

FASP 和 iST 酶切方案的比较分析： 蛋白质输入量的影响

陆 澳^{1,2}, 廖焕玥², 楚甜甜^{1,2}, 赵 洋², 江 游², 叶子弘¹, 孟 波²

(1. 中国计量大学生命科学学院, 浙江省生物计量及检验检疫技术重点实验室, 浙江 杭州 310018;

2. 中国计量科学研究院, 前沿计量科学中心, 国家市场监督管理总局技术创新中心(质谱), 北京 100029)

摘要: 微量蛋白样本的蛋白质组分析应用日益广泛。目前, 过滤辅助样品制备(filter aided sample preparation, FASP)和管内原位酶切(in-stage tip, iST)方案在微量蛋白样本制备中的使用效果尚未得到深入探讨。本研究针对 0.2~10 μg 微量蛋白样本, 探讨 2 种酶切方案的最佳蛋白输入量, 并在相同蛋白输入量下比较 2 种酶切方案。对于 iST 酶切方案, 当蛋白质输入量为 5 μg 时, 鉴定蛋白质的数目最多; 对于 FASP 酶切方案, 蛋白质的鉴定数目随着蛋白质输入量的增加而增加, 但增加幅度逐渐降低。同时, 分别从氨基酸序列覆盖率达到 $\geq 20\%$ 、鉴定频率为 3 以及变异系数 $< 5\%$ 的蛋白质数目 3 个方面评价各组实验方案的稳定性。结果表明, 当蛋白质输入量为 5 μg 时, iST 酶切方案鉴定的蛋白质数目均达到最大值, 稳定性达到最佳; 当蛋白质输入量从 0.2 μg 增加到 10 μg 时, FASP 酶切方案鉴定的蛋白质数目逐渐增大, 但增加幅度降低, 稳定性趋向最佳。在相同蛋白输入量下, 当蛋白质输入量 $\leq 1 \mu\text{g}$ 时, iST 酶切方案在鉴定蛋白质数目以及实验方案稳定性方面均优于 FASP 酶切方案; 而当蛋白质输入量 $> 1 \mu\text{g}$ 时, FASP 酶切方案更优。通过分析 iST 和 FASP 稳定鉴定的蛋白质, 发现 FASP 酶切方案稳定鉴定的蛋白质数目更多, 基因本体(gene ontology, GO)分析结果更丰富。

关键词: 液相色谱-串联质谱(LC-MS/MS); 蛋白质组学; 过滤辅助样品制备(FASP); 管内原位酶切(iST)

中图分类号: O657.63

文献标志码: A

文章编号: 1004-2997(2025)00-0001-09

DOI: 10.7538/zpxb.2024.0166

CSTR: 32365.14.zpxb.2024.0166

Comparison of FASP and iST Enzymatic Digestion Strategies: Impact of Protein Input Amount

LU Ao^{1,2}, LIAO Huan-yue², CHU Tian-tian^{1,2}, ZHAO Yang², JIANG You², YE Zi-hong¹, MENG Bo²

(1. Zhejiang Provincial Key Laboratory of Biometrology and Inspection & Quarantine, College of Life Sciences, China Jiliang University, Hangzhou 310018, China; 2. Technology Innovation Center of Mass Spectrometry for State Market Regulation, Center for Advanced Measurement Science, National Institute of Metrology, Beijing 100029, China)

Abstract: With the rapid development of mass spectrometry technology, it has been widely used in proteomic analysis of trace protein samples. To date, the effectiveness of filter-aided sample preparation (FASP) and in-stage tip (iST) protocols in microprotein sample preparation has not been thoroughly investigated. Although a number of researches have been performed to compare the capacity of the two enzyme digestion protocols for identifying the proteins, no studies have been conducted when the protein input is lower than 1 μg , and the depth of analysis is shallow in terms of

subsequent data analysis. In the present work, the optimal protein inputs of the two enzyme digestion schemes for 0.2-10 μg microprotein samples were studied, and the FASP and iST enzyme digestion protocols under the same protein input amount were compared. This study also systematically evaluated the two enzyme digestion protocols in terms of the numbers of identified proteins and peptides, the coverage of amino acid sequences, the frequency of identification, the coefficient of variation (CV), and the linear correlation within the groups of experiment. For iST, when the protein input was 5 μg , the number of identified proteins reached the highest value. For FASP, the number of identified proteins increased with increase in the amount of protein input, while the magnitude of the increase decreased after the protein input reaches to 5 μg . Further, the stability of each protocol of each group was evaluated from three aspects, including the number of identified proteins with amino acid sequence coverage of no less than 20%, 3 times in the triplicate samples, and a coefficient of variation of less than 5%. When the protein input was 5 μg , the number of proteins identified by iST in the three stability assays was maximum, meaning the stability reached to the best. When the protein input increased from 0.2 to 10 μg , the number of proteins identified in the three stability assays of FASP gradually increased with increase in protein inputs, while the magnitude of the increase decreased when the protein input was larger than 5 μg , and the stability of the protocol tended to the best. Under the same protein input, when the protein input amount was no more than 1 μg , the iST digestion protocol was better than FASP in terms of the number of proteins identified and the stability of the analysis procedure. When the protein input was more than 1 μg , the result was opposite. In summary, by comparing the numbers of proteins identified by iST and FASP with the protein input ranging from 0.2 to 10 μg , the FASP digestion protocol leads to identify more proteins with better stability, and gene ontology (GO) analysis can provide richer biological information.

