

# 精确修饰位点谱图库的建立与磷酸化蛋白质组的 DIA 解析

张伟 周岳

赛默飞世尔科技（中国）有限公司

## 关键词

翻译后修饰；DIA；谱图库；磷酸化；定量蛋白质组学

## 引言

数据非依赖采集（Data-Independent Acquisition, DIA）是当前最热门的质谱采集技术之一，它以非目标的方式将质量范围分为若干窗口，依次并循环采集窗口内所有母离子的二级碎片<sup>[1,2]</sup>。DIA 与 SRM 类似，也是基于子离子（transition）定量，相比传统蛋白质组学定量方法具有更好的选择性和更高的准确度。然而，目前DIA在翻译后修饰分析上仍有较大瓶颈。DIA 依赖于 DDA 建立谱图库，而 DDA 数据在搜库鉴定时，修饰位点定位错误的概率较高，特别是磷酸化修饰发生在常见的 S/T/Y 上，若肽段含有 2 个或以上位置接近的 S/T/Y，位点就容易找错。将含有错误位点信息的鉴定结果作为谱图库，就会导致翻译后修饰 DIA 解析结果的不可靠<sup>[3]</sup>。因此，DIA 尚难用于大规模的翻译后修饰样本分析。

针对修饰位点的打分算法使修饰位点的定位更加准确，Proteome Discoverer 软件中整合的 phosphoRS/ptmRS 模块<sup>[4]</sup>和 MaxQuant 软件的算法<sup>[5]</sup>都可以实现位点可信度（Site Probability）的计算，从而获得可靠的位点定位信息。本文基于上述软件对翻译后修饰 DDA 数据进行位点可信度分析，筛选具有准确位点定位的谱图建立谱图库，导入 Skyline 软件，进而实现可靠的翻译后修饰 DIA 解析。将该流程应用于磷酸化样本 DIA 数据分析，成功提取 6401 条高可信度的磷酸化肽段（ $Q < 0.01$ ），占谱图库肽段总数的 98.4%，其中可用于准确定量的肽段（ $CV < 20\%$ ）占 86.9%，有效解决了翻译后修饰 DIA 定量的难题。

## 实验条件

### 实验材料和方法

来源于大鼠组织富集的磷酸化样本，最终上样量为 700 ng/run，分别进行 DDA 和 DIA 采集，每种采集模式重复 3 遍。

### 色谱条件

纳流高效液相色谱仪：EASY-nLC 1000（Thermo Scientific）

分析柱：纳流 C18 色谱柱（长 15cm, ID 75  $\mu$ m, 粒径 3  $\mu$ m）

流动相：A：0.1% 甲酸水溶液；B：0.1% 甲酸乙腈溶液

梯度：0–3 min, 3–7% B；3–95 min, 7–22% B；95–113 min, 22–35% B；113–116 min, 35–90% B；116–120 min, 90% B

流速：300 nL/min

### 质谱条件

质谱仪：Orbitrap Fusion（Thermo Scientific）；

离子源：NanoFlex；离子模式：正离子；喷雾电压：1.8 kV；

毛细管温度：275° C；S-Lens RF：60%

DDA：分辨率：一级 120,000@ $m/z$  200，二级 30,000@ $m/z$  200；AGC：一级 2e5，二级 5e4；二级 Maximum Injection Time：100 ms；碰撞能量：HCD 30%

DIA：质量范围： $m/z$  400–1200；窗口：25  $m/z$ （窗口间 1  $m/z$  重叠，实际 Isolation window 设 26  $m/z$ ）；二级分辨率：30,000@ $m/z$  200；二级 AGC：1e5；二级 Maximum Injection Time：85 ms；碰撞能量：HCD 30%；每个 DIA 循环之间插入一次一级扫描。

## 数据处理

基于 Proteome Discoverer 2.0 软件进行 DDA 鉴定、位点筛选和谱图库建立，Proteome Discoverer 1.4 和 MaxQuant 也可以实现相同的工作。

搜库鉴定参数：Uniprot 大鼠蛋白数据库，母离子质量偏差：10 ppm；碎片离子质量偏差：0.02 Da；固定修饰：C 烷基化 (+57.021 Da)；动态修饰：M 氧化 (M+15.995 Da)；S/T/Y 磷酸化 (S/T/Y+79.966 Da)；酶：trypsin；Q 值 (Percolator)：< 0.01；ptmRS 模块：PhosphoRS mode: True; Use diagnostic ions: True

位点筛选和谱图库建立：对 PSM 表格 “Isoform Confidence Probability” 一项进行筛选，保留  $\geq 0.75$  的结果，导出 PepXML 格式，并将相应谱图导出为 mzML 格式。PepXML 文件和 mzML 文件同时导入 skyline 建立高可信磷酸化肽段谱图库。

DIA 数据 Skyline 解析：根据 DIA 采集参数设置隔离窗口，将谱图库中所有肽段及相应蛋白作为 Targets，每个肽段选取强度最高的 6 个 b/y 离子进行峰抽提，提峰结果使用 mProphet 进行假阳性评估，控制 Q 值 < 0.01。

## 实验结果

### 1. 精确修饰位点谱图库建立与 DIA 解析流程

由于肽段中含有多个可能发生翻译后修饰的位点，如果不对可能发生修饰的位点进行可信度打分，会造成翻译后修饰位点的错误匹配和定位。将错误匹配和定位的翻译后修饰肽段作为谱图库，就会导致 DIA 解析结果的错误。这是目前翻译后修饰 DIA 分析的瓶颈所在。

基于 Proteome Discoverer 的 phosphoRS/ptmRS 算法能够对可能发生修饰的位点进行打分 (Site Probability)，判断定位的准确性。通常认为 Probability  $\geq 75$  (100 分制) 或 0.75 (1 分制) 的位点定位准确、可靠；具有多个修饰位点的肽段，其所有位点的 Probability 均  $\geq 75$  (或 0.75)，则修饰位点及 isoform 唯一确定。通过这一方法，筛选出位点准确可信、isoform 唯一确定的 PSM 即 Isoform Confidence Probability  $\geq 75$  (或 0.75) 建立谱图库，实现精确的翻译后修饰 DIA 分析。整个流程如图 1 所示。

