

# 应用代谢组学批量数据分析工作流程表征细菌中的差异代谢物

## 应用简报

### 作者

Yuqin Dai 和 Steven M. Fischer  
安捷伦科技公司  
美国加利福尼亚州圣克拉拉市

### 摘要

本文应用发现代谢组学的精确质量 Q-TOF LC/MS 工作流程，研究了处于早稳定期和晚稳定期的细菌。所采用的软件可批量处理数据，使得数据分析更高效且实现了自动化。应用 Agilent MassHunter Profinder（批量特征提取软件）总共从正离子数据中获得了 488 个特征，从负离子数据中获得了 623 个特征。采用 Mass Profiler Professional (MPP) 进行的统计学分析揭示了细菌在早稳定期和晚稳定期丰度具有显著差异的特征。在正离子数据中，有 57 个特征在早稳定期中的丰度要高于晚稳定期。而在负离子数据中，有 52 个特征在早稳定期中的丰度要明显高于晚稳定期。为了理解这些代谢组学数据的生物学和生物化学背景，我们通过数据库搜索、精确质量 MS/MS 谱库匹配，以及 MS/MS 分子结构关联对超过 100 个差异特征进行了标注和鉴定。



**Agilent Technologies**

## 前言

人们对于细菌生长的研究已有超过一个世纪的历史，研究成果包括由 Jacob Monod 提出的细菌生长期的经典定义<sup>1</sup>。研究表明，细胞在晚稳定期依然存活且代谢活跃，但是晚稳定期的细胞对刺激或压力的响应与早稳定期的细胞并不相同。研究人员假设细菌在稳定期中产生的一种或多种代谢物会导致细菌在晚稳定期活力下降，或者，随着稳定期的发展，保护性代谢物会逐渐消失。据我们所知，至今还没有任何关于这些代谢物表征的信息。

在本文中，我们将介绍一种 Q-TOF LC/MS 代谢组学方法，用于研究细菌在早稳定期和晚稳定期的差异代谢物。安捷伦的全套数据处理软件简化了特征寻找、统计学分析和代谢物鉴定过程。它能迅速将复杂的原始数据转化为生物学相关的代谢物信息。

## 实验部分

### 样品前处理

分别收集早稳定期和晚稳定期的细胞培养物，用磷酸盐缓冲液清洗两次以去除残留的细胞外基质，然后用水清洗两次以去除缓冲液。然后将细胞悬浮液进行离心，使其形成细胞团块，倾倒入去除上清液。在 -80 °C 下向细胞团块中加入 0.5 mL MeOH，然后在 4 °C 下加入 1.5 mL 50% MeOH/H<sub>2</sub>O 溶液。对细胞进行机械裂解后，将粗提物在 48000 g 下离心 1 h，然后收集上清液。取 500 μL 上清液加入 3 kDa Nanosep 过滤器中，在 14000 g 下离心，收集滤液，然后进行真空干燥，最后将其保存于 -60 °C 下直至分析。

对于 LC/MS 分析，使用 500 μL 水/乙腈 (3:7 v/v) 溶液对样品进行重悬，然后涡旋混合并离心。将上清液从 Eppendorf 管转移至干净的 LC/MS 样品瓶中。

### 仪器

采用与 Agilent 6230 TOF 系统或 Agilent 6550 iFunnel Q-TOF 联用的 Agilent 1290 Infinity LC 系统进行 LC/MS 分析。液相色谱系统包括以下模块：

- Agilent 1290 Infinity 二元泵 (G4220A)
- 带恒温器的 Agilent 1290 Infinity 自动进样器 (G4226A)
- Agilent 1290 Infinity 柱温箱 (G1316C)
- 带 100:1 分流器 (部件号 G1607-60000) 的 Agilent 1100 系列等度泵 (G1310)

通过使用等度泵持续注入参比质量溶液进行动态质量轴校准。表 1 和表 2 汇总了优化的 LC 和 MS 条件。

### 数据分析软件

- Agilent MassHunter 定性分析 (Qual) B.07.00
- Agilent MassHunter Profinder B.06.00, service pack 1
- Agilent Mass Profiler Professional (MPP) B.13.0
- Agilent MassHunter 分子结构关联工具 (MSC) B.07.00
- Pathway to PCDL B.07.00

