植物标记和非标记定量蛋白质组学技术

高丰衣^{1,2}, 张文娜³, 许会敏¹, 梁馨琳¹, 陈艳梅^{1*}

1. 中国农业大学生物学院, 植物生理与生物化学国家重点实验室, 北京 100193;

2. 商丘师范学院生物与食品学院, 商丘 476000;

3. 中国农业大学园艺学院, 北京 100193

* 联系人, E-mail: chenyanmei@cau.edu.cn

2020-11-23 收稿, 2021-01-20 修回, 2021-01-20 接受, 2021-01-21 网络版发表

国家自然科学基金(U1904165, 31470345)、河南省科技攻关项目(212102310395)和国家重点研发计划(2019YFD1000300)资助

摘要 蛋白质是细胞中重要的有机大分子, 是生命活动的主要执行者. 植物在受到外界信号干扰时其主要反应是通过调节细胞内蛋白质的合成和降解来恢复稳态. 因而, 蛋白质的定量分析有助于人们更深入地揭示植物细胞中蛋白质的动态变化及其分子功能. 随着高分辨质谱技术的发展, 使用质谱方法对蛋白质进行高通量定量和定性分析的蛋白质组学技术已经成为研究蛋白质功能的重要分析手段, 同时也是植物生物学研究领域的热点之一. 结合本课题组取得的研究成果, 本文总结和概述了植物蛋白质组学定量方法的最新进展, 包括非标记定量和稳定同位素标记分析方法, 比较不同方法在蛋白定量的精确度、蛋白组覆盖率和稳定性等方面的优缺点, 分析其在植物生物学研究中的应用范围, 并对这方面的研究热点进行展望.

关键词 蛋白质组, 蛋白磷酸化, 非标记, 稳定性同位素标记, 细胞膜

蛋白质是构成生物体的重要组成成分之一, 在细胞生长发育和代谢过程中具有重要的分子功能和调节作用. 蛋白质组学是从整体水平研究细胞内动态变化的蛋白质组成、表达水平和修饰状态, 分析蛋白质之间的相互作用与联系, 揭示蛋白质功能和细胞生命活动规律的一门学科^[1,2]. 然而, 早期的蛋白质组学研究仅限于定性鉴定特定组织或细胞中所有蛋白质, 没有关于蛋白质表达量、定位或翻译后修饰等方面的研究. 定量策略则更多应用于使用基因扩增或者microarray技术分析基因在基因组或转录组水平的差异表达^[3]. 在植物生物学研究领域, 高通量基因组学或转录组学在一定程度上能够用来探讨植物生长发育过程中重要的分子机理^[4,5]. 但是, 转录水平的研究不能解析翻译后修饰调控, 也不能对蛋白质的不同异构体进行定量和亚

细胞定位分析^[6]. 由于细胞代谢、酶促反应和信号转导等过程最终取决于蛋白质活性和功能的变化, 生物体中蛋白质的表达量变化也是由蛋白质合成和降解两个机制调控. 所以, 蛋白质组学技术已成为分析复杂蛋白质的有力工具^[7-13].

早期, 双向电泳技术是蛋白质组学研究中重要的分离和定量方法, 可用于比较不同条件或组织的蛋白质样品之间的差异^[14,15]. 然而, 在传统的双向电泳实验中, 通过比较分析凝胶图像上斑点的差异并不能直接与蛋白质的含量相关联. 由于大多数情况下, 双向电泳凝胶上单个斑点包含不止一种蛋白质, 只有在切取差异斑点并通过质谱(mass spectrometry, MS)分析后才能确定具体的蛋白质种类. 随着20世纪90年代中期质谱测序技术的发展, 使用双向电泳进行蛋白分离和定量

引用格式: 高丰衣, 张文娜, 许会敏, 等. 植物标记和非标记定量蛋白质组学技术. 科学通报, 2021, 66: 3952-3964

Gao F Y, Zhang W N, Xu H M, et al. Labeling and label-free quantitative proteomics in plant biology (in Chinese). Chin Sci Bull, 2021, 66: 3952-3964, doi: 10.1360/TB-2020-1506

分析后可以与蛋白质质谱鉴定相结合^[14,16],使得基于双向电泳的蛋白定量方法仍然具有吸引力,并且过去的一段时间内仍被用于植物生物学研究领域^[14]。然而,在定量研究中,双向电泳分离蛋白方法由于重复性较差并且操作费时费力,已经逐渐被基于质谱技术的蛋白质组方法替代。

近年来,基于质谱技术的蛋白质组学已经发展成为常用而有效的高通量分析和鉴定蛋白质的技术^[8,9]。随着多个物种基因组测序的完成,使用质谱技术大规模鉴定不同物种的蛋白质组图谱已经达到了很高的深度^[10]。另外,离子源技术的发展,如电喷雾离子化(electron spray ionization, ESI)或基质辅助激光解吸离子化(matrix assisted laser desorption ionization, MALDI),进一步加速了质谱技术在蛋白质组学研究领域的广泛应用^[17]。也正是由于ESI和MALDI技术在生物大分子研究领域的突出贡献,其发明者约翰·芬恩(John Fenn)和田中耕一(Koichi Tanaka)于2002年被授予诺贝尔化学奖。此外,随着高分辨率质谱技术和大数据处理软件的开发,定量蛋白质组学方法也得到了快速发展,并被广泛应用于生物、医学、环境和农业等研究领域^[18]。因此,精确分析蛋白质表达量及其相关性质的动态变化,例如翻译后修饰(post-translational modification, PTM)、蛋白质周转率和亚细胞定位模式等有助于人们更加深入了解生命现象的本质。

定量蛋白质组学早期大多应用于研究酵母和动物细胞^[19-21]。近年来,在植物生物学研究领域,蛋白质组学也不再仅仅用作一个描述工具,已经成为一门研究蛋白质组成和分子功能的重要学科^[9,22]。质谱技术的发展也加速了定量蛋白质组学在植物蛋白复合体和蛋白互作领域的研究^[23],这些技术为研究植物细胞信号转导、生长调节和逆境机理等生物学过程提供了有力的手段^[24-26]。在蛋白质组学实验中,由于没有像microarray实验中的聚合酶链反应那样的扩增步骤,所以样品含量多少是实验成功的限制因素。因此,在进行蛋白质定量比较之前,关键步骤是对含量低的样品进行富集和纯化。目前,结合色谱分析、超速离心和蔗糖梯度等分离富集技术,定量蛋白质组学能进行如叶绿体、细胞核、线粒体、细胞膜和质膜上脂筏微区等亚细胞器蛋白的分析和定量研究^[27-29]。此外,亚细胞蛋白质组学技术也应用于研究不同发育阶段或胁迫条件下的磷酸化、糖基化和乙酰化等翻译后修饰程度变化^[30,31]。总之,定量蛋白质组学方法已经应用于研究植物生长发

育的所有领域。近10年来,已经发展了多种方法用来定量分析蛋白质含量,这些方法各有优缺点^[32-35]。因此,选择不同蛋白质组学定量策略时,首先要考虑该方法的重现性;在具有良好重现性的前提下,还必须考虑生物样品本身可能存在的变异以及实验成本等因素。