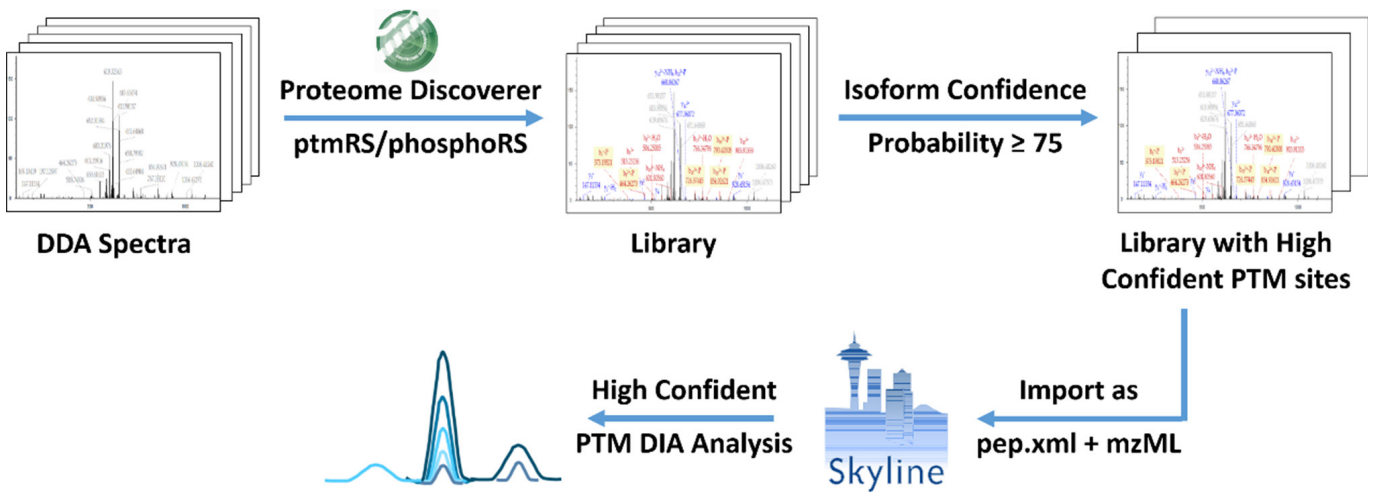


图 1. 精确修饰位点谱图库建立流程图

### 2. 磷酸化样本的 DDA 鉴定、可信度筛选和谱图库建立

磷酸化样本信息和色谱质谱参数见实验条件部分。3 针 DDA 数据按磷酸化检索流程使用 Proteome Discoverer 2.0 软件搜库鉴定 (S/T/Y+79.966 Da)，并使用 ptmRS 模块对位点打分 (图 2-1)。搜库完成后，打开结果，使用 Filter 功能对 PSM 列表 “Isoform Confidence Probability” 项进行筛选，保留得分大于等于 0.75 的 PSM (图 2-2)。然后，右键选择 “Check All-In This Table”，将符合条件的 PSM 选定。最后，在 “Spectra” 中导出 mzML 格式，在 “To PepXML” 中导出 PepXML 格式，并将两个文件置于同一文件夹中，即可作为谱图库导入 Skyline (图 2-3)。

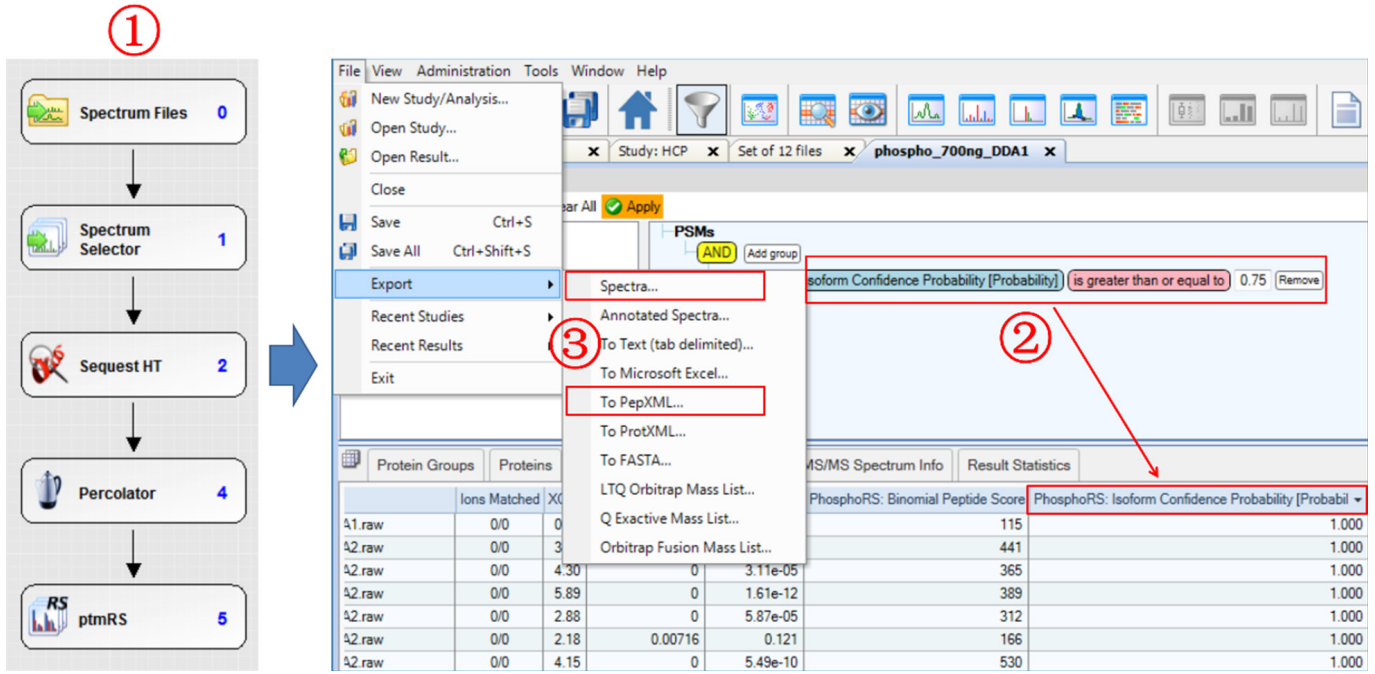


图 2. 翻译后修饰 DDA 数据检索、位点筛选和结果导出步骤

经 phosphoRS/ptmRS 计算得到每个磷酸化位点的打分 (Site Probability)，分数 > 75 (或 0.75) 则位点有确切的碎片支持，定位可靠。搜索引擎没有针对位点打分的功能，位点定位错误概率大。图 3 是一个典型的例子：RFSVTAEGGLTLEQVTDAR 肽段搜库鉴定出包含 1 个磷酸化位点的多个 PSM，3 号位丝氨酸和 5 号位苏氨酸均有匹配到磷酸化发生 (FDR < 1%)。而 ptmRS 的打分结果却分 3 种情况：1) ptmRS 得到唯一可靠的位点，且与搜索引擎得到的位点一致，此时 Isoform 100% 确定 (图 3-A)；2) ptmRS 得到两个位点都有可能，无法确定唯一位点，此时 Isoform 可靠度为 50% (图 3-B)；3) ptmRS 得到唯一可靠的位点，但与搜索引擎得到的位点不一致，此时 Isoform 可靠度为 0 (图 3-C)。最终，只筛选 Isoform 打分 (Isoform Confidence Probability)  $\geq 75$  (或 0.75)，即所有位点均明确、可靠的肽段建立谱图库。

