DDA 数据采集和 DIA 技术结合分析[68]。

随着测序技术的飞速发展，越来越多植物进行了全基因组测序，为植物蛋白质组学研究奠定了良好的数据基础。据不完全统计，约 66 种植物完成了全基因组测序工作，除了可以从各自的官网得到测序文件外，还可以从 PLAZA[69]、Phytozome[70]、Ensembl Plants[71]等植物数据库网站进行数据查询和下载。蛋白质组学可以广泛应用于植物研究中的作物改良[72]、信号响应、差异鉴定和系统生物学等各领域，是植物生物学研究过程中必须进行挖掘探索的环节。随着人力和物力资源的大量投入，基于质谱的植物蛋白质组学的研究方法和研究方向将日渐完善，有利于更好地理解植物的各项生命活动。

参考文献：

[1] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification[J]. *Nat Biotechnol*, 2008, 26(12): 1367-1372.

[2] Vaudel M, Barsnes H, Berven FS, Sickmann A, Martens L. SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X! Tandem searches[J]. *Proteomics*, 2011, 11(5): 996-999.

[3] Glinski M, Weckwerth W. The role of mass spectrometry in plant systems biology[J]. *Mass Spectrom Rev*, 2006, 25(2): 173-214.

[4] Giansanti P, Tsiatsiani L, Low TY, Heck AJ. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin[J]. *Nat Protoc*, 2016, 11(5): 993-1006.

[5] Uniprot C. UniProt: a hub for protein information[J]. *Nucleic Acids Res*, 2015, 43: 204-212.

[6] Moradian A, Kalli A, Sweredoski MJ, Hess S. The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications [J]. *Proteomics*, 2014, 14(4-5): 489-497.

[7] Hummel M, Wigger T, Brockmeyer J. Characterization of mustard 2S albumin allergens by bottom-up, middle-down, and top-down proteomics: a consensus set of isoforms of Sin a 1[J]. *J Proteome Res*, 2015, 14(3): 1547-1556.

[8] Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem[J]. *Mol Cell Proteomics*, 2005, 4(10): 1419-1440.

[9] Rauniyar N, Yates JR. Isobaric labeling-based relative quantification in shotgun proteomics[J]. *J Proteome Res*, 2014, 13(12): 5293-5309.

[10] Li YJ, Shu YW, Peng CC, Zhu L, Guo GY, *et al.* Absolute

quantitation of isoforms of post-translationally modified proteins in transgenic organism[J]. *Mol Cell Proteomics*, 2012, 11(8): 272-285.

[11] Stevenson SE, Mcclain S, Thelen JJ. Development of an isoform-specific tandem mass spectrometry assay for absolute quantitation of maize lipid transfer proteins[J]. *J Agric Food Chem*, 2015, 63(3): 821-828.

[12] Zhang GA, Neubert TA. Automated comparative proteomics based on multiplex tandem mass spectrometry and stable isotope labeling[J]. *Mol Cell Proteomics*, 2006, 5(2): 401-411.

[13] Liao LX, Song XM, Wang LC, Lü HN, Chen JF, *et al.* Highly selective inhibition of IMPDH2 provides the basis of antineuroinflammation therapy[J]. *Proc Natl Acad Sci USA*, 2017, 114(29): 5986-5994.

[14] Szymanski WG, Kierszniowska S, Schulze WX. Metabolic labeling and membrane fractionation for comparative proteomic analysis of *Arabidopsis thaliana* suspension cell cultures[J]. *J Vis Exp*, 2013, 79: e50535.

[15] Guo G, Li N. Relative and accurate measurement of protein abundance using 15N stable isotope labeling in *Arabidopsis* (SILIA)[J]. *Phytochemistry*, 2011, 72(10): 1028-1039.

[16] Hsu JL, Huang SY, Chow NH, Chen SH. Stable-isotope dimethyl labeling for quantitative proteomics[J]. *Anal Chem*, 2003, 75(24): 6843-6852.

[17] Unwin RD. Quantification of proteins by iTRAQ[J]. *Methods Mol Biol*, 2010, 658: 205-215.

[18] Dayon L, Sanchez JC. Relative protein quantification by MS/MS using the tandem mass tag technology[J]. *Methods Mol Biol*, 2012, 893: 115-127.

[19] Schenone M, Dancik V, Wagner BK, Clemons PA. Target identification and mechanism of action in chemical biology and drug discovery[J]. *Nat Chem Biol*, 2013, 9(4): 232-240.

[20] Wong JW, Cagney G. An overview of label-free quantitation methods in proteomics by mass spectrometry[J]. *Methods Mol Biol*, 2010, 604: 273-283.

[21] Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction[J]. *Mol Cell Proteomics*, 2014, 13(9): 2513-2526.

[22] Arsova B, Zauber H, Schulze WX. Precision, proteome coverage, and dynamic range of *Arabidopsis* proteome profiling using 15N metabolic labeling and label-free approaches[J]. *Mol Cell Proteomics*, 2012, 11(9): 619-628.

