

元蛋白质组分析 ——研究微生物生态功能的新途径*

牛 泽¹ 杨 慧² 刘 芳² 马荣才^{1,2} 高俊莲^{1,2,*}

(首都师范大学生命科学院 北京 100037)(北京市农林科学院农业生物技术研究中心 北京 100097)

摘要 随着对微生物纯化培养和元基因组学研究的不断深入,积累了大量的微生物基因组信息,元蛋白质组分析将促进我们对基因组功能的理解。目前,对微生物群落的元蛋白质组(metaproteome)的研究已成为后基因组时代进一步认识微生物生态功能的有效途径。对这一新技术的介绍结合元蛋白质组的提取和鉴定方法、元蛋白质组在研究微生物生态功能上的应用进行了综述,并对这一新的研究领域所面临的困难与挑战进行了讨论。

关键词 元蛋白质组,微生物生态,环境微生物,蛋白质分离纯化

中图分类号:Q93 文献标识码:A 文章编号:0253-2654(2007)04-0804-04

Metaproteome—a New Approach for Studying Microbial Ecology Functions*

NIU Ze¹ YANG Hui² LIU Fang² MA Rong-Cai^{1,2} GAO Jun-Lian^{1,2,*}

(Capital Normal University, Beijing 100037)(Beijing Agro-Biotechnology Research Centre, Beijing 100097)

Abstract The information of microbial genomics is accumulated with the progress of pure culture and metagenome research. Metaproteomic study will contribute to our understanding of genomic function. Recently, metaproteomic study of microbial community is becoming an effective approach to understand the function of microbial ecology in post-genome era. In this review, the new metaproteomic approach was introduced, including extraction and identification of metaproteomes and application in microbial ecology. Finally, the future perspectives were discussed.

Key words Metaproteome, Microbial ecology, Environmental microbes, Protein isolation and purification

20世纪70年代以来,随着分子生物学技术以及基于 rRNA 序列分析的生物分子系统进化理论的建立与发展,使人们有可能在纯培养的技术手段之外,采用非培养技术研究微生物群落结构组成^[1]。在过去的二十多年中,以 16S rDNA PCR 扩增为主的非培养技术的研究揭示了环境中庞大的微生物多样性^[2]。近年来发展起来的元基因组(metagenome)技术,进一步揭示出环境中微生物的遗传多样性和功能多样性^[3]。元基因组序列可提供环境中微生物潜在的有价值的功能性信息。但是,如果没有在特定条件下蛋白质合成方面的信息,欲从元基因组序列精确预测微生物的生态功能几乎是不可能的^[4]。由于蛋白质的表达是一个特定生态系统中微生物活性的反映,因此元蛋白质组技术具有进行微生物群落生态功能分析的巨大潜力。

1 元蛋白质组的概念

元蛋白质组(metaproteome)这个概念是由 Rodriguez-Valera 于 2004 年首先提出的^[5,6],目的是检测环境样品中高度表达的基因。元蛋白质组是利用物理或化学的手段,直接分离提取某一特定环境中的蛋白质组,并对其中的蛋白质成分进行大规模的鉴定,从而认识环境中微生物特定活性的方法。它主要是通过蛋白质组学技术对环境微生物在基因表达层面上进行研究。

2 元蛋白质组的提取与鉴定方法

元蛋白质组分析是一种从整体上研究环境微生物群落与动态功能关系的有效方法。应用这种分析技术依赖于从环境样品中高效的分离出蛋白

* 北京市自然科学基金项目(No. 5062010)北京市留学人员科技活动择优资助项目(2005年度)资助

** 通讯作者 010-51503834, E-mail: gaojunlian@baafs.net.cn

收稿日期:2006-12-07,修回日期:2007-02-28

质成分,从而获得一个蛋白质池(protein pool)。其利用的基本技术包括聚丙烯酰胺凝胶电泳(SDS-PAGE)和蛋白质谱鉴定技术,其中蛋白质谱鉴定主要应用肽质量图谱和肽片段的部分测序,但这些技术只在有限范围的环境微生物学研究中得到了应用^[7]。元蛋白质组研究流程可以参考图1。

