

基于质谱的单细胞蛋白质组学分析

范智瑞¹, 方群^{1,2}, 杨奕^{1,2}

(1. 浙江大学化学系, 杭州 310058; 2. 浙江大学杭州国际科创中心, 杭州 311215)

摘要 单细胞蛋白质组学分析能揭示细胞个体之间蛋白质的精细差异, 在诸多重要领域具有重要的应用价值, 已成为目前的研究热点; 其难点在于单细胞内的蛋白质极其微量, 需要解决样品处理过程中的损失问题、色谱质谱检测的灵敏度问题和低信号强度质谱数据的解析利用问题。本文从单细胞分选、样品处理、色谱质谱采集和数据分析等方面, 综合评述了目前基于质谱的单细胞蛋白质组学分析方法的研究进展及其在生物医学领域的应用, 并展望了其未来的发展前景。

关键词 质谱; 蛋白质组学; 单细胞; 微流控; 数据非依赖性采集

中图分类号 O657; Q516 文献标志码 A doi: 10.7503/cjcu20240294

Single Cell Proteomic Analysis by Mass Spectrometry

FAN Zhirui¹, FANG Qun^{1,2*}, YANG Yi^{1,2*}

(1. Department of Chemistry, Zhejiang University, Hangzhou 310058, China;

2. ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou 311215, China)

Abstract Single-cell proteomics allows revealing precisely the differences of proteins between individual cells, which has become a research hotspot showing indispensable application value in many important fields. Its difficulties lie in the fact that the proteins in a single cell are of extremely low abundance, which calls for ingenious solutions to the problems of sample loss during preparation, low sensitivity of chromatography-mass spectrometry detection, and insufficient analysis of spectral data with low signal intensities. This review summarizes the current research progress of mass spectrometry-based single-cell proteomic analysis, including single-cell sorting, sample preparation, chromatography-mass spectrometry acquisition, and data analysis, as well as its applications in biomedical fields. Its potential future development is also discussed.

Keywords Mass spectrometry; Proteomics; Single-cell; Microfluidics; Data-independent acquisition

蛋白质是生命活动的主要承担者, 对细胞中蛋白质种类、含量变化的精确认知是理解其在生理和病理过程中功能的先决条件。基于质谱的蛋白质组学技术可对细胞和组织中的成千上万种蛋白质进行全面的定性和定量分析^[1]。然而, 传统的蛋白质组学分析手段获得的是大量细胞群体的平均化结果, 忽视或掩盖了不同细胞个体的异质性。尽管单细胞基因组学和转录组学研究揭示了细胞异质性的存在^[2,3], 但在很多场景中 mRNA 水平无法直接反映蛋白质水平, 因此获取多层次组学数据对于完整地理解生物学过程必不可少^[4]。实现单个细胞蛋白质组的定性定量分析, 揭示细胞个体之间的精细差异, 从而深入了解细胞中蛋白质的生物功能, 在肿瘤异质性、免疫治疗、高通量药物筛选^[5]以及细胞分化和发育^[6]等领域具有重要的应用价值。

收稿日期: 2024-06-19. 网络首发日期: 2024-08-19.

联系人简介: 杨奕, 男, 博士, 主要从事质谱和蛋白质组信息学方面的研究. E-mail: y_yi@zju.edu.cn

方群, 男, 博士, 教授, 主要从事微流控、色谱、质谱和单细胞多组学方面的研究. E-mail: fangqun@zju.edu.cn

基金项目: 国家自然科学基金(批准号: 22304153, 22234007, 21827806)资助。

Supported by the National Natural Science Foundation of China(Nos.22304153, 22234007, 21827806).

单细胞蛋白质组学分析的难点主要在于单细胞内蛋白质极其微量,且不同种类蛋白质的动态范围极高.与核酸不同的是,蛋白质不能扩增,这给单细胞水平的分析带来难度.单细胞分离和样品制备过程中的样品损失问题、色谱质谱检测的灵敏度问题和低信号强度质谱数据的解析利用问题,使单细胞蛋白质组学研究面临巨大挑战.近年来,研究人员从单细胞分选、样品处理、色谱质谱采集和数据解析等方面努力突破,推动了单细胞蛋白质组学的发展.本文综合评述了目前基于质谱的单细胞蛋白质组学分析方法的研究进展及其在生物医学领域的应用,并展望了其未来的发展前景.

1 单细胞分选和样品处理方法

1.1 单细胞分选技术

单细胞蛋白质组学面临的首个技术挑战在于从生物样品中分离出单个细胞.目前,单细胞分选策略主要分为基于显微操作的方法、基于流式细胞术的方法、基于微流控芯片的方法和基于激光显微切割的方法.

显微操作是最简单直观的单细胞分选方法.将细胞悬液置于显微镜下,利用口吸管或显微操作仪即可完成可视化的单细胞挑选操作,但此方法需要熟练的操作者,易受人工干扰,且分选通量较低.为此,研究人员基于显微操作技术发展了自动化单细胞捕获与分选系统. Nelep 等^[7]发展的 CellCelector 可自动对细胞样品进行扫描,根据明场和荧光图像判断细胞形态和表面抗体的表达情况,筛选出目标细胞进行挑取和分离. Wang 等^[8,9]发展了基于机器视觉图像识别和微纳液体处理机器人的自动化单细胞分选系统,其可自动完成细胞阵列的扫描成像、目标细胞识别和捕获,以及单细胞液滴的生成.上述方法记录的显微图像保留了所挑取细胞的位置信息,可确认目标挑取正确且为单个细胞[图 1(A)].将自动化系统与多步分选策略结合,还可实现大量细胞样品中稀有细胞的分选^[8].

基于流式细胞术的单细胞分选方法主要采用荧光激活细胞分选(Fluorescence-activated cell sorting, FACS)技术,其主要原理是利用针对特定细胞类型的抗体偶联荧光染料标记细胞,引导单个细胞逐一通过光源,根据散射光和荧光信号对单个细胞进行多参数分析与分选[图 1(B)]^[10,11]. FACS 分选快速、可靠稳定,具有较高通量,是当前最成熟的细胞分选方法.但是该方法所需样品的初始体积较大,难以处理小体积或微量样品,且分选过程可能会损伤细胞^[12]. cellenONE 自动化细胞分选系统可实现低容量样品中细胞的温和分选. cellenONE 的基本工作流程是将细胞悬液吸入毛细管探针中,利用光学成像装置观察探针尖端的细胞种类和数量,根据细胞位置是否满足单细胞条件决定液滴的分配去向,从而精确地获得单细胞液滴. Vallone 等^[13]报道了使用 cellenONE 对人多能干细胞进行单细胞分选的操作方法.

微流控技术可在微观尺度上利用流体动力以及电、磁、声、光等外力对细胞进行精确操控.微流控装置中微米级的微通道或微结构与细胞大小匹配,其设计十分灵活,可以满足不同单细胞操控的要求,诸多基于微流控的单细胞分选方法已被报道^[14,15]. Gebreyesus 等^[16]发展了一体化单细胞蛋白质组芯片 iProChip,其内置细胞捕获和计数结构可捕获精确数量的细胞,并可在芯片内完成蛋白质组样品处理[图 1(C)]. Hu 等^[17,18]发展了一种基于大面积薄膜晶体管开关阵列的有源数字微流控平台 AM-DMF,结合电润湿数字微流控、薄膜晶体管和显示平板技术,利用电场和程序化的驱动电信号控制表面张力,实现了液滴的可编程操控、细胞悬液载入和分配以及单细胞液滴的生成和操控,并在芯片内完成单细胞蛋白质组样品的制备[图 1(D)].

与上述用于细胞悬液样品的方法不同,激光显微切割(Laser microdissection, LMD)是一种从实体组织样品中分离目标细胞的技术. LMD 利用激光切割组织切片中特定区域,然后利用重力或激光压力弹射分离和收集目标区域^[19]. 激光捕获显微切割(Laser capture microdissection, LCM)技术是 LMD 的一种变体,其采用低功率红外激光激活热塑性膜,使膜变黏并黏附目标区域,从而实现分离^[19,20]. LMD 和 LCM 技术最大的优势在于其可以从组织中直接提取目标细胞,保留单个细胞的空间位置信息.但该方法易受到相邻细胞的污染,且激光切割过程可能会破坏细胞的完整性. Mund 等^[21]提出了深度视觉蛋

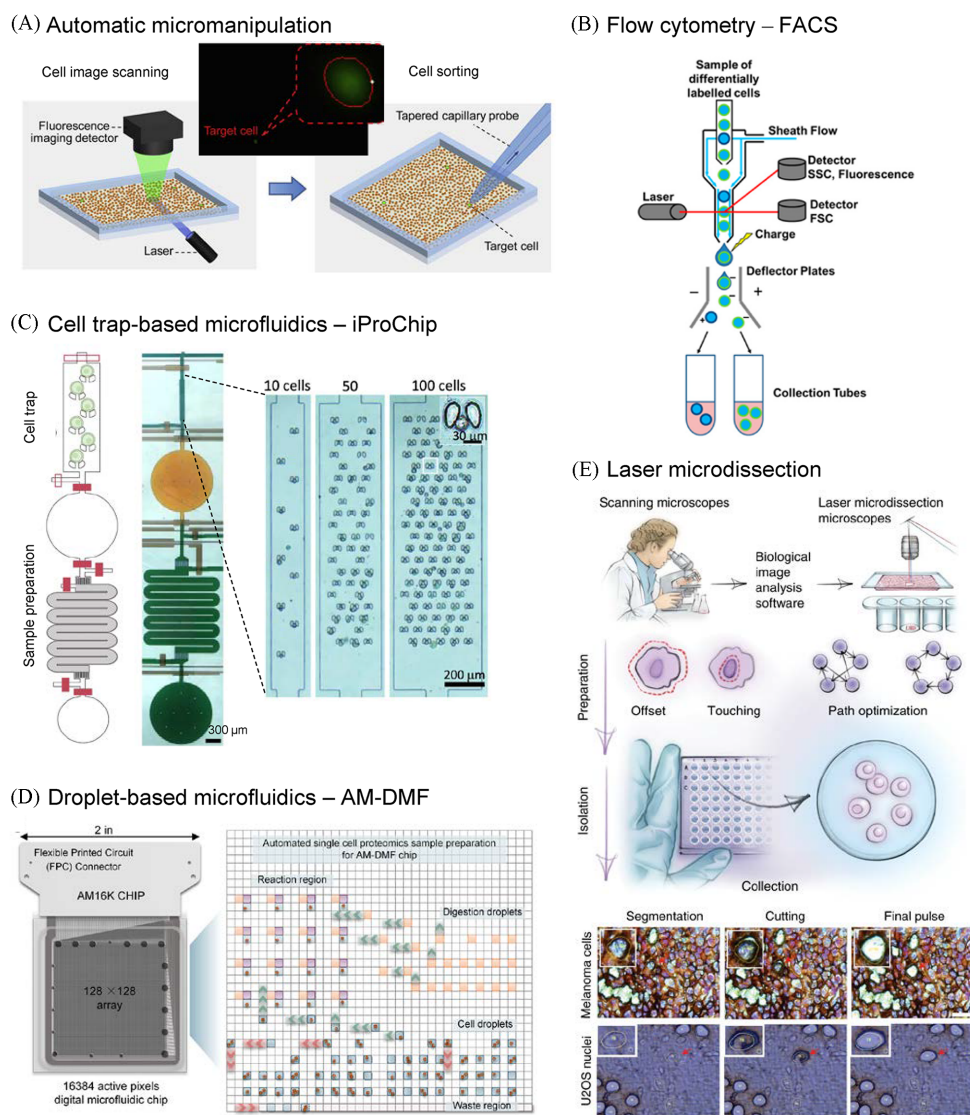


Fig. 1 Single cell sorting methods

(A) Automatic micromanipulation by machine vision^[8]; (B) fluorescence-activated cell sorting (FACS)^[12]; (C) integrated proteomics chip (iProChip)^[16]; (D) active matrix digital microfluidic chip (AMD-MF)^[18]; (E) deep visual proteomics by artificial intelligence-guided cell classification and laser microdissection^[21]. (A) Copyright 2021, Elsevier; (B) Copyright 2015, MDPI; (C) Copyright 2021, Springer Nature; (D) Copyright 2024, American Chemical Society; (E) Copyright 2022, Springer Nature.

