



# 基于核糖体蛋白质标志物及基质辅助激光解吸电离飞行时间质谱技术快速鉴定糖丝菌属放线菌

孙伟\*, 许杰, 张薇, 罗倩, 台萃

上海交通大学生命科学技术学院 仪器共享与技术服务平台, 上海 200240

孙伟, 许杰, 张薇, 罗倩, 台萃. 基于核糖体蛋白质标志物及基质辅助激光解吸电离飞行时间质谱技术快速鉴定糖丝菌属放线菌[J]. 微生物学报, 2023, 63(12): 4800-4813.

SUN Wei, XU Jie, ZHANG Wei, LUO Qian, TAI Cui. Rapid identification of *Saccharothrix* spp. based on ribosomal protein biomarkers and matrix-assisted laser desorption/ionization-time of flight mass spectrometry[J]. Acta Microbiologica Sinica, 2023, 63(12): 4800-4813.

**摘要:** 【目的】糖丝菌属(*Saccharothrix*)是一类丝状稀有放线菌, 在生物医药、工业酶制剂和环境修复等领域展现出应用价值。本研究尝试建立以核糖体蛋白质为标志物, 利用基质辅助激光解吸电离飞行时间质谱(matrix-assisted laser desorption/ionization-time of flight mass spectrometry, MALDI-TOF MS)技术鉴定糖丝菌属放线菌的方法。【方法】检索基因组数据库, 提取糖丝菌属测序菌株 15 种核糖体蛋白质的序列并计算理论分子量; 通过分子量比对分析糖丝菌属不同菌种之间及其模式菌株与邻近属菌种模式菌株之间 15 种核糖体蛋白质的匹配度, 提出鉴定至菌种及属的核糖体蛋白质匹配数标准; 选取目标属和非目标属菌种进行 MALDI-TOF MS 测试和分析并修正鉴定标准。【结果】将待测菌株的 MALDI-TOF 质谱峰与糖丝菌属各菌种模式菌株的 15 种核糖体蛋白质分别匹配, 通过最大匹配数、质谱峰强度模式及特征质谱峰鉴定至属或种。【结论】本研究建立了基于 15 种核糖体蛋白质标志物及 MALDI-TOF MS 技术鉴定糖丝菌属放线菌的方法, 可用于定向筛选和快速鉴定糖丝菌。

**关键词:** 基质辅助激光解吸电离飞行时间质谱; 核糖体蛋白质; 糖丝菌属; 放线菌; 鉴定

资助项目: 国家自然科学基金(81973230); 上海交通大学决策咨询课题(JCZXSJB-13)

This work was supported by the National Natural Science Foundation of China (81973230) and the Decision-making Consultation Project of Shanghai Jiao Tong University (JCZXSJB-13).

\*Corresponding author. Tel: +86-21-34207024, E-mail: bacterior@shtu.edu.cn

Received: 2023-03-06; Accepted: 2023-08-25; Published online: 2023-08-30

# Rapid identification of *Saccharothrix* spp. based on ribosomal protein biomarkers and matrix-assisted laser desorption/ionization-time of flight mass spectrometry

SUN Wei\*, XU Jie, ZHANG Wei, LUO Qian, TAI Cui

Core Facility and Technical Service Center, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

**Abstract:** [Objective] *Saccharothrix* is a genus of rare filamentous actinomycetes showing great application values in the biopharmaceutical field, industrial enzyme preparations, and environmental remediation. This study aims to establish a method to identify *Saccharothrix* spp. based on ribosomal protein biomarkers and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). [Methods] We searched two public genome databases to obtain the sequences of 15 ribosomal proteins from the genome-sequenced strains of *Saccharothrix* and calculated the theoretical molecular weights. We then compared the molecular weights to investigate the matching of 15 ribosomal proteins between different species of *Saccharothrix* and that between the type strains of *Saccharothrix* and its closely related three genera (*Umezawaea*, *Actinosynnema*, and *Lentzea*). The criteria for identifying *Saccharothrix* at the species and genus levels were proposed based on the number of ribosomal protein matches. Subsequently, the identification criteria were tested by MALDI-TOF MS with target and non-target organisms and improved. [Results] MALDI-TOF MS peaks from an unknown organism were compared with the weight list of 15 ribosomal protein biomarkers from each *Saccharothrix* type strain and the number of peaks that matched each strain's biomarkers was obtained. Based on the maximum matches, peak intensity mode, and characteristic peaks, the organism can be identified at the genus or species level. [Conclusion] This study established a method for identifying *Saccharothrix* spp. based on 15 ribosomal protein biomarkers and MALDI-TOF MS, which can be used in directed screening and rapid identification of *Saccharothrix* spp.

**Keywords:** matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS); ribosomal protein; *Saccharothrix*; actinomycetes; identification

糖丝菌属(*Saccharothrix*)建立于1984年<sup>[1]</sup>, 是一类典型的丝状稀有放线菌, 隶属于放线菌门(*Actinobacteria*)放线菌纲(*Actinobacteria*)假诺卡氏菌目(*Pseudonocardiales*)假诺卡氏菌科(*Pseudonocardiaceae*)。原核生物标准命名名录(list of prokaryotic names with standing in nomenclature, <https://lpsn.dsmz.de/>)显示, 糖丝菌

属现有23个有效描述种。糖丝菌能够产生种类多样的生物活性代谢产物<sup>[2-5]</sup>, 可作为新抗生素的筛选资源。此外, 糖丝菌能够产生活性酶, 可用于治疗真菌类感染疾病、降解纤维素和有机污染物<sup>[6-8]</sup>。近年来基于基因组分析发现糖丝菌基因组中含有丰富的次级代谢产物合成基因簇, 具有新骨架、新功能的代谢产物已通过基

因簇克隆得到发现<sup>[9]</sup>。这些研究显示了糖丝菌在生物医药、工业酶制剂和环境修复等领域的应用潜力。糖丝菌在自然界分布广泛,存在很大的新物种资源挖掘空间。目前糖丝菌的分离主要依赖于非靶向培养方法,以高通量测序结果为导向,从环境样品中大规模分离培养细菌或者放线菌,通过菌种鉴定发现其中的糖丝菌。传统的菌种鉴定方法中,16S rRNA 基因序列分析法应用广泛,是细菌鉴定的金标准,但在大规模菌种鉴定时工作量大,鉴定周期长。因此,开发一种简单快速的糖丝菌鉴定方法有助于促进这类放线菌资源的开发利用。

