

问题 1: 蛋白质谱鉴定不成功一般有哪些因素?

答: 质谱鉴定不成功主要可以分为两类, 一类是质谱效果不好, 一类是没有可参考的蛋白数据库。质谱效果不好的原因又可以细分为蛋白点太淡以至于蛋白含量过低(常见银染的点)、蛋白酶解及质谱操作失误等。要避免这些因素需要尽量取颜色深一些的蛋白点, 并且取的蛋白点面积需要适当大一些, 对于某些很小但很清晰的蛋白点, 取点的时候可适当选择孔径大一些的枪头, 把边上空白处适当取到一些也不要紧, 避免取点混入其他的蛋白点。其次, 由于没有合适的数据库导致的鉴定效果较差, 可以通过更换范围更大的数据库、近缘物种数据库、采用相关研究物种的蛋白、EST 数据库等构建本地数据库的方式进行解决。

问题 2: 凝胶内的蛋白样品是否可以长期保存? 如何运输? 对质谱鉴定有何影响?

答: 如果保存时间超过一周, 可以将蛋白点切取后放入 $-20^{\circ}\text{C}$ 或者 $-80^{\circ}\text{C}$ 冰箱保存; 如果小于一周, 则可以常温保持, 基本不会影响后续质谱鉴定效果, 运送时常温快递即可。

问题 3: iTRAQ/TMT 技术相对于电泳技术有什么优势?

答: 2-DGE 是伴随着 MALDI-TOF 技术发展起来的蛋白质分离技术, 理论上可以通过等电点(IEF)和分子量(SDS-PAGE)的差异将总蛋白质逐一分开, 实际情况并不理想, 因为低丰度蛋白无法染色成功而看不见, 蛋白翻译后修饰后偏离理论位置等等, 一张 2-DGE 能鉴定到的蛋白质总数最多 1000~2000 蛋白。相较于 2-DE 以及 DIGE 等基于电泳的技术, iTRAQ/TMT 分辨率高, 博苑生物进行的细胞样品最多有发现超过 6000 个蛋白, 并且绝大多数蛋白都有定量和定性信息; 其次 iTRAQ 通量高, 可以一次最多完成 8 个样品的实验, 如果使用 TMT 技术最多可以一次完成 10 个样品的实验, 特别适合于多组样品间的同时比较以及生物学过程的动态检测。最后, 鹿明生物提供的 iTRAQ 完整实验服务中包括 iTRAQ/TMT 实验中鉴定到的所有蛋白的定性和定量结果、差异筛选结果、并且免费提供差异蛋白的维恩图分析、聚类分析、热图分析、GO、Pathway、蛋白互作等生物信息学分析内容以及中英文双语的分析报告, 除此之外还可以根据客户的需求提供个性化的服务, 免去了数据分析的烦恼。

问题 4: iTRAQ/TMT 实验的主要步骤有哪些, 主要采用哪些仪器? 能得到什么数据及什么生物信息分析结果?

答: 实验步骤包括: 蛋白提取、浓度测定及 SDS-PAGE 凝胶检测, 质检报告; 蛋白酶解、同位素标记; 一维色谱分离, 冻干机冻干; 高分辨 LCMSMS 质谱检测; 搜库定性、定量分析; 差异蛋白筛选, 差异蛋白生物信息学分析; 完整报告整理。

目前常用的仪器有: AB 5600 Triple TOF、Thermo Q Exactive Orbitrap;

数据分析包括：统计学分析（火山图、韦恩图、热图、聚类）、GO 富集、Pathway 通路分析、PPI 网络分析。

问题 5：iTRAQ/TMT 实验的中“组分”或者“SCX 分级”代表什么？

答：iTRAQ/TMT 实验对肽段进行的是 2D-LC-MS，其中第一维的色谱分析主要采用的是 SCX 或者高 pH 值的 HPLC 对标记好的肽段混合样品进行初步分离，而将样品分离成 10 个组分，以此来达到降低每个组分中肽段样品的复杂程度，然后再将这十个组分分别在进行 LC-MS，以便能达到更好的质谱鉴定效果。