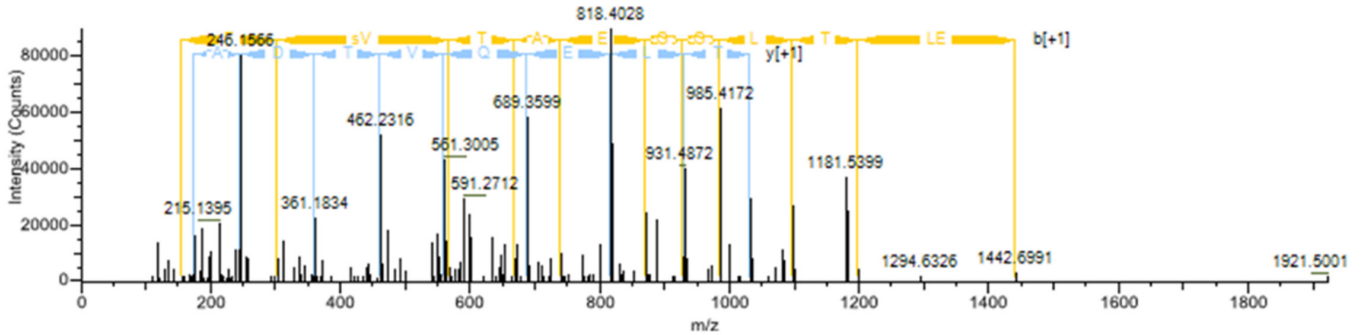
ptmRS结果

100

(A)

搜库结果 R]F]S]V]T]A]E]G]G]L]T]L]E]Q]V]T]D]A]R

Isoform Confident Probability: 100



ptmRS结果

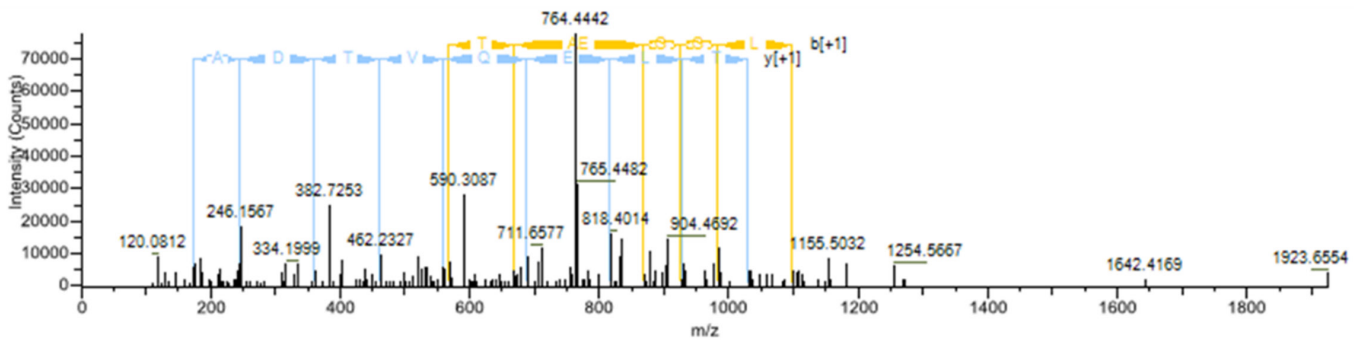
50

50

(B)

搜库结果 R]F]S]V]T]A]E]G]G]L]T]L]E]Q]V]T]D]A]R

Isoform Confident Probability: 50



ptmRS结果

100

(C)

搜库结果 R]F]S]V]t]A]E]G]G]L]T]L]E]Q]V]T]D]A]R

Isoform Confident Probability: 0

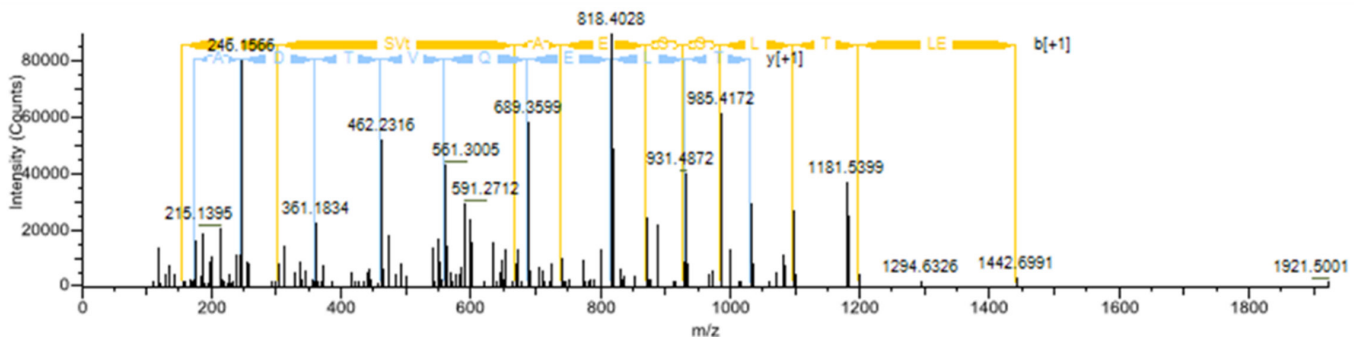


图3. 搜索引擎/ptmRS结果一致、部分一致、不一致三种情况示例

实验共鉴定 56617 张 PSM (FDR < 1%), 经过 Isoform Confident Probability 筛选, 共获得 30558 张精确定位、isoform 唯一的磷酸化肽 PSM (图 4), 结果导出 mzML 和 PepXML 格式作为谱图库。

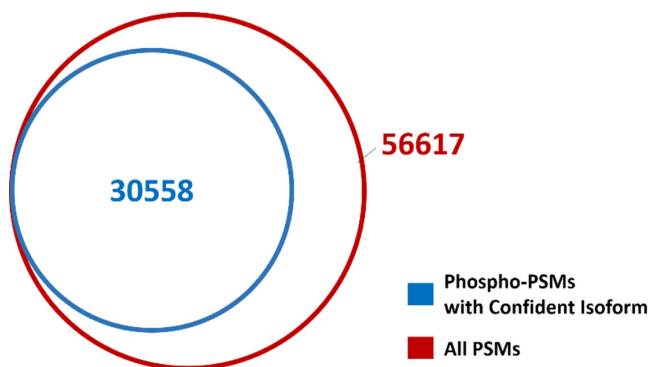


图 4. Isoform Confident Probability 筛选前后 PSM 比较

### 3. 基于精确谱图库的磷酸化样本 DIA 数据解析

将 mzML 和 PepXML 格式文件导入 Skyline 生成磷酸化肽谱图库, 用于磷酸化肽段 DIA 数据解析。Skyline 将谱图库中所有肽段作为 targets 对 DIA 数据中进行提峰, 每条肽段选取 3 个母离子 (单同位素, 单同位素 +1, 单同位素 +2) 和 6 个响应最高的 b/y 离子。提峰结果使用 mProphet 进行假阳性评估, Q 值 < 0.01 (即 FDR < 1%) 为可信的提峰结果。

结果显示, 从 DIA 数据中提取、定量到 6401 条可信的磷酸化肽, 占谱图库磷酸化肽总数 (6505 条) 的 98.4% (图 5)。磷酸化肽的丰度和离子化效率普遍较低, 本实验如此高的解析成功率表明, 基于 Orbitrap 的 DIA 数据具有极高的谱图质量和出色的灵敏度。