表 1. LC 参数

参数	Agilent 1290 Infinity 液相色谱系统
分析柱	Cogent Diamond Hydride 色谱柱, 2.1 × 150 mm, 4 μm, 100 A (部件号 70000-15P-2), Microsolv 科技公司
保护柱	Agilent ZORBAX SB-C8, 2.1 × 30 mm, 3.5 μm (部件号 827700-936)
柱温	60 °C
进样量	5 μL — Agilent 6230 TOF 分析实验 10 μL — Agilent 6550 Q-TOF 鉴定实验
自动进样器温度	4 °C
正离子模式的流动相	A) 50% 水/50% 异丙醇, 含有 0.1% 甲酸 B) 3% 水/97% 乙腈, 含有 0.1% 甲酸
负离子模式的流动相	A) 50% 水/50% 异丙醇/0.025% 甲酸及 5 μM EDTA B) 10% 水/90% 乙腈/5 mM 甲酸铵/5 μM EDTA, pH = 7
流速	0.6 mL/min
正离子模式的梯度	0 – 1 min 97% B 1 – 15 min 97% B 到 10% B
负离子模式的梯度	0 – 1 min 99% B 1 – 15 min 99% B 到 20% B
停止时间	15 min
后运行时间	5 min

## LC/MS 代谢组学工作流程

我们开发了一套高分辨率的 Q-TOF LC/MS 代谢组学工作流程 (图 1) 用于细菌早稳定和晚稳定期的无偏差代谢组学研究。工作流程如下:

1. 分别在正离子模式和负离子模式的 TOF 模式下采集提取物的数据
2. 采用 Agilent MassHunter Profinder<sup>2</sup> 批处理递归特征提取算法从原始数据中寻找特征
3. 采用 MPP 进行差异分析和标注
4. 然后采用 Q-TOF LC/MS 对目标差异特征进行靶向 MS/MS 分析
5. 根据 MS/MS 谱图匹配采用定性分析工具以鉴定代谢物, 或采用分子结构关联 (MSC) 软件对可能的代谢物进行结构解析

表 2. MS 参数

参数	Agilent 6230 TOF 系统
离子模式	正离子和负离子
离子源	安捷伦双 ESI 离子源
毛细管电压	3500 V ( $\pm$ )
干燥气温度	350 °C
干燥气流速	10 L/min
雾化器压力	45 psi
碎裂电压	100 V
MS 范围	25 – 1600 $m/z$
MS 采集速率	2 质谱图/秒
参比质量	正离子模式: 64.01577 和 922.009798 负离子模式: 68.9957 和 966.0007
参比泵流速	0.5 mL/min (正离子和负离子模式)
参比输送	带 100:1 分流器 (部件号 G1607-60000) 的 Agilent 1100 等度泵
仪器模式	扩展动态范围 (2 GHz)
参数	Agilent 6550 Q-TOF 系统
离子模式	正离子和负离子
离子源	安捷伦双 ESI 离子源
毛细管电压	3500 V ( $\pm$ )
干燥气温度	200 °C
干燥气流速	15 L/min
雾化器压力	45 psi
MS 范围	50 – 1700 $m/z$
MS 采集速率	3 质谱图/秒
MS/MS 范围	25 – 1700 $m/z$
MS/MS 采集速率	2 质谱图/秒, 靶向 MS/MS
分离峰宽	中等 (约 4 $m/z$ )
碰撞能量	10 – 40 eV
参比质量	正离子模式: 64.01577 和 922.009798 负离子模式: 68.9957 和 1033.9881
参比泵流速	0.6 mL/min (正离子模式) 0.5 mL/min (负离子模式)
参比输送	带 100:1 分流器 (部件号 G1607-60000) 的 Agilent 1100 等度泵
仪器模式	扩展动态范围 (2 GHz)