近年来,我们团队采用非标记定量和稳定性同位素标记定量蛋白质组学技术,结合磷酸化组学、生化和分子生物学等手段,对植物中一些重要的生物学问题,如逆境胁迫和营养缺失等不同环境条件下的蛋白组变化、亚细胞器如细胞质膜和细胞核的蛋白组变化及相应条件下的翻译后修饰等分子机理进行了较为深入系统的研究,取得了一些新成果。本文结合这些主要研究成果,综述蛋白质组学研究中常用的定量技术及其应用进展,评价每种方法的优缺点,分析其在植物生物学研究中的应用范围,为深入研究植物蛋白质的分子调控机理提供重要参考。

1 非标记定量蛋白质组学技术

基于质谱的定量蛋白质组方法有非标记定量(label free quantitation, LFQ)和同位素标记定量技术(stable isotope labeling)(图1)^[35,36]。LFQ方法包含两种:谱图计数(spectrum counting, SC)定量分析法和使用肽段离子强度的定量分析法。LFQ方法可来比较两个或两个以上的蛋白样品间的差异^[36]。这类方法要求待分析的所有样品在相同条件下连续进行LC-MS(liquid chromatography-MS)测定,以避免由于系统设置不同(例如柱特性、温度等)而导致样品之间的离子强度变化,因而才能提高蛋白样品之间的质谱数据的重复性^[37]。

LFQ方法不需要标记蛋白质或多肽,其实验成本低,并且可应用于任何生物材料(表1)。由于所有质谱检测到的蛋白质或者肽段都可以使用LFQ方法定量分析,因而该方法可以进行广谱和高通量蛋白质定量。此外,使用LFQ分析复杂混合蛋白样品时,不会因为样品的复杂程度而影响其测样难度。因此,LFQ方法分析蛋白质表达量具有较高的蛋白分析线性范围,使得该方法在应用于分析高通量、较复杂的蛋白质变化时具有很好的优势^[38]。

1.1 谱图计数定量分析方法

SC定量方法的原理是,通过比较同一蛋白质在多次质谱重复分析中产生的MS/MS谱图的数量,提供相对定量。这种定量方法基本特性是,在片段离子扫描

图 1 (网络版彩色)基于质谱的蛋白质组学定量方法. (a) 非标记定量技术. 非标记定量可以同时分析多个样品, 没有分析上限. (b) 代谢物标记的定量技术. 代谢物标记是在培养材料时添加稳定性同位素, 培养后同时取样, 材料混合后进行蛋白质组学实验分析. (c) 化学标记的定量技术. 化学标记可以在蛋白水平或者肽段水平标记同位素, 标记和未标记的样品混合后进行蛋白质组学实验

Figure 1 (Color online) Overview of different quantitative strategies in plant proteomic analysis. (a) Label-free quantitative methods allow comparative analysis of multiple samples analyzed in parallel. (b) Metabolic labeling methods allow mixing of plant samples before purification. (c) Chemical labeling methods introduce a label on either protein or peptide level and the quantification can be processed on survey ions or on fragment ions

中, 高丰度的肽段或蛋白质通常能检测到更多的肽段离子谱图计数(spectrum counting, SC), 低丰度的肽段或蛋白质则被检测到谱图计数较少, 也即蛋白质的含量与质谱检测到的肽段离子谱图数量直接相关. 该方法可对同一样品内不同蛋白质或者多个样品之间的蛋白质含量进行相对定量, 具有不需要标记、操作简单^[39,40]且定量分析的样品数量没有上限等优势(表1). 使用SC定量方法结合磷酸化蛋白质富集技术, 我们研究了外源五大类植物激素刺激拟南芥(*Arabidopsis thaliana*)细胞后的磷酸化信号网络^[39]. 该研究揭示了生长素通道蛋白PIN的磷酸化程度变化调控其在拟南芥细胞质膜上的极性定位. 此外, 我们还证实了ABA(abscisic acid)的信号网络, 即ABA引起SnRK1激酶的磷酸

化变化从而激活并磷酸化下游转录因子ABA responsive element-binding proteins(AREBs), AREB从而调控胞内ABA通路的基因表达^[39]. SC定量方法也应用于其他物种如马铃薯蛋白质组学研究中^[40,41].

然而, SC方法分析低丰度和肽段含量较少的蛋白质时的灵敏度较差, 因为这类蛋白的谱图计数较少(即存在检测下限), 会导致定量分析时出现假阳性误差. 为了克服这个问题, 研究者提出, 利用鉴定蛋白质的所有肽段的总离子数的平均值作为定量分析依据. 基本原理是蛋白质的离子图谱数量或肽段数与样品中该蛋白质的含量通常呈现线性关系, 这一特性可用于分析样品中每个蛋白质的绝对表达量(exponentially modified protein abundance index, emPAI)^[42]. emPAI是一个

表 1 基于质谱数据采集的定量蛋白质组学技术

Table 1 Quantitative proteomic approaches based on MS data acquisition

质谱离子扫描	定量方法	优点	缺点	应用样品材料
MS1	代谢标记 (如SILAC、 ^{15}N)	在植物培养过程中标记, 定量精确度高	最多只能标记3个样品, 数据分析复杂, 应用范围有限	仅限于细胞或者培养基上的植物材料
MS2	化学标记 (如TMT、iTRAQ)	标记蛋白或者多肽, 定量精确度高; 单次可标记3个以上样品; 多个样品同时检测, 可避免数据缺失值	同时分离的肽母离子会影响定量的准确性; 多个样本一次性检测, 不能进行样本个体检测	不同生长条件下的细胞和植物材料均可
MS3	化学标记 (如TMT、iTRAQ)	通过去除一起分离出的质量相似母离子后, 定量精确度更高	多肽鉴定覆盖率较低; 多个样本一次性检测, 不能进行样本个体检测	不同生长条件下的细胞和植物材料均可
MS2	非标记定量 (如图谱计数spectrum counting)	应用范围广, 分析样品无上限, 不需要标记材料, 价格低廉	对于低丰度蛋白定量精确度有限; 对仪器稳定性要求高	不同生长条件下的细胞和植物材料均可
MS1、MS2	非标记定量 (如离子强度定量法)	应用范围广, 检测样品无上限, 不需要标记材料, 价格低廉	实验样本之间误差大、重复性略低、对仪器稳定性要求高	不同生长条件下的细胞和植物材料均可

常用的衍生指数, 该数值是由每个蛋白的实际质谱检测到的谱图计数除以理论肽段数计算得出, 又称为蛋白丰度指数(PAI). 将PAI用以10为底数的指数转换后能得到一个指数修正值, 即emPAI指数. 这种方法可用来计算蛋白质的绝对表达指数(APEX)^[43]. 研究表明, 通过emPAI指数计算出的蛋白浓度与通过酶活性计算出的蛋白浓度有很好的相关性^[43]. emPAI和APEX都是根据质谱方法对给定样品中蛋白质进行绝对定量分析的方法. 利用该方法进行蛋白定量分析与基于标准蛋白染色的定量分析具有同样的准确度^[18]. 再通过数据校正后, 有效提高了蛋白质定量的范围. 虽然SC方法定量分析蛋白质线性范围大、适用范围广, 但其精确度有限.