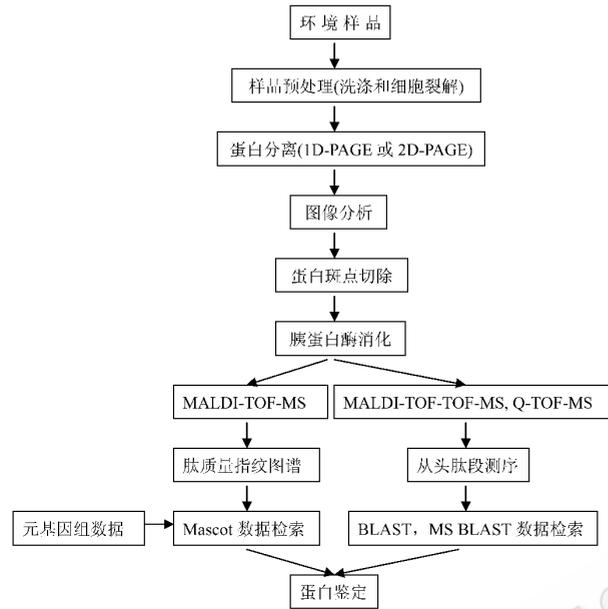


图1 元蛋白质组研究的流程图(部分)^[8]

注: MALDI-TOF-MS 基质辅助激光解吸/电离飞行时间质谱, MALDI-TOF-TOF-MS 基质辅助激光解吸/电离飞行时间串联质谱, Q-TOF-MS 四级杆飞行时间质谱

3 元蛋白质组在研究微生物生态功能上的应用

Ogunseitun 最先用蛋白质组的方法分析了不同环境样品中蛋白质混合物的情况^[9]。他用这种方法证明了所有污泥中都具有微生物多样性,并且证明其中发生了一系列复杂的生物化学变化。其后,随着元蛋白质组技术的不断进步,这种研究手段在环境微生物学领域显示出越来越显著的作用。

3.1 元蛋白质组能够有效的鉴定环境中蛋白质的来源及其功能

Wilmes 和 Bond^[8]应用元蛋白质组的方法研究了实验室规模的微生物混合群落。在这一研究中,首次应用了 2D-PAGE 的方法分离环境微生物中的蛋白质,并提取和纯化到整个混合群落的蛋白质组,应用质谱技术对获得的蛋白质成分进行了鉴定,研究了增强性生物除磷过程(EBPR)^[6]中特殊的多磷酸盐积累微生物(PAOs),例如与红环菌

(*Rhodocyclus*)有关的未培养微生物,并鉴定出了强表达的磷酸盐选择性膜蛋白、与聚苯氧树脂合成有关的蛋白和影响糖原合成与降解的蛋白^[6]。

Chen^[7]等人研究了 Chesapeake Bay 中特定大小的微生物群落(0.2 μ m ~ 3.0 μ m)的元蛋白质组。他们用 LC-MS/MS 和 MS-BLAST 对元蛋白质组进行测序和检索,鉴定出了 3 种蛋白质成分,其中的 CB1 与 Sargasso Sea 元基因组文库^[10]中的一个基因序列匹配;CB3 与 NADH 的亚基辅酶 Q 氧化还原酶有关, CB6 与一种氨基肽酶相似。这 3 种蛋白质都起源于海洋中的未培养微生物,包括拟杆菌(*Bacteroides*)和 α -变形细菌(α -proteobacteria)。

3.2 元蛋白质组能够帮助认识环境胁迫下微生物的代谢过程

元蛋白质组能够通过检测某些蛋白质是否存在、相对丰度和蛋白质的修饰状态,直接测量功能基因的表达情况。通过对具有关键酶活性蛋白质的鉴定,可以使人们进一步认识某些环境胁迫下微生物的代谢过程。

Wilmes 等人^[11]通过应用原位荧光杂交技术(FISH)研究具有 EBPR 和不具有 EBPR 的活性污泥,发现具有较强 EBPR 的污泥中占主要成分的微生物属于 β -变形菌纲,不具有 EBPR 的污泥中占主要成分的微生物属于 α -变形杆菌纲。比较这两种污泥在厌氧和好氧状态下蛋白质图谱的变化,发现不具有 EBPR 的蛋白质表达图谱具有明显的差异性,而具有较强 EBPR 的活性污泥的蛋白质表达图谱变化不大。一种可能的解释是,在较强 EBPR 的活性污泥中微生物表达了大量的补偿蛋白(equilibrating protein),这些蛋白质在能量代谢过程中为 PAOs 提供了更多的能源^[11]。