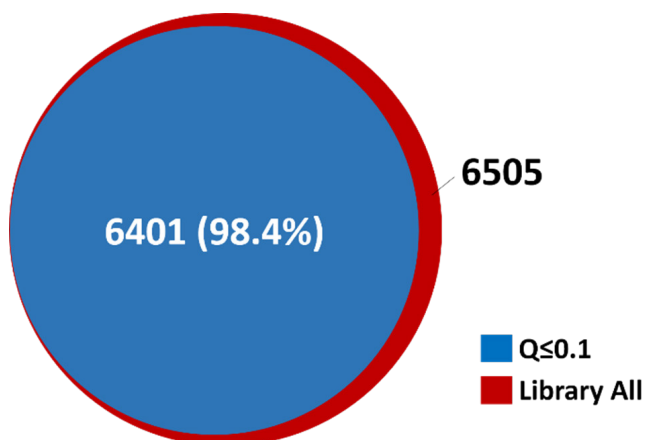


图 5. DIA 可靠解析的磷酸化肽占谱图库总数的 98.4%

进一步对 6401 条磷酸化肽 XIC 色谱峰面积的三针重现性进行统计。将每条肽的母离子和子离子峰面积分别加和, 选取两者中重现性最好 (即 CV 值最低) 的结果用于定量。一般认为, 峰面积 CV 值 < 20% 时定量结果可靠、准确。统计结果显示, 峰面积 CV 值 < 20% 的磷酸化肽占总数的 86.9% (图 6)。如此高的重现性说明 Orbitrap DIA 不仅具有出色的灵敏度, 同时具有优越的重现性, 胜任复杂的翻译后修饰定量。

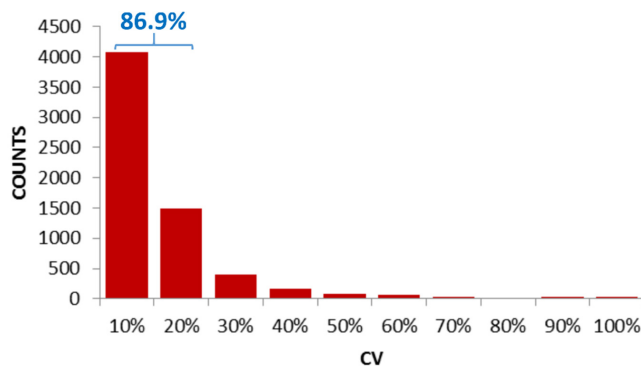


图 6. 磷酸化肽峰面积重现性 (CV 值) 统计

图 7 展示了一个典型的磷酸化肽 DIA 解析结果。该磷酸化肽有 2 种异构体, 分别在 3 号位和 9 号位的丝氨酸上发生了磷酸化。得益于高质量的 DIA 谱图和严格的谱图库建立, 即使这 2 种异构体保留时间非常接近, 也能成功、准确地分辨。DIA 获得的碎片丰度比与谱图库非常接近, 匹配打分的 dotp 值均在 0.95 以上。

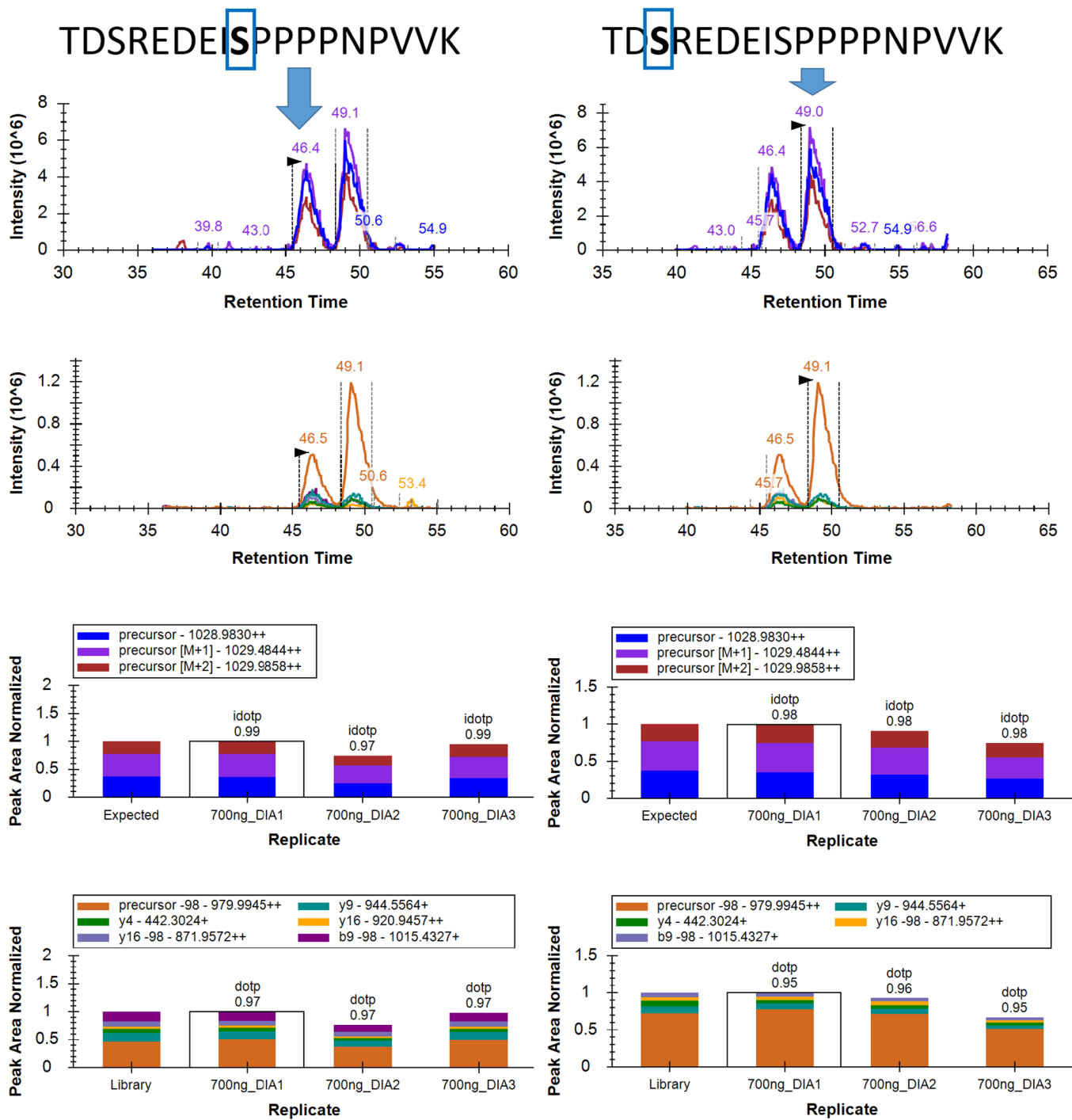


图 7. 磷酸化肽异构体在 DIA 中的分辨和匹配打分 ( dotp )