蛋白质组学 (Deep visual proteomics) 的概念, 将人工智能显微图像分析、自动化激光显微切割及超高灵敏度的质谱分析结合, 在保留空间背景的同时实现了蛋白质丰度与细胞表型的关联 [图 1(E)]. 该课题组^[22]进一步发展了单细胞分辨率的深度视觉蛋白质组学方法, 发现了肝脏组织中单细胞的蛋白质组空间异质性信息.

1.2 单细胞蛋白质组样品处理方法

单细胞蛋白质组样品处理的关键在于如何减少样品损失并提高蛋白质酶解效率. 为此, 研究人员发展了多种针对单细胞蛋白质样品的处理方法, 如构建毛细管、微流控芯片和液滴反应器等, 以达到最小化样品反应体积、减少蛋白质吸附损失及提高反应效率的目的.

毛细管的内径和体积与单细胞样品匹配, 其在提取、处理和检测单细胞中蛋白质方面的发展起步较早, 但早期的单细胞蛋白质分析方法的分离和检测能力有限^[23].

2015年, Zhang等^[24]发展了一种集成蛋白质组分析装置 iPAD-100, 实现了100个细胞的蛋白质组分析. 在此基础上, 该课题组^[25]发展了用于单细胞分析的集成装置 iPAD-1 [图 2(A)]. 该装置先将选中

的单个细胞吸入毛细管中，在线进行细胞裂解和蛋白质酶解；然后，将酶解后的样品直接转移至纳升级色谱柱中进行色谱-质谱分析，最多可从单个HeLa细胞中鉴定出328种蛋白质。

2023年，He等^[26]开发了一种毛细管烷基化微反应器，其可在还原烷基化的同时将蛋白质共价捕获至毛细管微反应器内壁。当洗脱样品中代谢物后，可在毛细管内对捕获的蛋白质进行酶解，从而实现同一个单细胞的蛋白质组和代谢组分析。该方法可从单个小鼠卵母细胞中鉴定出近3500种蛋白质和170余种代谢物。

此外，毛细管既可作为样品处理的反应器，又可作为毛细管电泳的分离通道。在同一根毛细管中进行在线样品处理和分离，可避免样品转移导致的样品损失。此策略已被应用于单个哺乳动物细胞中

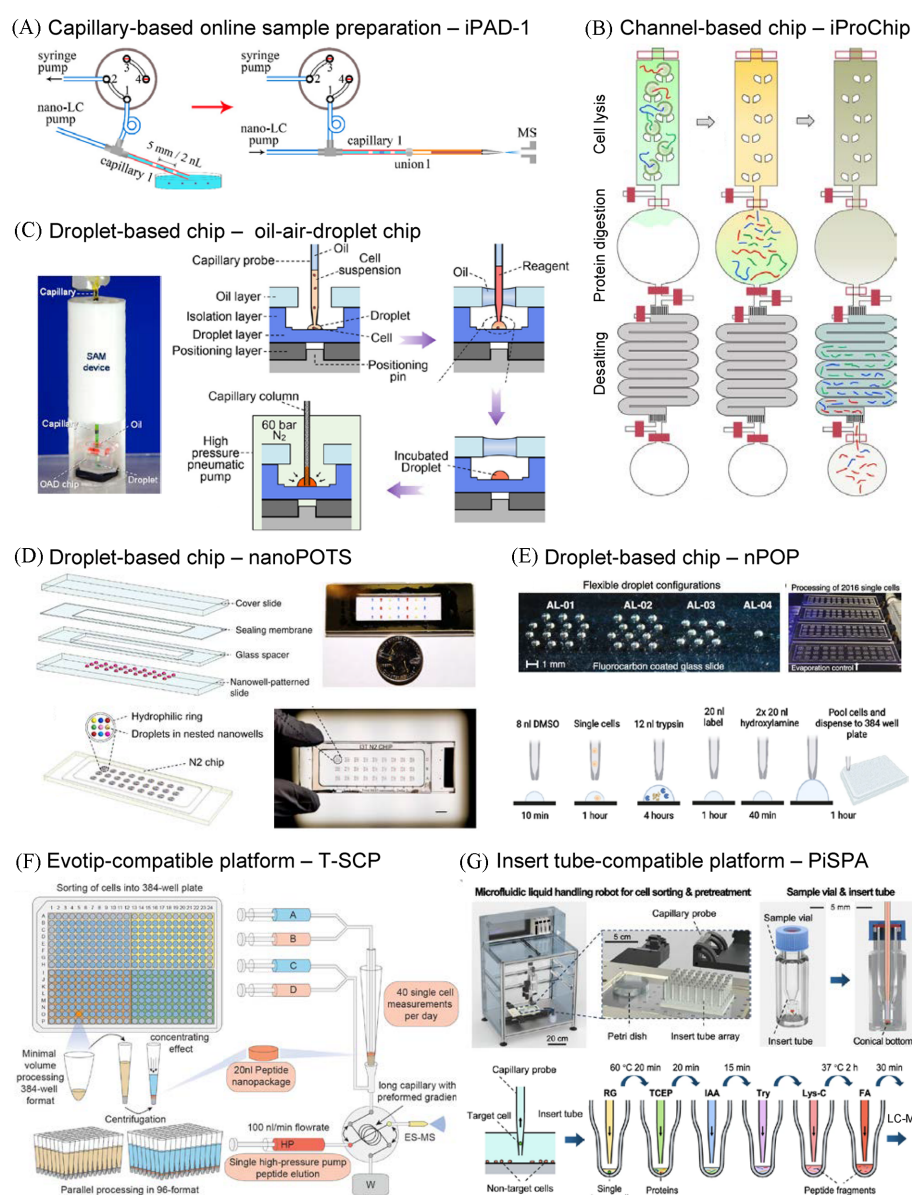


Fig. 2 Single cell proteome sample preparation methods

(A) Integrated proteome analysis device (iPAD-1)^[25]; (B) integrated proteomics chip (iProChip)^[16]; (C) oil-air-droplet chip^[30]; (D) nanodroplet processing in one pot for trace samples (nanoPOTS)^[31] and nested nanoPOTS (N2) chips^[33]; (E) nano-proteomic sample preparation (nPOP)^[35]; (F) true single-cell proteome (T-SCP) platform^[37]; (G) pick-up single-cell proteomic analysis (PiSPA) platform^[9]. (A) Copyright 2018, American Chemical Society; (B) Copyright 2021, Springer Nature; (C) Copyright 2018, American Chemical Society; (D) Copyright 2018, Springer Nature; Copyright 2021, the authors; (E) Copyright 2022, Springer Nature; (F) Copyright 2022, EMBO; (G) Copyright 2024, Springer Nature.

完整蛋白质^[27]和聚糖^[28]的分析.

Gebreyesus等^[16]发展的iProChip芯片是流动体系样品处理微流控装置的代表,其包含细胞捕获和裂解、蛋白质酶解、肽段脱盐以及肽段收集4个模块[图2(B)],可在芯片内实现细胞裂解和蛋白质样品处理,避免了多步骤样品转移过程中的样品损失.该课题组^[16]还发展了单细胞芯片SciProChip,其包括20个腔室,每个腔室捕获和处理1个单细胞,总体积从312 nL减少至78.5 nL, C₁₈脱盐柱的长度从2.5 cm减少至1 cm,有助于减少样品损失,与iProChip相比达到了更高的蛋白质鉴定覆盖率.

与流动体系相比,将样品封装在液滴中形成封闭稳定且相互隔离的环境,有利于减少表面吸附和污染.基于液滴的微流控系统通过集成各种控制模块,可实现微小体积(纳升乃至皮升级)液滴的注入、混合、分散和孵育等精确操控^[29],以快速精确地向样品中引入各种反应试剂,从而实现单细胞蛋白质组样品处理.

2018年,Fang等^[30]发展了一种纳升级油-空气-液滴微流控芯片[图2(C)].该方法采用固定式原位液滴微反应器,并结合基于毛细管探针的序控液滴阵列(Sequential operation droplet array, SODA)系统^[29],代替流动反应体系进行多步纳升级样品预处理操作,减少了样品损失,提升了反应效率.此外,该方法利用高压气泵将液滴直接注入到分析色谱柱内,避免了样品转移和经过液相仪器内复杂管路带来的损失.同年,Zhu等^[31]采用类似的固定式原位液滴微反应器和液滴操纵策略,建立了纳升级液滴痕量样品一锅法处理系统nanoPOTS[图2(D)].该方法将细胞分选至刻有微坑的玻璃芯片中,将纳升级试剂分配至微坑液滴中,逐步完成细胞裂解、蛋白质还原烷基化和酶解,然后将酶解后的样品吸取至毛细管中,连接固相萃取预柱进行除盐,最后连接预柱与分析色谱柱进行液相色谱-质谱检测,以减少样品转移和吸取造成的损失.在此基础上,该课题组^[32]结合商品化液体处理机器人和自动进样器,开发了autoPOTS全自动分析平台.该课题组^[33]进一步发展了嵌套式nanoPOTS N2芯片[图2(D)],其含有27组微坑阵列,每组9个微坑,单个芯片的容量增加至243个单细胞样品,每个样品反应体积减少至30 nL,适用于基于同位素标记的单细胞蛋白质组学分析.

Specht等^[34]发展了最小化蛋白质组样品制备方法mPOP.该方法将单个细胞分选至96或384孔板中,每孔接收1个细胞,采用优化的自热循环系统进行细胞裂解和蛋白质提取,然后利用自动液体处理系统进行蛋白质酶解.在此基础上,该课题组^[35]发展了nPOP方法[图2(E)],其可在载玻片上的微液滴中实现纳升级的样品处理,然后将样品转移至孔板中以进行色谱进样.该方法具有较高的自动化程度和通量,且实验耗材廉价易得.

2021年,Li等^[36]发展了一种单细胞蛋白质组和代谢组综合分析的方法.该方法采用表面亲水化修饰的微坑芯片,利用微坑表面涂料作为亲水相互作用色谱的固定相,实现了单细胞代谢物和蛋白质的分步提取,并在微坑中完成蛋白质酶解,实现了对同一单细胞的蛋白质组和代谢组分析.