Key words: liquid chromatography-tandem mass spectrometry (LC-MS/MS); proteomics; filter-aided sample preparation (FASP); in-stage tip (iST)

近些年, 质谱技术的快速发展使微量样本的蛋白质组学分析越来越受到重视^[1-2]。目前, “Bottom-up” 质谱是蛋白质组学分析的首选方法。该方法的操作步骤为: 首先使用特异性蛋白酶将蛋白质酶解为肽段; 然后根据肽段的亲水性, 使用液相色谱将肽段从 C18 分析柱上洗脱下来^[3]; 接着, 洗脱的肽段进入离子源内进行离子化处理; 最后, 离子化肽段通过离子传输管进入质谱。

样品制备作为蛋白质组学分析的基础会对蛋白质组的灵敏度、稳定性和重现性产生一定的影响^[4]。过滤辅助样品制备 (filter aided sample preparation, FASP)^[5]和管内原位酶切 (in-stage tip, iST)^[6]是目前蛋白质组学中常用的样本制备方案。FASP 最初由 Manza 等^[7]提出, 借助特定孔径的滤膜实现小分子成分和杂质的去除, 在过滤器中对蛋白质进行还原、烷基化和酶切, 最后在离心力的辅助作用下, 肽段穿过滤膜被分离; 该

方法可快速清除含有盐、酸碱缓冲液、洗涤剂的样品, 适用于不同裂解液溶解的蛋白质样品。Jiang 等^[8]在早期肝细胞癌新的治疗靶点研究中, 使用 FASP 样本制备方案酶切蛋白, 蛋白输入量为 500 μg , 每个样本分为 6 个馏分进行质谱分析, 平均每个癌组织样本鉴定出 5 953 个蛋白质, 每个癌旁组织样本鉴定出 5 114 个蛋白质, 但由于样本制备方案繁琐, 对于蛋白量较低的样本效果可能较差。Stage Tips (stop and go extraction tips) 是由 Rappsilber 等^[9]于 2007 年提出的, 其制作方法是 将 C18 膜、阳离子交换膜 (SCX) 或阴离子交换膜 (SAX) 等加入 100 μL 或 200 μL Tip 枪头中制成, 用于多肽的多维分馏和脱盐。2014 年, Kulak 等^[10]对 Stage Tip 进行进一步的开发, 提出了 iST 蛋白酶切策略, 并基于 iST-SCX 策略从 10 μg 蛋白中成功鉴定了 4 087 个酵母细胞蛋白质; 该策略能够在单个容器中完成蛋白质的消化、多肽的分离和脱盐, 达到减少操作步骤、

缩短操作时间和降低样品损失的目的。Kitata 等^[1]采用 iST 蛋白酶解策略处理 5 种非小细胞癌细胞系以及 22 个肿瘤组织合并的样本, 并使用 MaxQuant 构建了包含 12 344 个蛋白质的谱图库。针对 iST 和 FASP 酶切方案在微量蛋白样本中的使用效果, Malte 等^[4]评估了 2 种酶切方案在蛋白输入量为 1~20 μg 范围内鉴定的蛋白质组情况, 并从蛋白质鉴定数目、蛋白质丰度的变异系数(CV)以及相关性进行数据分析。

虽然上述文献比较了 2 种酶切方案鉴定蛋白质组的效果, 但后续的数据分析深度较浅, 且尚未见对于蛋白输入量低于 1 μg 的报道。因此, 本研究基于人胚胎肾细胞 293T(HEK293T)细胞系全蛋白设计梯度蛋白输入量实验方案(0.2~10 μg), 探讨 FASP 和 iST 酶切方案在微量蛋白样本中的使用效果, 并从蛋白质和肽段的鉴定量、氨基酸序列覆盖率、鉴定频率为 3 的蛋白质数目、变异系数(CV)<5%的蛋白质数目、各组实验的组内线性相关性等方面系统评估 2 种酶切方案。

1 实验部分

1.1 仪器与试剂

Orbitrap Fusion Lumos Tribrid 质谱仪、Easy-nLC 1200 纳升级液相色谱仪、NanoDrop OneC 微量紫外分光光度计、Multiskan sky 全波长酶标仪和 ThermoMixer C 金属浴: 美国 Thermo Scientific 公司产品; Vacufuge Plus 真空离心浓缩仪: 德国 Eppendorf 公司产品; JY92-IIN 超声波细胞粉碎

机: 宁波新芝生物科技公司产品; Nanosep 30 ku 超滤离心管(Omega 膜): 美国 Pall Corporation 公司产品。

尿素(UA)、碳酸氢铵(NH_4HCO_3)、蛋白酶抑制剂、氯乙酰胺(CAA): 均为分析纯, 德国 Sigma 公司产品; 胰蛋白酶(Trypsin, 25 $\mu\text{g}/\text{瓶}$, 纯度 $\geq 95\%$, 质谱级, 还原甲基化修饰): 北京酶知源生物科技有限公司产品; SDB-RPS 固相萃取盘: 美国 3M 公司产品; 三氟乙酸(TFA, 分析纯)、三(2-羧乙基)膦(TCEP, 分析纯)和 Tris-HCl(pH 8.0, 分析纯)、乙腈(ACN, 质谱纯): 美国 Thermo Scientific 公司产品。

1.2 实验方案设计

本研究基于 HEK293T 细胞系全蛋白设计了梯度蛋白输入量实验方案, 探讨 FASP 和 iST 2 种常用的酶切方案在微量蛋白样本中的使用效果, 示于图 1。使用细胞超声破碎仪处理同一批次的 3 盘 HEK293T 细胞样本, 并将获取的蛋白混合作为研究样本。2 种酶切方案有 0.2、1、2、5 和 10 μg 5 种蛋白输入量, 共进行 10 组实验, 每组实验重复 3 次。每个重复实验获得的肽段进行 1 次质谱分析, 使用 MaxQuant 解析质谱数据。

1.3 细胞培养与收集

HEK293T 细胞购自武汉普诺赛生命科技有限公司。细胞在 37 $^{\circ}\text{C}$ 、5% CO_2 培养箱中培养, 并添加杜氏改良 Eagle 培养基(DMEM)、10%胎牛血清(FBS)、1%青霉素-链霉素双抗(P/S)。从长满细胞的培养皿中吸出原培养液, 加入 2 mL 磷酸缓冲盐溶液(PBS), 轻轻晃动培养皿以润洗