基质辅助激光解吸电离飞行时间质谱(matrix-assisted laser desorption/ionization-time of flight mass spectrometry, MALDI-TOF MS)法鉴定微生物具有简单、快速、高通量等优势,已广泛应用于临床诊断中<sup>[10]</sup>。这一方法建立在测定微生物稳定表达的高丰度蛋白质的分子量基础上。处于对数生长期的细菌核糖体蛋白质丰度高,其菌体蛋白的质谱图中以核糖体蛋白质信号为主,可降低其他蛋白质,尤其是分子量接近的蛋白质的影响<sup>[11-12]</sup>。MALDI-TOF MS 法的鉴定准确性依赖于图谱数据库的完善性与可靠性。目前,商业数据库以收录病原微生物为主,收录非致病的环境微生物较少,比如布鲁克公司的 MALDI Biotyper 数据库(Bruker MBT DB7854)仅收录一种糖丝菌(*Saccharothrix mutabilis*)的 2 个亚种,无法满足对种类多样的糖丝菌属菌种的鉴定需求。按布鲁克公司提供的标准方法构建图谱数据库,需进行菌株收集、培养、前处理、图谱采集、质量控制和方法验证,操作繁琐,对技术要求严格,制约了 MALDI-TOF MS 法在环境微生物鉴定中的应用<sup>[13]</sup>。

在前期研究中提出了基于 15 种核糖体蛋白质标志物及 MALDI-TOF MS 技术鉴定放线

菌的设想<sup>[14]</sup>,之后通过对伦茨菌属放线菌的研究证实了其可行性<sup>[15]</sup>。这种基于核糖体蛋白质标志物的方法无需收集目标类群的菌种构建图谱数据库,而是利用基因组学技术构建目标类群菌种 15 种核糖体蛋白质的理论分子量数据库。基于核糖体蛋白质标志物构建数据库的方法操作简单,成本低廉,能够扩展 MALDI-TOF MS 法鉴定微生物的应用范围。为促进糖丝菌属菌种资源的挖掘,同时提供更多分析案例用于放线菌鉴定方法的开发,本研究尝试建立基于核糖体蛋白质标志物及 MALDI-TOF MS 技术鉴定糖丝菌属放线菌的方法,并展望其应用前景。

## 1 材料与方 法

### 1.1 糖丝菌属菌种核糖体蛋白质序列提取、分析及分子量计算

以“*Saccharothrix*”为检索词搜索美国国家生物技术信息中心(National Center for Biotechnology Information, NCBI)基因组数据库(<https://www.ncbi.nlm.nih.gov/genome/>)和美国联合基因组研究中心(Joint Genomics Institute, JGI)基因组数据库(<https://genome.jgi.doe.gov/portal/>),获得糖丝菌属测序菌种的信息,选取模式菌株和测序质量好的非模式菌株进行分析。非模式菌株的基因组序列提交至模式菌株基因组服务器(type strain genome server, <https://tygs.dsmz.de/>)进行鉴定<sup>[16]</sup>。

按前期研究所用方法提取各菌株 15 种核糖体蛋白质的序列,进行各种核糖体蛋白质序列的比对、分析和系统进化树构建以及核糖体蛋白质的理论分子量计算<sup>[15]</sup>。

### 1.2 糖丝菌属菌种间的核糖体蛋白质匹配分析

将各菌种代表菌株 15 种核糖体蛋白质的分子量两两比对,核糖体蛋白质间的相对分子

量之差小于  $500 \times 10^{-6}$  (线性模式的质量准确度范围)即认为二者匹配。统计各菌株间的核糖体蛋白质匹配数, 评估以 15 种核糖体蛋白质的组合区分菌种的可行性。基于最大匹配数建立鉴定至菌种的标准。

### 1.3 糖丝菌属菌种和邻近属菌种模式菌株间的核糖体蛋白质匹配分析

梅泽宾夫氏菌(*Umezawaea*)、束丝放线菌(*Actinosynnema*)和伦茨菌(*Lentzea*)是亲缘关系最接近糖丝菌的放线菌<sup>[17]</sup>, 区分糖丝菌属与这 3 个邻近属的菌种是准确鉴定至属的前提。按 1.1 所述方法提取这 3 个属菌种模式菌株 15 种核糖体蛋白质的序列并计算分子量, 与糖丝菌属各菌种模式菌株 15 种核糖体蛋白质的分子量依次比对, 统计核糖体蛋白质匹配数, 基于最大匹配数建立鉴定至糖丝菌属的标准。

### 1.4 糖丝菌属测试菌株

为测试鉴定标准对目标属菌种的鉴定结果, 收集 4 个糖丝菌属菌株进行测试和分析, 包括:*S. australiensis* CGMCC 4.1355<sup>T</sup>、*S. carnea* CGMCC 4.7097<sup>T</sup> 和 *S. texasensis* CGMCC 4.1713<sup>T</sup> (购自中国普通微生物菌种保藏管理中心); *Saccharothrix* sp. CCTCC AB 2015033 (购自中国典型培养物保藏中心)。取 28 °C 培养 3–4 d 的测试样本, 按前期研究所用方法提取菌株蛋白<sup>[14]</sup>。此外, MALDI Biotyper 数据库收录了 2 个糖丝菌属菌株 *S. mutabilis* subsp. *capreolus* HKI 264 HKJ 和 *S. mutabilis* subsp. *mutabilis* HKI 9771 HKJ, 调取其质谱峰列表用于后续分析。

### 1.5 菌株 CCTCC AB 2015033 的基因组测序、组装及鉴定

将菌株 *Saccharothrix* sp. CCTCC AB 2015033 接种于 ISP2 液体培养基中, 于 28 °C、180 r/min 转速下振荡培养 96 h。取 10 mL 培养液, 7 000×g 离心 10 min, 收集菌体送至苏州金唯智生物科