问题 6：对于 iTRAQ 下机的原始数据，使用什么软件进行质谱鉴定分析？如何判断差异蛋白？

答：我们采用与 Triple TOF 5600 plus 配套的 proteinpilot 5.0v。已经有大量使用该软件发表的蛋白质组相关文献，是一款被广泛认可的蛋白质组鉴定、定量软件。目前对于差异蛋白筛选的条件没有统一的标准，实际筛选过程中需要我们结果具体的实验结果进行灵活的变化，尽量控制每组的差异蛋白总数在 200 左右或者鉴定结果总蛋白数目的 10%左右，比较有利于我们进行后续的生物信息学分析。一般可信蛋白筛选可以根据 FDR、unused 和 unique peptide 来进行筛选。对于定量差异蛋白的评判标准，一般按照 Fold change 和 P-value 等值来进行筛选。

问题 7：蛋白质谱鉴定不成功一般有哪些因素？没有全基因组测序的物种能否做蛋白质组？

答：目前已经有大量的物种完成了全基因组测序，并预测了相应的基因，未测序的物种可以通过近缘物种的基因序列作为数据库进行蛋白质组的研究，如果没有较好的近缘物种也可以使用稍大一些的属或者科的数据库进行检索。除此之外，如果研究人员做过转录组亦可以使用相关的数据建立本地库进行蛋白的质谱鉴定数据库。

问题 8：为什么通过 elisa、WB 能检测到的蛋白在 itraq 实验中没有检测到？

答：人血浆中的蛋白是有上万种的，而一般质谱只能检测到五六百种，也就是说只占了其中的百分之几，这是普遍现象，说明咱们的方法是不存在问题的；其次，ELISA/WB 与质谱相对来说灵敏度不一样，前者具有放大效应，因为其间会用到相应的抗体，所以只要选对抗体，感兴趣的蛋白基本都会被检测到；因此，ELISA/WB 检测到的蛋白在组学中没有被检测到也是正常的。

问题 9：我们 itraq 的结果，客户做验证，用 elisa 验证了两个差异蛋白，发现都是阴性结果，客户问这个是什么情况，出现阴性的概率有多大？一般我们会提到这个概率吗？

答： iTRAQ 结果中，他选的蛋白是否得分够高，变化倍数大不大，选择的抗体特异性是否好，都可能会影响 WB 和 iTRAQ 的结果，我们之前有一些客户做过，如果以上三方面没有什么问题的话，基本上 80%还是方向一致的。

问题 10: iTRAQ、TMT、Labelfree 实验用的是什么质谱仪，其定量是基于一级质谱还是二级质谱？

答： 基于 LC-MSMS 的实验技术主要使用的是 ABI5600-Triple-TOF，同时公司也具有 Thermo Q-Exactive 质谱仪平台，部分实验项目使用的是 QE-Exactive 完成。其中标记定量技术 iTRAQ 和 TMT 是依据二级质谱的报告基团离子峰进行定量，而 Labelfree 是基于一级质谱峰进行定量。

问题 11: 蛋白组学实验文章一般发表在哪些期刊上，各有什么区别？

答： MCP (molecular & cellular proteomics, 影响因子 7 分左右), JPR(journal of proteome research, 影响因子 5 分左右), Proteomics (影响因子 4 分左右), Journal of Proteomics(影响因子 4 分左右), Electrophoresis(影响因子 4 分左右), Plos One(影响因子 3 分左右), 其它的杂志还包括 BMC Genomics, Expert Rev Proteomics, BMC Syst Biol, OMICS, Proteomics Clin Appl, Proteome Sci。

