

基于SWATH®的定量蛋白质组学在生物研究中的应用

Application of SWATH® Based Quantitative Proteomics in Biological Study

陈晨, 罗继, 郭立海, 靳文海

Chen Chen, Luo Ji, Guo Lihai, Jin Wenhai

SCIEX China

Key Words: SWATH®, TripleTOF®, Eksigent®, ProteinPilot™, SWATH® Acquisition MicroApp 2.0, DIA

引言

蛋白质组学是研究特定时间和空间内,所有的蛋白质及其相关变化的学科,而定量蛋白质组学是蛋白质组学研究中非常重要的领域。目前定量蛋白质组学研究中使用的最多,最广泛的方法是使用高分辨质谱仪平台进行研究。目前为止发表文章最多的定量蛋白质组学研究策略,使用的是稳定同位素标记技术(iTRAQ®试剂,或同类试剂)。这种技术使用不同的稳定同位素标签,对不同的样品进行标记。之后将样品进行等量混合,通过质谱仪所识别到的来自于不同标签分子的报告离子的相对信号强度,来实现肽段的相对定量。这种定量策略的准确性非常好,后期数据处理也有很多非常成熟的软件进行支持,所以一直是科研领域内很受欢迎的策略。但是由于其自身的同位素标记的特性,也存在着一一定的限制,比如所有的稳定同位素标记试剂都有数量限制,无法直接对大量样品进行相对定量,同时由于同位素标记试剂自身的价格,也令实验成本显著增加。与此同时,蛋白质组学样品中,肽段的种类通常情况下为几十万种,甚至上百万种。稳定同位素标记策略一般使用数据依赖型采集模式(IDA)进行数据的采集,而数据依赖型采集模式自身的特点——每采一次一级谱图,从其中挑选若干个母离子进行二级谱图的采集——注定了这种采集方法无法采集到样品中所有的肽段信息,这是因为采集的速度远远不及肽段的种类。

为此SCIEX创新的研发了SWATH®数据采集模式,这种数据采集模式是一种数据非依赖型采集模式。与传统的IDA采集方法不同,SWATH®不需要从一级谱图中挑选目标母离子,而是将母离子的总质量范围划分为若干较小的窗口,每一次扫描,通过四极杆的选择功能,让其中一个窗口内的母离子通过四极杆,打碎之后采集这个窗口内所有母离子的碎片信息。这样在每个扫描循环中,就

能够将整个母离子的总扫描范围内所有母离子的碎片信息全部采集完成。基于这种方法,我们能够从蛋白样品中分析出更多的生物学信息和生物学意义,助力生物学研究的发展。

硬件与软件搭配特色

1. TripleTOF® 5600+质谱系统拥有最快的扫描速度,能够非常高效的采集复杂样品中(如蛋白质组学样品)的不同肽段的二级谱图并依靠软件进行鉴定。
2. ProteinPilot™是当前蛋白质组领域内最优秀的数据分析软件之一,其特有的算法能够从数据中鉴定到更多的肽段二级谱图,从而提高蛋白鉴定的效率。在本实验中使用ProteinPilot™软件进行IDA数据分析。
3. SWATH® Acquisition MicroApp 2.0插件是一款对SWATH®数据进行分析的软件,能够同时对大量SWATH®数据进行蛋白质组学相对定量分析。在本实验中使用SWATH® Acquisition MicroApp 2.0软件对SWATH®数据进行分析。

实验思路

为了对比稳定同位素标记策略的定量蛋白质组学与基于SWATH®策略的定量蛋白质组学的各自优势。我们使用相同的样品,相同的液质联用体系对使用两种不同策略的定量结果进行数据的对比,来比较两种策略各自的优势。

在本实验中,我们使用小鼠的细胞系作为分析样品,分别取某种疾病前与疾病后的细胞系作为对照组与实验组。对于稳定同位素标记实验,在标记后使用离线分级的策略,将样品分级为12个组分,分别进行IDA分析,最后使用ProteinPilot™软件进行整理分析,同时为SWATH®分析建立必须的数据库Library。每组样品进行SWATH®采集后直接使用SWATH® Acquisition MicroApp 2.0进行分析。