技有限公司, 通过 Illumina NovaSeq 6000 测序平台进行基因组框架图测序。组装后的基因组序列提交至模式菌株基因组服务器进行比对。提取基因组中的 16S rRNA 基因序列, 在 EzBioCloud (<https://www.ezbiocloud.net/>) 数据库中进行序列比对。

### 1.6 MALDI-TOF MS 测试、MALDI 质谱峰匹配及菌株鉴定

基质辅助激光解吸高分辨飞行时间质谱成像系统购自布鲁克公司。按前期研究所用方法采集测试样本的质谱图并进行质谱峰匹配<sup>[14]</sup>。相对质量误差设为小于  $500 \times 10^{-6}$ 。统计测试菌株匹配各模式菌株核糖体蛋白质的峰数量, 获得最大匹配数及对应的模式菌株, 按鉴定至糖丝菌属和种的标准鉴定并评估准确性。

按前期研究所用方法分析糖丝菌属菌株核糖体蛋白质的 MALDI 质谱峰强度模式<sup>[15]</sup>。

### 1.7 糖丝菌属外菌种的 MALDI 质谱峰匹配

为测试鉴定标准对非目标属菌种的鉴定结果, 从 MALDI Biotyper 数据库选取糖丝菌属外的假诺卡氏菌目的 7 个菌种及放线菌纲其他目的 73 个菌种, 调取其质谱峰列表进行分析。相对质量误差设为小于  $500 \times 10^{-6}$ 。按 1.6 所述方法对测试菌株依次比对和统计, 获得最大匹配数, 按鉴定至糖丝菌属的标准鉴定并评估准确性。

## 2 结果与分析

### 2.1 糖丝菌属菌种核糖体蛋白质序列提取、分析及分子量计算

截至 2023 年 1 月, 从 NCBI 和 JGI 基因组数据库检索到糖丝菌属 16 个菌种的模式菌株以及鉴定到属的 6 个菌株。基于基因组比对的鉴定结果见表 1。菌株 *Saccharothrix* sp. 6-C 鉴定为 *S. texasensis*, 另外 5 个菌株均未归入 16 个

表 1 六个糖丝菌属菌株的基因组鉴定结果

Table 1 Genome-based identification of six *Saccharothrix* strains

Query	Closest type strain	DDH value (%)	Identification result
<i>Saccharothrix</i> sp. ALI-22-I	<i>S. deserti</i> BMP B8144 <sup>T</sup>	44.1	Potential new species
<i>Saccharothrix</i> sp. CB00851	<i>S. carnea</i> CGMCC 4.7097 <sup>T</sup>	60.0	Potential new species
<i>Saccharothrix</i> sp. NRRL B-16314	<i>S. ecbatanensis</i> DSM 45486 <sup>T</sup>	45.6	Potential new species
<i>Saccharothrix</i> sp. NRRL B-16348	<i>S. texasensis</i> DSM 44231 <sup>T</sup>	43.5	Potential new species
<i>Saccharothrix</i> sp. S26	<i>S. saharensis</i> DSM 45456 <sup>T</sup>	41.7	Potential new species
<i>Saccharothrix</i> sp. 6-C	<i>S. texasensis</i> DSM 44231 <sup>T</sup>	86.6	<i>S. texasensis</i>

模式菌株所属菌种。从 22 个菌株的基因组均注释出 15 种目标核糖体蛋白质，其理论分子量见表 2。

序列分析发现 15 种核糖体蛋白质的序列保守性差异显著，大致分 3 类：高度保守序列(L36)，该序列在分析菌株中一致；中度保守序列(L33、S19、S18、S14-Z、L34、L28、L30 和 L32)，一般有 1–3 个位点发生氨基酸替换；高度变化序列(L31、S20、S15、L27、L35 和 L29)，一般有 3 个以上位点发生氨基酸改变，主要是氨基酸替换，有时是氨基酸插入或缺失，数据已提交国家微生物科学数据中心，编号：NMDCX0000220。这些序列的组合为菌种区分和鉴定奠定了基础。

## 2.2 糖丝菌属菌种间的核糖体蛋白质匹配分析

对代表糖丝菌属不同菌种的 21 个菌株进行核糖体蛋白质匹配分析，21 个菌株间 15 种核糖体蛋白质的匹配结果见表 3，详细匹配信息已提交国家微生物科学数据中心，编号：NMDCX0000220。

表 3 显示 21 个菌株间的核糖体蛋白质最大匹配数在 5–11 之间，表明 15 种核糖体蛋白质的组合理论上可以区分糖丝菌属不同菌种。基于最大匹配数将鉴定至菌种的标准设为核糖体蛋白质匹配数 $\geq 12$ 。如测试菌株与表 2 中某模式菌株的核糖体蛋白质匹配数 $\geq 12$ ，则鉴定为该模

式菌株所属菌种。

菌株 *Saccharothrix* sp. 6-C 与 *S. texasensis* DSM 44231<sup>T</sup> 15 种核糖体蛋白质的理论分子量均一致，匹配数为 15，按上述标准鉴定为 *S. texasensis*，与基于基因组比对的鉴定结果一致。

## 2.3 糖丝菌属菌种与邻近属菌种模式菌株间的核糖体蛋白质匹配分析

截至 2023 年 1 月，从 NCBI 和 JGI 基因组数据库检索到梅泽宾夫氏菌属 2 个、束丝放线菌属 2 个及伦茨菌属 19 个有效描述种，这些菌种模式菌株 15 种核糖体蛋白质的理论分子量数据已提交国家微生物科学数据中心，编号：NMDCX0000220。分子量比对发现糖丝菌与 3 个邻近属菌种间一些核糖体蛋白质的分子量一致，如 L36 (4 403.38 Da)、L33 (6 363.33 Da)、S14-Z (6 831.18 Da)、S18 (8 874.42 Da) 和 S19 (10 507.27 Da) 等。此外，一些核糖体蛋白质间的分子量差值在仪器质量误差范围内，可能产生错配。梅泽宾夫氏菌、束丝放线菌及伦茨菌与糖丝菌属各模式菌株间的核糖体蛋白质匹配结果见表 4，详细匹配信息已提交国家微生物科学数据中心，编号：NMDCX0000220。表 4 显示糖丝菌与 3 个邻近属菌种模式菌株间的核糖体蛋白质最大匹配数均为 6。因此，将鉴定至糖丝菌属的标准设为核糖体蛋白质匹配数 $\geq 7$ 。如测试菌株与表 2 中某模式菌株的核糖体蛋白质匹配数 $\geq 7$ ，则鉴定为糖丝菌属。