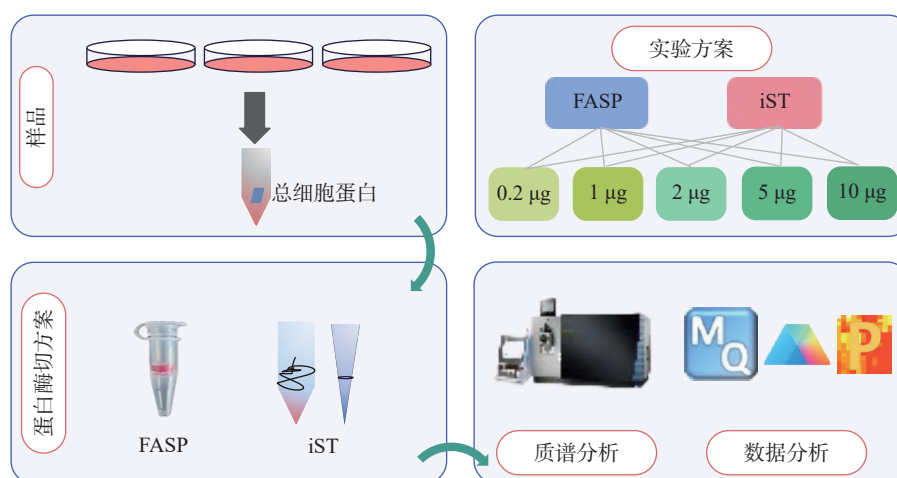


图 1 实验方案设计

Fig. 1 Experimental design

细胞;加入3 mL胰蛋白酶消化5 min,再加入等量的培养基终止消化;将其转移至15 mL离心管中,以1 500 r/min离心3 min,再用PBS洗涤3次,于-80 °C冰箱保存。

1.4 细胞蛋白提取

将收集的HEK293T细胞转移至1.5 mL离心管中,并置于冰面上。向离心管中加入570 μ L蛋白裂解液(8 mol/L UA, 100 mmol/L Tris-HCl, pH 8)、30 μ L蛋白酶抑制剂(1 $\times$),于冰上静置10 min,待细胞充分裂解。随后进行细胞超声破碎,设置参数为功率200 W、超声2 s、暂停2 s、超声总时长5 min。超声后在冰上静置10 min,然后在4 °C离心机中以17 000 r/min离心20 min,取上清液至新的离心管中,即得HEK293T细胞蛋白。采用Bradford方案测定提取的细胞蛋白浓度^[12]。

1.5 FASP 方案

将不同蛋白量(0.2、1、2、5、10 μ g)的样品加载至超滤管中,离心去除废液,用8 mol/L UA(100 mmol/L Tris-HCl, pH 8)冲洗柱子3次,离心后弃废液;然后加入5 mmol/L TCEP和25 mmol/L CAA混合液,于37 °C恒温箱中孵育1 h;再用50 mmol/L NH_4HCO_3 冲洗柱子3次;更换套管,加入Trypsin(酶:蛋白=1:10 (m/m)),37 °C恒温箱中酶切16 h;酶切结束后离心收集酶切液,用50 mmol/L NH_4HCO_3 洗涤2次,合并离心液;在45 °C、VAQ模式下的真空离心浓缩仪中热干肽段,随后保存于-80 °C冰箱,待后续LC-MS/MS分析。实验中离心转速为16 000 r/min,每次15 min。

1.6 iST 样本制备方案

用25 mmol/L Tris-HCl稀释蛋白样品,并置于95 °C金属浴上孵育10 min;将不同蛋白量(0.2、1、2、5、10 μ g)样品加至载有SDB-RPS膜的Tip柱中(10 μ g蛋白使用2层SDB-RPS膜,其余实验组均使用1层膜),加入5 mmol/L TCEP、25 mmol/L CAA和25 mmol/L Tris-HCl使总体积为20 μ L,37 °C金属浴1 h,转速1 500 r/min;还原烷基化结束后,加入Trypsin(酶:蛋白=1:10 (m/m)),10%ACN/25 mmol/L Tris-HCl至总体积为100 μ L,37 °C金属浴16 h,转速1 500 r/min;然后加入1 μ L TFA终止酶切,4 000 r/min离心1 min,弃废液;再加入100 μ L 0.2%TFA,洗涤肽段1次,

4 000 r/min离心1 min;最后使用80%乙腈(ACN)/5% NH_4HCO_3 洗脱肽段,并在45 °C、VAQ模式下的真空离心浓缩仪中热干肽段,随后保存于-80 °C冰箱。

1.7 实验条件

1.7.1 色谱条件 自制的C18毛细管分析柱(30 cm $\times$ 100 μ m, 1.9 μ m)。流动相:A为0.1%甲酸(FA),B为80%ACN-0.1%FA。使用流动相A溶解肽段,用NanoDrop OneC在205 nm吸光度下测得肽段浓度。取1 μ g肽段经流动相A加载至毛细管柱中,流速600 nL/min,梯度洗脱程序为0~5 min(4%~10%B),5~48 min(10%~22%B),48~66 min(22%~35%B),66~76 min(35%~90%B),76~78 min(90%B)。

1.7.2 质谱条件 一级和二级质谱分析均采用Orbitrap质量分析器进行检测,一级质谱分辨率为120 000,最大离子注入时间50 ms,质量扫描范围 m/z 350~1 500;二级质谱分辨率为15 000,最大离子注入时间30 ms,高能碰撞解离(HCD)模式,碰撞能量30%,隔离窗口1.6 u,动态排除时间18 s,循环时间3 s。