## 结论

翻译后修饰由于肽段上修饰位点的多重性和不确定性，难以获得可靠的 DIA 结果，这一问题是 DIA 发展的瓶颈所在。另一方面，phosphoRS/ptmRS 和 MaxQuant 等算法和软件的发展，使翻译后修饰位点的 DDA 鉴定、定位更加准确。本文基于上述算法和软件对磷酸化样本的 DDA 鉴定结果进行位点打分，并筛选出位点定位准确、可靠的谱图建立谱图库用于 DIA 分析，成功从磷酸化样本 DIA 数据中提取 6401 条可靠的磷酸化肽 ( $Q < 0.01$ )，占谱图库中磷酸化肽总数的 98.4%。其中，86.9% 的肽段峰面积  $CV < 20\%$ ，可用于精确的磷酸化定量。该策略有效解决了翻译后修饰 DIA 定量的难题，证明基于 Orbitrap 超高分辨质谱技术的 DIA 流程兼具出色的灵敏度和优越的重现性，是复杂样本定量特别是翻译后修饰样本定量的最佳选择。

## 参考文献

- [1] Multiplexed peptide analysis using data-independent acquisition and Skyline, Nat. Protoc., 2015, 10(6): 887-903
- [2] Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues, Mol. Cell. Proteomics, 2015, 14(5): 1400-1410
- [3] MS1 Peptide Ion Intensity Chromatograms in MS2 (SWATH) Data Independent Acquisitions. Improving Post Acquisition Analysis of Proteomic Experiments, Mol. Cell. Proteomics, 2015, 14(9): 2405-2419
- [4] Universal and confident phosphorylation site localization using phosphoRS, J. Proteome Res., 2011, 10(12): 5354-5362
- [5] MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, Nat. Biotechnol., 2008, 26(12): 1367-1372



Orbitrap 组学俱乐部



赛默飞小分子质谱应用技术群

赛默飞世尔科技(中国)有限公司

[www.thermofisher.com](http://www.thermofisher.com)

**全国服务热线: 800 810 5118**  
**400 650 5118 (支持手机用户)**

**Thermo**  
SCIENTIFIC

A Thermo Fisher Scientific Brand

# Q Exactive HF 一小时 DDA 鉴定 DIA 定量 4000 个 HeLa 蛋白

周岳；张伟

赛默飞世尔科技（中国）有限公司

## 关键词

QE HF; DDA; DIA; HeLa; 蛋白质鉴定; 定量蛋白质组学

## 引言

数据非依赖性的扫描模式 (data-independent acquisition, DIA) 是近几年来发展的一种新的质谱数据采集方式<sup>[1]</sup>。它的理念是用二级碎片离子进行蛋白相对/绝对定量。DIA 扫描模式中, 超高分辨质谱对特定质量范围内的所有母离子进行碎裂, 采集所有母离子的碎片离子, 并快速地依次扫描相邻的母离子宽口内的所有碎片离子。DIA 的数据中包含了所有碎片离子的保留时间和强度信息。用非常小的质量偏差宽口 (如 10 ppm) 目标性地抽提同一肽段的多个子离子, 计算子离子的强度, 就能对该肽段进行鉴定和定量。DIA 定量相比传统的基于母离子强度的 DDA 定量有选择性好, 定量准确等优点<sup>[1]</sup>, 所以 DIA 成为定量蛋白质组学新的发展方向。

Q Exactive HF 是赛默飞世尔科技在 2014 年的 ASMS 上推出的全新静电场轨道阱超高分辨质谱仪 (图 1)<sup>[2,3]</sup>。Q Exactive HF 采用了分段式四极杆技术 (Advanced Quadrupole Technology, AQT) 使离子传输效率至少提高了 2 倍; 超高场 Orbitrap 技术, 提高了 Orbitrap 扫描速度, 在 15000 分辨率时, 二级谱图的扫描速度是 20 Hz。这两项技术提高了 QE HF 进行 DDA、DIA 数据的采集能力。本文用 1 小时快速色谱梯度对 QE HF 的 DDA 鉴定能力和 DIA 定量能力进行考察, 同时从定量肽段数目和 CV 两方面对 DDA 定量和 DIA 定量能力进行比较。

## 实验条件

### 实验材料和方法

Pierce HeLa Protein Digest Standard (货号: 88329), 稀释至 500 ng/ $\mu$ l, EASY-nLC 进样 1 $\mu$ l, 500 ng 进行 DDA、DIA 数据采集, 每种采集模式重复 3 遍。

### 高效液相色谱分离

高效液相色谱仪: EASY-nLC 1000 (Thermo Scientific™)

分析柱: 实验室自制 C18, 15 cm, ID 75  $\mu$ m, 3  $\mu$ m

流动相: A: 0.1% 甲酸水溶液; B: 0.1% 甲酸乙腈溶液

梯度: 60 min, 3/0–6/2–22/48–40/53–80/55–80/60 (%B/min)

流速: 300 nL/min

### 质谱分析

DDA 数据采集:

质谱仪: Q Exactive HF (Thermo Scientific™);

离子源: NanoFlex 离子源; 离子模式: 正离子;

喷雾电压: 1.8 kV; 毛细管温度: 275° C; S-Lens RF: 55%;

分辨率: 一级 120000@m/z 200, 二级 15000@m/z 200; 一级

AGC: 3e6, Maximum IT: 50ms;

碰撞能量: NCE 27%; Fixed first mass: 110 m/z

DIA数据采集:

质谱仪: Q Exactive HF (Thermo Scientific™);

离子源: NanoFlex 离子源; 离子模式: 正离子;

喷雾电压: 1.8 kV; 毛细管温度: 275°C; S-Lens RF: 55%;

目标  $m/z$  窗口: 400–1000; isolation window: 12Da;

碰撞能量: 27%; fixed first mass: 200  $m/z$ ; AGC target: 1e6;

Maximum ion injection time: atuo; loop count: 50

## 数据处理

Proteome Discoverer 蛋白鉴定流程: 人蛋白数据库 (uniprot human\_201309), 母离子质量偏差: 10 ppm; 碎片离子质量

偏差: 0.02 Da; 固定修饰: 半胱氨酸烷基化 (+57.021 Da); 动态修饰: 甲硫氨酸氧化 (+15.995 Da); 天冬酰胺和谷氨酰胺脱氨基化 (+0.984 Da); 酶: trypsin; 漏切位点: 2; FDR < 0.01

Skyline DDA、DIA 蛋白定量流程: DDA 定量用 skyline MS1 filtering 功能, 对每个肽段强度最高的 3 个母离子同位素峰进行抽提, idotp  $\geq$  0.8; DIA 定量用 skyline DIA 功能, 设置隔离窗口 12 Da, 对给个肽段强度最高的 5 个子离子进行峰抽提, mProphet 打分, 控制 FDR < 0.01。