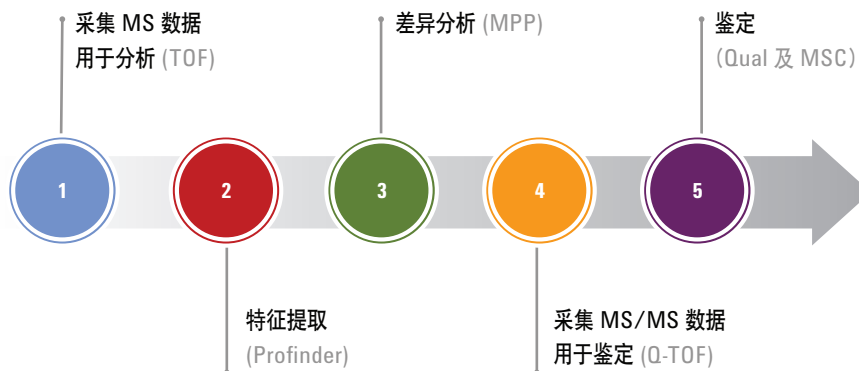


图 1. 分析和鉴定细菌代谢物的 LC/MS 代谢组学工作流程

## 结果与讨论

### 利用 TOF LC/MS 进行代谢物分析

理论上, 采用 TOF 或 Q-TOF 均可进行代谢物分析。如果实验室只有一台可用仪器, Q-TOF 会是更好的选择, 因为它既能进行代谢物分析, 又能进行用于鉴定的 MS/MS 分析。但是, 如果实验室样品体积需要多个仪器, TOF 和 Q-TOF 的组合更加经济有效。在该研究中, 我们采用 6230 TOF 进行代谢物分析, 采用 6550 Q-TOF 进行鉴定。

通过比较 10 个早稳定期细胞提取物和 9 个晚稳定期细胞提取物的代谢物分析结果, 我们评价了细菌代谢组在这两个不同阶段的差异。我们采用 6230 TOF LC/MS 在正离子和负离子模式下对这些样品进行分析。为了最大化极性代谢物的分析范围, 色谱分离采用了水相正相 (ANP) 色谱柱。图 2 展示了细胞提取物在负离子模式下的代谢物分析结果。来自 19 个细胞提取物的 38 张总离子流色谱图 (TIC) 的叠加图表明, 色谱图之间的保留时间差异很小。保留时间的重现性在发现代谢组学研究中非常重要, 因为只有保留时间具有重现性时, 才能进行不同样品间的色谱峰匹配。

### 采用 Agilent MassHunter Profinder 进行批量数据特征提取

需要耗费大量时间从大量复杂的原始数据中提取特征是代谢组学数据分析的瓶颈之一。MassHunter Profinder 软件能够自动化批量处理数据, 专为提高通量而设计。在本研

究中, Profinder 根据用户自定义的化合物筛查条件和手动修正从正离子数据中获得了 488 个特征, 从负离子模式中获得了 623 个特征 (图 3)。所获得的特征接下来被导入 MPP (B.13.0) 进行差异分析。