问题 12: 有两个 Mix 混样，在计算差异表达蛋白时如何使用？

答： 计算差异表达蛋白时，将 WT 组 (113、114、115) 与 Mix1 (119) 和 Mix2 (121) 的比值取平均值，APO1 组 (116、117、118) 与 Mix1 (119) 和 Mix2 (121) 的比值取平均值，平均值再进行差异筛选。

问题 13: 生物学重复和技术重复的区别？

答： 生物学重复是指不同生物个体或者不同生物群体的样品之间进行地相同处理而进行的实验，而技术重复是指同一样品进行的多次相同实验。举例：对来自三只健康小鼠 A、B、C 的血清样品分别各进行一次双向电泳，这三次双向电泳就是生物学重复；而取其中一只小鼠的血清样品进行三次双向电泳，或者将三只小鼠的血清混合后的样品进行三次双向电泳，这三次实验就是技术重复。总结生物学重复的样品来自不同的个体或者是群体，重复之间样品不同；而技术重复的几次实验的样品完全相同，可以来自单一个体，也可以是多个个体的混合样品，但用于重复实验的样品完全一样。

问题 14: 结果报告中怎么没有差异多肽的定量？

答： 多肽组质谱数据进行搜库匹配后的结果会进行归一化整理并寻找差异多肽，所以说差异多肽的定量结果是有的，且以 Excel 发给客户，报告中一般会提及差异多肽的个数，具体的定量信息如果客户有需求也是可以添加的。

问题 15: 哪些技术需要进行重复实验, 以及具体采用生物学重复还是技术重复来完成?

答: **Label free** 需要进行三次重复, 而 **iTRAQ**、**TMT** 一般建议客户做 3 次实验重复, 如果实验方案后期设计了大量的验证实验如 **WB** 或 **ELISA** 等, 最好做也能做两次重复; 如果不进行重复实验后期的验证实验也少, 在后期发表文章的时候结果可能会受到一些审稿人的质疑, 从而影响文章的发表, 特别是一些高分杂志较为突出, 对于立志于发表高水平文章的研究人员建议客户最好进行重复实验设计。如果是进行技术重复, 各实验结果之间的数据重复性会更好, 而生物学重复由于多种原因可能导致重复性与偏差或者不好, 一般情况下 **Label free** 最好使用技术重复, 这样有利于后续数据的分析, 而其他是实验可以采用生物学重复或者技术重复。代谢组根据样品的不同选择合适的数量进行生物学重复。

问题 16: 如何来判定哪些蛋白是和自己研究相关的蛋白, 挑选什么样的蛋白来进行后期的验证实验?

答: 需要研究人员根据自己研究的相关生物学背景, 再结合生物信息学分析的结果, 以及查看相关的文献资料来确认哪些蛋白可以用于后续深入的研究。

问题 17: 样品寄送过程, 哪些需要低温? 哪些需要干冰? 哪些常温就可以?

答: 细胞、组织、蛋白溶液、菌体等样品最好采用干冰运输, 对于 **SDS-PAGE** 条带和 **2D** 点以及酶解好的肽段样品直接常温运输即可。可以理解为需要提取蛋白或提取好的蛋白样品(易降解)都需要低温运输, 而处于凝胶中的蛋白或者是酶解好的肽段可常温运输。

问题 18: 哪些样品需要去除高丰度蛋白? 怎么去除? 是否需要费用?

答: ①一般情况下, 血清、血浆样品中都会存在高丰度蛋白(白蛋白等免疫蛋白), 如果客户是此类样品, 需提前跟客户沟通好, 一定需要去除高丰度蛋白; 目前有成熟的试剂盒可以去除血清中的高丰度蛋白。

②体液样品, 细胞培养液等样品, 出现高丰度蛋白的情况也比较高, 这类样品, 如果出现高丰度蛋白, 需要先通过质谱鉴定是否是常见免疫蛋白, 然后再采用高丰度蛋白去除试剂盒进行去除。

③其他样品存在高丰度的情况不常见, 如在实验过程中发现高丰度蛋白, 解决方案如上。其他样品可能会出现鉴定出来的高丰度蛋白没有合适的试剂盒可以去除, 这时候需要跟客户提前告知风险。

④去除高丰度蛋白的费用详见鹿明生物蛋白组报价。

问题 19: 蛋白质组结果一般需要做哪些方面的生物信息学分析?