1.8 数据分析

从Uniprot下载用于检索的智人蛋白质数据库,通过MaxQuant^[13](version 2.0.3,下载自<https://www.maxquant.org/>)对原始质谱文件进行定性定量分析。参数设置:蛋白水解酶为Trypsin,每个肽段最多允许有2个位点漏切,母离子、子离子的质量偏差分别为 2×10^{-5} 、 4.5×10^{-6} ;以carbamidomethyl(C)为固定修饰,甲硫氨酸氧化和N-乙酰化为可变修饰;错误发现率(FDR)<1%;其他参数设置默认不变。使用Perseus(下载自<https://www.maxquant.org/perseus/>)^[14]、Excel和GraphPad Prism(v.10.1.2,下载自<https://graphpad-prism.cn/>)等工具完成数据分析。基因本体(gene ontology, GO)富集分析使用Hplot Pro(<https://hiplot.com.cn/>)中的GO/KEGG完成,选择 $P<0.05$ 为显著条目,使用GraphPad Prism完成可视化。

2 结果与讨论

2.1 各组实验鉴定蛋白质和肽段的数目

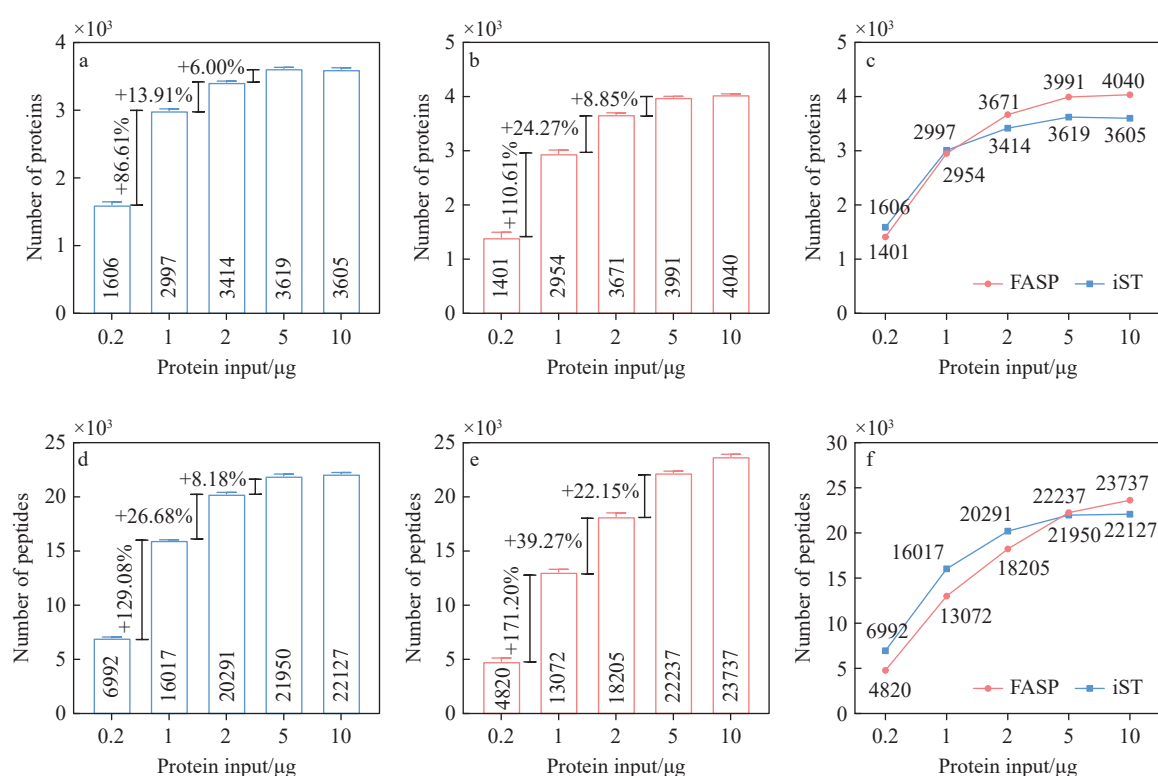
本研究比较了同一种酶切方案在不同蛋白输入量下鉴定蛋白质的数目。对于iST酶切方案,当蛋白输入量从0.2 μ g增加到5 μ g,鉴定蛋

白质数目的均值从 1 606 增加到 3 619 个; 当输入量从 5 μg 增加到 10 μg , 蛋白质的鉴定数目基本保持不变, 示于图 2a。对于 FASP 酶切方案, 当蛋白输入量从 0.2 μg 增加到 5 μg , 鉴定蛋白质数目的均值从 1 401 增加到 3 991 个; 当输入量从 5 μg 增加到 10 μg , 蛋白质的鉴定数目基本保持不变, 示于图 2b。在相同蛋白质输入量下, 比较 2 种酶切方案鉴定蛋白质的数目, 示于图 2c。当蛋白输入量为 0.2 μg 时, iST 酶切方案鉴定的蛋白质数目比 FASP 多 205 个 (14.54%); 当蛋白输入量为 1 μg 时, 2 种酶切方案鉴定蛋白质的数目相近; 当蛋白输入量为 2、5、10 μg 时, FASP 鉴定的蛋白质数目比 iST 多 257 个 (7.53%)、372 个 (10.28%)、435 个 (12.07%)。可知, 当蛋白输入量超过 1 μg , 随着蛋白输入量的增加, FASP 鉴定的蛋白质数目比 iST 多, 但增加幅度减小。在肽段层面, iST 酶切方案鉴定肽段的数目变化趋势与鉴定蛋白质数目基本一致, 示于图 2d; FASP 酶切方案鉴定肽段的数目随蛋白输入量的增加

而增加, 但增加幅度逐渐降低, 示于图 2e。在相同的蛋白输入量下, 当蛋白输入量 $>5 \mu\text{g}$ 时, FASP 鉴定肽段数目比 iST 多, 示于图 2f。