目前, APEX指数已被用于拟南芥定量蛋白质组学研究^[44]. emPAI指数也被用于分析昼夜循环中代谢相关酶和核糖体蛋白的绝对含量^[45]. 这类方法已经被应用于苜蓿(*Medicago truncatula*)根瘤抗旱蛋白组研究^[46]、植物磷酸化蛋白信号网络, 以及比较叶肉和维管束鞘中叶绿体蛋白质含量的研究^[47]. 通过与高分辨质谱技术相结合, SC方法在鉴定马铃薯(*Solanum tuberosum*)品种间特异性蛋白质组的变化方面也发挥了很大作用^[48].

1.2 根据多肽离子强度进行蛋白定量的分析方法

离子强度定量的分析方法是, 通过提取肽段离子的色谱图, 从中再提取色谱峰的高度或面积; 使用统计

学方法分析每个多肽色谱峰的离子强度为依据来定量. 理论上, 蛋白质组实验中LC-MS检测到所有的离子都能提取出相应的色谱图, 其峰面积就可以用来做定量比较^[18]. 但在实际定量分析中, 由于不同肽段的离子化效率不同, 只有不同蛋白样品的同一个肽段都检测到才可以相互比较. 当使用总离子强度作为比较时, 这些因素显著增加了数据分析的难度.

由于基于离子强度的定量分析是使用多肽完整的一级质谱图(MS^1)信息比较, 而肽段鉴定则必须使用二级质谱图(MS^2)信息^[49]. 在典型的串联质谱实验中, 一级离子通常要被进一步打碎成片段(MS^2), 这个过程中会丢失部分碎片离子; 其次, 二级质谱分析时会或多或少改变离子在检测过程中的碎片数目, 这些因素均会影响定量分析的准确度. 解决这些问题的策略是多次扫描二级质谱碎片离子(MS^2), 以获取复杂混合样品分析中所需的更多肽段信息, 也就能检测到更多的肽段和蛋白质^[36]. 由于在提高二级质谱肽段数量和增加一级质谱图信息之间会竞争仪器扫描时间, 因此必须在这两个目标之间取得平衡. 所以, 要提高定量的精确度不可避免地会降低蛋白质组鉴定的覆盖度, 反之亦然.

我们团队把经过优化的LFQ方法应用于研究单细胞体系花粉发育分子机理, 探索了花粉管发育和极性生长过程中对蔗糖、硼和钙离子等营养吸收的分子机理^[11]. 此外, 我们也利用LFQ方法结合tandem MOAC技术解析拟南芥根系的营养吸收机制. 通过定量磷酸化蛋白组学技术发现, 硼缺失后植物中大量转运体蛋白

的磷酸化水平被不同程度调控,例如硼转运体蛋白(BOR2和BOR3)^[12]。后期的功能研究发现,硼信号通过激活植物MPK激酶信号通路,从而磷酸化BOR2和BOR3,激活其转运活性来调控物质的运输^[12]。这些研究为探索植物营养吸收的分子机制提供了新的方法和理论依据。

在蛋白质组学实验中,定量分析的前提是不同样品之间都检测到同一个肽段的离子才可以相互比较。然而,同一蛋白混合物在不同批次的LC-MS实验中检测到的肽段离子中大约只有60%重叠,导致两个样本之间共有离子的覆盖度很差。当比较两个以上样本时,这一问题更加明显。为了克服这些缺点,研究人员开发了蛋白质关联分析方法(protein correlation profile)^[50],该方法是根据相同肽段碎片的离子种类(即肽序列检测)和它们在色谱分析过程中的洗脱时间相关联。利用这一特征,在一个样本中鉴定到碎片肽段也可以在另一个样品的质谱数据中找到(即使另外一个样品中没有该碎片信息)。因此,依据稳定的色谱洗脱时间和高质量精确度的肽段离子,就可以在两个样品之间提取离子色谱图用来定量比较。该方法适合于比对不同样本的总离子色谱图。实验证明,使用这种方法能把用于定量分析的蛋白质数量提升至少40%。然而,离子色谱提取的计算时间不一、不同色谱的图谱分析和缺失数据较复杂等特性,都会导致数据分析的难度提高。但是,提高质谱分析样品过程中洗脱时间的重现性,对于提升肽段离子强度的非标记定量分析的准确度至关重要。尽管蛋白质关联分析也广泛用于研究拟南芥等植物磷酸化蛋白质组学^[12]。相比SC方法,该方法通常需要进一步手动检查和核验。但是,当与高分辨率质谱仪联用时,蛋白质关联分析方法可以区分质量相似的共洗脱肽段。因此,相比较而言,基于离子强度定量的方法比SC更灵敏、更准确。

随着仪器分辨率的提升,这两种LFQ方法已成为蛋白质组定量研究的首选方法。因为使用这些策略所定义的每个离子峰的宽度更窄、质量精确度更高,所以鉴定和定量分析的蛋白质数据可信度也更强^[48]。

通过比较不同LFQ方法发现,在定量分析中每个蛋白质若使用4个以上质谱峰用以定量比较,则样品之间的差异性会降低至少两倍,且样品之间蛋白质的相关性也能控制在95%可信度内^[12]。此外,去除样本中的高丰度蛋白质也可以提高定量分析的重复性和线性范围^[39]。基于离子强度的分析方法早期应用于人细胞中

心体的蛋白质组学研究,如今,该方法已经被广泛应用于植物蛋白质定量分析^[46,51~53]。结合生物信息学手段,LFQ还可用于蛋白互作组学的定量研究^[23]。大量的研究表明,这几种LFQ方法在定量蛋白质方面优势互补^[54],因此,研究人员可根据使用经验和研究目的在不同方法间选择。

综上所述,LFQ方法相比标记方法有以下优点^[54,55]:(1)检测范围大。能分析识别更多数量的蛋白质,具有更宽的动态检测范围;(2)成本较低。不需要特殊的介质或昂贵的试剂盒来进行实验;(3)操作简单。不需要确定标签效率^[56];(4)应用范围广。适于定量比较任何生物材料,而标记方法在生物材料应用范围方面有所限制。

2 同位素标记的定量分析方法

使用同位素标记的定量方法原理是,用不同的稳定性同位素试剂标记生物材料或蛋白样品,将标记的样品合并成一个样品后使用质谱分析鉴定(图1)。由于同位素标记的肽段分子量与未标记的肽段分子量有差异,质谱分析时根据分子量差异即可区分肽段所来自的样品,后续数据分析中通过提取质谱中的成对离子来进行蛋白质相对或绝对定量分析。通常标记方法定量蛋白质灵敏度和精确性高,可同时分析多个样品。

常用的稳定性同位素标记方法有两类:代谢物标记和化学标记法(图1)。代谢标记法是在细胞培养过程中标记整个细胞或生物材料,化学标记法则通过化学反应标记蛋白质或胰蛋白酶切后的肽段。此外,还可通过添加带标记的标准合成肽到胰蛋白酶消化的样品中进行靶向定量蛋白质组研究。随着高通量蛋白测序技术重复性的提高和大数据处理方法的提升,这些定量方法在植物生物学研究中也得到广泛应用。

2.1 代谢物标记: ^{15}N 代谢物标记和稳定性同位素标记

代谢物标记的技术有 ^{15}N 代谢物标记(^{15}N -metabolic labeling,通常用 K^{15}NO_3 或 $\text{NH}_4^{15}\text{NO}_3$)和稳定性同位素标记(stable isotope labeling in cell cultures, SILAC)氨基酸(通常为赖氨酸和精氨酸)的技术。代谢标记方法的原理是,使用含有同位素标记的元素加入培养基中培养细胞或植物材料,经过一段时间培养后,细胞里蛋白质的氮元素或特定氨基酸就被稳定性同位素替代。标记的植物材料和没标记的材料混合后提取蛋白质进行

质谱分析. 由于有同位素作为内标, 这种方法可以消除样品制备过程中产生的实验差异, 从而提高蛋白定量的准确性. 因此, 代谢标记方法被认为是蛋白质组定量的黄金技术. 此外, 稳定同位素标记技术还可以用来研究细胞内蛋白质合成和转运过程中在时间和空间上的动态变化.