## 实验结果

### 1. DIA 数据采集以及分析流程

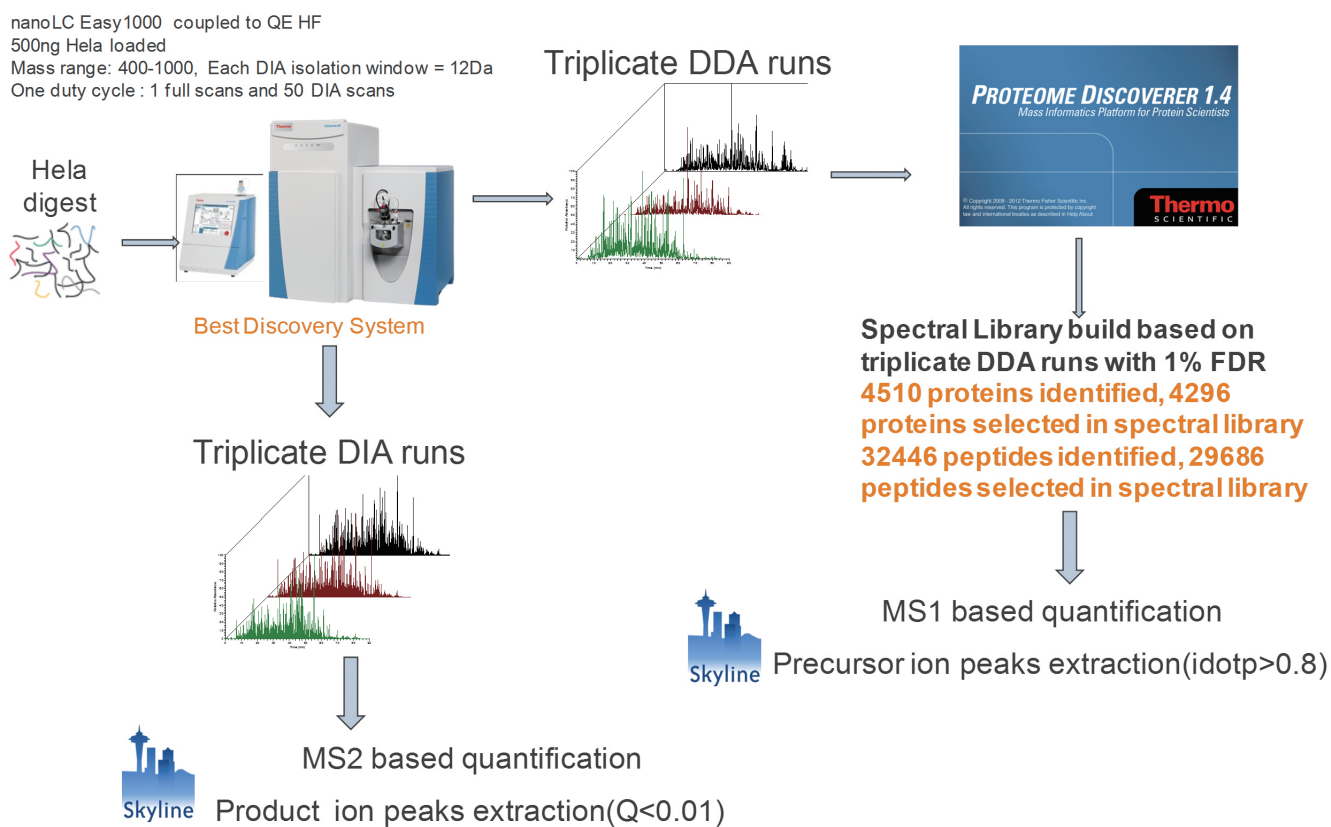


图 1. DDA 定量和 DIA 定量实验流程

500 ng Hela 细胞裂解液进行 3 次 1 h 梯度 DDA 分析, Proteome Discoverer 1.4 数据库检索。将鉴定到的蛋白和肽段信息导入 skyline 作为候选定量蛋白和肽段。基于母离子的定量用 skyline 中 MS1 filtering 功能对 DDA 数据进行母离子强度抽提。500 ng Hela 细胞裂解液进行 3 次目标  $m/z$  窗口 400–1000, 隔离窗口 12 Da 的 DIA 分析。基于二级子离子的定量用 skyline DIA 功能对子离子进行峰抽提, mProphet 打分, 筛选 Q value < 0.01 的肽段为定量肽段。

在分析 DIA 数据之前，需要建立谱图库，谱图库中包含所有蛋白在质谱中鉴定到的肽段，以及肽段的保留时间、碎片离子质荷比、碎片离子强度等信息。数据依赖性扫描是最好的建立谱图库的数据采集方式。500 ng Hela 细胞裂解液进行三次 DDA 数据采集。原始数据经 Proteome Discoverer 检索并控制 FDR < 1%。将三次 DDA 鉴定结果合并，导入 skyline 建立谱图库。DDA 数据除了能给出蛋白和肽段的鉴定信息，还能基于母离子的强度/峰面积进行定量。Skyline 中 MS1 filtering 功能对母离子的多个同位素峰进行抽提，并且根据同位素分布进行打分 (idotp 值)<sup>[4]</sup>。Idotp 代表测定的同位素分布和理论同位素分布的相似度。筛选 idotp 值大于 0.8 的母离子作为可信的定量肽段。

500 ng Hela 细胞裂解液用相同的色谱柱，相同的色谱梯度，将 QE HF 切换至 DIA 扫描模式进行 3 次 DIA 数据采集。在一次扫描循环中，目标母离子质核比范围 400-1000，四极杆的隔离窗口为 12 Da，包含 50 次 MS/MS 扫描。每一张 MS/MS 谱图中包含了 12 Da 窗口内的所有母离子的碎片离子信息。QE HF 使用了超高场的 Orbitrap，扫描速度是 20 Hz，所以每次循环所用的时间约为 2.4-3 s，与色谱兼容 (图 2)。Skyline 处理 DIA 数据时从谱图库中选取强度最高的多个碎片离子进行色谱峰抽提。Skyline 中嵌入的 mProphet 软件会根据同一肽段的多个子离子峰的 feature 进行打分。Feature 包括子离子共流出峰形、保留时间偏差、dotp 值、信噪比等。为了区分假阳性的子离子峰，mProphet 可以建立 decoy 库，也可以将打分排名第二的肽段作为 decoy，计算 FDR<sup>[5,6]</sup>。筛选 FDR < 0.01 的肽段段作为可信的定量肽段 (图 1)。