单细胞样品处理方法还应与色谱进样系统适配,以进一步减少样品转移损失,并提高自动化程度和分析通量.为此,研究人员建立了采用标准化耗材的单细胞蛋白质组学分析平台.

2022年,Brunner等^[37]发展了一种适配于Evosep系统的单细胞样品处理平台T-SCP. Evosep是一种特殊的色谱仪,区别于利用样品瓶的传统进样方式,其采用一次性的EvoTip(装有C₁₈填料的移液吸头)作为装载样本的捕集柱^[38],以减少进样过程中的损失.该方法将细胞分选至384孔板中,利用液体工作站在微升级体积下进行样品处理,然后转移至EvoTip中进行脱盐和色谱进样[图2(F)].

同年,Ctorteccka等^[39]发展了基于液滴的单细胞样品处理芯片proteoCHIP.该芯片可与cellenONE细胞分选系统和标准的色谱自动进样器兼容.2023年,Ye等^[40]发展了将cellenONE, proteoCHIP和Evosep结合的分析流程,其核心是专为单细胞样品制备而设计的proteoCHIP EVO 96芯片.该方法可同时制备96个纳升级样品,并可与Evosep色谱系统兼容,从而简化了样品转移过程.2024年,该课题组^[41]发展了基于EvoTip的一锅法样品处理方法One-Tip,进一步简化了样品处理过程.

2024年,Wang等^[9]利用SODA系统^[29]搭建了单细胞分离和处理工作站,发展了“点取式”(Pick-up)单细胞蛋白质组学分析平台PiSPA.该方法以样品瓶内插管作为反应容器,根据显微图像将目标细胞

挑选至内插管中,利用自动化液体操控程序完成了纳升级细胞裂解、蛋白质还原烷基化和酶解等样品处理步骤[图2(G)].载有样品的内插管可直接用于色谱自动进样.

2 单细胞蛋白质组色谱质谱分析方法

2.1 蛋白质组样品的分离

单细胞蛋白质含量少且种类多,质谱分析前的分离技术将影响质谱检测效果,从而影响蛋白质鉴定数.

纳升级液相色谱是蛋白质组学研究最常用的分离方法.在常规蛋白质组学分析中,主流的商品化纳升级色谱柱内径为50或75 μm ^[42].针对单细胞蛋白质组学分析,研究人员发展了更小内径的色谱柱,以提供更高的分离能力. Shao等^[25]发展的iPAD-1采用22 μm 内径的色谱柱和3 μm 内径的喷嘴,从单个哺乳动物细胞中鉴定出120余种蛋白质. Cong等^[43]比较了30 μm 和20 μm 内径的填充柱在单个哺乳动物细胞蛋白质组学分析中的性能,发现采用20 μm 内径色谱柱的蛋白质可鉴定数比采用30 μm 内径色谱柱多41%. 该课题组^[44]进一步将20 μm 内径的填充柱与高灵敏质谱仪结合,从单个哺乳动物细胞中鉴定出1000余种蛋白质.

此外,研究人员还在色谱柱结构方面进行了改进. Liang等^[45]发展了一种 C_{18} 功能化胺桥杂化整体柱.胺桥杂化可减少硅羟基上的非特异性吸附,以减少峰拖尾并缩小半峰宽.采用该柱分析单细胞当量(200 pg)的多肽样品,可鉴定出1000余种蛋白质. Xiang等^[46]发展了一种2 μm 内径的皮升级开管柱,采用该柱分析低于单细胞当量(75 pg)的多肽样品,可鉴定出约1000种蛋白质.利用微加工技术在色谱柱内设置形状规整的微结构阵列作为固定相,可提高色谱柱的均一性,以获得较高的分离效率.研究表明,微柱阵列柱(Micropillar array column, μPAC)可实现低于单细胞当量(50 pg)样品的定量分析^[47].目前, μPAC 已在单细胞蛋白质组学研究中得到实际应用^[48,49].

除液相色谱外,毛细管电泳是单细胞蛋白质组学分析中的另一种分离技术,但其应用不如液相色谱广泛. Qin等^[50]对毛细管电泳在单细胞蛋白质组学中的应用作了较详细的介绍.

2.2 质谱数据的采集

蛋白质组学中的质谱数据采集模式主要分为数据依赖性采集(Data-dependent acquisition, DDA)和数据非依赖性采集(Data-independent acquisition, DIA). DDA模式的每次循环先进行一次一级质谱(MS1)扫描,然后逐一选择 MS1 谱图中峰强度最高的一部分母离子进行二级质谱(MS2)扫描[图3(A)].母离子选择过程的随机性会阻碍低丰度肽段的检测,降低每次进样分析结果的重现性^[51],这在单细胞蛋白质组学中尤为明显.与DDA模式相比,DIA模式将整个母离子质量范围划分为一系列有序的隔离窗口(Isolation window),令每个隔离窗口中的所有母离子共同发生碎裂,确保采集到全部母离子的 MS2 信息^[52].这种无偏向性采集有利于检测样品中的低丰度肽段.研究表明,在分析十几个细胞组成的微量样品时,DIA的蛋白质鉴定数可达DDA的2倍^[16].

除了常规的顺序窗口DIA外,DIA在不同类型仪器上还存在形态各异的变体[图3(B)]^[53-55].通常DIA窗口越宽, MS2 谱图越复杂,肽段鉴定越困难.为此,组合型轨道离子阱(Orbitrap)质谱仪上的气相分馏(Gas-phase fractionation, GPF)方法将整个母离子质量范围划分为多个较宽的区域^[56],同一个样品每次进样时仅对其中一个区域进行窄窗口DIA,以减少共碎裂母离子的数量.由于需要多次重复进样,GPF-DIA通常不直接用于单细胞样品分析,但可用于构建多个细胞样品的谱图库,提高单细胞DIA分析的覆盖深度^[57].在Orbitrap质谱仪上,更高的分辨率需要更长的扫描时间,如果每次循环所采集的窗口数量保持不变,色谱峰的数据点数量会减少.为此,WISH-DIA^[48]采用宽窗口 MS2 和高分辨率 MS1 采集相结合的方案,将每次循环中的 MS2 扫描分为若干组,每组之间插入一次 MS1 扫描,保证 MS1 色谱峰有足够的用于定量分析.此外,研究表明,适当扩大DDA的隔离窗口宽度,可在保留部分DDA母离子选择性的同时实现类似DIA的采集特性,以达到更好的定量精确度^[49,58].最近出现的非对称轨道无损(Aymmetric track lossless, Astral)质量分析器的 MS2 扫描速度比Orbitrap显著提高,使DIA

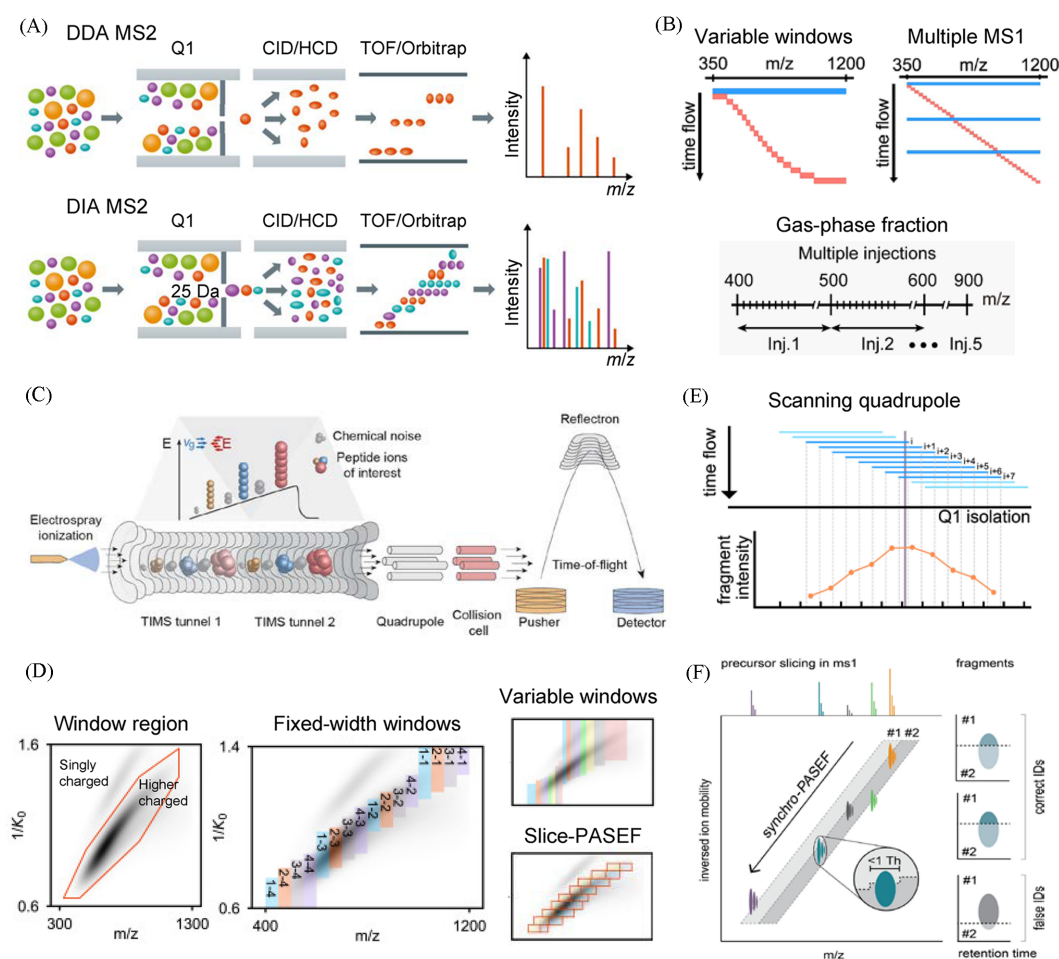


Fig. 3 Mass spectrometry data acquisition methods

(A) Data-dependent acquisition (DDA) and data-independent acquisition (DIA)^[52]; (B) variant DIA schemes^[55]; (C) DIA with parallel accumulation-serial fragmentation (diaPASEF)^[37]; (D) various window arrangements for diaPASEF^[55]; (E) scanning quadrupole DIA^[55]; (F) Synchro-PASEF, where quadrupole scanning is synchronized with ion release in the mobility dimension^[70]. (A) Copyright 2018, EMBO; (B, D, E) Copyright 2024, Elsevier; (C) Copyright 2022, EMBO; (F) Copyright 2023, Elsevier.

窗口能缩小到与 DDA 相当的水平 (2 Da). 采用组合型 Orbitrap Astral 质谱仪的窄窗口 DIA 可从单个哺乳动物体细胞中鉴定超过 4500 种蛋白质^[59]. 最近, Bubis 等^[60]报道了 Orbitrap Astral 质谱仪从单个哺乳动物体细胞中鉴定出 5300 种蛋白质的覆盖深度.