^{15}N 代谢物标记最初是用于酵母蛋白分析, 后也应用于植物材料中. 具体方法是: 一组培养基里含 ^{15}N 标记的化合物, 另一组含有 ^{14}N 标记的化合物, 培养好的植物混合后提取蛋白质用于质谱分析. 根据同一多肽同位素的质量数不一样, 质谱会产生特定质量差异的多肽对, 多肽离子的信号强度即可用于比较相应蛋白质的表达量^[59]. 该方法定量准确度高, 在细胞材料中更有优势. Yang等人^[60]使用 ^{15}N 代谢物标记方法定量研究乙烯信号通路中磷酸化蛋白质变化, 发现水通道蛋白的磷酸化位点在乙烯信号通路中有很重要的功能.

我们团队使用 ^{15}N 代谢物标记方法报道了拟南芥应对环境胁迫后的信号转导通路变化机理^[13]. 该研究表明, 盐胁迫下钠、钾离子通道蛋白的磷酸化信号变化呈现协同调控模式, 揭示了盐胁迫后引起离子动态平衡适应环境变化的重要性. 我们的研究也证实逆境胁迫后植物通过改变蛋白激酶和光系统组分蛋白亚基的磷酸化程度, 从而影响光合作用机制. 此外, 代谢组学数据证明, 光合作用过程中的代谢物如蔗糖、果糖和葡萄糖等合成在胁迫处理后有不同程度的增加^[13]. 该研究把 ^{15}N 标记技术应用于磷酸化蛋白质组学和代谢组学相结合的多组学定量研究中, 并验证了蛋白磷酸化变化与植物代谢调控之间的相关性, 为进一步阐明植物抵抗逆境胁迫信号转导机理提供了新的理论依据^[13]. ^{15}N 标记技术是最早开发的代谢标记方法, 定量精确度高, 但由于经过同位素标记后的质谱数据比较复杂, 增加了定量分析时数据处理的难度.

稳定性同位素标记技术(SILAC)是在培养基中使用稳定同位素标记的氨基酸(通常是赖氨酸和精氨酸)进行培养细胞或植物, 进而可对蛋白质进行定量分析. 在SILAC蛋白质组实验中, 经过胰蛋白酶消化后的多肽混合物中除C端肽外, 其他片段肽都被已知质量差异的氨基酸标记上. 质谱检测后, 同一个肽段的不同标记会产生质量差异的肽段对^[61]. 这种方法已广泛用于研究动物细胞和酵母的定量蛋白质组学^[62]. 研究报道, SILAC在标记氨基酸缺陷型细菌、动物细胞或酵母菌株上应用效果较好^[63,64]. 在植物材料中, SILAC方法只

能标记上大约70%的蛋白质, 因而不能满足大规模蛋白质组定量研究的需求. 目前, 仅在拟南芥细胞中实现了较高效率的标记^[32].

由于SILAC实验中氨基酸同位素的导入会提高 MS^1 质谱中母离子信息复杂度, 此方法最多只能标记3种元素(三标), 也即只能同时定量分析3种不同蛋白样品. 为了提高定量范围, 根据稳定同位素之间的细微质量差异, 开发出了中子编码SILAC(NeuCode SILAC), 这种方法可以实现多达6个氨基酸的标记(六标)^[63]. 但是, 由于该方法使用质量差为36 mDa甚至6 mDa的母离子进行定量, MS^1 分辨率提高到480k或960k, 远高于常用的30k分辨率, 从而降低了质谱分析时的扫描速度, 限制了定量蛋白质的覆盖深度. Ong^[64]通过添加额外的氨基酸同位素的方法进一步优化了这种标记技术, 可对多达9个不同样品进行高精度的定量分析(九标). SILAC方法定量精确度高, 但大多数应用于动物细胞或酵母研究中(表1).

2.2 化学标记法

化学标记法有基于母离子和基于报告离子定量的两种分析法. 基于母离子分析的方法是通过化学反应, 分别将轻同位素(如 ^{12}C)和重同位素(如 ^{13}C)合成的标签, 连接到蛋白质或者肽段的基团上来实现. 两种标记的肽段样品经过混合后, 带有轻同位素(称为轻链)或重同位素标签(称为重链)相同肽段在质谱检测时能够形成质量差异的多肽对. 之后, 根据每个肽对质谱出峰时的离子强度来提取色谱图进行蛋白质定量分析. 该标记方法应用范围广, 适用于多种样品的定量, 包括动植物细胞、组织和动物体液. 化学标记法中常用的官能基团包括胺基(赖氨酸和氨基末端)和巯基(半胱氨酸).

同位素编码亲和标记(isotope-coded affinity tags, ICAT)是第一个化学标记的商业试剂, 其原理是标签能与肽段中半胱氨酸的巯基发生反应结合而标记多肽, 从而进行蛋白质表达量的比较^[65]. 因而, ICAT不适合分析不含半胱氨酸的蛋白质, 限制了该方法的适用范围. 后来, 研究者根据蛋白质的氨基酸能与甲醛和氰基酸钠特异性相结合这一性质而开发的“二甲基标记法”克服了ICAT的局限性, 这种方法适合于定量分析所有蛋白质. 且该方法具有标记效率高、反应条件温和、标记试剂成本低、定量范围广等优点. 然而, 由于二甲基标记法要求同位素峰产生大于4 Da质量差才能区分重链和轻链, 所以增加了数据分析的复杂性, 后经过改

进后,该方法定量蛋白质能同时标记12个样品。

基于母离子定量分析的化学标记还可以通过酶催化标记(^{18}O 标记)而实现定量分析^[66]。 ^{18}O 标记不仅适用于所有蛋白质组样品,而且易于与其他标记技术结合使用,该方法可用于糖基化修饰多肽的定量分析。尽管可以通过抑制蛋白酶活性的方法来避免 ^{18}O 掺入不完全,但该方法的主要局限性是标记通量低。除了 ^{18}O 标记, Pan等人^[67]使用胰蛋白酶将同位素标记的精氨酸引入到有机相的多肽上,这为引入稳定的同位素用于蛋白质组的相对定量提供了一种新方法。但是,要满足更高通量的蛋白质组定量需求的前提是提高样品的标记效率。

以上几种基于母离子的蛋白质定量方法是通过比较多肽的 MS^1 信息来进行定量。由于母离子具有特定的质荷比(m/z),所以这类方法定量特异性和精度很高。但是,大多数蛋白质组样品复杂程度高,同时洗脱的不同肽在 MS^1 质谱图中具有相同的母离子质量数,并且多肽母离子的信噪比(S/N)低于 MS^2 质谱图中的信噪比。往往低丰度蛋白质定量分析时数据更容易受到干扰峰的影响,导致定量覆盖率低和动态范围狭窄。然而,系统生物学研究中需要同时比较多个样品,因此,能同时标记多个样品且定量准确度高的方法逐渐成为定量蛋白质组学表达分析的首选技术。