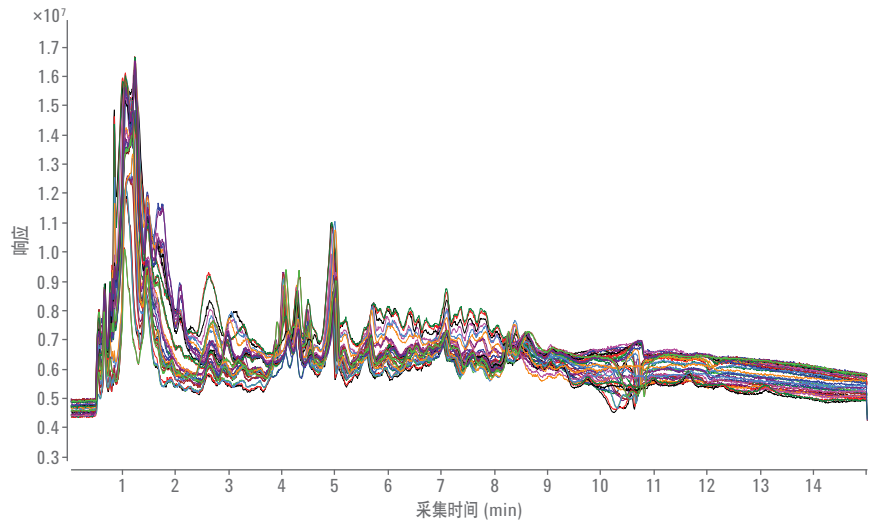


图 2. ESI 负离子模式下, 19 个细胞提取物的 38 张总离子流色谱图 (TIC) 的叠加图。每个样品进样分析两次

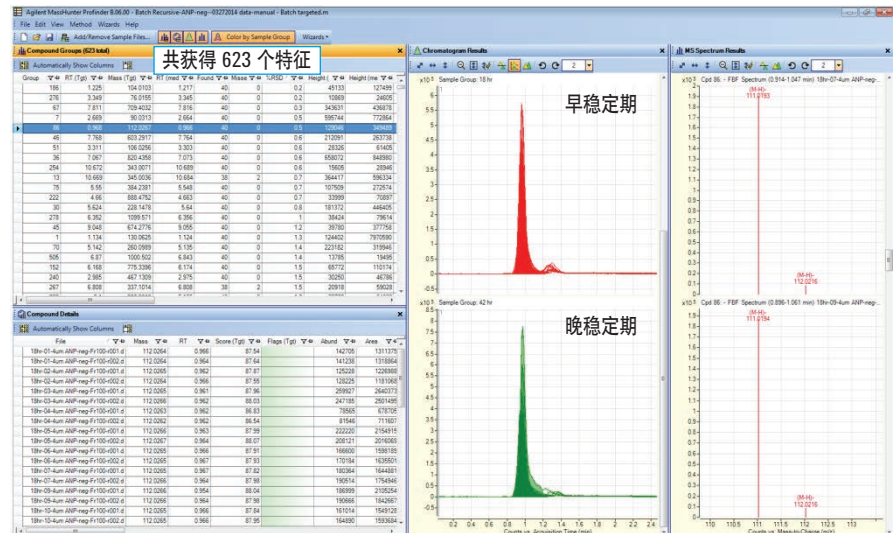


图 3. 采用 Profinder 批量处理递归特征提取算法从早稳定期和晚稳定期的细胞提取物的负离子质谱数据中获得的非靶向特征提取结果

## 采用 MPP 进行差异分析和标注

为了确定细菌早稳定期和晚稳定期之间的差异特征，我们采用 MPP 中的火山图筛查算法对所获得的特征进行了评估，其中，临界值为  $P < 0.005$ ，倍数变化 (FC)  $> 2$ 。通过统计学分析获得了细菌在早稳定期和晚稳定期丰度具有显著差异的特征。由正离子数据获得的 488 个特征中，有 98 个特征表现出统计显著性，并且其中的 57 个特征在早稳定期的丰度要高于晚稳定期。同样的，由负离子数据获得的 623 个特征中，有 152 个特征具有显著差异 (图 4)，其中的 52 个特征在早稳定期的丰度要高于晚稳定期 (蓝色方块)。

为了理解这些代谢组学数据的生物学和生物化学背景，我们采用 MPP 中的 ID 浏览器以及允许质量数误差为 10 ppm 的精确质量数数据库匹配对这些差异特征进行了标注。用于查询的 3 个数据库分别为 Agilent-Metlin、KEGG 和 BioCyc。KEGG 和 BioCyc 数据库均由 Agilent Pathway to PCDL 软件创建。图 5 展示了在负离子模式下获得的一组差异特征的 ID 浏览器结果。在化合物列表中点击某种化合物 (例如这里的 L-亮氨酸)，即可显示相应的质谱图结果、详细的同位素峰信息，以及化学结构。用户可以快速浏览并根据精确质量数、同位素模式、保留时间以及匹配分值等参数对结果进行排序，从而选择最佳的结果用于标注。