Ram 等人应用“鸟枪”质谱蛋白质组学(‘shotgun’MS proteomics)与群落微生物基因组分析相结合的方法,成功进行了大规模的元蛋白质组研究^[12],共得到了大约 2033 个不同的蛋白质成分,其中 48% 起源于钩端螺旋菌(*Leptospirillum*)。鉴定到大量在极端环境中与微生物生存直接相关的蛋白质,如伴侣蛋白、硫氧还蛋白和还原酶类。也有很大一部分蛋白质不能直接归于某种功能类别,它们可能是新基因的表达产物。胞外部分包含 52% 的特殊蛋白(与任何已知蛋白质都不明显匹配)和 14% 的保守蛋白。其中一种蛋白质被鉴定为一种新

的细胞色素 Cyt579 是铁氧化过程和 AMD 代谢过程的核心^[12]。

Pierre-Alain^[4]等人研究了受重金属离子汞和镉胁迫下水体微生物所表达蛋白质的变化。发现在重金属离子的胁迫下,环境微生物表达了大量特异性的蛋白质,在元蛋白质组上形成了特殊的条带。众所周知,金属离子对微生物群落有毒害作用,蛋白质图谱的变化表明环境中微生物的群落功能发生了变化。Singleton^[13]等人用液氮快速冻融裂解土壤中细菌提取蛋白质的方法,比较了在重金属镉污染条件下细菌表达蛋白量的变化。发现从镉污染的土壤中提取得到的蛋白质总量比没被污染的对照土壤要少很多。此外,Ogunseitan^[14]等人通过研究有机污染物胁迫下水体中的元蛋白质组,也发现了水体微生物群落蛋白质指纹图谱会随环境条件的变化而发生变化。

3.3 元蛋白质组能够认识不同环境中微生物的群落组成与分布

随着从头肽段测序(*de novo* peptide sequence)技术的发展,元蛋白质组学研究已经能够进行蛋白质鉴定以及系统发生和起源分析,通过确定生态系统中各个蛋白质之间的系统发生关系,可以进一步认识不同环境中微生物的群落特点。

Schulze^[15]等人用质谱联用(MS-MS)技术研究了四处不同水环境中的蛋白质成分,这些水源中都含有高浓度的有机溶解物。通过分析所有样品的蛋白质成分,发现在森林沥出液中,起源于植物、真菌和脊椎动物的蛋白质数量大约是湖水中数量的2倍;与复杂分子降解有关的过氧化物酶类只存在于森林的水源中,而与甲烷产生有关的转移酶类却只存在于泥炭沼中的水源里。尽管这一研究只检测了环境样品中表达的一小部分蛋白质,它还是反映出了不同分类地位的微生物的存在与活性。

Chen^[7]等人应用元蛋白质组的方法对 Chesapeake Bay 横断面的微生物群落蛋白质的组成情况进行了分析。发现在不同地点收集得到的蛋白质组存在着明显的不同。对海湾中部的元蛋白质组进行重复检测发现大约92%的蛋白质位点是一致的,但是与上游和下游海湾的元蛋白质组相比,只有30%和70%的相似性。

元蛋白质组技术与元基因组技术结合,能更好的认识环境微生物。在大型的元基因组数据库中,

元蛋白质组图谱无疑提供了确定基因表达水平的依据,这样有助于研究高度表达的未知功能基因。另外,通常从海量的元基因组信息中获得某种微生物的全基因组费时费力,而通过元蛋白质组序列做蛋白质的鉴定相对来讲要简单易行。当环境微生物的基因组序列信息缺失时,可以辅助应用从头肽段测序技术和质谱信息搜索工具,鉴定环境样品中获得的蛋白质^[16]。这样就可以不需要完整的元基因组序列,直接认识复杂环境中的生理生化现象。

4 元蛋白质组研究面临的困难与挑战

在环境微生物学领域,尽管用蛋白质组分析的方法鉴定了为数不少的蛋白质,但是由于蛋白质本身呈现出不同的生物和物理构造,并且没有统一的提取方法可以参照^[17],对复杂微生物生态系统的研究还存在着极大的困难。尤其对于海水和土壤环境,应用已有的研究手段,最多只能还原整个元蛋白质组的一部分^[7]。另外,在一个样品中微生物的种类和蛋白质的表达水平都处于很强的动态变化中。有人研究发现,一个细胞内蛋白质的表达水平在不同的条件下就可能相差6个数量级之多^[18]。再有,一些研究主要关注的是胞内蛋白,这就要求对元蛋白质组的细胞成分进行分离(胞内的、外周胞质的、全细胞的、可溶的和膜蛋白成分),再对某些功能性蛋白做细胞定位。这就对蛋白质的分离工作提出了更高的要求。