通过串联质谱完成定量也即报告离子定量的化学标记法可同时标记多个样品,如等质量标签($\text{ISOBARIC MASS TAGS}$)(图1)。这种方法通过改变报告基团和平衡基团的质量数,设计相同分子量的标签。带不同标签的多肽混合后同时检测,由于来自不同样品带标签的肽段母离子具有相同的质量,在质谱中分子量完全相同,可提高同一个肽段的峰强度。在质谱完成扫描阶段后,进行碰撞诱导解离时,携带不同标签的多肽在 MS^2 分析中断裂生成报告离子和质量平衡基团,定量分析是通过比较报告离子的强度来实现蛋白质定量^[68]。

等质量标签是利用相同质量的标签与蛋白质衍生标记后,相等质量的分子能够在色谱中同时洗脱下来。由于相同质量的标签引入到多肽上,所以即使样品中含有多个标签,携带不同标签的同一种肽段只产生一个一级质谱离子峰(母离子),也即一个 m/z 值。而在二级质谱(MS^2)上可区分带不同标签的肽段。该方法定量是基于肽段碎片离子(MS^2)信息,而不是在一级质谱层面上的定量。所以定量的精确度取决于母离子的分离宽度,母离子分离的碎片宽度越大定量越精确。目前,

TRAQ和TMT是常用的等质量标签方法, iTRAQ已经广泛应用于植物蛋白质组学研究,并且商业化的iTRAQ试剂盒可以同时标记8个样品(八标)^[57]。

相比之下,使用报告离子定量的化学标记法具有以下优点^[36](表1): (1) 应用范围广。化学标记方法能应用于任何生物材料的蛋白质样品的分析。(2) 高通量。分析样品的质量差在数道尔顿和数毫道尔顿之间,因而可同时分析多个样品(商业化的TMT试剂可同时分析11个样品)。使用该方法能有效缩短分析时间,也能用于研究蛋白质动态合成和降解等机理。(3) 高灵敏度。检测蛋白质多肽的灵敏度高,有助于提高蛋白质组定量深度,因而在一个实验中可以定量约10000种蛋白质。(4) 高质量精确度。与母离子定量方法相比,由于报告离子的方法中 MS^2 质谱图的干扰峰较少,因此定量精度更高。(5) 高离子化程度。TMT试剂能提高多肽的离子化效率,从而增加母离子电离化程度和在串联质谱中碎片化程度,也有助于PTM的分析。

综上所述,在以上几种同位素标记方法中,使用报告离子法定量精确度更高,且可用于多个样品的定量。然而,稳定同位素标记法定量分析复杂样品时,由于质量相似的离子会同时从色谱中洗脱下来,从而干扰目标蛋白的检测,最终影响定量方法的准确性。提高定量精度的关键是将目标肽与干扰肽分离,例如,通过延长色谱分析时间或在LC-MS分析之前预分馏样品组分,减少样本复杂度。生物信息学方法也可以解决这个问题,例如,评估 MS^1 质谱数据中的干扰峰和检查母离子的洗脱时间重叠性即可去除干扰混合物的 MS/MS 数据。此外,由新一代质谱仪引入的“离子淌度”(ion mobility)和数据非依赖性(data independent acquisition, DIA)技术可以有效地减少母离子分离时干扰峰的发生^[69,70]。因此,在不影响定量灵敏度和精确度的前提下,在未来的研究中,蛋白质定量的线性范围还可以进一步提高。

3 蛋白质翻译后修饰的定量分析

PTM能调控细胞中蛋白质的功能、定位、降解、合成和信号转导等多个过程。据报道,植物细胞中含有大约200多种不同蛋白质PTM,研究较多的有磷酸化、糖基化、泛素化、甲基化和乙酰化^[12,71]。高通量质谱技术的发展加速了翻译后修饰蛋白质组学在植物生物学中的研究。最近研究表明,使用磷酸化蛋白质组学方法在拟南芥细胞中能够检测到超过40000个磷酸化位点^[10]和数百个糖基化位点^[72,73]。因而,大规模定量分析

蛋白质PTM对研究其在植物生物学中的调控机理具有重要意义。本文所述的非标记定量方法和大多数标记定量方法都适用于蛋白质PTM的定量分析。我们团队^[12]把非标记定量技术与磷酸化蛋白富集的新方法——tandem MOAC相结合,定性和定量分析了植物磷酸化信号网络,并对特异激酶和磷酸化位点进行了功能验证。为定量蛋白质组学方法在翻译后修饰中的应用研究提供了新的思路。

与普通蛋白质组的定量相比,提高蛋白质PTM定量的精确度和灵敏度更有技术难度^[74],主要有以下三方面原因:(1)细胞中修饰的蛋白含量较低,在质谱分析之前需要先富集PTM肽段。(2)可用于定量的质谱数据信息较少。标准蛋白质组学的定量是整合蛋白质所有肽段的数据信息进行定量分析。但蛋白质PTM分析需仅对特定肽段的特定位点进行定量。(3)由于甲基化和乙酰化修饰发生在赖氨酸残基上,二甲基标记和TMT定量方法不能适用于定量分析这类型修饰。Schweppe等人^[58]使用TMT多重标记方法,通过优化MS²水平的定量分析策略,结合离子淌度技术后,在磷酸化蛋白质组研究中很大程度上提高了磷酸化多肽定量精度和覆盖度。

蛋白质PTM的生物学调控功能由其在细胞中的化学计量多少而决定,即PTM修饰位点的数量与蛋白质多肽数总量的比值,也称为位点占有率(site occupancy)^[37]。修饰蛋白质组学定量分析是分析发生修饰的蛋白表达量相对变化和分析修饰氨基酸位点在细胞中的含量。因此,高通量研究修饰蛋白质组学有助于探讨特定修饰位点在植物细胞中的分子功能。在不同修饰组学研究中,磷酸化蛋白质组学研究表明,一些重要的功能氨基酸位点常发生磷酸化,其位点占有率也高,而赖氨酸乙酰化的位点占有率在细胞中相对较低^[9]。因此,当分析乙酰化修饰时,100%化学乙酰化肽和自然乙酰化肽之间的通量差异可能太大而无法准确定量。此外,修饰蛋白质组学也可以在单细胞状态下用蛋白酶处理和同位素标记进行分析。

4 定量蛋白质组学研究中的数据分析

定量蛋白质组学实验旨在分析对照组与实验组植物样品间蛋白质表达量的变化差异;或者分析蛋白质复合物或特定亚蛋白质组的变化差异。而数据分析的目的是寻找“偏离”数据,也就是在所研究的生物学背景下表达量发生变化或被诱导的蛋白质^[75]。当涉及蛋

白复合物或不同亚细胞器定位的蛋白时,与其他表达量不一样的蛋白质相比,表达量具有相似变化趋势的蛋白质通常是来自同一复合物或亚细胞蛋白质组分。由于植物材料的细胞结构和蛋白成分较动物细胞复杂,增加了蛋白质数据分析时的难度。通常来说,使用已知细胞器的marker蛋白或亚细胞器蛋白数据库作为参考能有效提高数据分析的精确性。