[23] Sun QQ, Zhang N, Wang JF, Cao YY, Li XS, *et al.* A label-free differential proteomics analysis reveals the effect of melatonin on promoting fruit ripening and anthocyanin accumulation upon postharvest in tomato[J]. *J Pineal Res*, 2016, 61(2): 138-153.

[24] Ponnala L, Wang Y, Sun Q, van Wijk KJ. Correlation of mRNA and protein abundance in the developing maize leaf[J]. *Plant J*, 2014, 78(3): 424-440.

[25] Wang YD, Wang X, Ngai SM, Wong YS. Comparative proteomics analysis of selenium responses in selenium-enriched rice grains[J]. *J Proteome Res*, 2013, 12(2): 808-820.

[26] Harlan R, Zhang H. Targeted proteomics: a bridge between discovery and validation[J]. *Expert Rev Proteomics*, 2014, 11(6): 657-661.

[27] Liu Y, Huttenhain R, Collins B, Aebersold R. Mass spectrometric protein maps for biomarker discovery and clinical research[J]. *Expert Rev Mol Diagn*, 2013, 13(8): 811-825.

[28] Wu C, Shi T, Brown JN, He J, Gao Y, *et al.* Expediting SRM assay development for large-scale targeted proteomics experiments[J]. *J Proteome Res*, 2014, 13(10): 4479-4487.

[29] Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics[J]. *Mol Cell Proteomics*, 2012, 11(11): 1475-1488.

[30] Rauniyar N. Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry[J]. *Int J Mol Sci*, 2015, 16(12): 28566-28581.

[31] Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis[J]. *Mol Cell Proteomics*, 2012, 11(6): O111.016717.

[32] Johnson D, Boyes B, Fields T, Kopkin R, Orlando R. Optimization of data-dependent acquisition parameters for coupling high-speed separations with LC-MS/MS for protein identifications[J]. *J Biomol Tech*, 2013, 24(2): 62-72.

[33] Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, *et al.* Building high-quality assay libraries for targeted analysis of SWATH MS data[J]. *Nat Protoc*, 2015, 10(3): 426-441.

[34] Collins BC, Hunter CL, Liu Y, Schilling B, Rosenberger G, *et al.* Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry[J]. *Nat Commun*, 2017, 8(1): 291.

[35] Rost HL, Rosenberger G, Navarro P, Gillet L, Miladinovic SM, *et al.* OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data[J]. *Nat Biotechnol*, 2014, 32(3): 219-223.

[36] Egertson JD, Maclean B, Johnson R, Xuan Y, Maccoss MJ. Multiplexed peptide analysis using data-independent acquisition and Skyline[J]. *Nat Protoc*, 2015, 10(6): 887-903.

[37] Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, *et al.* DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics[J]. *Nat Methods*, 2015, 12(3): 258-264.

[38] Bruderer R, Bernhardt OM, Gandhi T, Reiter L. High-precision iRT prediction in the targeted analysis of data-independent acquisition and its impact on identification and quantitation[J]. *Proteomics*, 2016, 16(15-16): 2246-2256.

[39] Navarro P, Kuharev J, Gillet LC, Bernhardt OM, Maclean B, *et al.* A multicenter study benchmarks software tools for label-free proteome quantification[J]. *Nat Biotechnol*, 2016, 34(11): 1130-1136.

[40] Zhu FY, Chan WL, Chen MX, Kong RP, Cai C, *et al.* SWATH-MS quantitative proteomic investigation reveals a role of jasmonic acid during lead response in *Arabidopsis*[J]. *J Proteome Res*, 2016, 15(10): 3528-3539.

[41] Lambert JP, Ivosev G, Couzens AL, Larsen B, Taipale M, *et al.* Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition[J]. *Nat Methods*, 2013, 10(12): 1239-1245.

[42] Collins BC, Gillet LC, Rosenberger G, Rost HL, Vichalkovski A, *et al.* Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system[J]. *Nat Methods*, 2013, 10(12): 1246-1253.

[43] Zhang H, He D, Yu J, Li M, Damaris RN, *et al.* Analysis of dynamic protein carbonylation in rice embryo during germination through AP-SWATH[J]. *Proteomics*, 2016, 16(6): 989-1000.

[44] Jia H, Sun W, Li M, Zhang Z. Integrated analysis of protein abundance, transcript level, and tissue diversity to reveal developmental regulation of maize[J]. *J Proteome Res*, 2018, 17(2): 822-833.

[45] Rodriguez-Suarez E, Whetton AD. The application of quantification techniques in proteomics for biomedical research[J]. *Mass Spectrom Rev*, 2013, 32(1): 1-26.

[46] Schirle M, Bantscheff M, Kuster B. Mass spectrometry-based proteomics in preclinical drug discovery[J]. *Chem Biol*, 2012, 19(1): 72-84.

[47] Molina DM, Jafari R, Ignatushchenko M, Seki T, Larsson EA, *et al.* Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay[J]. *Science*, 2013, 341(6141): 84-87.

[48] Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, *et al.* Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination[J]. *Proc Natl Acad Sci USA*, 2006, 103(43): 15835-15840.