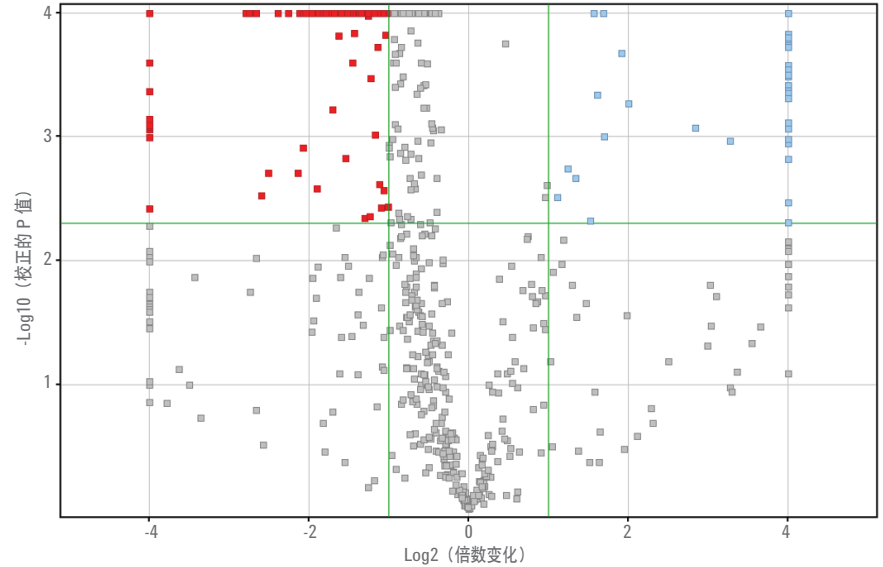


图 4. 早稳定期和晚稳定期特征的火山图 (负离子数据)。红色的点表示这些特征在早稳定期的丰度低于晚稳定期的丰度 ( $P < 0.005$ ,  $FC > 2$ )。蓝色的点表示这些特征在早稳定期的丰度高于晚稳定期的丰度 ( $P < 0.005$ ,  $FC > 2$ )

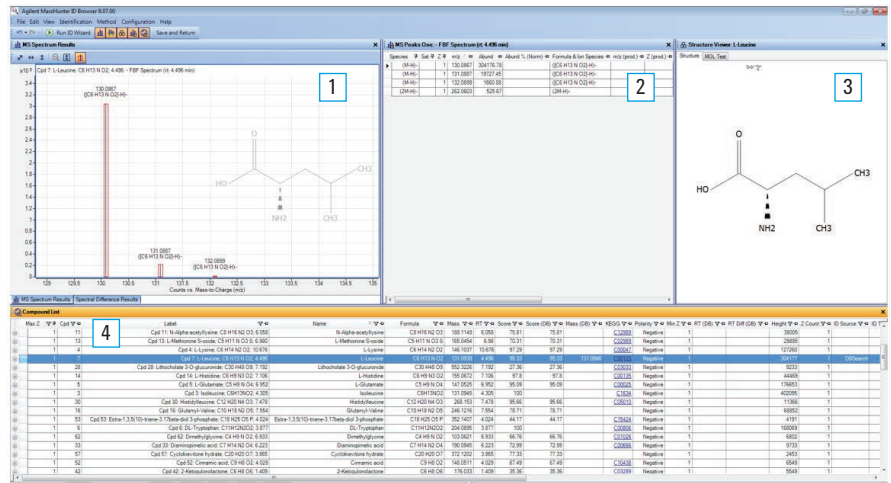


图 5. 在负离子模式下获得的一组差异特征的 ID 浏览器结果。(1) 质谱图结果, (2) 同位素峰的详细信息, (3) 标注化合物的结构, 和 (4) 化合物列表

## 化合物鉴定和软件辅助的结构解析

为了验证初始标注，我们采用 6550 Q-TOF LC/MS 对目标差异特征进行了靶向 MS/MS 实验。采用两种方法进行化合物鉴定和结构解析。首先，采用 MassHunter 定性分析软件中的 *Find by Targeted MSMS* 工具处理 LC/MS/MS 数据。然后将所得的 MS/MS 谱图与包含精确质量 MS/MS 信息的 Agilent-Metlin 代谢物数据库进行比对。图 6 和图 7 展示了一个被鉴定为烟酸的差异特征，并且可信度极高。由镜像图谱（中图）可以看出，无论是  $m/z$  值还是母离子和碎片离子的丰度，采集到的样品谱图（上图）与谱库标准谱图（下图）的匹配度都极高，其中正离子数据的反向匹配分值为 97，负离子数据的反相匹配分值为 99。