无论选择任何一种定量方法,定量蛋白质组学都会面临数据不完整的困难。即使是最先进的质谱技术也不可能采集到所有肽段的离子,往往最终只能获得样本中一部分肽段和蛋白质的信息。此外,受多种因素影响,能够实际用于定量分析的多肽和蛋白质比实验中检测到的少得多^[10]。因此,为了获取高质量的蛋白质组数据,必须在蛋白质分离、富集和纯化等方面进行合理的实验设计。在实验设计时,需要慎重考虑以下几方面的问题:(1)影响肽段定量精确度的因素。例如,仪器的分辨率、灵敏度、扫描速度和色谱仪(影响峰宽度)都会影响蛋白定量的精确性。因此,使用高分辨率和高精度精确度的质谱仪,可以在离子扫描过程中获得更窄峰宽的质谱图谱,以减少背景干扰峰(噪音),从而提高蛋白质检测和定量分析的可信度。(2)质谱检测的出峰数量。肽段水平或报告离子的定量准确性取决于二级质谱出峰的数量,质谱检测的峰越多定量数据可信度越高。(3)单个肽段的离子强度。离子强度越高的肽段定量越准确,而离子强度低的肽段会导致数据误差更大。(4)使用高分辨质谱能减少样品分析过程中的干扰峰。(5)重链和轻链标记的肽段在色谱分离时的离子化效率。稳定性同位素标记方法中重链和轻链标记的肽段对在色谱分离时会有不同的离子化效率,从而影响肽段离子强度的比较分析。因而,必须分别确定每个肽段的峰面积,与提取的离子色谱图的总峰面积计算比值。(6)低丰度蛋白在检测谱图计数少,有时会导致SC方法定量分析时计数为0,但在另一样品中可能会检测到,从而影响定量的准确性,并引起数据中有缺失值。在稳定同位素标记中也会发生仅检测到肽对中的其中一个(如仅标记的肽或仅未标记的肽),导致无法计算比值,并且不能准确确定单个肽段的实际丰度。(7)离子数据的质量程度。蛋白质定量分析的准确性主要取决于质谱检测到的肽段离子数据的质量程度,还取决于每个需要定量的蛋白质的理论肽段数量和实验酶切肽段数量。一般情况下,质谱检测到的肽段数量越多定量准确度越高。为避免错误率,仅一个肽段离子进行

定量分析时, 必须特别验证。

在稳定性同位素标记的实验中, 可以使用反相重复实验来减少由于实验标记和样品之间的差异性而引入的实验误差。 ^{15}N 标记的反向重复实验已用于定量研究植物细胞磷酸化蛋白质变化^[13]和脂筏蛋白变化^[76], 该方法很大程度上提高了蛋白质定量的精确度。

5 总结和展望

定量蛋白质分析技术的应用有助于人们深入了解复杂生物系统中蛋白质的分子功能和调控机制。随着质谱仪的采集速度、分辨率和灵敏度的巨大提高, 定量蛋白质组方法在定量的准确性、精确度、多重能力和覆盖度等方面得到了迅速发展。不同方法中, 用于定量的离子特性是其定量性能的先决条件。近年来, 使用报告离子和离子强度比值定量的方法得到了大力发展和广泛应用, 并在定量准确度、精密度和动态范围方面具有很强的优越性。此外, 我们前期综述了基于数据非依赖性采集(data-independent acquisition, DIA)的

SWATH技术也日益成为定量蛋白质组学研究的核心技术^[2]。然而, 由于细胞中蛋白质具有多样性和高度动态性等特性, 任何一种蛋白质定量方法都面临很大挑战。因此, 需要研究针对特定蛋白质性质的定量方法, 如分析新合成蛋白质、蛋白质周转率、乙酰化、棕榈酰化翻译后修饰等较高和较宽动态范围的定量方法; 提高定量覆盖率和多重能力也能提高蛋白定量的应用范围。在开发定量方法时, 需要全面掌握权衡不同特性, 包括定量覆盖度、多重能力、准确度、精密度、动态范围和仪器分析时间等。

另外, 由于质谱仪分析的随机性, 会导致样品中的部分肽段检测不到。例如, 低丰度肽和离子化效率低的肽经常仅在个别样品中检测到, 这些因素均会导致定量数值的缺失。因此, 定量蛋白质组研究的最大挑战是解决肽段缺失值而导致的蛋白质组覆盖不完全。在今后的研究中, 开发有效的生物信息学方法可填充缺失值; 改进多肽的分离方法、预分馏多肽和优化色谱方法均能够提高蛋白质定量的覆盖率。

参考文献

- Müller J B, Geyer P E, Colaço A R, et al. The proteome landscape of the kingdoms of life. *Nature*, 2020, 582: 592–596
- Gao F Y, Chen Y M. Decoding protein phosphorylation signaling networks by mass spectrometry (in Chinese). *Chin Sci Bull*, 2021, 66: 2529–2541 [高丰衣, 陈艳梅. 质谱技术探索蛋白磷酸化信号机制的研究进展. *科学通报*, 2021, 66: 2529–2541]
- Zhang W, Liu J, Zhang Y, et al. A high-quality genome sequence of alkaligrass provides insights into halophyte stress tolerance. *Sci China Life Sci*, 2020, 63: 1–14
- Xiong Y, McCormack M, Li L, et al. Glucose-TOR signalling reprograms the transcriptome and activates meristems. *Nature*, 2013, 496: 181–186
- Menz J, Li Z, Schulze W X, et al. Early nitrogen-deprivation responses in *Arabidopsis* roots reveal distinct differences on transcriptome and (phospho-) proteome levels between nitrate and ammonium nutrition. *Plant J*, 2016, 88: 717–734
- Schmollinger S, Mühlhaus T, Boyle N R, et al. Nitrogen-sparing mechanisms in *Chlamydomonas* affect the transcriptome, the proteome, and photosynthetic metabolism. *Plant Cell*, 2014, 26: 1410–1435
- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*, 2003, 422: 198–207
- Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. *Nature*, 2016, 537: 347–355
- Chen Y, Weckwerth W. Mass spectrometry untangles plant membrane protein signaling networks. *Trends Plant Sci*, 2020, 25: 930–944
- Mergner J, Frejno M, List M, et al. Mass-spectrometry-based draft of the *Arabidopsis* proteome. *Nature*, 2020, 579: 409–414
- Chen Y, Liu P, Hoehenwarter W, et al. Proteomic and phosphoproteomic analysis of *Picea wilsonii* pollen development under nutrient limitation. *J Proteome Res*, 2012, 11: 4180–4190
- Chen Y, Hoehenwarter W. Rapid and reproducible phosphopeptide enrichment by tandem metal oxide affinity chromatography: Application to boron deficiency induced phosphoproteomics. *Plant J*, 2019, 98: 370–384
- Chen Y, Hoehenwarter W. Changes in the phosphoproteome and metabolome link early signaling events to rearrangement of photosynthesis and central metabolism in salinity and oxidative stress response in *Arabidopsis*. *Plant Physiol*, 2015, 169: 3021–3033
- Chen Y, Chen T, Shen S, et al. Differential display proteomic analysis of *Picea meyeri* pollen germination and pollen-tube growth after inhibition of actin polymerization by latrunculin B. *Plant J*, 2006, 47: 174–195
- Chen T, Wu X, Chen Y, et al. Combined proteomic and cytological analysis of Ca^{2+} -calmodulin regulation in *Picea meyeri* pollen tube growth. *Plant Physiol*, 2009, 149: 1111–1126
- Fey S J, Larsen P M. 2D or not 2D. *Curr Opin Chem Biol*, 2001, 5: 26–33