离子淌度谱技术提供了在色谱和质谱之间新的分离维度, 其可在气相中将肽段的多电荷离子与非肽段污染物的单电荷离子分离. 新型 Orbitrap 质谱仪提供了高场不对称波形离子淌度谱 (Field asymmetric ion mobility spectrometry, FAIMS) 接口, 通过设置不同的补偿电压可对进入质谱仪的离子进行筛选, 以降低背景干扰^[61,62]. 研究表明, FAIMS 可将单细胞蛋白质鉴定数提高至未启用 FAIMS 时的 2.3 倍^[44]. 捕集离子淌度谱 (Trapped ion mobility spectrometry, TIMS) 是另一种离子淌度谱技术, 其原理是利用气流推动和反向电场在特定的位置捕获离子, 然后逐渐降低电场, 使离子逐个释放^[63]. 在 TIMS-四级杆-飞行时间质谱仪上, TIMS 分离可与四极杆质量选择同步进行 [图 3(C)], 此方法称为平行累积连续碎裂 (Parallel accumulation-serial fragmentation, PASEF)^[64], 其与 DIA 结合的 diaPASEF 技术在理论上可实现近 100% 的离子利用率^[65]. 2022 年, Brunner 等^[37]构建的 T-SCP 平台将 diaPASEF 与微量样品制备技术、极低流量色谱结合, 使蛋白质组学分析的灵敏度提高至常规方法的 10 倍, 实现了单细胞中超过 2000 种蛋白质的鉴定和定量分析. 2024 年, Wang 等^[9]实现了 diaPASEF 对单个哺乳动物体细胞中高达 3000 种蛋白质的定量深度.

常规 diaPASEF 采集方案是按 m/z 轴划分一系列不重复的窗口, 这些窗口在离子淌度轴上存在重叠, 以覆盖整个待采集的目标区域[图 3(D)]. 根据离子淌度轴的重叠情况, 将这些窗口分成若干组, 每组内窗口的离子淌度互不重叠, 每组窗口的采集需要一次离子淌度循环(Frame). Skowronek 等^[66]提出了一种规划 diaPASEF 窗口的策略, 该策略根据母离子在 m/z -离子淌度空间中的密度分布设置了尺寸可变的窗口, 以在有限的离子淌度循环次数内实现完整覆盖. Szyrwiel 等^[67]提出了 Slice-PASEF 采集方案, 该方案划分了一系列 m/z 重复而离子淌度不重复的窗口, 以实现在一次离子淌度循环中完成采集. 在设置多次离子淌度循环的条件下, 这些窗口可在 m/z 方向被分割成多块, 每次离子淌度循环可采集每个窗口中的一块. 与常规 diaPASEF 相比, Slice-PASEF 提高了 MS2 采集的占比.

由于 DIA 产生的 MS2 谱图为混合谱图, 所以判断碎片离子对应的母离子是 DIA 数据解析中的重要问题. 四级杆-飞行时间质谱仪的滑动窗口扫描(Scanning quadrupole)可提供一个额外的数据维度——四级杆窗口位置(Q1)^[68,69]. 在 Q1 轴上建立碎片离子峰强度的变化曲线, 其峰值理论上对应于母离子 m/z , 由此可对碎片离子的来源进行归属[图 3(E)], 提高 MS2 分析的选择性. 基于与滑动窗口扫描类似的方式, Skowronek 等^[70]和 Distler 等^[71]对 diaPASEF 进行改进, 分别提出了 Synchro-PASEF 和 midiaPASEF, 以不同窗口对母离子和碎片离子信号进行分割切片, 利用信号在不同窗口切片中的分布推断出碎片离子对应的母离子[图 3(F)]. 尽管上述方法在单细胞分析中的应用尚未见报道, 但其在常规样品分析方面显示出优于 diaPASEF 的性能. 从原理上看, 上述方法将有利于单细胞样品的检测.

2.3 多通道稳定同位素标记

非标记分析是单细胞蛋白质组学分析最简单的方法, 但其分析通量有限. 为此, 研究人员将稳定同位素标记引入蛋白质组学分析, 在单次色谱质谱进样中可同时实现多个样品的定量分析^[72]. 基于同位素标记的定量方法可分为等重标记和非等重标记. 最常用的等重标记试剂包括 TMT(或 TMTpro)^[73], iTRAQ^[73]和 iBT^[74], 其基本结构包含一个报告基团、一个平衡基团和一个反应基团[图 4(A)], 编码基团因其同位素组成不同而具有不同质量, 平衡基团的同位素组成与报告基团相对应而使整个标记质量相等, 反应基团可与肽段 N 端或赖氨酸的氨基发生特异性反应, 从而使标记连接至肽段上. 标记后的肽段在 MS1 中的母离子质量相同, 因而不会增加 MS1 谱图的复杂性. 在 MS2 中报告基团与肽段脱离, 会在低质量区域产生不同质量的报告离子峰以用于定量分析.

Budnik 等^[75]将等重同位素标记用于单细胞蛋白质组学分析, 开发了 SCoPE-MS 技术. SCoPE-MS 利用 TMT(或 TMTpro)实现了最多 18 个通道的定量分析, 其中一个 TMT 通道以 200 个细胞样品填充作为

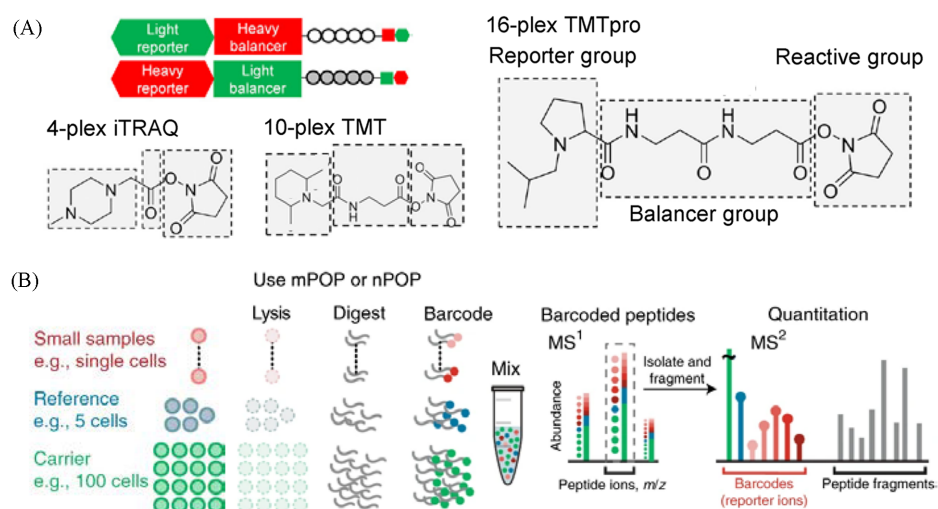


Fig. 4 Isobaric labeling for reporter ion-based quantification

(A) Representative labeling reagents^[72]; (B) multiplexed single-cell proteomics using SCoPE2^[77]. (A) Copyright 2021, Wiley; (B) Copyright 2021, Springer Nature.

载体(Carrier), 相邻通道留空以避免干扰, 其余通道用于单细胞样品分析. 载体的引入一方面可减少样品损失, 另一方面可增强DDA的MS1信号以触发更多MS2扫描, 为肽段鉴定提供丰富的碎片离子信息. 该课题组^[76,77]改进了微量样品处理 and 数据分析方法, 发展了SCoPE2[图4(B)], 在10 d的机时内定量分析了近1500个单核细胞和巨噬细胞中的3000多种蛋白质^[77].

基于报告基团的方法无法区分不同肽段的报告离子, 不适用于DIA分析, 在DDA中也可能因肽段发生共碎裂而导致定量比率失真. 与之相比, 非等重标记在MS1中产生的肽段母离子质量不同, 且在MS2中产生的一部分肽段碎片离子质量不同, 因此DDA可利用MS1定量, DIA可同时利用MS1和MS2信息实现定量分析. 在利用肽段母离子或碎片离子进行定量分析时, 为了避免肽段离子本身天然同位素分布的干扰, 每个通道之间通常需要至少4 Da的间隔^[72].

Derks等^[78]将非等重标记试剂mTRAQ^[79](iTRAQ的非等重版本)与DIA结合, 发展了一种称为plexDIA的多通道DIA方法, 可同时分析3个单细胞样品, 并定量检测哺乳动物单细胞中约1000种蛋白质. 该方法每次分析需要30 min, 平均每天可分析144个细胞. Thielert等^[80]将二甲基化标记与DIA结合, 采用常规的胰蛋白酶(Trypsin)消化时可实现3个样品的同时分析. 对于赖氨酸氨基端肽键水解酶(Lys-N)消化的样品, 肽段N端的两个氨基同时带上标记, 可使同时分析的通道数增加至5. 将其中一个通道填充较大数量样品以作为参考通道(类似于SCoPE-MS的载体), 可在单细胞中定量分析近4000种蛋白质. 该方法平均每天可分析80个细胞.

3 单细胞蛋白质组数据分析方法

3.1 DDA质谱数据的解析

DDA数据分析的主流方法为序列数据库搜索, 其基本思路是按照蛋白质序列数据库理论酶切产生肽段, 根据肽段在MS2中的碎裂规律产生理论碎片, 将实验采集到的谱图与肽段理论碎片进行匹配, 按照一定的规则进行打分, 选出匹配最佳的肽段, 然后根据肽段鉴定结果推断出样品中存在的蛋白质, 并对鉴定结果进行质控^[81-83]. 在鉴定基础上, 提取MS1色谱峰(非标记定量和不等重标记定量)或MS2报告离子峰(等重标记定量), 根据峰面积或峰强度进行肽段和蛋白质定量分析. MaxQuant^[84], FragPipe^[85,86]和pFind^[87]等软件工具整合了上述分析流程, 使蛋白质组数据分析更便利.

目前, 几乎所有肽段鉴定算法和软件工具均参照大量样品的谱图开发, 相比于大量样品, 单细胞样品进入质谱仪的离子显著减少, 可能会使MS2谱图发生变化. Boekweg等^[88]研究了单细胞样品的MS2谱图特征[图5(A)], 发现其与大量样品谱图的主要区别在于3个方面: (1) 一些本应注释到的碎片离子发生了缺失; (2) 离子强度降低导致样品的质谱信号与背景之间的界限变模糊; (3) 碎片离子的峰强度模式具有明显区别. 这些特征是肽段鉴定中的关键因素, 因此有必要针对单细胞数据对已有算法做出适应性调整.

针对单细胞样品, Huffman等^[89]开发的DO-MS可利用MaxQuant分析结果对SCoPE-MS分析的重要指标进行展示, 诊断分析中的问题, 并指导色谱质谱参数的优化. Chen等^[90]开发的DART-ID利用预测的肽段保留时间等信息改进肽段鉴定的可靠性, 提高了单细胞蛋白质鉴定的数量, 减少了定量缺失值. Wang等^[91]开发的DeepSCP利用深度学习模型可预测肽段的保留时间和碎片离子强度, 以构建肽段谱图匹配的一系列特征, 并可使用优化的集成学习模型区分正确和错误的鉴定结果, 以提高单细胞蛋白质鉴定数.

运行间匹配(Match-between-runs, MBR)可减少蛋白质定量结果中的缺失值[图5(B)], 其基本原理是将一次进样中由MS/MS鉴定到的肽段, 根据其母离子保留时间和 m/z , 在其它进样数据中查找出对应的MS1色谱峰, 以得到更完整的定量结果. Yu等^[92]开发的IonQuant对MBR过程中的错误率进行了可靠质控, 提高了单细胞数据MBR的灵敏度和精确度. Kalxdorf等^[93]开发的IceR利用离子流信息进行混合肽段鉴定的MBR, 适合于处理因肽段丰度低而难以从MS1中提取出同位素峰特征的情况, 有利于单细胞等低样品量数据分析.