2.2 各组实验鉴定蛋白质的准确性评价

肽段序列覆盖率能够有效反映蛋白质鉴定的准确性。本实验比较了各组实验鉴定蛋白质的氨基酸序列覆盖率。对于 iST 酶切方案, 当蛋白输入量分别为 1、2、5、10 μg 时, 鉴定氨基酸序列覆盖率 $\geq 20\%$ 的蛋白质数目是蛋白输入量分别为 0.2、1、2、5 μg 的 2.39、1.30、1.04、1.07 倍, 示于图 3a。可知, 当蛋白输入量 $\leq 5 \mu\text{g}$ 时, 随着蛋白输入量的增加, 鉴定氨基酸序列覆盖率 $\geq 20\%$ 的数目增加, 但增加幅度降低。对于 FASP 酶切方案, 当蛋白输入量分别为 1、2、5、10 μg 时, 鉴定氨基酸序列覆盖率 $\geq 20\%$ 的蛋白质数目是蛋白输入量分别为 0.2、1、2、5 μg 的 2.65、1.47、1.26、1.14 倍, 示于图 3b。可知, 随着蛋白输入量的增加, 鉴定氨基酸序列覆盖率 $\geq 20\%$ 的数目增加, 但增加幅度降低。在相同蛋白输入量



注: a. 基于 iST 酶切方案, 不同的蛋白输入量鉴定蛋白质的数目; b. 基于 FASP 酶切方案, 不同的蛋白输入量鉴定蛋白质的数目; c. 相同蛋白输入量下, FASP 和 iST 酶切方案鉴定蛋白质的数目; d. 基于 iST 酶切方案, 不同的蛋白输入量鉴定肽段的数目; e. 基于 FASP 酶切方案, 不同的蛋白输入量鉴定肽段的数目; f. 相同蛋白输入量下, FASP 和 iST 酶切方案鉴定肽段的数目

图 2 各组实验鉴定蛋白质和肽段数目的比较

Fig. 2 Comparison of the number of experimentally identified proteins and peptides in each group

下比较 FASP 和 iST 酶切方案鉴定蛋白质的氨基酸序列覆盖率, 结果示于图 3c。可见, 当蛋白输入量 $\leq 5 \mu\text{g}$ 时, iST 酶切方案鉴定到氨基酸序列覆盖率 $\geq 20\%$ 的蛋白质数目比 FASP 酶切方案多; 当蛋白输入量为 $10 \mu\text{g}$ 时, 则相反。

2.3 各组实验鉴定蛋白质的稳定性评价

在 3 次重复实验中均能鉴定到的蛋白质可以作为评估样本制备方案重复性的关键指标。对于 iST 酶切方案, 蛋白输入量从 $0.2 \mu\text{g}$ 增加到 $5 \mu\text{g}$, 鉴定频率为 3 的蛋白质数目逐渐增加, 但增加幅度逐渐降低; 当蛋白输入量增加至 $10 \mu\text{g}$ 时, 其数目与 $5 \mu\text{g}$ 的基本一致, 示于图 4a。对于 FASP 酶切方案, 随着蛋白输入量的增加, 鉴定频率为 3 的蛋白质数目的变化趋势与 iST 酶切方案基本一致, 示于图 4b。在相同蛋白输入量下比较 2 种酶切方案鉴定蛋白质的频率, 结果示于图 4c。当蛋白输入量 $\leq 1 \mu\text{g}$ 时, iST 酶切方案鉴定到频率为 3 的蛋白质数目比 FASP 多; 当蛋白输入量 $\geq 2 \mu\text{g}$ 时, FASP 酶切方案鉴定到频率为 3 的蛋白质数目比 iST 多。

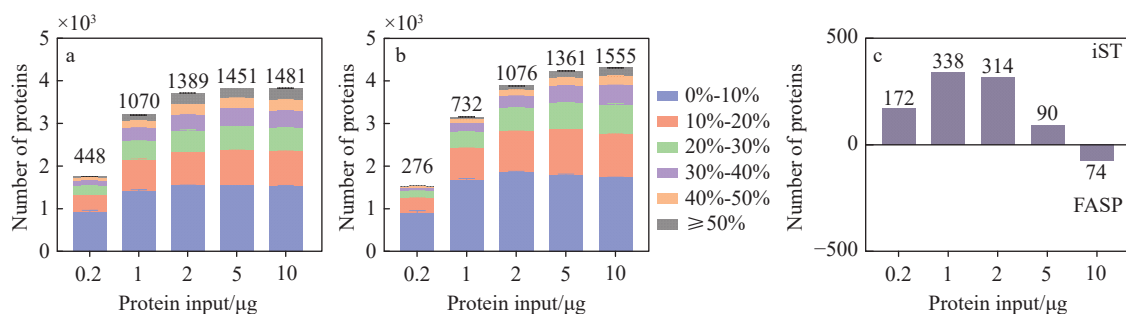
对各组可定量蛋白质依次进行中位数转换和全局最小值插值处理, 再计算各组鉴定蛋白质的变异系数(CV)。对于 iST 酶切方案, 蛋白输入量从 $0.2 \mu\text{g}$ 增加到 $5 \mu\text{g}$, $CV < 5\%$ 的蛋白质数目逐渐增加, 但增加幅度逐渐降低; 当蛋白输入量增加至 $10 \mu\text{g}$ 时, $CV < 5\%$ 的蛋白质数目与 $5 \mu\text{g}$ 的基本一致, 示于图 4d。对于 FASP 酶切方案, 随着蛋白输入量的增加, $CV < 5\%$ 的蛋白质数目变化趋势与 iST 酶切方案基本一致, 示于图 4e。在相同蛋白输入量下比较 2 种酶切方案鉴定蛋白质的 CV, 结果示于图 4f。当蛋白

输入量 $\leq 1 \mu\text{g}$ 时, iST 酶切方案鉴定到 $CV < 5\%$ 的蛋白质数目比 FASP 多; 当蛋白输入量 $\geq 2 \mu\text{g}$ 时, FASP 酶切方案鉴定到 $CV < 5\%$ 的蛋白质数目比 iST 多。

对各组实验进行线性相关性分析, 均表现出较好的组内线性关系, 除 FASP_0.2 μg 的相关系数(R^2)均值为 0.83 外, 其他各组 R^2 均值均大于 0.90, 示于图 4g。当输入量为 0.2 μg 时, FASP 和 iST 鉴定结果之间的 R^2 均值仅为 0.71, 但随着蛋白输入量增加, 其相关性也随之增强。