表 2 22 个糖丝菌属菌株 15 种核糖体蛋白质的理论分子量表

Table 2 The list of the masses predicted for 15 ribosomal proteins from 22 *Saccharothrix* strains

No. of strains	L36 (Da)	L34 (Da)	L33 (Da)	L32 (Da)	L30 (Da)	S14-Z (Da)	L28 (Da)	L35 (Da)	L31 (Da)	S18 (Da)	L29 (Da)	L27 (Da)	S20 (Da)	S15 (Da)	S19 (Da)
01 <i>S. algeriensis</i> DSM 44581 <sup>T</sup>	4 403.38	5 349.28	6 389.37	6 587.61	6 804.87	6 872.23	6 723.74	6 936.24	7 828.92	8 883.39	9 401.60	9 099.39	9 368.92	10 233.82	10 507.27
02 <i>S. australiensis</i> DSM 43800 <sup>T</sup>	4 403.38	5 349.28	6 363.33	6 587.61	6 848.92	6 858.20	6 750.77	6 946.27	7 827.93	8 883.39	9 359.52	9 143.44	9 322.87	10 233.82	10 523.27
03 <i>S. carnea</i> CGMCC 4.709 <sup>T</sup>	4 403.38	5 319.25	6 363.33	6 570.62	6 868.96	6 831.18	6 723.74	6 934.20	7 696.80	8 874.42	9 116.26	9 099.39	9 396.97	10 217.82	10 507.27
04 <i>S. coeruleofusca</i> DSM 43679 <sup>T</sup>	4 403.38	5 349.23	6 333.35	6 613.65	6 818.90	6 872.23	6 723.74	6 924.16	7 714.77	8 860.35	9 093.22	8 997.30	9 221.76	10 260.84	10 507.27
05 <i>S. deserti</i> BMP B8144 <sup>T</sup>	4 403.38	5 320.23	6 363.33	6 570.62	6 818.90	6 831.18	6 704.70	6 934.20	7 708.81	8 874.42	9 116.26	9 087.37	9 405.02	10 187.84	10 507.27
06 <i>S. ecbatanensis</i> DSM 45486 <sup>T</sup>	4 403.38	5 320.23	6 363.33	6 658.69	6 823.92	6 831.18	6 704.70	6 962.26	7 753.85	8 874.42	9 116.26	9 103.37	9 336.86	10 163.83	10 493.24
07 <i>S. espanaensis</i> DSM 44229 <sup>T</sup>	4 403.38	5 349.23	6 363.33	6 557.59	6 804.87	6 817.15	6 706.76	6 966.26	7 813.91	8 883.39	9 132.26	9 127.44	9 334.89	10 260.84	10 507.27
08 <i>S. longispora</i> DSM 43749 <sup>T</sup>	4 403.38	5 319.25	6 363.33	6 600.65	6 809.89	6 831.18	6 731.72	6 991.26	7 684.74	8 874.42	9 146.28	9 156.48	9 350.89	10 204.78	10 507.27
09 <i>S. luteola</i> NEAU-S10 <sup>T</sup>	4 403.38	5 319.25	6 363.33	6 628.66	6 868.96	6 858.20	6 704.70	6 918.20	7 716.80	8 874.42	9 116.26	9 087.37	9 310.84	10 217.82	10 465.19
10 <i>S. obliqua</i> SC076 <sup>T</sup>	4 403.38	5 319.25	6 389.37	6 547.55	6 834.90	6 831.18	6 750.77	6 934.20	7 809.87	8 902.43	9 116.26	9 207.48	9 334.93	10 233.82	10 507.27
11 <i>S. saharensis</i> DSM 45456 <sup>T</sup>	4 403.38	5 319.25	6 363.33	6 570.62	6 809.89	6 872.23	6 723.74	6 934.20	7 696.80	8 874.42	9 144.31	9 099.39	9 340.86	10 187.84	10 507.27
12 <i>S. syringae</i> NRRL B-16468 <sup>T</sup>	4 403.38	5 349.28	6 363.33	6 598.64	6 818.90	6 831.18	6 760.81	6 896.11	7 700.74	8 902.43	9 116.26	9 025.35	9 253.82	10 231.85	10 507.27
13 <i>S. tamarrassetensis</i> CECT 8640 <sup>T</sup>	4 403.38	5 349.23	6 363.33	6 557.59	6 818.90	6 831.18	6 750.77	6 936.24	7 800.86	8 883.39	9 403.57	9 143.44	9 350.89	10 263.85	10 507.27
14 <i>S. texasensis</i> DSM 44231 <sup>T</sup>	4 403.38	5 292.18	6 363.33	6 614.64	6 868.96	6 858.20	6 704.70	6 918.20	7 700.80	8 874.42	9 116.26	9 089.35	9 326.84	10 146.74	10 493.24
15 <i>S. variisporae</i> DSM 43911 <sup>T</sup>	4 403.38	5 349.28	6 389.37	6 557.59	6 818.90	6 888.27	6 750.77	6 948.23	7 698.77	8 902.43	9 003.10	9 073.35	9 338.93	10 260.89	10 507.27
16 <i>S. violaceirubra</i> DSM 45084 <sup>T</sup>	4 403.38	5 320.23	6 335.32	6 619.61	6 862.91	6 831.18	6 760.81	6 904.18	7 700.74	8 828.29	8 973.07	9 059.32	9 336.90	10 217.86	10 523.27
17 <i>Saccharothrix</i> sp. ALI-22-I	4 403.38	5 319.21	6 363.33	6 557.59	6 818.90	6 831.18	6 704.70	6 934.20	7 730.83	8 874.42	9 172.37	9 087.37	9 378.94	10 217.86	10 493.24