当采集到的 MS/MS 谱图与 Agilent-Metlin MS/MS 谱库匹配度很低或无匹配时，我们可采用另一种使用 MSC 软件的方法。MSC<sup>3</sup> 利用系统性键断裂方法提供预测分子结构的理论碎片质量。然后将这些理论碎片质量与观察所得的目标化合物的精确质量碎片离子关联起来。MSC 根据每个碎片离子的单独得分计算总体相关性得分（拟合优度）；然后给出一个或多个可能结构的建议。

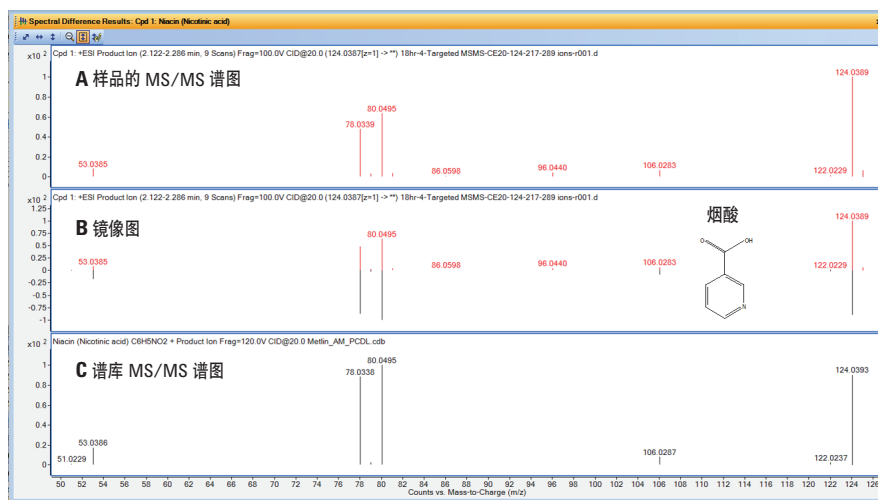


图 6. 使用 MS/MS 谱图匹配所得的化合物结构鉴定结果，其中正 ESI 模式下烟酸的反向匹配分值为 97。反向匹配分值反映了谱库中标准品的离子峰与样品谱图中离子峰的匹配度。(A) 样品 MS/MS 谱图，(B) 镜像图，(C) 谱库 MS/MS 谱图。采集 MS/MS 谱图所用的碰撞能量 (CE) 为 20 eV