2.4 FASP 和 iST 稳定鉴定蛋白质的交集分析及特征鉴定蛋白质的功能分析

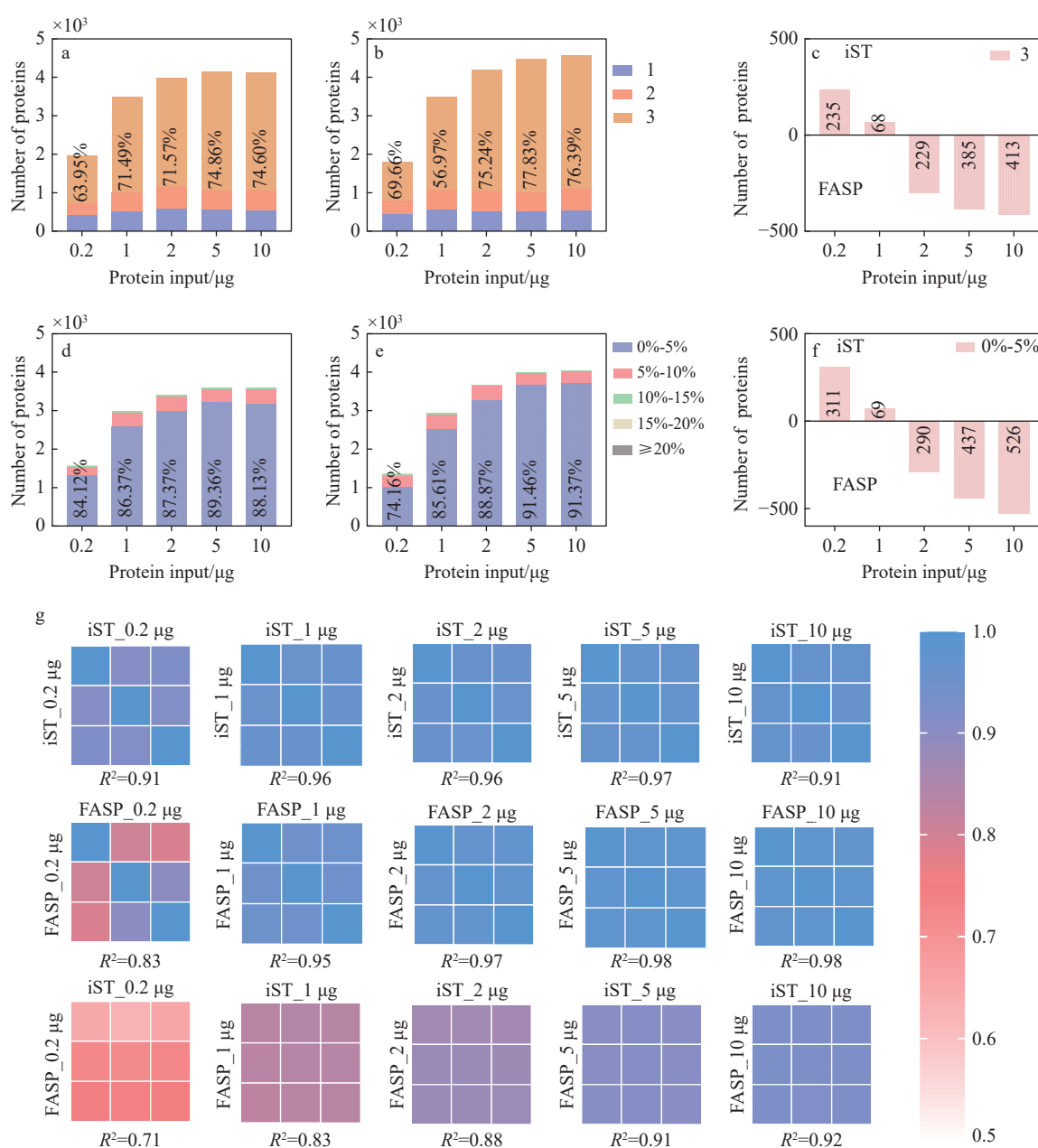
对于 iST 酶切方案, 当蛋白输入量为 5 和 $10 \mu\text{g}$ 时, 2 种输入量下同时鉴定的蛋白质有 3 834 个, 占鉴定所有蛋白质的 86.12%, 标记此数据集为 iST_A, 示于图 5a, 表明 iST_5 μg 和 iST_10 μg 鉴定蛋白质的差异较小。对于 FASP 酶切方案, 当蛋白输入量为 5 和 $10 \mu\text{g}$ 时, 2 种输入量同时鉴定的蛋白质有 4 261 个, 占鉴定所有蛋白质的 88.68%, 标记此数据集为 FASP_A, 示于图 5b, 表明 FASP_5 μg 和 FASP_10 μg 鉴定蛋白质的差异较小。以上结果表明, 随着蛋白输入量的增加, iST 和 FASP 酶切鉴定蛋白质的种类逐渐趋于稳定。筛选出 iST_A 和 FASP_A 中鉴定频率为 3 的蛋白质, 分别标记为 iST_B(图 5c)和 FASP_B(图 5d)。通过 iST_B 与 FASP_B 的交集分析发现, iST_B 和 FASP_B 特异性鉴定的蛋白质分别有 217、657 个, 示于图 5e。因此, 对于 2 种酶切方案, 当蛋白鉴定数目随着蛋白输入量的增加逐渐趋于稳定时, FASP 酶切方案可以稳定鉴定更多的蛋白质。



注: a. iST 酶切方案在各蛋白输入量下鉴定蛋白质的氨基酸序列覆盖率; b. FASP 酶切方案在各蛋白输入量下鉴定蛋白质的氨基酸序列覆盖率; c. 相同蛋白输入量, iST 和 FASP 鉴定氨基酸序列覆盖率 $\geq 20\%$ 的蛋白质数目差异