答：目前一般文献中的分析主要为维恩图分析、热图分析、火山图分析、层次聚类分析、蛋白质功能分析（GO 分类为主）、蛋白质所属代谢通路富集分析（基于 KEGG 的 pathway）、差异蛋白互作网络构建、转录因子预测等方面。其它一些比较特异的高级分析包括多组数据关联分析、多 pathway 延伸分析、磷酸蛋白激酶预测分析、互作网络及生物调控模型构建分析。

问题 20：小鼠脑脊液做 label free 是否需要去除高丰度蛋白，有没有案例或者介绍下蛋白组这块哪些样本需要去高丰度蛋白的；

答：需要根据样本提取蛋白后 SDS-PAGE 结果评估。一般血清血浆都需要除去高丰度蛋白，试剂盒成熟。其余需要根据跑胶结果来判断。

问题 21：iTRAQ 定量蛋白质组做完实验后，能拿到什么结果呢？可以直接用来发文章么？

答：我们目前的 iTRAQ 定量蛋白质组的实验流程和结果都是受到业内顶级杂志 JPR 和 MCP 公认的，可以直接用于文章的发表。客户不仅可以拿到 iTRAQ 定量蛋白质组中的差异蛋白、蛋白的 GO 和 KEGG 归类信息、COG 归类、不同样本的差异蛋白的聚类分析、差异蛋白的互作蛋白信息等，还可以提供客户指定的分析内容，提供个性化的服务要求，比如转录组和蛋白质组的关联分析。

问题 22：物种没有测序，能否做蛋白质组呢？

答：因为当前大量的物种已经测序，预测了相应的基因，未测序的物种可以通过近缘物种的基因序列作为数据库进行蛋白质组的研究。因此，一般来说，未测序物种可以考虑近缘物种的基因组序列作为数据库，可以做蛋白质组。

问题 23：关于“样品重复”，技术重复和上机重复有什么区别？发表的文章一般要求几次重复呢？

答：一般建议客户做 3 次实验重复，如此可以在发表文章的过程中不被 editor 挑刺，从而影响发文章的速度。一般的文章会要求 3 次实验重复，如果 iTRAQ 后续做了大量的 WB 或 ELISA 验证，那么，2 次实验重复也是必要的；仅作 1 次实验有风险，因为如果是实验过程出现问题，缺乏结果的评估，往往说服力不够，对文章发表可能构成影响，建议客户最好不要只作 1 次实验。

技术重复一般和上机重复非常类似，即上质谱的重复，即测试仪器的稳定性，判断系统误差等问题。这一说法主要和质谱仪器的发展进程相关，早期的仪器不是很稳定，同一个样本上机 3 次，结果差异较大，当前的上机重复仍然在一些 JPR 和 MCP 文章中可见，即沿用了早期的习惯。而应用型的文章中，上机重复或技术重复一般没有硬性规定，可以不做，如果做了当然最好，editor 就不会再此处找你的问题。

问题 24：为什么有些差异表达基因在蛋白层面没有相应的差异蛋白？

答：这是学术性的问题，不是技术的问题。表达层次即 RNA 层次，RNA 层次和 protein 层次的不一致性是一个公认的事实，通常相关性只有~0.5 左右，即 RNA 和 protein 往往是两个不一样的概念。引起 RNA 和 protein 差异的原因很多，比如 miRNA、lincRNA、circle RNA 和 E3 ligase 等等