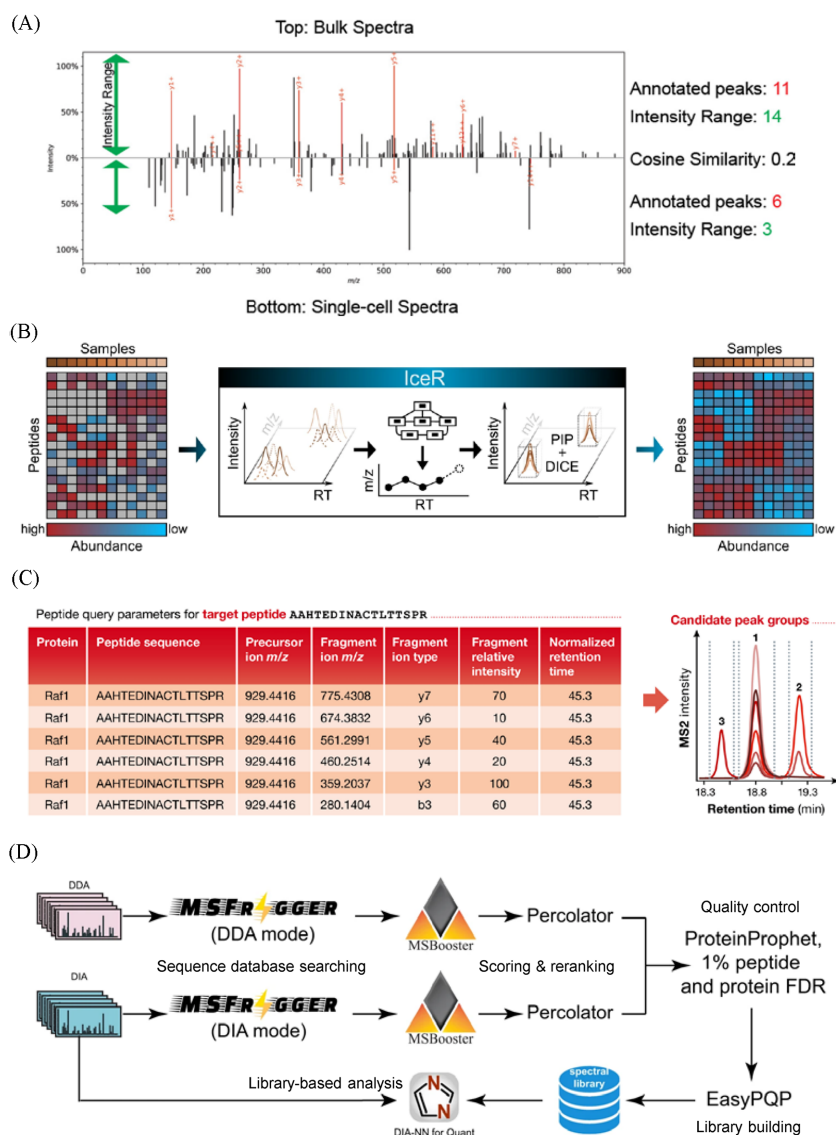


Fig. 5 Mass spectrometry data analysis methods for proteomics

(A) Spectral features of single-cell proteome samples^[88]; (B) match-between-runs by IceR for reducing missing values^[93]; (C) library-based DIA analysis by targeted data extraction^[52]; (D) integrated DIA analysis workflow by FragPipe and DIA-NN^[112]. (A) Copyright 2022, American Chemical Society; (B) Copyright 2021, Springer Nature; (C) Copyright 2018, EMBO; (D) Copyright 2023, Springer Nature.

3.2 DIA 质谱数据的解析

DIA 数据中不同肽段母离子的碎片离子可存在于同一张 MS2 谱图中, 如何解析混合 MS2 谱图是 DIA 数据解析的关键技术难点. 主流的 DIA 数据分析方法可分为基于谱图库 (Library-based) 和不依赖谱图库 (Library-free) 的策略. DIA-Umpire^[94] 为不依赖谱图库的分析工具的代表, 其根据母离子和碎片离子的色谱峰相似度将 DIA 谱图重构为模拟的 DDA 谱图, 然后采用常规 DDA 数据库搜索方法进行解析. 该方法依赖于 MS1 和 MS2 中离子的色谱峰曲线的解卷积, 对质谱信号质量要求较高, 限制了分析的灵敏度. OpenSWATH^[95] 等基于谱图库的工具则根据谱图库中的肽段保留时间、离子淌度和碎片离子信息从 DIA 中提取谱图信号 [图 5(C)]. Hou 等^[54] 和 Lou 等^[55] 分别对 DIA 数据分析方法作了更全面准确的分类总结, Ludwig 等^[52] 详细介绍了基于谱图库的分析方法的原理和步骤.

在 DIA 分析方法发展早期, 谱图库主要由 DDA 实验构建. 近年来, 利用深度学习技术实现了谱图库的高精度预测. 深度学习在谱图预测中的首次应用可追溯到 2017 年的 pDeep^[96], 此后各种谱图预测模型及其在 DIA 中的应用方法纷纷涌现^[97-101]. Gessulat 等^[98] 开发了谱图库预测算法 Prosit, 利用双向门

控循环单元网络模型结合注意力机制,对肽段的理论 MS2 谱图和保留时间进行了预测. 2020 年, Yang 等^[100]开发了 DeepDIA, 结合卷积神经网络和长短记忆神经网络预测肽段谱图库, 并使用另一个深度网络模型预测了肽段的可检测性. Lou 等^[101]开发了 DeepPhospho, 利用 Transformer 网络实现了磷酸化修饰肽段的谱图库预测和 DIA 分析. 该课题组^[102]将预测谱图与实验谱图结合, 提出了构建混合谱图库的方法, 先采用不依赖谱图库的方法, 由 DIA 数据鉴定肽段得到实验谱图, 用其训练谱图库预测模型, 再将预测谱图与实验谱图混合建库以提高 DIA 分析的覆盖深度. 分析 diaPASEF 产生的四维 DIA 数据时, 需充分利用其中每个维度的信息^[103]. Chen 等^[104]开发了一种多功能深度学习模型 Deep4D, 实现了未修饰肽段和磷酸化肽段的色谱保留时间、碰撞截面积、碎片离子强度和电荷状态的高精度预测, 从而建立了高覆盖四维 DIA 常规蛋白质组学和磷酸化蛋白质组学工作流程. Zeng 等^[105]开发了模块化的谱图库预测工具 AlphaPeptDeep, 支持肽段碰撞截面积预测(适用于 diaPASEF 数据分析), 还提供了对非编程专家用户友好的图形界面和模型商店. Yang 等^[106]详细总结了深度学习方法在 DIA 数据分析中的应用.

谱图库不仅提供了肽段保留时间和碎片离子信息, 还决定了 DIA 分析的搜索空间. 过小的谱图库覆盖范围不足, 限制了蛋白质鉴定定量的深度; 而过大的谱图库对样品的特异性不够, 不利于数据分析中的统计质控, 限制了鉴定定量的灵敏度^[107]. 因此, 采用实验或预测方式, 单细胞样品谱图库的构建均需考虑各种复杂因素. Siyal 等^[108]研究发现, 对 5 个细胞(约 0.75 ng)的样品采用较小的谱图库(含有约 3000 种蛋白质)可得到较高的 DIA 蛋白质鉴定数, 而随着谱图库的增大, DIA 蛋白质鉴定数明显减少.

除了常规的谱图库外, Searle 等^[109]提出了色谱库的概念并开发了 EncyclopeDIA 软件. 色谱库的构建方法是先采集 GPF-DIA 数据, 再用预测谱图库或不依赖谱图库的分析方法鉴定肽段, 获得可检测肽段的碎片离子色谱峰形和保留时间, 并对原有的预测谱图库进行校正^[56], 从而构建出在特定色谱质谱条件下的特异性谱图库. 在单细胞数据上, 采用 EncyclopeDIA 软件, 以 GPF-DIA 建库的方法比不依赖谱图库直接进行 DIA 分析的方法可达到更高的蛋白质鉴定数^[48].

Demichev 等^[110]开发了 DIA-NN 软件, 将深度学习进一步应用于 DIA 数据分析中的统计质控. DIA-NN 既支持单独使用, 用于构建预测谱图库以直接进行 DIA 数据分析, 也可整合至 FragPipe 流程中, 先构建实验谱图库, 再进行 DIA 分析^[111]. FragPipe 可先利用 DDA 或 DIA 数据构建谱图库, 再用 DIA-NN 分析 DIA 数据[图 5(D)]. 与单独使用 DIA-NN 直接进行 DIA 分析相比, 该方法可提高单细胞样品的蛋白质鉴定数^[112]. DIA 还支持 plexDIA 同位素标记定量^[78]. 商业软件 Spectronaut 既支持基于实验谱图库的分析策略^[113], 又提供了利用直接分析 DIA 数据的 directDIA 模块(结合深度学习升级为 directDIA+)^[114], 两种方式可混合使用, 以达到较高的蛋白质覆盖深度^[115]. 凭借性能和易用性方面的优势, DIA-NN 和 Spectronaut 为目前单细胞蛋白质组领域最受欢迎的两种 DIA 数据分析方案.

4 单细胞蛋白质组学的应用

目前, 单细胞蛋白质组学的应用主要涉及发育、免疫、细胞周期和药物反应等领域, 集中于细胞异质性相关研究. Tsai 等^[116]利用基于 TMT 标记和 nanoPOTS 的单细胞蛋白质组学分析方法, 从 100 余个急性髓系白血病(Acute myeloid leukemia, AML)单细胞中定量检测到约 1500 种蛋白质, 并区分出 3 种 AML 细胞. Schoof 等^[117]利用类似 SCoPE-MS 的方法表征了 AML 细胞的层次结构. Specht 等^[77]利用 SCoPE2 定量分析了近 1500 个单核细胞和巨噬细胞中的 3000 多种蛋白质, 解释了在缺乏极化细胞因子条件下巨噬细胞的异质性. Leduc 等^[35]利用 nPOP 分析了黑色素瘤细胞, 根据耐药相关标志物表达情况, 揭示了蛋白质组变化与糖酵解和氧化磷酸化等生物功能的关联. Wang 等^[9]利用 PiSPA 平台研究了肿瘤细胞迁移过程中蛋白质表达的异质性. Orsburn 等^[118]首次将单细胞蛋白质组学应用于药物机制研究, 表征细胞对小分子药物抑制剂急性扰动的反应, 发现了与经典血管内皮生长因子通路相关的蛋白质的表达变化. 此外, 该工作首次在单细胞中检测了乙酰化、磷酸化、甲基化、二甲基化、琥珀酰化、

羟基丁基化和巴豆酰化等翻译后修饰。

最近, 依托新型高灵敏质谱仪的单细胞蛋白质组学已显示出其在生物医学研究中的应用潜力。Petrosius 等^[119]采用 Orbitrap Astral 质谱仪和优化的 DIA 方法, 研究了细胞周期动力学和癌细胞异质性。Bubis 等^[60]利用 Orbitrap Astral 质谱仪研究了细胞大小与细胞周期阶段之间的关系。此外, 该课题组^[60]研究了体外模拟人囊胚的细胞谱系, 揭示了上胚层与滋养外胚层细胞之间的蛋白质表达差异。