样品前处理

细胞样品使用RIPA裂解进行裂解，使用超声破碎仪打碎其中的DNA后，离心取上清。使用BCA定量试剂盒对上清溶液进行定量。取250 μg蛋白溶液，50 mM碳酸氢铵稀释至150 μl，用10K超滤管14000g离心后向超滤管中加入100 μl 7M 盐酸胍和4 μl 1M 二硫苏糖醇(DTT)，55°C孵育1小时。之后加入10 μl 1M 碘乙酰胺，避光孵育30分钟。14000 g离心15分钟，加入100 μl 50 mM碳酸氢铵后继续14000 g离心15分钟，加入100 μl 50 mM 碳酸氢铵和5 μg胰蛋白酶，在37 °C酶解过夜。第二天收集酶解好的肽段样品。

对于稳定同位素标记实验，将酶解好的肽段样品进行iTRAQ®标记实验，之后将所有样品等量混合。采用MudPIT分级策略，将样品分成12个组分，采用IDA数据采集模式对每个组分进行数据采集，最后通过ProteinPilot™软件进行数据分析。

对于SWATH®实验，将第二天收集到的酶解好的肽段样品直接取5 μg，使用质谱仪进行SWATH®数据采集。之后使用SWATH® Acquisition MicroApp 2.0软件对SWATH®数据进行分析。

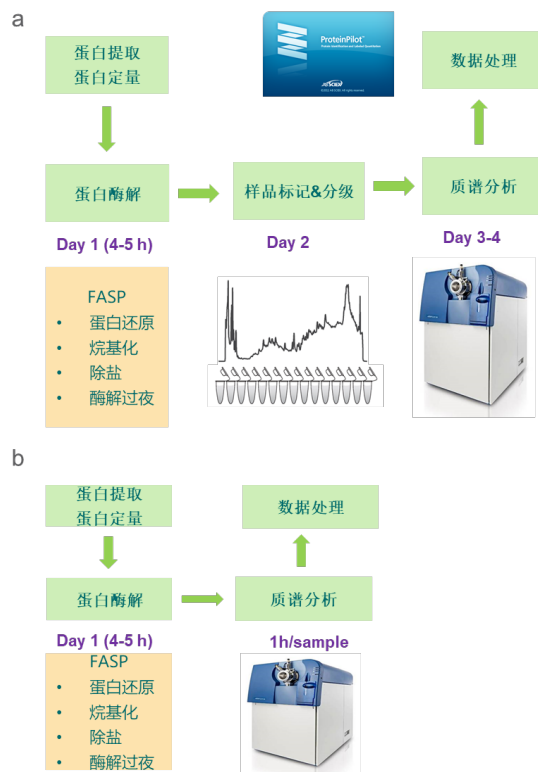


图1. a.稳定同位素标记定量蛋白质组学的样品前处理流程示意图； b.基于SWATH®的定量蛋白质组学样品前处理流程示意图。

液相条件

液相系统: Eksigent® NanoLC™ 425 HPLC system with tray cooling (Micro体系)

色谱柱: Micro column 300 μm×15 cm ChromXP™ C18-CL 3 μm 120 Å, Eksigent®

预柱: Micro Trap column 300 μm×6 mm ChromXP™ C18-CL 3 μm 120 Å, Eksigent®

流动相: A: 95% 水 + 2% 乙腈 + 0.1% 甲酸
B: 95% 乙腈 + 2% 水 + 0.1% 甲酸

流速: 5 μL/min;

柱温: 35 °C;

进样体积: 5 μL

梯度:

Time	%B
0	5
1	8
65	28
77	50
78	80
83	80
83.5	5
90	5

IDA数据采集

DuoSpray™离子源参数:

喷针: 50 mm内径 Micro喷针
IS电压: 5500 V
雾化气Gas1: 18 psi

气帘气CUR: 30 psi
源温度: 350 °C
辅助气Gas2: 20 psi

TOF-MS:

扫描范围: 350-1000
扫描时间: 0.25 s

TOF-MSMS:

扫描范围: 100-1500
扫描时间: 0.05 s

High Sensitivity

IDA设置:

Rolling CE: On Dynamic Exclusion: On
After 1 occurrence for 14 s

SWATH®数据采集

DuoSpray™离子源参数:

喷针: 50 mm内径 Micro喷针 气帘气CUR: 30 psi
IS电压: 5500 V 源温度: 350 °C
雾化气Gas1: 18 psi 辅助气Gas2: 20 psi

TOF-MS:

扫描范围: 350-1000 扫描时间: 0.25 s

TOF-MSMS:

扫描范围: 100-1500 扫描时间: 0.035 s

High Sensitivity

SWATH®的具体窗口划分使用SWATH® Variable Window Calculator_V1.1进行计算, 并导入Analyst®生成可变窗口的SWATH®采集方法。

IDA蛋白定量结果

稳定同位素标记方法标记并等量混匀后的样品, 采用IDA的数据采集模式, 并使用ProteinPilot™软件进行分析, 在不分级的情况下, 一针进样定量到了2587个蛋白质, 24942条肽段。在采用MudPIT策略, 分级的情况下, 每个组分的样品单独采集IDA数据, 最后整合分析, 定量到蛋白7120个。而与本次研究疾病模型相关的核心蛋白SOX2与OCT4, 我们从单针IDA的结果中鉴定到了OCT4, 但是SOX2则并未鉴定到。

SWATH®蛋白定量结果

样品在酶解成为肽段后, 每个样品使用SWATH®数据采集模式进行数据采集。所有的SWATH®数据使用SWATH® Acquisition MicroApp 2.0软件进行统一分析。使用预先建立好的Library (本次实验使用了两个SWATH®数据库, 一个数据库包含6400个小鼠蛋白, 另一个包含8600个小数蛋白), 在所有的SWATH®数据中定量到了6202个蛋白质, 其中包括了IDA方法中鉴定到的OCT4, 同时也包括IDA方法中未鉴定到的SOX2。相比于稳定同位素标记方法的结果,

SWATH®新鉴定到3739个蛋白, 仅有123个蛋白是稳定同位素标记方法结果中独有的。充分说明了SWATH®在蛋白鉴定层面的能力。

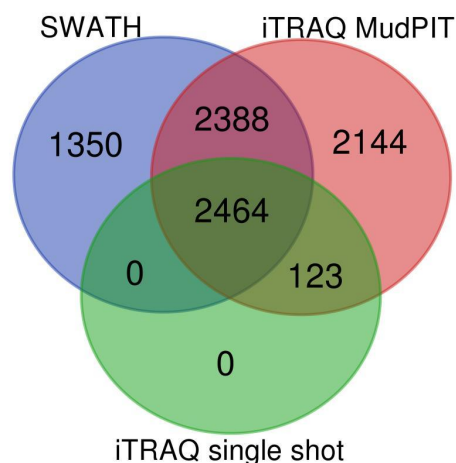


图2. 基于SWATH®方法的定量蛋白质组学在同样为一针进样的前提下, 能够定量到更多的蛋白。

SWATH®鉴定结果分析

SWATH®数据对于肽段的鉴定是建立在已有数据库的基础之上的, 该数据库为已经完成的蛋白质的IDA鉴定结果。所以这个数据库是肽段实际的二级谱图, 以此为库进行的匹配, 其准确度和假阳性率相对于传统蛋白质组分析使用的理论序列库来讲都会有更好的结果。

样品前处理比对

SWATH®的样品前处理仅需用FASP方法进行胰蛋白酶酶解后即可进样采集, 而稳定同位素标记方法的样品在得到肽段样品后需要额外增加一步同位素标记实验, 相对更繁琐。

蛋白数目比对

使用相同样品, 同样在一针90分钟有效梯度下进行的质谱数据采集。SWATH®结果能够定量到6202个蛋白, 而稳定同位素标记方法的数据中仅能够定量到2587个蛋白。其中SWATH®新鉴定到3739个蛋白, 仅有123个蛋白是稳定同位素标记方法的结果中独有的。而使用离线分级的方法, 将稳定同位素标记的样品分为12个组分, 分别采样再整合分析的情况下, 能够定量到7120个蛋白, 蛋白个数与SWATH®单针定量数目并无显著差异。说明SWATH®在

蛋白鉴定和定量能力上，比相同条件下的稳定同位素标记方法更加有效。

重现性评估

为了评估SWATH®定量结果的重现性，我们对每组样品进行了三次重复采集。经过对结果的分析，发现SWATH®定量结果中，CV<20%的蛋白多于80%，体现了SWATH®定量良好的稳定性和重现性。