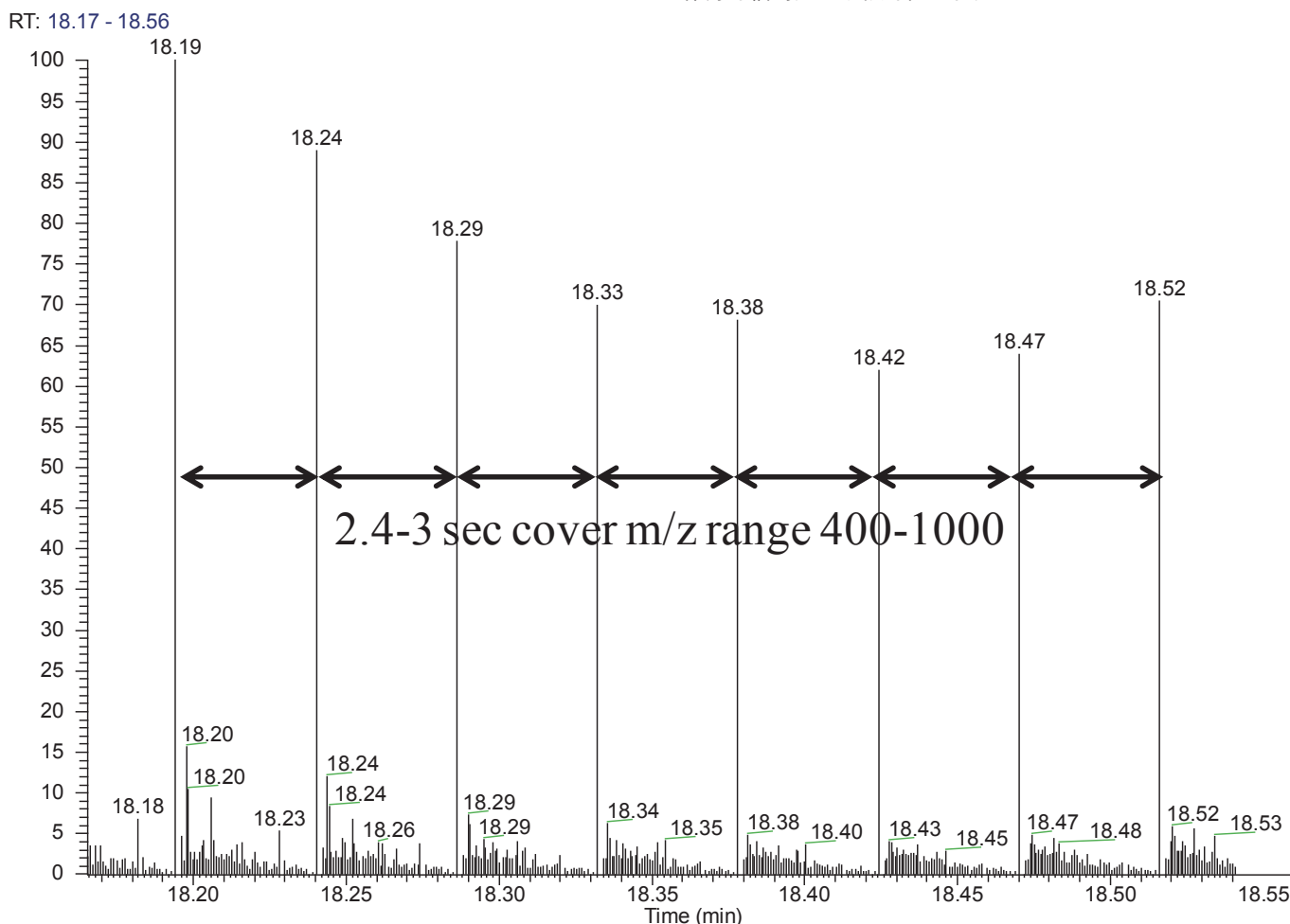


图 2. QE HF 扫描目标质核比窗口 400-1000 采集 50 张 MS/MS 所需时间在 2.4-3 s

## 2. DDA 鉴定结果以及谱图库建立

500 ng Hela 细胞进行 60 min 色谱梯度进行 DDA 数据采集。在 60 min 内，三次重复，QE HF 分别采集到 49065、49031、48885 张谱图，鉴定到 22030、21820、21654 个肽段，对应到约 3909、3900、3878 个蛋白。将三次的肽段和蛋白合并，一共鉴定到 32446 个肽段，4510 个蛋白（图 3）。蛋白和肽段鉴定结果导入 skyline 建立谱图库。通过该谱图库，在 skyline 中建立需定量的候选肽段和蛋白。为了提高蛋白定量的准确性，在 skyline 中设置一些肽段的限制条件， $m/z$  400-1000，母离子电荷  $2^+$ - $4^+$ ，无漏切位点。32446 个肽段中符合这些条件的肽段有 29686 个，对应 4296 个蛋白。在 DDA 定量实验中，为每个候选肽段添加强度最高的 3 个同位素峰 ( $M$ 、 $M+1$ 、 $M+2$ ) 作为定量离子，共生成 89027 个定量离子。在 DIA 定量实验中，为每个肽段添加 5 个强度最高的子离子 ( $b_3$ - $b_n$ ， $y_3$ - $y_n$ )，共生成 145736 个定量离子。

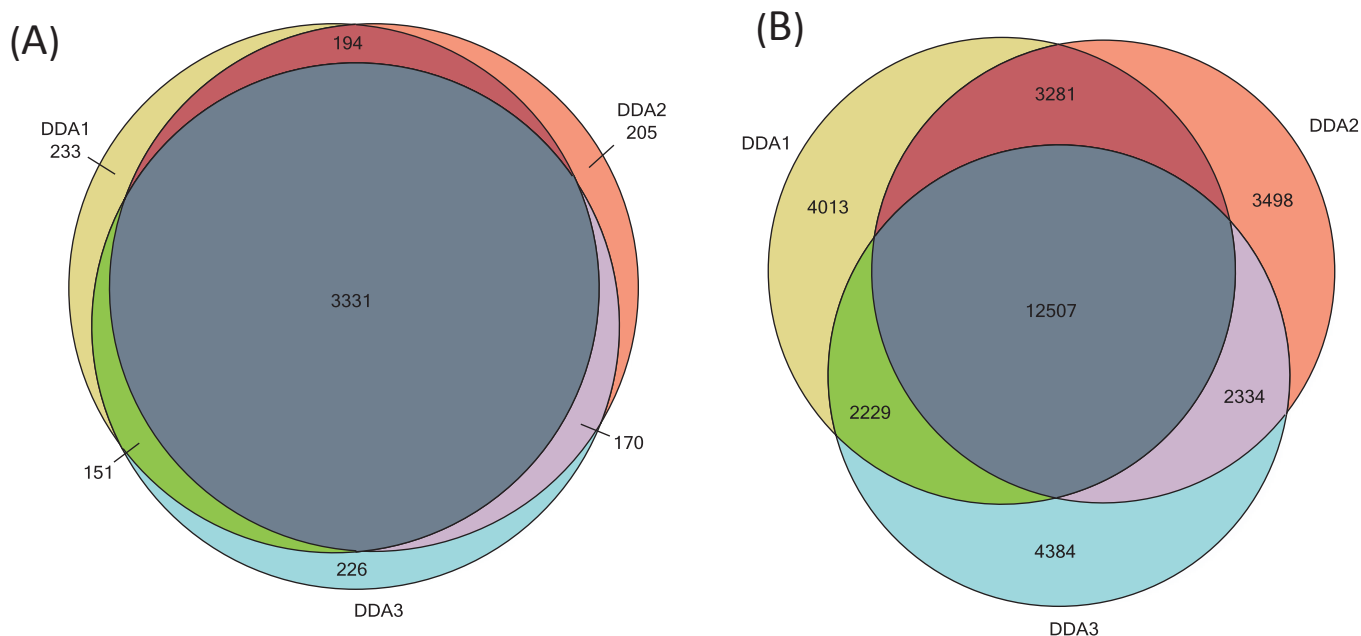


图 3. 500 ng Hela 细胞裂解液 3 次 DDA 鉴定结果 (A) 鉴定到的蛋白交盖图 (B) 鉴定到的肽段的交盖图

## 3. 基于母离子的 DDA 定量和基于子离子的 DIA 定量结果比较

对于 DDA 定量，skyline 会从 DDA 数据一级谱图中进行母离子同位素峰抽提。筛选同位素分布与理论同位素分布较好的肽段作为可靠的定量肽段，即  $\text{idotp} > 0.8$ 。由于 DDA、DIA 实验中用的是相同的色谱梯度，DDA 实验中肽段的保留时间信息可以传递给 DIA 实验。所以在 DIA 定量时，限制 skyline 只抽提该肽段谱库中保留时间 5 min 范围内的子离子。通过保留时间限制可以降低 DIA 数据处理的复杂度和提高定量的准确度。同时筛选  $Q \text{ value} < 0.01$  的肽段最为可靠的定量肽段。