图 3 各组实验鉴定蛋白质的准确性评价

Fig. 3 Accuracy evaluation of identified proteins in each group



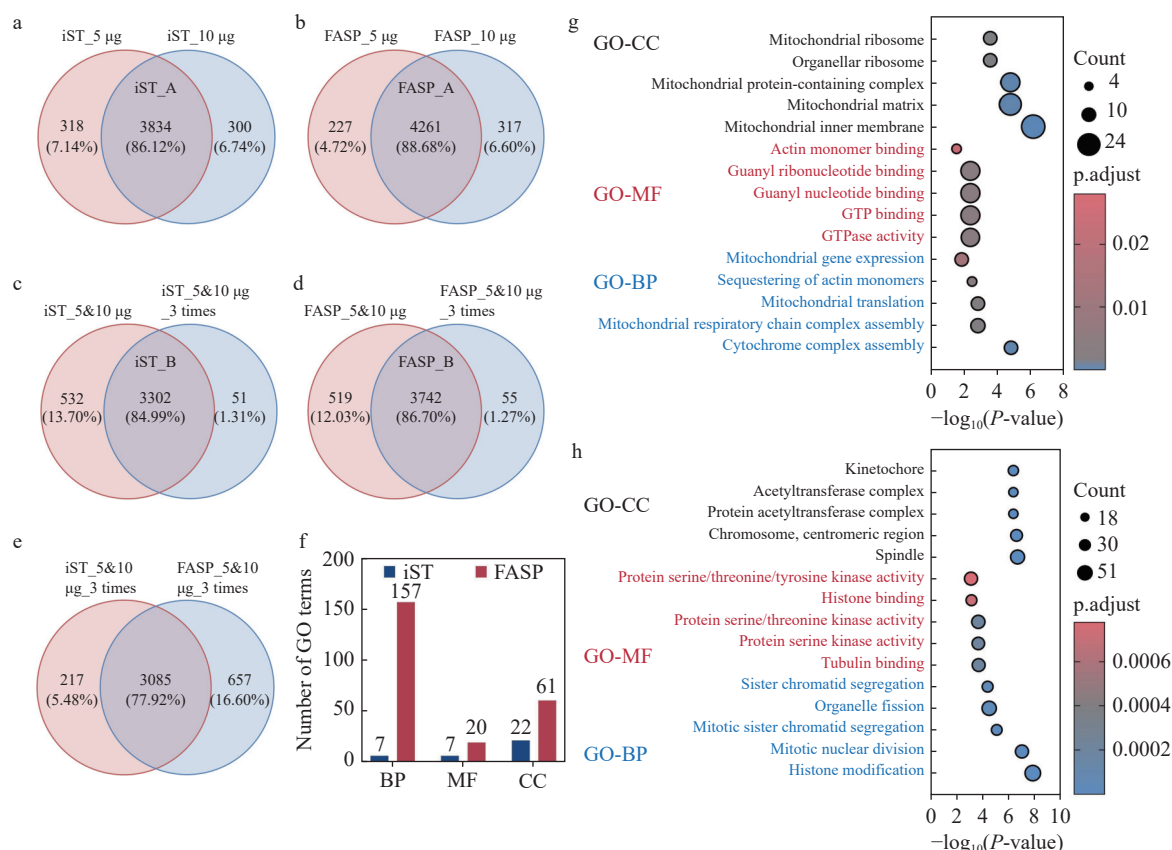
注: a. iST 酶切方案在各蛋白输入量下鉴定蛋白质的频率; b. FASP 酶切方案在各蛋白输入量下鉴定蛋白质的频率;
c. 相同蛋白输入量, iST 和 FASP 鉴定频率为 3 的蛋白质数目差异; d. iST 酶切方案在各蛋白输入量下鉴定蛋白质的变异系数;
e. FASP 酶切方案在各蛋白输入量下鉴定蛋白质的变异系数; f. 相同蛋白输入量, iST 和 FASP 鉴定蛋白质的变异系数在 0%~5% 之间的数目差异; g. 各组实验的线性相关性

图 4 各组实验鉴定蛋白质的稳定性评价

Fig. 4 Stability evaluation of identified proteins in each group

分别对 iST_B 与 FASP_B 特异性鉴定蛋白质进行 GO 富集分析, FASP_B 特异性鉴定蛋白质的 GO 分析结果更丰富, 示于图 5f。iST_B 特异性鉴定蛋白质主要定位于线粒体内膜、线粒体基质、含线粒体蛋白的复合体、细胞器核糖体和线粒体核糖体, 参与的生物过程包括细胞色素复合体组装、线粒体呼吸链复合体组装、线粒体

翻译、肌动蛋白单体的隔离和线粒体基因表达, 示于图 5g。FASP_B 特异性鉴定的蛋白质主要定位于纺锤体、染色体、着丝粒区域、蛋白质乙酰转移酶复合体、乙酰转移酶复合体和动粒, 参与的生物过程包括组蛋白修饰、有丝分裂核分裂、有丝分裂姐妹染色单体分离、细胞器分裂以及姐妹染色单体分离, 示于图 5h。



注: a. iST 酶切方案在蛋白输入量为 5 和 10 μ g 时鉴定蛋白质的交集分析; b. FASP 酶切方案在蛋白输入量为 5 和 10 μ g 时鉴定蛋白质的交集分析; c. iST 酶切方案在 5 和 10 μ g 共同鉴定的蛋白质与 5 和 10 μ g 鉴定频率为 3 的蛋白质的交集分析; d. FASP 酶切方案在 5 和 10 μ g 共同鉴定的蛋白质与 5 和 10 μ g 鉴定频率为 3 的蛋白质的交集分析; e. iST-B 与 FASP-B 的交集分析; f. iST 和 FASP 稳定鉴定且特异性鉴定蛋白质的 GO 分析结果; g. iST 酶切方案稳定鉴定且特异性鉴定蛋白质的 GO 富集结果; h. FASP 酶切方案稳定鉴定且特异性鉴定蛋白质的 GO 富集结果