在蛋白质鉴定方面,目前主要应用的是质谱技术,但是不同方法获得的肽质量指纹图谱与现存的数据库比对往往不能得到相似性很高的结果。一般来讲,MALDI-TOF需要所查询的蛋白质与目标蛋白的氨基酸序列具有97%以上的一致性才能明显匹配^[19]。由于目前存在的蛋白质序列数据库大都来源于可培养微生物,因此对于环境样品中获得的蛋白质,要与这些数据库中的序列达到如此高的一致性是不太可能的。另外,蛋白质的翻译后修饰作用也增加了鉴定的难度^[6,7]。而其它的鉴定技术,如N-末端测序技术,由于从环境中获得的大多数蛋白质丰度太低,也不适用于进行鉴定。同时,获得的蛋白质匹配信息可能存在相对较高的假阳性,所以每种蛋白质需要两条以上的肽段能够与鉴定结果匹配才能确定鉴定信息^[7]。

元蛋白质组研究环境微生物群落仍存在一

些障碍,包括:如何判断在2D-SDS-PAGE凝胶上某一蛋白质斑点是来自于环境样品中的一种蛋白质还是几种蛋白质;用什么信息判断来自不同样品上的斑点是否是同一的;元蛋白质组对环境微生物群落组成和群落中微生物生理状态的变化的灵敏度如何;怎样进一步验证元蛋白质组学提供的蛋白质功能性信息。很明显,要解决这些问题需要开展更多的研究工作和改善目前的技术环节。

5 结语

目前关于元蛋白质组的研究虽然不多,它还是在环境微生物学领域显示出了重要作用。由以上综述可知,尽管应用蛋白质组对更加复杂环境的微生物群落进行研究还面临着更大的挑战,但是元蛋白质组在阐明未培养微生物的功能性信息、鉴定新的蛋白质成分和认识不同环境微生物群落组成及其生化功能方面有着巨大的应用潜力。

参考文献

- [1] Riesenfeld C S, Schloss P D, Handelsman J. *Annu Rev Genet*, 2004, **38**: 525 ~ 552.
- [2] 王 涛, 柴丽红, 郭春雷, 等. *微生物学通报*, 2003, **30**: 38 ~ 42.
- [3] 阎 冰, 洪 葵, 许 云, 等. *微生物学通报*, 2005, **32**: 113 ~ 117.
- [4] Pierre-Alain M, Christophe M, Se' verine S, *et al.* *Microb Ecol*, 2006, Epub ahead of print.
- [5] Rodriguez-Valera F. *FEMS Microbiol Lett*, 2004, **231**: 153 ~ 158.
- [6] Wilmes P, Bond P L. *Environ. Microbiol*, 2004, **6**: 911 ~ 920.
- [7] Kan J, Hanson T E, Chen F, *et al.* *Saline Systems*, 2005, **1**: 7.
- [8] Wilmes P, Bond P L. *Trends Microbiol*, 2006, **14**: 92 ~ 97.
- [9] Ogunseitani, O A. *J Microbiol Methods*, 1993, **17**: 273 ~ 281.
- [10] Venter J C, Remington K, Heidelberg J F, *et al.* *Science*, 2004, **304**: 66 ~ 74.
- [11] Wilmes P, Bond P L, *Water Sci Technol.* 2006, **54**: 217 ~ 26.
- [12] Ram R J, Verberkmoes N C, Thelen M P, *et al.* *Science*, 2005, **308**: 1915 ~ 1920.
- [13] Singleton I, Merrington G, Colvan S, *et al.* *Appl Soil Ecol*, 2003, **23**: 25 ~ 32.
- [14] Ogunseitani O A. *Microb Ecol*, 1996, **31**: 291 ~ 304.
- [15] Schulze W X, Gleixner G, Kaiser K, *et al.* *Oecologia*, 2005, **142**: 335 ~ 343.
- [16] Shevchenko A, Sunyaev S, Loboda A, *et al.* *Anal Chem*, 2001, **73**: 1917 ~ 1926.
- [17] Ogunseitani O A. *Molecular Microbial Ecology Manual*, 2004, Second. pp. Edition: 1027 ~ 1046.
- [18] Pandey A, Mann M. *Nature*, 2000, **405**: 837 ~ 846.
- [19] Shevchenko A, Sunyaev S, Liska A, *et al.* *Methods Mol Biol*, 2003, **211**: 221 ~ 234.