单细胞蛋白质组学还可与单细胞转录组学和代谢组学等其它组学互为补充, 以实现多组学联合分析。Brunner 等^[37]利用基于离子淌度质谱的单细胞分析平台研究了不同细胞周期的蛋白质表达水平, 与转录组数据进行比较, 发现细胞存在一个稳定表达的核心蛋白质组。Jiang 等^[120]对鼠卵母细胞同时进行了单细胞蛋白质组和转录组分析, 比较了卵母细胞减数分裂过程中蛋白质和 RNA 的变化模式。Wu 等^[121]对肿瘤细胞系同时进行了单细胞蛋白质组和代谢组分析, 解释了药物作用下细胞蛋白质和代谢物含量变化的原因。

5 总结与展望

近年来, 单细胞蛋白质组学在样品制备、质谱检测和数据解析等环节取得了重要进展。在样品制备方面, 发展了多种毛细管、微流控芯片和液滴反应器, 以减少样本损失, 提高反应效率; 在质谱检测方面, 发展了新型质谱采集技术, 以提高采集范围和灵敏度; 在数据解析方面, 发展了以深度学习为关键技术的算法和软件工具, 以提升数据解析利用率。表 S1 (见本文支持信息) 列出了近年来具有代表性的单细胞蛋白质组学研究结果。这些技术进步使单细胞蛋白质组学在生物医学领域具备了应用潜力。

目前, 单细胞蛋白质组学分析技术距离满足生物医学应用需求还存在一定差距。大多数单细胞蛋白质组学分析方法对一般哺乳动物细胞系鉴定深度不足 2000 种蛋白质, 少数结果方法可达 3000~5000 种蛋白质^[9,40], 尚不能覆盖多种重要生物功能和代谢通路。单细胞蛋白质组学分析通量较低, 远低于单细胞转录组分析通量。同位素标记技术是提高覆盖深度和通量的一种可能方向。此外, 现有技术虽然实现了在蛋白质层面上的鉴定和定量分析, 以及少数翻译后修饰的检测^[36,118], 但难以提供全面的单细胞修饰组学信息。上述问题的解决需要样品制备、质谱检测和数据解析等方面相互配合的改进和优化。

单细胞蛋白质组学研究产出的数据逐渐增多, 为促进单细胞蛋白质组数据的共享和利用, 需要建立实验设计、数据解释和结果报告的基本规范^[122], 以符合可发现、可访问、可互操作和可重用 (Findable, accessible, interoperable, and reusable; FAIR) 原则^[123]的方式管理和开放数据。一方面, 统一标准有利于不同实验室结果的可比性和可重复性; 另一方面, 这有利于实现不同来源的大规模复杂数据的整合^[124], 有利于与人工智能和生物信息学等领域的跨学科合作和交叉应用^[125,126], 以实现更全面的细胞信息获取和综合分析。

上述方面的发展将推动单细胞蛋白质组学技术在生物医学领域的应用, 为深入理解细胞功能、疾病机制和药物作用提供强大工具, 为精准医疗提供支持。

支持信息见 <http://www.cjcu.jlu.edu.cn/CN/10.7503/cjcu20240294>。

参 考 文 献

- [1] Wu Q., Sui X. T., Tian R. J., *Chin. J. Chromatogr.*, **2021**, 39(2), 112—117 (吴琼, 隋欣桐, 田瑞军. 色谱, **2021**, 39(2), 112—117)
- [2] Regev A., Teichmann S. A., Lander E. S., Amit I., Benoist C., Birney E., Bodenmiller B., Campbell P., Carninci P., Clatworthy M., Clevers H., Deplancke B., Dunham I., Eberwine J., Eils R., Enard W., Farmer A., Fugger L., Göttgens B., Hacohen N., Haniffa M., Hemberg M., Kim S., Klenerman P., Kriegstein A., Lein E., Linnarsson S., Lundberg E., Lundeberg J., Majumder P., Marioni J. C., Merad M., Mhlanga M., Nawijn M., Netea M., Nolan G., Pe'er D., Phillipakis A., Ponting C. P., Quake S., Reik W., Rozenblatt-Rosen O., Sanes J., Satija R., Schumacher T. N., Shalek A., Shapiro E., Sharma P., Shin J. W., Stegle O., Stratton M., Stubbington M. J. T., Theis F. J., Uhlen M., van Oudenaarden A., Wagner A., Watt F., Weissman J., Wold B., Xavier R., Yosef N., *eLife*, **2017**, 6, e27041

- [3] Ahmad R., Budnik B., *Anal. Bioanal. Chem.*, **2023**, 415(28), 6889—6899
- [4] Liu Y., Beyer A., Aebbersold R., *Cell*, **2016**, 165(3), 535—550
- [5] Labib M., Kelley S. O., *Nat. Rev. Chem.*, **2020**, 4(3), 143—158
- [6] Zhu Y., Scheibinger M., Ellwanger D. C., Krey J. F., Choi D., Kelly R. T., Heller S., Barr-Gillespie P. G., *eLife*, **2019**, 8, e50777
- [7] Nelep C., Eberhardt J., *Cytom. A*, **2018**, 93(12), 1267—1270
- [8] Wang Y., Wang D. F., Wang H. F., Wang J. W., Pan J. Z., Guo X. G., Fang Q., *Talanta*, **2021**, 226, 122136
- [9] Wang Y., Guan Z. Y., Shi S. W., Jiang Y. R., Zhang J., Yang Y., Wu Q., Wu J., Chen J. B., Ying W. X., Xu Q. Q., Fan Q. X., Wang H. F., Zhou L., Wang L., Fang J., Pan J. Z., Fang Q., *Nat. Commun.*, **2024**, 15(1), 1279
- [10] Nagamoto-Combs K., *Animal Models of Allergic Disease: Methods and Protocols*, Springer US, New York, **2021**, 169—182
- [11] Adan A., Alizada G., Kiraz Y., Baran Y., Nalbant A., *Crit. Rev. Biotechnol.*, **2017**, 37(2), 163—176
- [12] Gross A., Schoendube J., Zimmermann S., Steeb M., Zengerle R., Koltay P., *Int. J. Mol. Sci.*, **2015**, 16(8), 16897—16919.
- [13] Vallone V. F., Telugu N. S., Fischer I., Miller D., Schommer S., Diecke S., Stachelscheid H., *Curr. Protoc. Stem Cell Biology*, **2020**, 55(1), e123
- [14] Song Y., Lin J. M., *Sci. Sin. Chim.*, **2023**, 53(8), 1472—1493(宋扬, 林金明. 中国科学: 化学, **2023**, 53(8), 1472—1493)
- [15] Pan T., Wu Y. Y., Guo G. S., Wang X. Y., *Chin. J. Anal. Chem.*, **2023**, 51(6), 934—944(潘婷, 武园园, 郭广生, 汪夏燕. 分析化学, **2023**, 51(6), 934—944)
- [16] Gebreyesus S. T., Siyal A. A., Kitata R. B., Chen E. S. W., Enkhbayar B., Angata T., Lin K. I., Chen Y. J., Tu H. L., *Nat. Commun.*, **2022**, 13(1), 37
- [17] Hu S., Ye J., Shi S., Yang C., Jin K., Hu C., Wang D., Ma H., *Anal. Chem.*, **2023**, 95(17), 6905—6914
- [18] Yang Z., Jin K., Chen Y., Liu Q., Chen H., Hu S., Wang Y., Pan Z., Feng F., Shi M., Xie H., Ma H., Zhou H., *JACS Au*, **2024**, 4(5), 1811—1823
- [19] Guo W., Hu Y., Qian J., Zhu L., Cheng J., Liao J., Fan X., *J. Genet. Genom.*, **2023**, 50(9), 641—651
- [20] Espina V., Heiby M., Pierobon M., Liotta L. A., *Expert. Rev. Mol. Diagn.*, **2007**, 7(5), 647—657
- [21] Mund A., Coscia F., Kriston A., Hollandi R., Kovács F., Brunner A. D., Migh E., Schweizer L., Santos A., Bzorek M., Naimy S., Rahbek-Gjerdum L. M., Dyring-Andersen B., Bulkescher J., Lukas C., Eckert M. A., Lengyel E., Gnann C., Lundberg E., Horvath P., Mann M., *Nat. Biotechnol.*, **2022**, 40(8), 1231—1240
- [22] Rosenberger F. A., Thielert M., Strauss M. T., Schweizer L., Ammar C., Mädler S. C., Metousis A., Skowronek P., Wahle M., Madden K., Gote-Schniering J., Semenova A., Schiller H. B., Rodriguez E., Nordmann T. M., Mund A., Mann M., *Nat. Methods*, **2023**, 20(10), 1530—1536
- [23] Liu J., Liu Z., *Chin. J. Chromatogr.*, **2016**, 34(1), 1154—1160(刘佳, 刘震. 色谱, **2016**, 34(1), 1154—1160)
- [24] Chen Q., Yan G., Gao M., Zhang X., *Anal. Chem.*, **2015**, 87(13), 6674—6680
- [25] Shao X., Wang X., Guan S., Lin H., Yan G., Gao M., Deng C., Zhang X., *Anal. Chem.*, **2018**, 90(23), 14003—14010
- [26] He Y., Yuan H., Liang Y., Liu X., Zhang X., Ji Y., Zhao B., Yang K., Zhang J., Zhang S., Zhang Y., Zhang L., *Chem. Sci.*, **2023**, 14(46), 13495—13502
- [27] Johnson K. R., Gao Y., Greguš M., Ivanov A. R., *Anal. Chem.*, **2022**, 94(41), 14358—14367
- [28] Marie A. L., Gao Y., Ivanov A. R., *Nat. Commun.*, **2024**, 15(1), 3847
- [29] Zhu Y., Zhang Y. X., Cai L. F., Fang Q., *Anal. Chem.*, **2013**, 85(14), 6723—6731
- [30] Li Z. Y., Huang M., Wang X. K., Zhu Y., Li J. S., Wong C. C. L., Fang Q., *Anal. Chem.*, **2018**, 90(8), 5430—5438
- [31] Zhu Y., Piehowski P. D., Zhao R., Chen J., Shen Y., Moore R. J., Shukla A. K., Petyuk V. A., Campbell-Thompson M., Mathews C. E., Smith R. D., Qian W. J., Kelly R. T., *Nat. Commun.*, **2018**, 9(1), 882
- [32] Liang Y., Acor H., McCown M. A., Nwosu A. J., Boekweg H., Axtell N. B., Truong T., Cong Y., Payne S. H., Kelly R. T., *Anal. Chem.*, **2021**, 93(3), 1658—1666
- [33] Woo J., Williams S. M., Markillie L. M., Feng S., Tsai C. F., Aguilera-Vazquez V., Sontag R. L., Moore R. J., Hu D., Mehta H. S., Cantlon-Bruce J., Liu T., Adkins J. N., Smith R. D., Clair G. C., Pasa-Tolic L., Zhu Y., *Nat. Commun.*, **2021**, 12(1), 6246
- [34] Specht H., Harmange G., Perlman D. H., Emmott E., Niziolek Z., Budnik B., Slavov N., *bioRxiv*, **2018**, 399774[2018-08-25]. <https://doi.org/10.1101/399774>
- [35] Leduc A., Huffman R. G., Cantlon J., Khan S., Slavov N., *Genome Biol.*, **2022**, 23(1), 261
- [36] Li Y., Li H., Xie Y., Chen S., Qin R., Dong H., Yu Y., Wang J., Qian X., Qin W., *Anal. Chem.*, **2021**, 93(42), 14059—14067
- [37] Brunner A. D., Thielert M., Vasilopoulou C., Ammar C., Coscia F., Mund A., Hoerning O. B., Bache N., Apalategui A., Lubeck M., Richter S., Fischer D. S., Raether O., Park M. A., Meier F., Theis F. J., Mann M., *Mol. Syst. Biology*, **2022**, 18(3), e10798
- [38] Krieger J. R., Wybenga-Groot L. E., Tong J., Bache N., Tsao M. S., Moran M. F., *J. Proteome Res.*, **2019**, 18(5), 2346—2353
- [39] Hartlmayr D., Ctortocka C., Seth A., Mendjan S., Tourniaire G., Mechtler K., *bioRxiv*, **2021**, 2021.04.14.439828. <https://doi.org/10.1101/2021.04.14.439828>
- [40] Ye Z., Sabatier P., van der Hoeven L., Phlairaharn T., Hartlmayr D., Izaguirre F., Seth A., Joshi H. J., Bekker-Jensen D. B., Bache N., Olsen J. V., *bioRxiv*, **2023**, 2023.11.27.568953[2023-11-28]. <https://doi.org/10.1101/2023.11.27.568953>
- [41] Ye Z., Sabatier P., Martin-Gonzalez J., Eguchi A., Lechner M., Østergaard O., Xie J., Guo Y., Schultz L., Truffer R., Bekker-Jensen