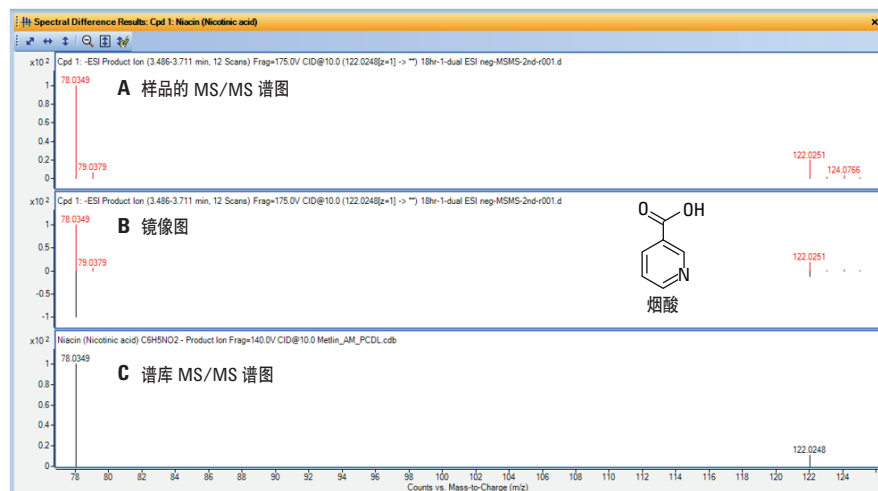


图 7. 使用 MS/MS 谱图匹配所得的化合物结构鉴定结果，其中负 ESI 模式下烟酸的反向匹配分值为 99。反向匹配分值反映了谱库中标准品的离子峰与样品谱图中离子峰的匹配度。(A) 样品 MS/MS 谱图，(B) 镜像图，(C) 谱库 MS/MS 谱图。采集 MS/MS 谱图所用的碰撞能量 (CE) 为 10 eV

在 MSC B.07.00 中批量处理多个数据文件，即分别在 Agilent-Metlin、由 Pathway to PCDL 生成的 KEGG 和 Biocyc，以及基于网络的 PubChem 和 ChemSpider 数据库中搜索这些数据。可搜索多个数据库的能力提高了找到匹配结构的几率，但会降低选择性。如图 8 和图 9 所示，首先搜索 Agilent-Metlin 数据库，获得了  $m/z$  为 187.1114、203.1041 和 307.1512 的 3 种化合物的匹配结构（图 8）。对于  $m/z$  为 187.1114 的离子，为其找到了 9 种可能的匹配结构。采用相同的 MSC 筛查标准在 PubChem 数据库中进行搜索，所有 6 种化合物均获得了匹配结构（图 9），而  $m/z$  为 187.1114 的离子获得了 300 种可能的匹配结构。 $m/z$  为 187.1114 的离子的 MS/MS 谱图结构解析如图 10 所示。所有碎片离子都被标记了 MSC 认为最合理的碎片结构。采用这种方法，我们可以快速解析所得的 MS/MS 谱图，从而对这些显著特征进行标注。