图 5 FASP 和 iST 稳定鉴定蛋白质的交集分析及特征鉴定蛋白质的功能分析

Fig. 5 Intersection analysis of FASP and iST stable identified proteins and functional analysis of characterized identified proteins

3 结论

本研究基于 HEK293T 细胞系全蛋白, 探讨了蛋白输入量对 FASP 和 iST 酶切方案的影响, 以及在相同蛋白输入量下 2 种酶切方案的差异。对于 iST 酶切方案, 蛋白输入量为 5 μ g 时, 蛋白质鉴定数目最大, 氨基酸序列覆盖率 $\geq 20\%$ 的蛋白质数目与蛋白输入量为 10 μ g 的差异较小, 鉴定频率为 3 的蛋白质数目最多, $CV < 5\%$ 的蛋白质数目最多。因此, 在蛋白质输入量 $\geq 5 \mu$ g 时, iST 酶切方案才能获得最佳的鉴定结果。对于 FASP 酶切方案, 蛋白质的鉴定数目随着蛋白输入量的增加而增加, 但增加幅度逐渐降低; 蛋白输入量为 5 μ g 和 10 μ g 时, 鉴定蛋白质的数目相差较小; 氨基酸序列覆盖率 $\geq 20\%$ 、鉴定频率为 3、以及 $CV < 5\%$ 的蛋白质数目均在此研究中最大蛋白输入量 (10 μ g) 时达到最佳。因此, 对

于 0.2~10 μ g 微量样本, 蛋白输入量越大, FASP 酶切方案鉴定蛋白质数目越多, 实验方案越稳定。从鉴定蛋白质数目、氨基酸序列覆盖率 $\geq 20\%$ 、鉴定频率为 3 和 $CV < 5\%$ 的蛋白质数目等方面分析, 当蛋白输入量 $\leq 1 \mu$ g 时, iST 酶切方案优于 FASP; 当蛋白输入量 $> 1 \mu$ g 时, FASP 酶切方案优于 iST。FASP 酶切方案稳定鉴定的蛋白质数目比 iST 多, GO 分析结果更丰富。

参考文献:

- [1] LI Y, WEN Y, LI Y, TAN X, GAO S, FAN P, TIAN W, WONG C C L, CHEN Y. Rab10-CAV1 mediated intraluminal vesicle transport to migrasomes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2024, 121(30): e2319267121.
- [2] FRIEDRICH C, SCHALLENBERG S, KIRCHNER M,

- ZIEHM M, NIQUET S, HAJI M, BEIER C, NEUDECKER J, KLAUSCHEN F, MERTINS P. Comprehensive micro-scaled proteome and phosphoproteome characterization of archived retrospective cancer repositories[J]. *Nature Communications*, 2021, 12(1): 3 576.
- [3] DUONG V A, LEE H. Bottom-up proteomics: advancements in sample preparation[J]. *International Journal of Molecular Sciences*, 2023, 24(6): 5 350.
- [4] SIELAFF M, KUHAREV J, BOHN T, HAHNBROCK J, BOPP T, TENZER S, DISTLER U. Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range[J]. *Journal of Proteome Research*, 2017, 16(11): 4 060-4 072.
- [5] ZHANG Z, DUBIAK K M, HUBER P W, DOVICHIN J. Miniaturized filter-aided sample preparation (MICRO-FASP) method for high throughput, ultrasensitive proteomics sample preparation reveals proteome asymmetry in xenopus laevis embryos[J]. *Analytical Chemistry*, 2020, 92(7): 5 554-5 560.
- [6] ZHANG X, SADOWSKI P, PUNYADEERA C. Evaluation of sample preparation methods for label-free quantitative profiling of salivary proteome[J]. *Journal of Proteomics*, 2020, 210: 103 532.
- [7] MANZA L L, STAMER S L, HAM A L, CODREANU S G, LIEBLER D C. Sample preparation and digestion for proteomic analyses using spin filters[J]. *Proteomics*, 2005, 5(7): 1 742-1 745.
- [8] JIANG Y, SUN A, ZHAO Y, YING W, SUN H, YANG X, XING B, SUN W, REN L, HU B, LI C, ZHANG L, QIN G, ZHANG M, CHEN N, ZHANG M, HUANG Y, ZHOU J, ZHAO Y, LIU M, ZHU X, QIU Y, SUN Y, HUANG C, YAN M, WANG M, LIU W, TIAN F, XU H, ZHOU J, WU Z, SHI T, ZHU W, QIN J, XIE L, FAN J, QIAN X, HE F. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma[J]. *Nature*, 2019, 567(7 747): 257-261.
- [9] RAPPILBER J, MANN M, ISHIHAMA Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips[J]. *Nature Protocols*, 2007, 2(8): 1 896-1 906.
- [10] KULAK N A, PICHLER G, PARON I, NAGARAJ N, MANN M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells[J]. *Nature Methods*, 2014, 11(3): 319-324.
- [11] KITATA R B, CHOONG W K, TSAI C F, LIN P Y, CHEN B S, CHANG Y C, NESVIZHSHKII A I, SUNG T Y, CHEN Y J. A data-independent acquisition-based global phosphoproteomics system enables deep profiling[J]. *Nature Communications*, 2021, 12(1): 2 539.
- [12] KIELKOPF C L, BAUER W, URBATSCH I L. Bradford assay for determining protein concentration[J]. *Cold Spring Harbor Protocols*, 2020, 2020(4): 102 269.
- [13] COX J, MANN M. MaxQuant enables high peptide identification rates, individualized p.p.b. -range mass accuracies and proteome-wide protein quantification[J]. *Nature Biotechnology*, 2008, 26(12): 1 367-1 372.
- [14] TYANOVA S, TEMU T, SINITSYN P, CARLSON A, HEIN M Y, GEIGER T, MANN M, COX J. The Perseus computational platform for comprehensive analysis of (prote)omics data[J]. *Nature Methods*, 2016, 13(9): 731-740.

(收稿日期: 2024-09-13; 修回日期: 2025-02-18)