(待续)

(续表 2)

No. of strains	L36 (Da)	L34 (Da)	L33 (Da)	L32 (Da)	L30 (Da)	S14-Z (Da)	L28 (Da)	L35 (Da)	L31 (Da)	S18 (Da)	L29 (Da)	L27 (Da)	S20 (Da)	S15 (Da)	S19 (Da)
18 <i>Saccharothrix</i> sp. CB00851	4 403.38	5 319.25	6 363.33	6 570.62	6 818.90	6 858.20	6 723.74	6 934.20	7 757.86	8 874.42	9 116.26	9 099.39	9 396.97	10 231.85	10 507.27
19 <i>Saccharothrix</i> sp. NRRL B-16314	4 403.38	5 320.23	6 363.33	6 658.69	6 859.95	6 858.20	6 704.70	6 948.23	7 771.88	8 874.42	9 116.26	9 089.35	9 334.88	10 163.83	10 493.24
20 <i>Saccharothrix</i> sp. NRRL B-16348	4 403.38	5 292.18	6 363.33	6 628.66	6 868.96	6 831.18	6 704.70	6 934.20	7 700.80	8 874.42	9 116.26	9 103.37	9 338.89	10 217.86	10 507.27
21 <i>Saccharothrix</i> sp. S26	4 403.38	5 319.25	6 363.33	6 598.64	6 818.90	6 858.20	6 723.74	6 934.20	7 740.77	8 902.43	9 088.25	9 099.39	9 350.89	10 188.82	10 507.27
22 <i>Saccharothrix</i> sp. 6-C	4 403.38	5 292.18	6 363.33	6 614.64	6 868.96	6 858.20	6 704.70	6 918.20	7 700.80	8 874.42	9 116.26	9 089.35	9 326.84	10 146.74	10 493.24

表 3 21 个糖丝菌属菌株间的核糖体蛋白质匹配数

Table 3 Count of ribosomal proteins matched to 15 ribosomal proteins from 21 *Saccharothrix* strains

No. of strains	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21
01 <i>S. algeriensis</i> DSM 44581 <sup>T</sup>	6	7	5	4	2	5	2	2	2	5	6	4	6	2	4	1	2	7	1	5	5
02 <i>S. australiensis</i> DSM 43800 <sup>T</sup>	6	2	2	2	2	4	3	3	3	3	3	4	6	4	4	2	2	4	5	2	3
03 <i>S. carnea</i> CGMCC 4.7097 <sup>T</sup>	7	2	4	9	7	3	6	7	6	11	5	5	5	3	4	7	11	5	10	7	7
04 <i>S. coeruleofusca</i> DSM 43679 <sup>T</sup>	5	2	4	3	1	5	2	3	2	4	4	4	5	4	5	2	2	4	2	3	4
05 <i>S. deserti</i> BMP B8144 <sup>T</sup>	4	2	9	3	7	5	6	7	6	8	6	7	6	3	3	9	9	7	8	8	8
06 <i>S. ecbatanensis</i> DSM 45486 <sup>T</sup>	2	2	7	1	7	4	5	6	5	6	4	3	6	2	4	7	6	10	8	4	4
07 <i>S. espanaensis</i> DSM 44229 <sup>T</sup>	5	4	3	5	5	4	3	3	3	3	5	8	3	7	2	5	4	4	5	4	4
08 <i>S. longispora</i> DSM 43749 <sup>T</sup>	2	3	6	2	6	5	3	4	4	4	7	5	6	3	2	3	5	5	4	5	6
09 <i>S. luteola</i> NEAU-S10 <sup>T</sup>	2	3	7	3	7	6	3	4	3	5	3	2	9	1	3	7	6	9	8	5	5
10 <i>S. obliqua</i> SC076 <sup>T</sup>	5	3	6	2	6	5	3	4	3	4	6	5	2	6	4	4	6	4	6	5	5
11 <i>S. saharensis</i> DSM 45456 <sup>T</sup>	6	3	11	4	8	6	3	7	5	4	3	5	4	4	4	3	5	9	4	8	8
12 <i>S. syringae</i> NRRL B-16468 <sup>T</sup>	4	4	5	4	6	4	5	5	3	6	3	6	4	6	4	4	4	6	3	6	6
13 <i>S. tamanrassetensis</i> CECT 8640 <sup>T</sup>	6	6	5	5	7	3	8	6	2	5	5	6	2	7	2	6	5	2	5	6	6
14 <i>S. texasensis</i> DSM 44231 <sup>T</sup>	2	4	5	4	6	6	3	3	9	2	4	4	2	2	2	6	5	9	8	4	4
15 <i>S. variispora</i> DSM 43911 <sup>T</sup>	4	4	3	5	3	2	7	2	1	6	4	6	7	2	3	3	3	3	3	4	4
16 <i>S. violaceirubra</i> DSM 45084 <sup>T</sup>	1	2	4	2	3	4	2	3	3	4	3	4	2	2	3	4	2	4	5	2	2
17 <i>Saccharothrix</i> sp. ALI-22-I	2	2	7	2	9	7	5	5	7	4	5	4	6	6	3	4	6	7	7	6	6
18 <i>Saccharothrix</i> sp. CB00851	7	4	11	4	9	6	4	5	6	6	9	6	5	5	3	2	6	7	7	9	9
19 <i>Saccharothrix</i> sp. NRRL B-16314	1	4	5	2	7	10	4	4	8	4	4	3	2	8	3	4	7	6	6	5	5
20 <i>Saccharothrix</i> sp. NRRL B-16348	5	2	10	3	8	8	5	5	8	6	8	6	5	8	4	5	7	7	6	5	5
21 <i>Saccharothrix</i> sp. S26	5	3	7	4	8	4	4	6	5	5	8	6	6	4	4	2	6	9	6	5	5

Cells that have the maximum value in each row are colored gray.