三次 DDA 实验从一级谱图中分别定量到 25214、25515、24873 个肽段；3969、3975、3917 个蛋白；定量到的肽段占总候选肽段的 84%；峰面积的 CV 值在 20% 以下的占总定量肽段的 58.81%；CV 值在 10% 以下的占 25.25%（图 4）。三次 DIA 定量分别定量到 27558、27604、27483 个肽段；4067、4080、

4073 个蛋白。定量到的肽段占总候选肽段的 93%；峰面积的 CV 值在 20% 以下的占总定量肽段的 90.3%；CV 值在 10% 以下的占 69.01%（图 5）。可以看出 DIA 能够定量到更多的肽段，而且定量肽段的峰面积 CV 值远远小于 DDA。

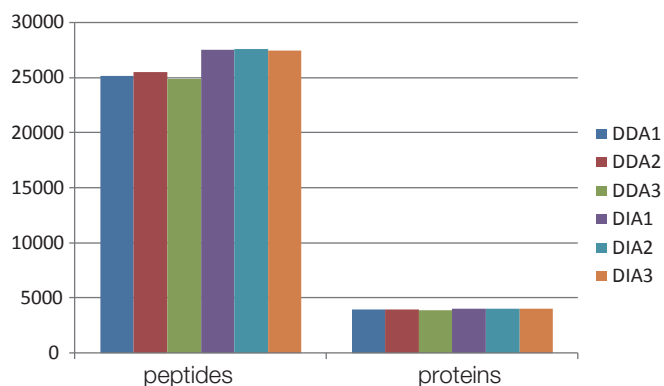


图 4. 3 次 DDA 实验和 3 次 DIA 实验定量到的肽段和蛋白

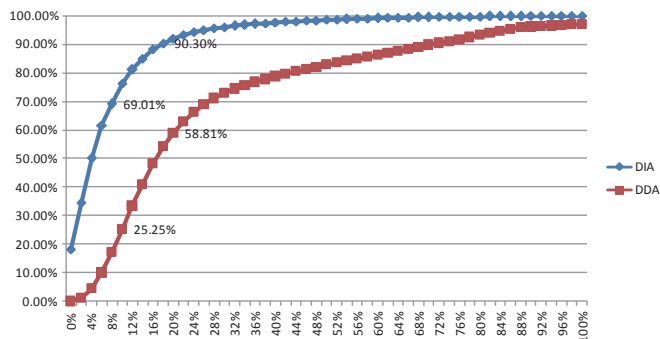


图 5. DDA 定量和 DIA 定量母离子和子离子峰面积 CV

## 结论

DDA 一般用于蛋白/肽段鉴定，同时可以基于母离子进行的定量，在非常复杂的基质中，非常容易受到其他肽段的干扰，导致定量不准确，CV 值过大。DIA 定量是基于二级子离子的定量，用子离子定量会提高定量的选择性，降低其他肽段的干扰，定量会比基于一级峰面积的定量更加准确，CV 值较小。该实验中，用 500 ng Hela 细胞裂解液评价 DDA 定量和 DIA 定量能力。3 次 DDA 实验鉴定到 32446 个肽段，4510 个蛋白。将鉴定到的蛋白和肽段导入 skyline 建立谱图库，筛选出 4296 个候选定量蛋白，29686 个候选定量肽段，作为 DDA 和 DIA 共同定量目标。DDA 定量到了 84% 的肽段，92% 的蛋白；而 DIA 定量到了 93% 的肽段，95% 的蛋白。DIA 比 DDA 能定量到更多的肽段和蛋白，同时 DIA 定量子离子峰面积的 CV 远远小于 DDA 定量母离子峰面积的 CV。结果表明，基于二级子离子的定量要优于基于一级母离子的定量。

QE HF 在一小时梯度内就能从 500 ng Hela 细胞裂解液鉴定并定量 4000 个蛋白，这得益于 QE HF 较好的离子传输效率，超快的扫描速度。目前 DIA 定量基于一维反向色谱质谱联用分析还不适用于二维色谱质谱联用分析。但是可以通过延长一

维反向色谱的梯度，提高分离效率达到细胞蛋白质组的全覆盖。Matthias Mann 研究小组 2014 年用一维反向色谱在小鼠 NSC-34 和 N2a 细胞裂解液 240 min 色谱图梯度，就鉴定到了 8000 多蛋白，基本上达到细胞系蛋白组全覆盖[7]。通过 DDA 鉴定肽段蛋白，建立谱图库，DIA 进行蛋白定量可能成为定量蛋白质组学新的发展方向。

## 参考文献

- Gillet, L C. et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics*, 2012, 11(6): O111 016717.
- Kelstrup, C D. et al. Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field orbitrap mass spectrometer. *J Proteome Res*, 2014, 13(12):6187-6195.
- Scheltens R A. et al. The Q Exactive HF, a benchtop mass spectrometer with a pre-filter, high performance quadrupole and an ultra-high field orbitrap analyzer. *Mol Cell Proteomics*, 2014, 13(12):3698-3708
- Abbatiello S E. et al. Design, implementation, and multi-site evaluation of a system suitability protocol for the quantitative assessment of instrument performance in LC-MRM-MS. *Mol Cell Proteomics*, 2013, 12(9):2623-2639
- Reiter, L. et al. mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods*. 2011, 8 (5):430-435
- Röst H.L. et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS-data. *Nat. Biotechnol.* 2014, 32(3):219-223
- Horburg D. et al. Deep Proteomic Evaluation of Primary and Cell Line Motoneuron Disease Models Delineates Major Differences in Neuronal Characteristics. *Mol Cell Proteomics*, 2014, 13(12):3410-3420



Orbitrap 组学俱乐部



赛默飞小分子质谱应用技术群

赛默飞世尔科技（中国）有限公司

[www.thermoscientific.cn](http://www.thermoscientific.cn)

**全国服务热线：800 810 5118**  
**400 650 5118 (支持手机用户)**

AN\_C\_LCMSMS\_17\_201511Y

**Thermo**  
SCIENTIFIC  
A Thermo Fisher Scientific Brand