- D. B., Bache N., Olsen J. V., *Nat. Commun.*, **2024**, *15*(1), 2474
- [42] Wilson S. R., Olsen C., Lundanes E., *Analyst*, **2019**, *144*(24), 7090—7104
- [43] Cong Y., Liang Y., Motamedchaboki K., Huguet R., Truong T., Zhao R., Shen Y., Lopez-Ferrer D., Zhu Y., Kelly R. T., *Anal. Chem.*, **2020**, *92*(3), 2665—2671
- [44] Cong Y., Motamedchaboki K., Misal S. A., Liang Y., Guise A. J., Truong T., Huguet R., Plowey E. D., Zhu Y., Lopez-Ferrer D., Kelly R. T., *Chem. Sci.*, **2021**, *12*(3), 1001—1006
- [45] Liang Y., Wang C., Liang Z., Zhang L., Zhang Y., *Anal. Chem.*, **2022**, *94*(16), 6084—6088
- [46] Xiang P., Zhu Y., Yang Y., Zhao Z., Williams S. M., Moore R. J., Kelly R. T., Smith R. D., Liu S., *Anal. Chem.*, **2020**, *92*(7), 4711—4715
- [47] Stejskal K., Op de Beeck J., Dürnberger G., Jacobs P., Mechtler K., *Anal. Chem.*, **2021**, *93*(25), 8704—8710
- [48] Petrosius V., Aragon-Fernandez P., Üresin N., Kovacs G., Phlairaharn T., Furtwängler B., Op De Beeck J., Skovbakke S. L., Goletz S., Thomsen S. F., Keller U. A. D., Natarajan K. N., Porse B. T., Schoof E. M., *Nat. Commun.*, **2023**, *14*(1), 5910
- [49] Matzinger M., Schmücker A., Yelagandula R., Stejskal K., Krššáková G., Berger F., Mechtler K., Mayer R. L., *Nat. Commun.*, **2024**, *15*(1), 1019
- [50] Qin S. J., Bai Y., Liu H. W., *Chin. J. Chromatogr.*, **2021**, *39*(2), 142—151 (秦少杰, 白玉, 刘虎威. 色谱, **2021**, *39*(2), 142—151)
- [51] Tabb D. L., Vega-Montoto L., Rudnick P. A., Variyath A. M., Ham A. J. L., Bunk D. M., Kilpatrick L. E., Billheimer D. D., Blackman R. K., Cardasis H. L., Carr S. A., Clauser K. R., Jaffe J. D., Kowalski K. A., Neubert T. A., Regnier F. E., Schilling B., Tegeler T. J., Wang M., Wang P., Whiteaker J. R., Zimmerman L. J., Fisher S. J., Gibson B. W., Kinsinger C. R., Mesri M., Rodriguez H., Stein S. E., Tempst P., Paulovich A. G., Liebler D. C., Spiegelman C., *J. Proteome Res.*, **2010**, *9*(2), 761—776
- [52] Ludwig C., Gillet L., Rosenberger G., Amon S., Collins B. C., Aebersold R., *Mol. Syst. Biology*, **2018**, *14*(8), e8126
- [53] Zhang F., Ge W., Ruan G., Cai X., Guo T., *Proteomics*, **2020**, *20*(17/18), 1900276
- [54] Hou X. H., Zhou P. Y., Gong P. Y., Fu J. L., Liu C., Wang H. P., *Prog. Biochem. Biophys.*, **2022**, *49*(12), 2364—2386 (侯鑫行, 周丕宇, 宫鹏云, 付嘉乐, 刘超, 王海鹏. 生物化学与生物物理进展, **2022**, *49*(12), 2364—2386)
- [55] Lou R., Shui W., *Mol. Cell. Proteomics*, **2024**, *23*(2), 100712
- [56] Searle B. C., Swearingen K. E., Barnes C. A., Schmidt T., Gessulat S., Küster B., Wilhelm M., *Nat. Commun.*, **2020**, *11*(1), 1548
- [57] Phlairaharn T., Ye Z., Krismer E., Pedersen A. K., Pietzner M., Olsen J. V., Schoof E. M., Searle B. C., *Anal. Chem.*, **2023**, *95*(26), 9881—9891
- [58] Matzinger M., Müller E., Dürnberger G., Pichler P., Mechtler K., *Anal. Chem.*, **2023**, *95*(9), 4435—4445
- [59] Guzman U. H., Martinez-Val A., Ye Z., Damoc E., Arrey T. N., Pashkova A., Renuse S., Denisov E., Petzoldt J., Peterson A. C., Harking F., Østergaard O., Rydbirk R., Aznar S., Stewart H., Xuan Y., Hermanson D., Horning S., Hock C., Makarov A., Zabrouskov V., Olsen J. V., *Nat. Biotechnol.*, **2024**[2024-02-01]. <https://doi.org/10.1038/s41587-023-02099-7>
- [60] Bubis J. A., Arrey T. N., Damoc E., Delanghe B., Slovákova J., Sommer T. M., Kagawa H., Pichler P., Rivron N., Mechtler K., Matzinger M., *bioRxiv*, **2024**, 2024.02.01.578358[2024-02-27]. <https://doi.org/10.1101/2024.02.01.578358>
- [61] Bekker-Jensen D. B., Martínez-Val A., Steigerwald S., Rütther P., Fort K. L., Arrey T. N., Harder A., Makarov A., Olsen J. V., *Mol. Cell. Proteomics*, **2020**, *19*(4), 716—729
- [62] Reilly L., Lara E., Ramos D., Li Z., Pantazis C. B., Stadler J., Santiana M., Roberts J., Faghri F., Hao Y., Nalls M. A., Narayan P., Liu Y., Singleton A. B., Cookson M. R., Ward M. E., Qi Y. A., *Cell Rep. Methods*, **2023**, *3*(10), 100593
- [63] Ridgeway M. E., Lubeck M., Jordens J., Mann M., Park M. A., *Int. J. Mass Spectrom.*, **2018**, *425*, 22—35
- [64] Vasilopoulou C. G., Sulek K., Brunner A. D., Meitei N. S., Schweiger-Hufnagel U., Meyer S. W., Barsch A., Mann M., Meier F., *Nat. Commun.*, **2020**, *11*(1), 331
- [65] Meier F., Brunner A. D., Frank M., Ha A., Bludau I., Voytik E., Kaspar-Schoenefeld S., Lubeck M., Raether O., Bache N., Aebersold R., Collins B. C., Röst H. L., Mann M., *Nat. Methods*, **2020**, *17*(12), 1229—1236
- [66] Skowronek P., Thielert M., Voytik E., Tanzer M. C., Hansen F. M., Willems S., Karayel O., Brunner A. D., Meier F., Mann M., *Mol. Cell. Proteomics*, **2022**, *21*(9), 100279
- [67] Szyrwiel L., Sinn L., Ralser M., Demichev V., *bioRxiv*, **2022**, 2022.10.31.514544[2022-10-31]. <https://doi.org/10.1101/2022.10.31.514544>
- [68] Moseley M. A., Hughes C. J., Juvvadi P. R., Soderblom E. J., Lennon S., Perkins S. R., Thompson J. W., Steinbach W. J., Geromanos S. J., Wildgoose J., Langridge J. I., Richardson K., Vissers J. P. C., *J. Proteome Res.*, **2018**, *17*(2), 770—779
- [69] Messner C. B., Demichev V., Bloomfield N., Yu J. S. L., White M., Kreidl M., Egger A. S., Freiwald A., Ivosev G., Wasim F., Zelezniak A., Jürgens L., Suttorp N., Sander L. E., Kurth F., Lilley K. S., Müllerer M., Tate S., Ralser M., *Nat. Biotechnol.*, **2021**, *39*(7), 846—854
- [70] Skowronek P., Krohs F., Lubeck M., Wallmann G., Itang E. C. M., Koval P., Wahle M., Thielert M., Meier F., Willems S., Raether O., Mann M., *Mol. Cell. Proteomics*, **2023**, *22*(2), 100489
- [71] Distler U., Łacki M. K., Startek M. P., Teschner D., Brehmer S., Decker J., Schild T., Krieger J., Krohs F., Raether O., Hildebrandt A., Tenzer S., *bioRxiv*, **2023**, 2023.01.30.526204[2023-02-02]. <https://doi.org/10.1101/2023.01.30.526204>