表 4 梅泽宾夫氏菌、束丝放线菌及伦茨菌与糖丝菌属模式菌株间的核糖体蛋白质匹配数

Table 4 Count of ribosomal proteins from *Umezawaea*, *Actinosynnema*, and *Lentzea* strains matched to 15 ribosomal proteins from 16 *Saccharothrix* type strains

Organism	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
<i>Umezawaea endophytica</i> DSM 103496 <sup>T</sup>	3	3	4	2	4	3	6	4	2	3	3	4	5	2	2	2
<i>Umezawaea tangerina</i> DSM 44720 <sup>T</sup>	3	4	4	2	4	3	6	4	2	3	3	4	5	3	2	2
<i>Actinosynnema mirum</i> DSM 43827 <sup>T</sup>	3	3	5	3	6	3	3	4	3	5	4	3	6	2	5	3
<i>Actinosynnema pretiosum</i> DSM 44131 <sup>T</sup>	3	3	5	3	6	3	3	4	3	5	4	3	6	2	5	3
<i>Lentzea aerocolonigenes</i> NBRC 13195 <sup>T</sup>	3	2	4	3	3	2	3	4	3	3	4	4	3	3	3	2
<i>Lentzea alba</i> NEAU-D13 <sup>T</sup>	3	1	3	3	2	2	2	2	2	3	3	2	2	2	2	2
<i>Lentzea albida</i> DSM 44437 <sup>T</sup>	4	1	4	4	2	1	2	3	2	2	4	2	2	2	2	1
<i>Lentzea albidocapillata</i> DSM 44073 <sup>T</sup>	3	1	3	3	2	1	3	3	2	2	3	2	2	2	2	1
<i>Lentzea atacamensis</i> DSM 45479 <sup>T</sup>	3	3	3	2	3	2	4	3	2	2	3	3	4	2	3	1
<i>Lentzea californiensis</i> DSM 43393 <sup>T</sup>	3	2	4	3	4	3	4	3	4	2	4	3	3	4	2	1
<i>Lentzea cavernae</i> CGMCC 4.7367 <sup>T</sup>	3	3	4	3	3	3	4	4	3	3	4	3	3	3	2	2
<i>Lentzea flava</i> JCM 3296 <sup>T</sup>	3	3	5	2	5	4	3	5	4	4	5	3	3	3	3	3
<i>Lentzea flaviverrucosa</i> DSM 44664 <sup>T</sup>	3	2	4	3	3	2	3	4	3	2	4	3	3	3	2	1
<i>Lentzea fradiae</i> CGMCC 4.3506 <sup>T</sup>	3	3	5	3	4	3	3	4	4	4	5	4	3	3	3	2
<i>Lentzea guizhouensis</i> DHS C013 <sup>T</sup>	3	1	3	3	2	1	2	3	2	2	3	2	2	2	2	3
<i>Lentzea jiangxiensis</i> CGMCC 4.6609 <sup>T</sup>	3	3	4	3	3	3	4	3	3	3	4	3	3	3	2	1
<i>Lentzea kentuckyensis</i> NRRL B-24416 <sup>T</sup>	3	4	4	3	3	2	3	4	3	2	4	3	3	3	2	3
<i>Lentzea nigeriaca</i> DSM 45680 <sup>T</sup>	2	2	5	2	5	4	3	5	4	3	5	3	3	3	3	3
<i>Lentzea pudingi</i> CGMCC 4.7319 <sup>T</sup>	3	2	4	3	3	3	3	4	3	2	4	3	3	3	2	1
<i>Lentzea terrae</i> NEAU-LZS 42 <sup>T</sup>	3	3	4	3	3	2	3	3	3	3	4	4	3	3	3	1
<i>Lentzea tibetensis</i> FXJ1.1311 <sup>T</sup>	3	2	4	3	3	3	4	4	2	3	6	3	4	3	5	2
<i>Lentzea waywayandensis</i> DSM 44232 <sup>T</sup>	3	3	4	3	3	2	4	4	3	2	4	3	3	3	2	2
<i>Lentzea xinjiangensis</i> CGMCC 4.3525 <sup>T</sup>	3	3	4	3	3	2	3	4	3	2	4	3	3	3	2	2

Cells that have the maximum value are colored gray.

### 2.4 菌株 CCTCC AB 2015033 的基因组组装及鉴定

菌株 CCTCC AB 2015033 的基因组组装后获得 79 个 contigs, 基因组大小为 8 463 194 bp, 平均 GC 含量为 73.51%。序列提交至国家微生物科学数据中心基因组数据库(<https://nmdc.cn/resource/genomics/genome/>), 数据编号为 NMDC60064245。16S rRNA 基因序列比对表明, 其与 *Saccharothrix lopnurensis* YIM LPA2h<sup>T</sup> 的 16S rRNA 基因序列一致性最高, 为 99.93%, 鉴定为糖丝菌属菌种。但 *S. lopnurensis* 尚无基因组序列报道。与已有基因组比对表明, 菌株 CCTCC AB 2015033 与

*S. syringae* NRRL B-16468<sup>T</sup> 的 DNA-DNA 杂交 (DNA-DNA hybridization, DDH) 值最高, 为 32.9%, 符合属内不同种间的 DDH 值范围。

### 2.5 糖丝菌属测试菌株的 MALDI 质谱峰匹配及菌株鉴定

本研究收集 4 个菌株的 MALDI-TOF MS 图谱见图 1, 6 个菌株(含 MALDI Biotyper 数据库中的 2 个)的质谱峰列表已提交国家微生物科学数据中心, 编号: NMDCX0000220, 与糖丝菌属各模式菌株的核糖体蛋白质匹配结果见表 5, 详细匹配信息已提交国家微生物科学数据中心, 编号: NMDCX0000220。