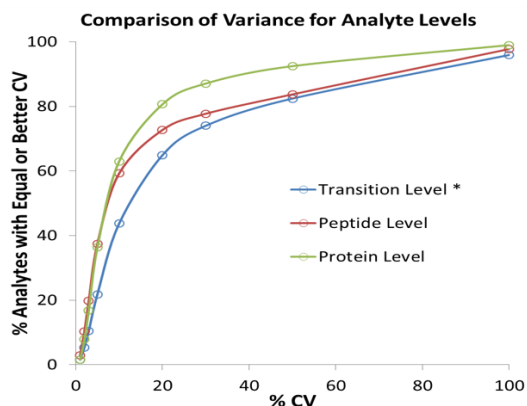


图3. SWATH®采集三次重复下，蛋白、肽段和碎片离子的重现性情况。

生物学分析对比

本次生物学模型使用的是小鼠细胞系的疾病模型，在该模型中，关键的生物过程包括细胞周期，细胞分化等。我们将两种蛋白质组学策略得到的差异蛋白分别使用DAVID进行GO分析。在同位素标记定量的结果中发现了14个与细胞周期相关的蛋白质，而在SWATH®结果中发现了65个与细胞周期相关的蛋白质。同时在SWATH®结果中发现了14个与细胞分化相关的蛋白质，而这类蛋白在同位素标记定量策略的蛋白质组学数据中只有两个。这个结果说明通过SWATH®数据采集得到的数据除了一针能够得到更多的蛋白质之外，也能够非常准确的标示出生物模型内蛋白质的真实变化。为我们进行生物学研究提供更加全面的生物学信息。

SCIEX临床诊断产品线仅用于体外诊断。仅凭处方销售。这些产品并非在所有国家地区都提供销售。

获取有关具体可用信息，请联系当地销售代表或查阅 <https://sciex.com.cn/diagnostics>。所有其他产品仅用于研究。不用于临床诊断。

本文提及的商标和/或注册商标的所有权，归属于AB Sciex Pte. Ltd. 或在美国和/或某些其他国家地区的各权利所有人。AB SCIEX™ 商标经许可使用。

© 2019 DH Tech. Dev. Pte. Ltd. RUO-MKT-02-10624-ZH-A



SCIEX中国公司

北京分公司
地址：北京市朝阳区酒仙桥中路24号院
1号楼5层
电话：010-5808 1388
传真：010-5808 1390

上海公司及中国区应用支持中心
地址：上海市长宁区福泉北路518号
1座502室
电话：021-2419 7200
传真：021-2419 7333

广州分公司
地址：广州市天河区珠江西路15号
珠江城1907室
电话：020-8510 0200
传真：020-3876 0835

全国免费垂询电话：800 820 3488, 400 821 3897 网址：sciex.com.cn 官方微信：ABSciex-China

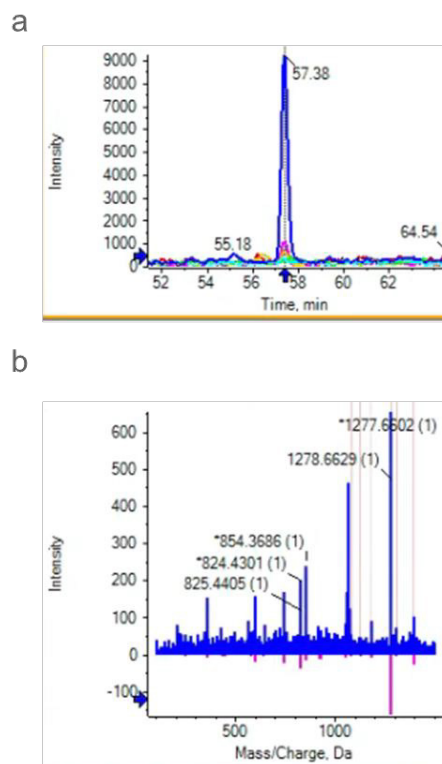


图4. SWATH®定量蛋白质组数据中SOX2的肽段的鉴定结果。图a. 蛋白质SOX2的肽段的碎片离子XIC图。图b. 蛋白质SOX2的肽段的SWATH®谱图与Library中谱图的比对结果。

结论

本实验通过相同的样品，相同的液质体系，相同的液相梯度，直接对比了SWATH®与稳定同位素标记方法进行蛋白质组学研究的差异。通过比较，发现SWATH®定量蛋白质组学在蛋白质的数量，通量和适用性上优于基于稳定同位素标记的定量蛋白质组学方案。