[49] Savitski MM, Reinhard FB, Franken H, Werner T, Savitski MF, *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome[J]. *Science*, 2014, 346(6205): 1255784.

[50] Dejonghe W, Russinova E. Plant chemical genetics: from phenotype-based screens to synthetic biology[J]. *Plant Physiol*, 2017, 174(1): 5-20.

[51] Ye J, Zhang Z, Long H, Zhang Z, Hong Y, *et al.* Proteomic and phosphoproteomic analyses reveal extensive phosphorylation of regulatory proteins in developing rice anthers[J]. *Plant J*, 2015, 84(3): 527-544.

[52] Wang K, Zhao Y, Li M, Gao F, Yang MK, *et al.* Analysis of phosphoproteome in rice pistil[J]. *Proteomics*, 2014, 14(20): 2319-2334.

[53] Lü DW, Zhu GR, Zhu D, Bian YW, Liang XN, *et al.* Proteomic and phosphoproteomic analysis reveals the response and defense mechanism in leaves of diploid wheat *T. monococcum* under salt stress and recovery[J]. *J Proteomics*, 2016, 143: 93-105.

[54] Facette MR, Shen Z, Bjornsdottir FR, Briggs SP, Smith LG. Parallel proteomic and phosphoproteomic analyses of successive stages of maize leaf development[J]. *Plant Cell*, 2013, 25(8): 2798-2812.

[55] Zhang Z, Chen J, Lin S, Li Z, Cheng R, *et al.* Proteomic and phosphoproteomic determination of ABA's effects on grain-filling of *Oryza sativa* L. inferior spikelets[J]. *Plant Sci*, 2012, 185-186: 259-273.

[56] Yu B, Li J, Koh J, Dufresne C, Yang N, *et al.* Quantitative proteomics and phosphoproteomics of sugar beet monosomic addition line M14 in response to salt stress[J]. *J Proteomics*, 2016, 143: 286-297.

[57] Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, *et al.* A draft map of the human proteome[J]. *Nature*, 2014, 509(7502): 575-581.

[58] Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, *et al.* Mass-spectrometry-based draft of the human proteome[J]. *Nature*, 2014, 509(7502): 582-587.

[59] Zhang B, Wang J, Wang X, Zhu J, Liu Q, *et al.* Proteogenomic characterization of human colon and rectal cancer[J]. *Nature*, 2014, 513(7518): 382-387.

[60] Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, *et al.* Integrated proteogenomic characterization of human high-grade serous ovarian cancer[J]. *Cell*, 2016, 166(3): 755-765.

[61] Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, *et al.* Genome scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics[J]. *Science*, 2008, 320(5878): 938-941.

[62] Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, *et al.* A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont *Sinorhizobium meliloti*[J]. *Nat Biotechnol*, 2016, 34(11): 1198-1205.

[63] Walley JW, Sartor RC, Shen Z, Schmitz RJ, Wu KJ, *et al.* Integration of omic networks in a developmental atlas of maize[J]. *Science*, 2016, 353(6301): 814-818.

[64] Ulrich-Merzenich G, Zeitler H, Jobst D, Panek D, Vetter H, *et al.* Application of the "-Omic-" technologies in phytomedicine[J]. *Phytomedicine*, 2007, 14(1): 70-82.

[65] Shingaki-Wells RN, Huang S, Taylor NL, Carroll AJ, Zhou W, *et al.* Differential molecular responses of rice and wheat coleoptiles to anoxia reveal novel metabolic adaptations in amino acid metabolism for tissue tolerance[J]. *Plant Physiol*, 2011, 156(4): 1706-1724.

[66] Boekel J, Chilton JM, Cooke IR, Horvatovich PL, Jagtap PD, *et al.* Multi-omic data analysis using Galaxy[J]. *Nat Biotechnol*, 2015, 33(2): 137-139.

[67] Mcqueen P, Spicer V, Schellenberg J, Krokhin O, Sparling R, *et al.* Whole cell, label free protein quantitation with data independent acquisition: quantitation at the MS2 level [J]. *Proteomics*, 2015, 15(1): 16-24.

[68] Bromilow SN, Gethings LA, Langridge JI, Shewry PR, Buckley M, *et al.* Comprehensive proteomic profiling of wheat gluten using a combination of data-independent and data-dependent acquisition[J]. *Front Plant Sci*, 2016, 7: 2020.

[69] Bel MV, Diels T, Vancaester E, Kreft L, Botzki A, *et al.* PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics[J]. *Nucleic Acids Res*, 2018, 46(D1): 1190-1196.

[70] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, *et al.* Phytozome: a comparative platform for green plant genomics[J]. *Nucleic Acids Res*, 2012, 40: 1178-1186.

[71] Bolser D, Staines DM, Pritchard E, Kersey P. Ensembl Plants: integrating tools for visualizing, mining, and analyzing plant genomics data[J]. *Methods Mol Biol*, 2016, 1374: 115-140.

[72] Eldakak M, Milad SI, Nawar AI, Rohila JS. Proteomics: a biotechnology tool for crop improvement[J]. *Front Plant Sci*, 2013, 4: 35.

(责任编辑：周 媛)