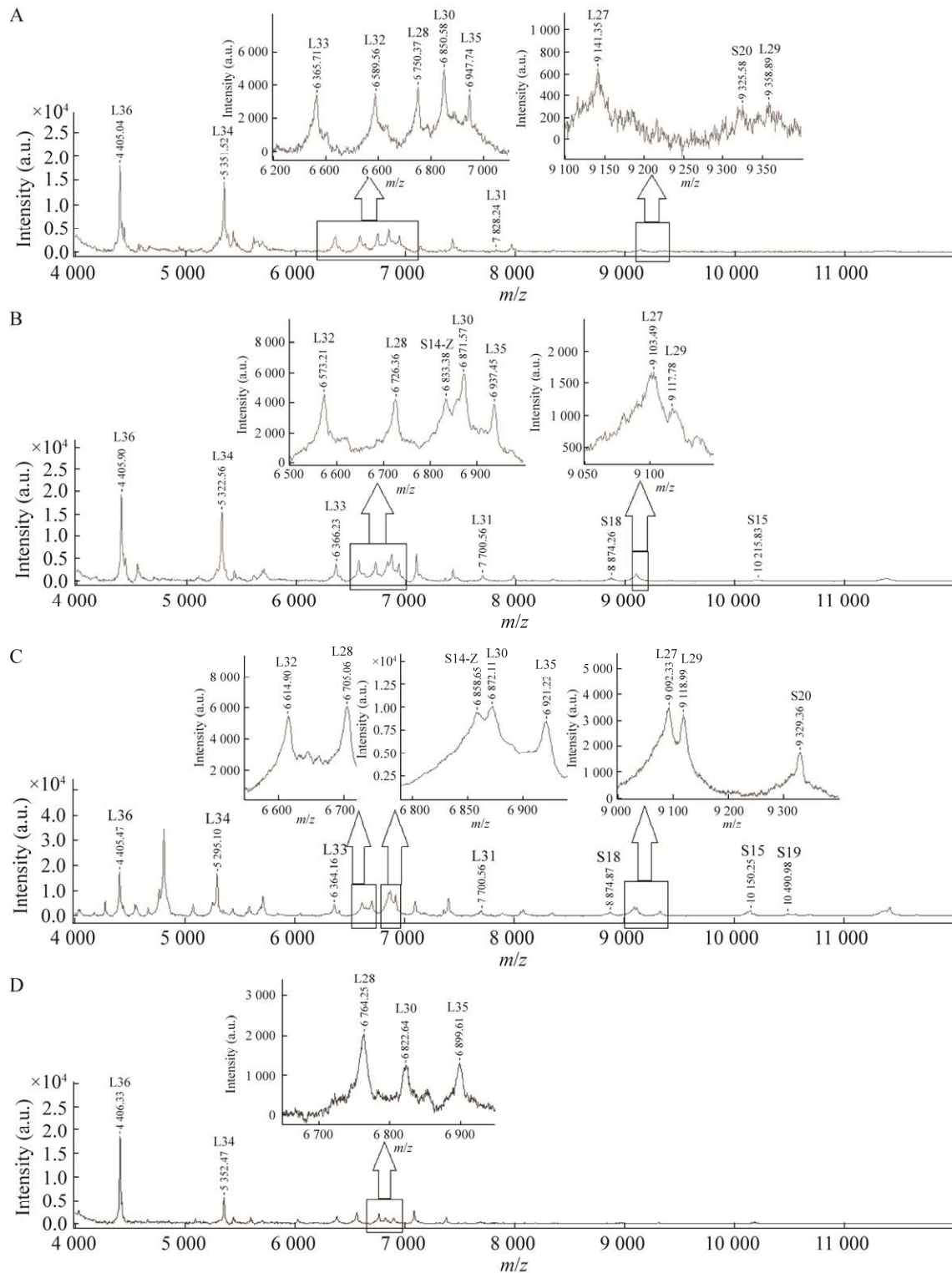


图 1 糖丝菌属测试菌株的 MALDI-TOF MS 图谱

Figure 1 MALDI-TOF MS spectra of *Saccharothrix* test strains. A: *S. australiensis* CGMCC 4.1355<sup>T</sup>. B: *S. carnea* CGMCC 4.7097<sup>T</sup>. C: *S. texasensis* CGMCC 4.1713<sup>T</sup>. D: *Saccharothrix* sp. CCTCC AB 2015033. Peaks that match ribosomal biomarkers are marked.

表 5 糖丝菌属测试菌株的 MALDI 质谱峰与糖丝菌属模式菌株的核糖体蛋白质匹配数

Table 5 Count of MALDI spectrum peaks from *Saccharothrix* test strains matched to ribosomal proteins from 16 *Saccharothrix* type strains

Test strain	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
<i>S. australiensis</i> CGMCC 4.1355 <sup>T</sup>	4	11	2	2	2	2	3	2	2	2	3	3	5	3	4	1
<i>S. carnea</i> CGMCC 4.7097 <sup>T</sup>	5	2	13	3	9	7	2	5	7	6	9	5	5	6	2	5
<i>S. texasensis</i> CGMCC 4.1713 <sup>T</sup>	3	4	6	5	6	6	4	3	9	2	5	5	3	15	3	2
<i>S. mutabilis</i> subsp. <i>capreolus</i> HKI 264 HKJ	4	4	4	1	2	3	2	1	2	5	5	5	4	2	7	2
<i>S. mutabilis</i> subsp. <i>mutabilis</i> HKI 9771 HKJ	5	3	3	5	3	2	7	2	1	5	4	5	7	1	13	2
<i>Saccharothrix</i> sp. CCTCC AB 2015033	2	3	1	4	3	2	2	1	2	1	2	5	3	1	3	3

Cells that have the maximum value in each row are colored gray.