- 17 Fenn J B, Mann M, Meng C K, et al. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 1989, 246: 64–71
- 18 Schulze W X, Usadel B. Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol*, 2010, 61: 491–516
- 19 Washburn M P, Wolters D, Yates III J R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol*, 2001, 19: 242–247
- 20 de Godoy L M F, Olsen J V, Cox J, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, 2008, 455: 1251–1254
- 21 Old W M, Meyer-Arendt K, Aveline-Wolf L, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics*, 2005, 4: 1487–1502
- 22 Ludwig C, Gillet L, Rosenberger G, et al. Data-independent acquisition-based SWATH-MS for quantitative proteomics: A tutorial. *Mol Syst Biol*, 2018, 14: e8126
- 23 Van Leene J, Han C, Gadeyne A, et al. Capturing the phosphorylation and protein interaction landscape of the plant tor kinase. *Nat Plants*, 2019, 5: 316–327
- 24 Reiland S, Finazzi G, Endler A, et al. Comparative phosphoproteome profiling reveals a function of the STN8 kinase in fine-tuning of cyclic electron flow (CEF). *Proc Natl Acad Sci USA*, 2011, 108: 12955–12960
- 25 Koh J, Chen G, Yoo M J, et al. Comparative proteomic analysis of *Brassica napus* in response to drought stress. *J Proteome Res*, 2015, 14: 3068–3081
- 26 Uhrig R G, Schlöpfer P, Roschitzki B, et al. Diurnal changes in concerted plant protein phosphorylation and acetylation in *Arabidopsis* organs and seedlings. *Plant J*, 2019, 99: 176–194
- 27 Goto C, Hashizume S, Fukao Y, et al. Comprehensive nuclear proteome of *Arabidopsis* obtained by sequential extraction. *Nucleus*, 2019, 10: 81–92
- 28 Kar U K, Simonian M, Whitelegge J P. Integral membrane proteins: Bottom-up, top-down and structural proteomics. *Expert Rev Proteomics*, 2017, 14: 715–723
- 29 Stanislas T, Bouyssié D, Rossignol M, et al. Quantitative proteomics reveals a dynamic association of proteins to detergent-resistant membranes upon elicitor signaling in tobacco. *Mol Cell Proteomics*, 2009, 8: 2186–2198
- 30 Arrivault S, Guenther M, Florian A, et al. Dissecting the subcellular compartmentation of proteins and metabolites in *Arabidopsis* leaves using non-aqueous fractionation. *Mol Cell Proteomics*, 2014, 13: 2246–2259
- 31 Hsu J L, Wang L Y, Wang S Y, et al. Functional phosphoproteomic profiling of phosphorylation sites in membrane fractions of salt-stressed *Arabidopsis thaliana*. *Proteome Sci*, 2009, 7: 42
- 32 Schutz W, Hausmann N, Krug K, et al. Extending SILAC to proteomics of plant cell lines. *Plant Cell*, 2011, 23: 1701–1705
- 33 Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem*, 2011, 80: 273–299
- 34 Kadota Y, Liebrand T W H, Goto Y, et al. Quantitative phosphoproteomic analysis reveals common regulatory mechanisms between effector- and PAMP-triggered immunity in plants. *New Phytol*, 2019, 221: 2160–2175
- 35 Ankney J A, Muneer A, Chen X. Relative and absolute quantitation in mass spectrometry-based proteomics. *Annu Rev Anal Chem*, 2018, 11: 49–77
- 36 Ankney J A, Muneer A, Chen X. Relative and absolute quantitation in mass spectrometry-based proteomics. *Annu Rev Anal Chem*, 2018, 11: 49–77
- 37 Millar A H, Heazlewood J L, Giglione C, et al. The scope, functions, and dynamics of posttranslational protein modifications. *Annu Rev Plant Biol*, 2019, 70: 119–151
- 38 Cox J, Hein M Y, Lubner C A, et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*, 2014, 13: 2513–2526
- 39 Chen Y, Hoehenwarter W, Weckwerth W. Comparative analysis of phytohormone-responsive phosphoproteins in *Arabidopsis thaliana* using TiO₂-phosphopeptide enrichment and mass accuracy precursor alignment. *Plant J*, 2010, 63: 1–17
- 40 Weckwerth W, Wienkoop S, Hoehenwarter W, et al. From proteomics to systems biology: Mapa, mass western, promex, and covain as a user-oriented platform. *Methods Mol Biol*, 2014, 1072: 15–27
- 41 Hoehenwarter W, Larhlmi A, Hummel J, et al. Mapa distinguishes genotype-specific variability of highly similar regulatory protein isoforms in potato tuber. *J Proteome Res*, 2011, 10: 2979–2991
- 42 Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (EMPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*, 2005, 4: 1265–1272
- 43 Lu P, Vogel C, Wang R, et al. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol*, 2007, 25: 117–124
- 44 Baerenfaller K, Grossmann J, Grobei M A, et al. Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics.