- [72] Tian X., Permentier H. P., Bischoff R., *Mass Spectrom. Rev.*, **2023**, *42*(2), 546—576
- [73] Li J., Van Vranken J. G., Pontano Vaites L., Schweppe D. K., Huttlin E. L., Etienne C., Nandhikonda P., Viner R., Robitaille A. M., Thompson A. H., Kuhn K., Pike I., Bomgardner R. D., Rogers J. C., Gygi S. P., Paulo J. A., *Nat. Methods*, **2020**, *17*(4), 399—404
- [74] Ning X., Li Q., Zi J., Mei Z., Liu J., Zhang Y., Bi M., Ren Y., Liu X., Lv C., Yao H., Sun J., Rao F., Li S., Liu S., *Anal. Chem.*, **2023**, *95*(13), 5788—5795
- [75] Budnik B., Levy E., Harmange G., Slavov N., *Genome Biol.*, **2018**, *19*(1), 161
- [76] Petelski A. A., Emmott E., Leduc A., Huffman R. G., Specht H., Perlman D. H., Slavov N., *Nat. Protoc.*, **2021**, *16*(12), 5398—5425
- [77] Specht H., Emmott E., Petelski A. A., Huffman R. G., Perlman D. H., Serra M., Kharchenko P., Koller A., Slavov N., *Genome Biol.*, **2021**, *22*(1), 50
- [78] Derks J., Leduc A., Wallmann G., Huffman R. G., Willetts M., Khan S., Specht H., Ralser M., Demichev V., Slavov N., *Nat. Biotechnol.*, **2023**, *41*(1), 50—59
- [79] Kang U. B., Yeom J., Kim H., Lee C., *J. Proteome Res.*, **2010**, *9*(7), 3750—3758
- [80] Thielert M., Itang E. C. M., Ammar C., Rosenberger F. A., Bludau I., Schweizer L., Nordmann T. M., Skowronek P., Wahle M., Zeng W. F., Zhou X. X., Brunner A. D., Richter S., Levesque M. P., Theis F. J., Steger M., Mann M., *Mol. Syst. Biology*, **2023**, *19*(9), e11503
- [81] Nesvizhskii A. I., *J. Proteomics*, **2010**, *73*(11), 2092—2123
- [82] Mallick P., Kuster B., *Nat. Biotechnol.*, **2010**, *28*(7), 695—709
- [83] Noor Z., Ahn S. B., Baker M. S., Ranganathan S., Mohamedali A., *Brief. Bioinform.*, **2021**, *22*(2), 1620—1638
- [84] Tyanova S., Temu T., Cox J., *Nat. Protoc.*, **2016**, *11*(12), 2301—2319
- [85] Kong A. T., Leprevost F. V., Avtonomov D. M., Mellacheruvu D., Nesvizhskii A. I., *Nat. Methods*, **2017**, *14*(5), 513—520
- [86] da Veiga Leprevost F., Haynes S. E., Avtonomov D. M., Chang H. Y., Shanmugam A. K., Mellacheruvu D., Kong A. T., Nesvizhskii A. I., *Nat. Methods*, **2020**, *17*(9), 869—870
- [87] Chi H., Liu C., Yang H., Zeng W. F., Wu L., Zhou W. J., Wang R. M., Niu X. N., Ding Y. H., Zhang Y., Wang Z. W., Chen Z. L., Sun R. X., Liu T., Tan G. M., Dong M. Q., Xu P., Zhang P. H., He S. M., *Nat. Biotechnol.*, **2018**, *36*(11), 1059—1061
- [88] Boekweg H., Van Der Watt D., Truong T., Johnston S. M., Guise A. J., Plowey E. D., Kelly R. T., Payne S. H., *J. Proteome Res.*, **2022**, *21*(1), 182—188
- [89] Huffman R. G., Chen A., Specht H., Slavov N., *J. Proteome Res.*, **2019**, *18*(6), 2493—2500
- [90] Chen A. T., Franks A., Slavov N., *Plos Comput. Biol.*, **2019**, *15*(7), e1007082
- [91] Wang B., Wang Y., Chen Y., Gao M., Ren J., Guo Y., Situ C., Qi Y., Zhu H., Li Y., Guo X., *Brief. Bioinform.*, **2022**, *23*(4), bbac214
- [92] Yu F., Haynes S. E., Nesvizhskii A. I., *Mol. Cell. Proteomics*, **2021**, *20*, 100077
- [93] Kalxdorf M., Müller T., Stegle O., Krijgsveld J., *Nat. Commun.*, **2021**, *12*(1), 4787
- [94] Tsou C. C., Avtonomov D., Larsen B., Tucholska M., Choi H., Gingras A. C., Nesvizhskii A. I., *Nat. Methods*, **2015**, *12*(3), 258—264
- [95] Röst H. L., Rosenberger G., Navarro P., Gillet L., Miladinović S. M., Schubert O. T., Wolski W., Collins B. C., Malmström J., Malmström L., Aebersold R., *Nat. Biotechnol.*, **2014**, *32*(3), 219—223
- [96] Zhou X. X., Zeng W. F., Chi H., Luo C., Liu C., Zhan J., He S. M., Zhang Z., *Anal. Chem.*, **2017**, *89*(23), 12690—12697
- [97] Cox J., *Nat. Biotechnol.*, **2023**, *41*(1), 33—43
- [98] Gessulat S., Schmidt T., Zolg D. P., Samaras P., Schnatbaum K., Zerweck J., Knaute T., Rechenberger J., Delanghe B., Huhmer A., Reimer U., Ehrlich H. C., Aiche S., Kuster B., Wilhelm M., *Nat. Methods*, **2019**, *16*(6), 509—518
- [99] Tiwary S., Levy R., Gutenbrunner P., Salinas Soto F., Palaniappan K. K., Deming L., Berndl M., Brant A., Cimercancic P., Cox J., *Nat. Methods*, **2019**, *16*(6), 519—525
- [100] Yang Y., Liu X., Shen C., Lin Y., Yang P., Qiao L., *Nat. Commun.*, **2020**, *11*(1), 146
- [101] Lou R., Liu W., Li R., Li S., He X., Shui W., *Nat. Commun.*, **2021**, *12*(1), 6685
- [102] Lou R., Tang P., Ding K., Li S., Tian C., Li Y., Zhao S., Zhang Y., Shui W., *iScience*, **2020**, *23*(3), 100903
- [103] Hao Y., Chen M., Huang X., Xu H., Wu P., Chen S., *Anal. Chem.*, **2023**, *95*(37), 14077—14085
- [104] Chen M., Zhu P., Wan Q., Ruan X., Wu P., Hao Y., Zhang Z., Sun J., Nie W., Chen S., *Anal. Chem.*, **2023**, *95*(19), 7495—7502
- [105] Zeng W. F., Zhou X. X., Willems S., Ammar C., Wahle M., Bludau I., Voytik E., Strauss M. T., Mann M., *Nat. Commun.*, **2022**, *13*(1), 7238
- [106] Yang Y., Lin L., Qiao L., *Expert. Rev. Proteomics*, **2021**, *18*(12), 1031—1043
- [107] Rosenberger G., Bludau I., Schmitt U., Heusel M., Hunter C. L., Liu Y., MacCoss M. J., MacLean B. X., Nesvizhskii A. I., Pedrioli P. G. A., Reiter L., Röst H. L., Tate S., Ting Y. S., Collins B. C., Aebersold R., *Nat. Methods*, **2017**, *14*(9), 921—927
- [108] Siyal A. A., Chen E. S. W., Chan H. J., Kitata R. B., Yang J. C., Tu H. L., Chen Y. J., *Anal. Chem.*, **2021**, *93*(51), 17003—17011
- [109] Searle B. C., Pino L. K., Egerton J. D., Ting Y. S., Lawrence R. T., MacLean B. X., Villén J., MacCoss M. J., *Nat. Commun.*, **2018**, *9*(1), 5128
- [110] Demichev V., Messner C. B., Vernardis S. I., Lilley K. S., Ralser M., *Nat. Methods*, **2020**, *17*(1), 41—44

- [111] Demichev V., Szyrwił L., Yu F., Teo G. C., Rosenberger G., Niewianda A., Ludwig D., Decker J., Kaspar-Schoenefeld S., Lilley K. S., Müller M., Nesvizhskii A. I., Ralser M., *Nat. Commun.*, **2022**, *13*(1), 3944
- [112] Yu F., Teo G. C., Kong A. T., Fröhlich K., Li G. X., Demichev V., Nesvizhskii A. I., *Nat. Commun.*, **2023**, *14*(1), 4154
- [113] Bruderer R., Bernhardt O. M., Gandhi T., Miladinović S. M., Cheng L. Y., Messner S., Ehrenberger T., Zanotelli V., Butscheid Y., Escher C., Vitek O., Rinner O., Reiter L., *Mol. Cell. Proteomics*, **2015**, *14*(5), 1400—1410
- [114] Bekker-Jensen D. B., Bernhardt O. M., Hogrebe A., Martinez-Val A., Verbeke L., Gandhi T., Kelstrup C. D., Reiter L., Olsen J. V., *Nat. Commun.*, **2020**, *11*(1), 787
- [115] Muntel J., Gandhi T., Verbeke L., Bernhardt O. M., Treiber T., Bruderer R., Reiter L., *Mol. Omics*, **2019**, *15*(5), 348—360
- [116] Tsai C. F., Zhao R., Williams S. M., Moore R. J., Schultz K., Chrisler W. B., Pasa-Tolic L., Rodland K. D., Smith R. D., Shi T., Zhu Y., Liu T., *Mol. Cell. Proteomics*, **2020**, *19*(5), 828—838
- [117] Schoof E. M., Furtwängler B., Üresin N., Rapin N., Savickas S., Gentil C., Lechman E., Keller U. a. d., Dick J. E., Porse B. T., *Nat. Commun.*, **2021**, *12*(1), 3341
- [118] Orsburn B. C., Yuan Y., Bumpus N. N., *Nat. Commun.*, **2022**, *13*(1), 7246
- [119] Petrosius V., Aragon-Fernandez P., Arrey T. N., Üresin N., Furtwängler B., Stewart H., Denisov E., Petzoldt J., Peterson A. C., Hock C., Damoc E., Makarov A., Zabrouskov V., Porse B. T., Schoof E. M., *bioRxiv*, **2023**, 2023.06.06.543943[2023-06-08]. <https://doi.org/10.1101/2023.06.06.543943>
- [120] Jiang Y. R., Zhu L., Cao L. R., Wu Q., Chen J. B., Wang Y., Wu J., Zhang T. Y., Wang Z. L., Guan Z. Y., Xu Q. Q., Fan Q. X., Shi S. W., Wang H. F., Pan J. Z., Fu X. D., Wang Y., Fang Q., *Cell Rep.*, **2023**, *42*(11), 113455
- [121] Wu J., Xu Q. Q., Jiang Y. R., Chen J. B., Ying W. X., Fan Q. X., Wang H. F., Wang Y., Shi S. W., Pan J. Z., Fang Q., *Anal. Chem.*, **2024**, *96*(14), 5499—5508
- [122] Gatto L., Aebersold R., Cox J., Demichev V., Derks J., Emmott E., Franks A. M., Ivanov A. R., Kelly R. T., Khoury L., Leduc A., MacCoss M. J., Nemes P., Perlman D. H., Petelski A. A., Rose C. M., Schoof E. M., Van Eyk J., Vanderaa C., Yates J. R., Slavov N., *Nat. Methods*, **2023**, *20*(3), 375—386
- [123] Caulfield J. H., Fu J., Wang D., Guevara-Gonzalez V., Wang W., Ping P., *J. Proteome Res.*, **2021**, *20*(5), 2182—2186
- [124] Wang F., Liu C., Li J., Yang F., Song J., Zang T., Yao J., Wang G., *Nucleic Acids Res.*, **2024**, *52*(D1), D562-D571
- [125] Li W., Yang F., Wang F., Rong Y., Liu L., Wu B., Zhang H., Yao J., *Nat. Methods*, **2024**, *21*(4), 623—634
- [126] Wang F., Yang F., Huang L., Li W., Song J., Gasser R. B., Aebersold R., Wang G., Yao J., *Nat. Mach. Intell.*, **2023**, *5*(11), 1236—1249

(Ed.: X, N, K)