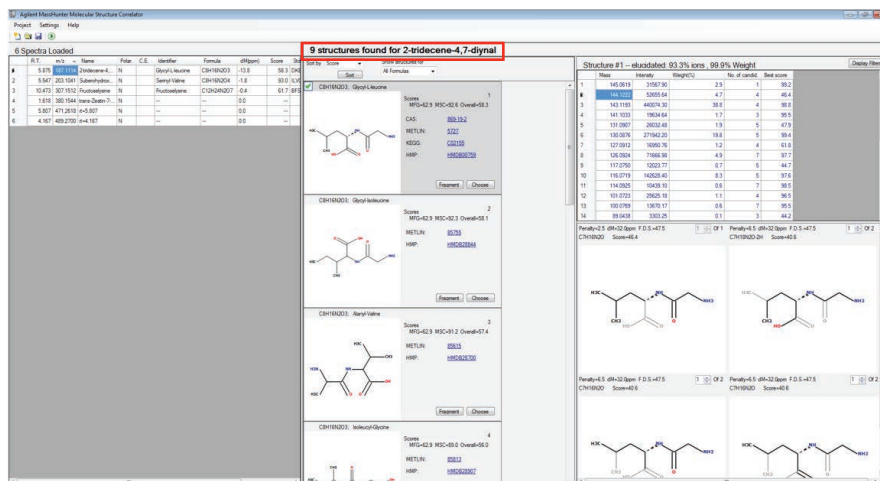


图 8. 搜索 Agilent-Metlin 数据库所得的 MSC 结果

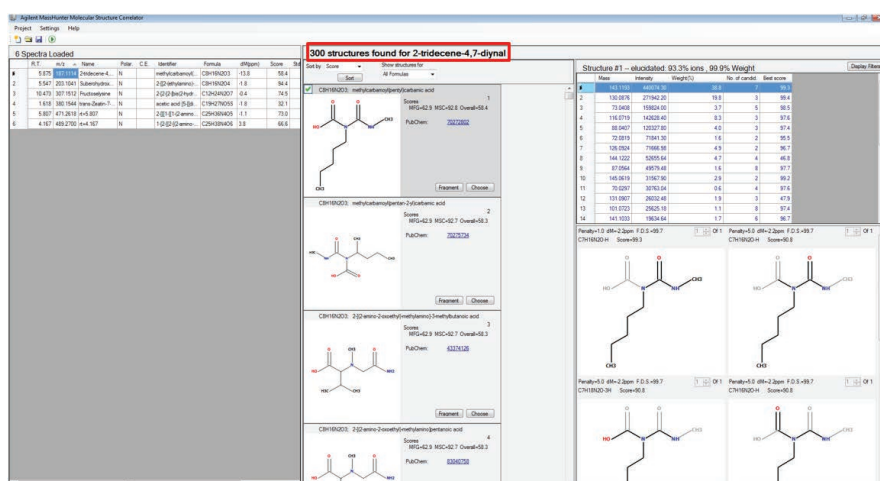


图 9. 搜索基于网络的 PubChem 数据库所得的 MSC 结果

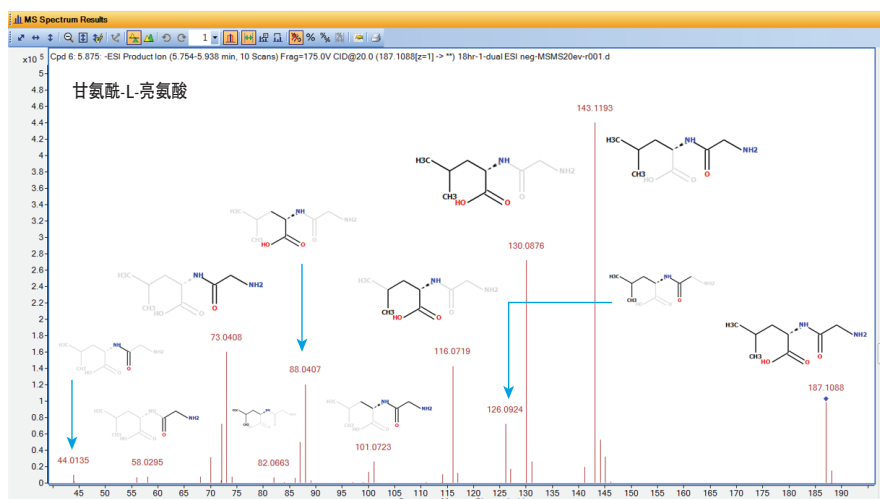


图 10. 采用 MSC 和 Agilent-Metlin 数据库所得的  $m/z$  为 187.1088 的离子的 MS/MS 谱图结构解析结果

## 结论

本文展示了一种简单易用的基于 Q-TOF LC/MS 的工作流程，用于细菌的发现代谢组学研究。安捷伦软件工具（Profiler、MPP 和 MSC）实现了高效、高质量的特征提取、统计学分析、标注和鉴定。本研究的结果揭示了细菌早稳定和晚稳定期代谢物的一些有趣的变化。

## 参考文献

1. Monod, J. The Growth of Bacterial Cultures. *Annual Review of Microbiology* **1949**, *3*, pp 371-394.
2. MassHunter Profinder: 用于质谱数据高质量特征提取的批处理软件, *安捷伦科技公司技术概览*, 出版号 5991-3947CHCN, **2014 年**
3. S. Joseph, Y. Dai, 安捷伦 Q-TOF LC/MS 和先进的 MassHunter 数据处理软件结合使用进行药物杂质的鉴定和分析, *安捷伦科技公司应用简报*, 出版号 5991-1375CHCN, **2012 年**

查找当地的安捷伦客户中心:

[www.agilent.com/chem/contactus-cn](http://www.agilent.com/chem/contactus-cn)

免费专线:

**800-820-3278, 400-820-3278 (手机用户)**

联系我们:

[LSCA-China\\_800@agilent.com](mailto:LSCA-China_800@agilent.com)

在线询价:

[www.agilent.com/chem/erfq-cn](http://www.agilent.com/chem/erfq-cn)

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

本文中的信息、说明和指标如有变更, 恕不另行通知。

© 安捷伦科技(中国)有限公司, 2015

2015 年 4 月 28 日, 中国印刷

5991-5706CHCN



**Agilent Technologies**