- Science, 2008, 320: 938–941
- 45 Piques M, Schulze W X, Höhne M, et al. Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in *Arabidopsis*. *Mol Syst Biol*, 2009, 5: 314
- 46 Larrainzar E, Wienkoop S, Weckwerth W, et al. *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol*, 2007, 144: 1495–1507
- 47 Majeran W, Zybailov B, Ytterberg A J, et al. Consequences of C₄ differentiation for chloroplast membrane proteomes in maize mesophyll and bundle sheath cells. *Mol Cell Proteomics*, 2008, 7: 1609–1638
- 48 Hoehenwarter W, van Dongen J T, Wienkoop S, et al. A rapid approach for phenotype-screening and database independent detection of CSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics*, 2008, 8: 4214–4225
- 49 Marshall A G, Hendrickson C L. High-resolution mass spectrometers. *Annu Rev Anal Chem*, 2008, 1: 579–599
- 50 Krahmer N, Hilger M, Kory N, et al. Protein correlation profiles identify lipid droplet proteins with high confidence. *Mol Cell Proteomics*, 2013, 12: 1115–1126
- 51 Haj Ahmad F, Wu X N, Stintzi A, et al. The systemin signaling cascade as derived from time course analyses of the systemin-responsive phosphoproteome. *Mol Cell Proteomics*, 2019, 18: 1526–1542
- 52 Hoehenwarter W, Thomas M, Nukarinen E, et al. Identification of novel *in vivo* map kinase substrates in *Arabidopsis thaliana* through use of tandem metal oxide affinity chromatography. *Mol Cell Proteomics*, 2013, 12: 369–380
- 53 Niittylä T, Fuglsang A T, Palmgren M G, et al. Temporal analysis of sucrose-induced phosphorylation changes in plasma membrane proteins of *Arabidopsis*. *Mol Cell Proteomics*, 2007, 6: 1711–1726
- 54 Dowle A A, Wilson J, Thomas J R. Comparing the diagnostic classification accuracy of iTRAQ, peak-area, spectral-counting, and emPAI methods for relative quantification in expression proteomics. *J Proteome Res*, 2016, 15: 3550–3562
- 55 Park S K, Venable J D, Xu T, et al. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat Methods*, 2008, 5: 319–322
- 56 Tyanova S, Temu T, Cox J. The maxquant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*, 2016, 11: 2301–2319
- 57 Zeng W Y, Sun Z D, Cai Z Y, et al. Proteomic analysis by iTRAQ-MRM of soybean resistance to lamprosema indicate. *BMC Genomics*, 2017, 18: 444
- 58 Schweppe D K, Rusin S F, Gygi S P, et al. Optimized workflow for multiplexed phosphorylation analysis of TMT-labeled peptides using high-field asymmetric waveform ion mobility spectrometry. *J Proteome Res*, 2020, 19: 554–560
- 59 Engelsberger W R, Erban A, Kopka J, et al. Metabolic labeling of plant cell cultures with K¹⁵NO₃ as a tool for quantitative analysis of proteins and metabolites. *Plant Methods*, 2006, 2: 14
- 60 Yang Z, Guo G, Zhang M, et al. Stable isotope metabolic labeling-based quantitative phosphoproteomic analysis of *Arabidopsis* mutants reveals ethylene-regulated time-dependent phosphoproteins and putative substrates of constitutive triple response 1 kinase. *Mol Cell Proteomics*, 2013, 12: 3559–3582
- 61 Cox J, Matic I, Hilger M, et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc*, 2009, 4: 698–705
- 62 de Godoy L M F, Olsen J V, de Souza G A, et al. Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol*, 2006, 7: R50
- 63 Geiger T, Cox J, Ostasiewicz P, et al. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat Methods*, 2010, 7: 383–385
- 64 Ong S E. The expanding field of SILAC. *Anal Bioanal Chem*, 2012, 404: 967–976
- 65 Hagglund P, Bunkenborg J, Maeda K, et al. Identification of thioredoxin disulfide targets using a quantitative proteomics approach based on isotope-coded affinity tags. *J Proteome Res*, 2008, 7: 5270–5276
- 66 Nelson C J, Hegeman A D, Harms A C, et al. A quantitative analysis of *Arabidopsis* plasma membrane using trypsin-catalyzed ¹⁸O labeling. *Mol Cell Proteomics*, 2006, 5: 1382–1395
- 67 Pan C, Kumar C, Bohl S, et al. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol Cell Proteomics*, 2009, 8: 443–450
- 68 Rauniyar N, Yates III J R. Isobaric labeling-based relative quantification in shotgun proteomics. *J Proteome Res*, 2014, 13: 5293–5309
- 69 Gabelica V, Shvartsburg A A, Afonso C, et al. Recommendations for reporting ion mobility mass spectrometry measurements. *Mass Spec Rev*, 2019, 38: 291–320
- 70 Röst H L, Rosenberger G, Navarro P, et al. Openswath enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotechnol*, 2014, 32: 219–223
- 71 Wong M M, Bhaskara G B, Wen T N, et al. Phosphoproteomics of *Arabidopsis* highly ABA-induced1 identifies AT-Hook-Like10 phosphorylation required for stress growth regulation. *Proc Natl Acad Sci USA*, 2019, 116: 2354–2363

- 72 Scheys F, Van Damme E J M, Pauwels J, et al. N-glycosylation site analysis reveals sex-related differences in protein N-glycosylation in the rice brown planthopper (*Nilaparvata lugens*). [Mol Cell Proteomics](#), 2020, 19: 529–539
- 73 Ruiz-May E, Hucko S, Howe K J, et al. A comparative study of lectin affinity based plant N-glycoproteome profiling using tomato fruit as a model. [Mol Cell Proteomics](#), 2014, 13: 566–579
- 74 Bekker-Jensen D B, Bernhardt O M, Högberg A, et al. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. [Nat Commun](#), 2020, 11: 787
- 75 McWhite C D, Papoulas O, Drew K, et al. A pan-plant protein complex map reveals deep conservation and novel assemblies. [Cell](#), 2020, 181: 460–474.e14
- 76 Kierszniowska S, Seiwert B, Schulze W X. Definition of *Arabidopsis* sterol-rich membrane microdomains by differential treatment with methyl- β -cyclodextrin and quantitative proteomics. [Mol Cell Proteomics](#), 2009, 8: 612–623

Summary for “植物标记和非标记定量蛋白质组学技术”

Labeling and label-free quantitative proteomics in plant biology

Fengyi Gao^{1,2}, Wenna Zhang³, Huimin Xu¹, Xinlin Liang¹ & Yanmei Chen^{1*}¹ State Key Laboratory of Plant Physiology and Biochemistry, College of Biological Sciences, China Agricultural University, Beijing 100193, China;² School of Biology and Food, Shangqiu Normal University, Shangqiu 476000, China;³ College of Horticulture, China Agricultural University, Beijing 100193, China* Corresponding author, E-mail: chenyanmei@cau.edu.cn

Proteins not only construct cellular structures but also play regulatory functions in plant development and growth. Accurate quantification of protein dynamics, turnover rates, and post-translational modifications provide crucial insights into the molecular mechanisms of biological systems. While early studies only consisted of listing proteins identified in given organs or under specific conditions. There is growing need to apply quantitative proteomic techniques to be biased or unbiased studies of plant protein dynamic changes. With rapid improvements in mass spectrometry (MS) technologies, various proteomic quantification methods have been developed and widely used in biological research. In the meantime, MS-based quantitative proteomics has emerged as a powerful approach for gaining insights into functional aspects of proteins in biological systems. MS is not inherently quantitative; thus, several strategies have been developed to assist MS for protein quantification. These strategies can be distinguished not only by the method of quantification, such as labeled or label-free method, but also by the quantitative goal, such as relative or absolute quantification. They can be also distinguished by the scopes of the experiments performed, such as unbiased or targeted studies.

Even within these different research goals, numerous options exist, since each strategy has advantages and disadvantages. For example, labeling methods can tag proteins/peptides with chemical, metabolic, or enzymatic reagents, whereas label-free quantitation methods use spectra ion counting, or peak intensity-based analysis. In comparison with labeling strategies, label-free methods are widely used because of the dynamic quantitative range, costing less, and without labeling steps. However, label-free methods are limited in the sample multiplex capacity and quantitative accuracy. For labeling-based quantitation, materials or samples are labeled by different isotope reagents, the paired MS spectra ions are extracted for relative or absolute quantification. In combination with high accuracy MS instruments, various labeling strategies can be used in quantitative proteome analysis, these methods can be distinguished by quantification ions, such as precursor ion-based, reporter ion-based, peptide-coupled reporter ion-based, and fragment ion-based quantitation. For stable isotope labeled methods, proteins or peptides labeled with different stable isotopes are combined into one vial for LC-MS analysis. These strategies have gained popularity due to the diversity and flexibility of stable isotope reactions, the power to analyze multiple samples simultaneously, and the high quantification accuracy. In this review, we discuss the most widely used label-free and stable isotope-labeling-based proteomic strategies for relative and absolute quantitative proteomics in plants. We compare their strengths and weaknesses in terms of quantitation accuracy, proteome coverage and robustness. We summarize the most promising recent developments and their applications for solving plant biological complexities.

proteome, protein phosphorylation, label-free, stable isotope-labeling, cell membranedoi: [10.1360/TB-2020-1506](https://doi.org/10.1360/TB-2020-1506)