问题 25：对于 iTRAQ 下机数据，使用什么软件分析？如何判断差异蛋白？

答：采用与 Triple TOF 5600 plus 配套的 proteinpilot。已经有大量使用该软件发表的蛋白质组相关文献，是一款被广泛认可的蛋白质组鉴定、定量软件。我们对鉴定结果要求  $FDR \leq 1\%$ ，每个蛋白至少包含 1 个独有肽段。对于定量差异蛋白的评判标准，我们一般按照  $FC \geq 1.5$ ，同时  $p\text{-value} \leq 0.05$  的标准过滤。

问题 26：iTRAQ 与其他蛋白质组学技术的区别是什么？

答：双向电泳是最早、最经典的蛋白质组技术，适用于大部分样本，只是对强酸或强碱性蛋白，以及分子量太大或太小的蛋白质分离效果不好；另外由于各个样本需要单独跑胶，很难避免操作误差对结果的影响，可能导致定量不准确；但是双向电泳的服务价格相对便宜，可以用于样本差异的初步筛选；

DIGE 在双向电泳基础上做了改进，引入了荧光标记物和内标，可以使定量更加准确；但该技术也是依据蛋白的等电点和分子量分离，因此对强酸或强碱性蛋白，以及分子量太大或太小的蛋白质分离效果也不好。

iTRAQ、TMT、label-free 和 SILAC 都是利用液质联用技术来寻找差异蛋白。其中，iTRAQ 和 TMT 的原理和检测方法基本一致，只是使用了不同的标记物；TMT 有 10 种标记物，可对多达 10 种不同样本进行分析，但是由于其标记物的质量数非常相近，因此必须使用超高分辨率的质谱才可以满足检测，这也必然会导致信号干扰更多，可能会影响定量的准确性；label-free 是非标记的蛋白质组学技术，由于没有标记物，其定量需要依赖于操作和质谱仪的稳定性，该技术鉴定到的蛋白数目会少于 iTRAQ 技术，并且定量的准确性较差，优势是服务费用较低；SILAC 是基于稳定同位素标记的细胞培养技术的蛋白质组学方法，利用不同的同位素培养基来培养不同组别的细胞，达到体内标记的目的，其定量准确性要比其他蛋白质组学技术更高；但是由于缺乏合适的同位素培养基，该方法基本只能用于可传代的哺乳动物细胞系。

从定量准确性、应用性等各方面综合考虑，iTRAQ 技术已成为近年来利用最广泛的蛋白质组学技术。

问题 27：iTRAQ 实验对样本有何要求？

答：任何类型的样本都可以，如果是罕见物种，可以用近缘物种的数据库检索，也可提供转录组数据库来检索；一次上机，每组样本需要用 200 $\mu$ g 蛋白，但由

于需要蛋白定量检测和预实验,因此每组样本所需的蛋白量大概在 500 $\mu$ g 左右; 我公司负责蛋白质的提取, 客户只需要提供足够的原始样本。

问题 28: 不同类型的样本做 iTRAQ 能否一次上机?

答: 不能。如果样品来源于不同物种, 检索时就无法选择数据库, 也就无法分析数据; 即使是同一物种的样本, 它们的蛋白种类和丰度也可能有很大差别(比如植物的根和叶), 将这些酶切肽段混合上机, 其数据会相互干扰, 导致蛋白定性和定量的不准确。

问题 29: iTRAQ 实验的数据采用不同数据库检索时, 各组间的比值为何是不同的?

答: iTRAQ 技术中不同标记物之间的比值与检索数据库、质谱电信号的噪音以及采集质谱时的随机性都有关。所以在使用不同数据库检索时, 由于匹配肽段情况的不同, 定量结果也有些许差异, 但是总体变化趋势应是一致的。

问题 30: 检索时为什么同一个编号对应了很多蛋白质?

答: 因为一个蛋白家族中常常含有多个同源性很高的蛋白质, 这些蛋白质被酶切之后, 很可能产生很多相同的肽段, 软件在进行定性分析时, 就会将这些肽段均归于得分最高的那种蛋白质, 其他同源蛋白不计分, 并且这些同源蛋白都采用同一编号。