### 2.5.1 糖丝菌属菌株鉴定

菌株 CGMCC 4.1355<sup>T</sup> 与表 2 中 02 号菌株的核糖体蛋白质匹配数最大(11), 但低于 2.2 部分鉴定至菌种的标准, 未能鉴定为 *S. australiensis*, 可按 2.3 部分鉴定至属的标准鉴定为糖丝菌。菌株质谱图中未检测到 4 种核糖体蛋白质(S14-Z、S18、S15 和 S19)的峰(图 1A), 导致其低于鉴定至菌种的标准。

菌株 CGMCC 4.7097<sup>T</sup> 与表 2 中 03 号菌株的核糖体蛋白质匹配数最大(13)且满足鉴定至菌种的标准, 鉴定为 *S. carnea*, 鉴定结果准确。菌株质谱图中未检测到 2 种核糖体蛋白质(S20 和 S19)的峰(图 1B), 但满足鉴定至菌种的标准。

菌株 CGMCC 4.1713<sup>T</sup> 与表 2 中 14 号菌株的核糖体蛋白质匹配数最大(15)且满足鉴定至菌种的标准, 鉴定为 *S. texasensis*, 鉴定结果准确。菌株质谱图中检测到 15 种核糖体蛋白质的峰(图 1C), 表明菌株实际测试的匹配数与理论匹配数一致。

菌株 HKI 264 HKJ 与表 2 中 15 号菌株的核糖体蛋白质匹配数最大(7)但低于鉴定至菌种的标准, 可按鉴定至属的标准鉴定为糖丝菌。

菌株 HKI 9771 HKJ 与表 2 中 15 号菌株的核糖体蛋白质匹配数最大(13)且满足鉴定至菌种的标准, 鉴定为 *S. variisporea*, 种水平鉴定

错误。原因是 *S. mutabilis* 的基因组尚未公开, 进行糖丝菌属菌种间的核糖体蛋白质匹配时缺少对 *S. mutabilis* 的分析。菌种 *S. variisporea* 和 *S. mutabilis* 的亲缘关系接近<sup>[18]</sup>, 核糖体蛋白质匹配数达到 13, 提示 2.2 部分鉴定至菌种的标准设定偏低, 无法区分个别亲缘关系接近的菌种。为确保种水平鉴定准确性, 将鉴定至菌种的标准修正为核糖体蛋白质匹配数 $\geq 14$ 。按修正后的标准, 质谱中有 2 个及以上核糖体蛋白质峰未检出的菌株无法鉴定至种。

菌株 HKI 264 HKJ 和 HKI 9771 HKJ 同为 *S. mutabilis* 的亚种, 但其最大匹配数差异显著(7 和 13)。HKI 264 HKJ 的峰列表中有 6 个峰与核糖体蛋白质的理论  $m/z$  偏差在  $550 \times 10^{-6}$ – $650 \times 10^{-6}$  之间, 超出设定的误差范围, 未判定为核糖体蛋白质峰。鉴于 6 个峰偏移方向一致, 通过质量轴校正, 该菌株与 15 号菌株的核糖体蛋白质匹配数也可达到 13。

综上所述, 菌株实际测试的核糖体蛋白质匹配数小于理论匹配数的主要原因是其质谱峰未检出或超出误差范围。即使受此影响, 5 个测试菌株的最大匹配数均 $\geq 7$ , 仍满足鉴定至属的标准。

菌株 CCTCC AB 2015033 与表 2 中各模式菌株的核糖体蛋白质最大匹配数为 5, 低于 2.3

部分鉴定至糖丝菌属的标准, 未能鉴定为糖丝菌, 表明与表 2 中的菌种亲缘关系较远的糖丝菌, 按最大匹配数 $\geq 7$ 的标准无法鉴定至属。

### 2.5.2 糖丝菌属菌株核糖体蛋白质的质谱峰强度分析

图 1 显示 4 个测试菌株核糖体蛋白质的 MALDI 质谱峰呈现一致的强度模式, 随  $m/z$  升高峰强度呈阶梯式下降。L36 的峰强度最高, L34 次之, 其次是  $m/z$  6 300–7 000 的 6 种核糖体蛋白质峰, 之后是  $m/z$  7 700–9 400 的 5 种核糖体蛋白质峰,  $m/z$  10 000 以上的峰强度最低, 这种峰强度模式可作为鉴定糖丝菌属菌种的必要条件。测试菌株中分子量 7 000 Da 以下的 8 种核糖体蛋白质的质谱峰均以高强度稳定出现, 缺失的大多数是分子量 8 800 Da 以上的核糖体蛋白质信号, 主要是峰强度低于检出限所致。

### 2.6 糖丝菌属外菌种的 MALDI 质谱峰匹配

假诺卡氏菌目测试菌株与糖丝菌属各模式菌株的核糖体蛋白质匹配结果见表 6, 详细匹配信息已提交国家微生物科学数据中心, 编号: NMDCX0000220, 另外 7 个目的测试菌株的匹配结果已提交国家微生物科学数据中心, 编号: NMDCX0000220。

表 6 中假诺卡氏菌目 7 个测试菌株的核糖

体蛋白质最大匹配数为 5, 不满足 2.3 部分鉴定至糖丝菌属的标准, 表明假诺卡氏菌目内与糖丝菌亲缘关系较远的放线菌类群一般不会鉴定为糖丝菌。假诺卡氏菌目外 7 个目 73 个测试菌株(数据已提交国家微生物科学数据中心, 编号: NMDCX0000220)的核糖体蛋白质最大匹配数为 4, 不满足鉴定至糖丝菌属的标准, 表明与糖丝菌亲缘关系更远的假诺卡氏菌目外的放线菌类群一般不会鉴定为糖丝菌。

### 2.7 糖丝菌属菌种鉴定方法建立

如前文所述, 个别糖丝菌属菌种的最大匹配数为 5, 不满足匹配数 $\geq 7$ 的标准, 未能鉴定为糖丝菌。如将标准调整为匹配数 $\geq 5$ , 邻近属一些菌种同样可能满足(表 4)。为了区分这些菌种, 尝试寻找质谱中稳定出现的属水平特征峰。L34 的分子量区间在糖丝菌(5 292.18–5 349.28 Da)与梅泽宾夫氏菌(5 378.27 Da)及束丝放线菌(5 379.30 Da)间差异显著, 可通过 L34 的质谱峰区分糖丝菌与这 2 个属。L32 的分子量区间在糖丝菌(6 547.55–6 658.69 Da)与伦茨菌(6 457.51–6 517.56 Da)间差异显著, 可通过 L32 的质谱峰区分糖丝菌与伦茨菌。以菌株 CCTCC AB 2015033 为例, 其质谱中  $m/z$  5 352.47 匹配糖丝菌 L34,  $m/z$  6 562.18 位于糖丝菌 L32 的分子量区间, 因此将菌株鉴

表 6 假诺卡氏菌目测试菌株的 MALDI 质谱峰与糖丝菌属模式菌株的核糖体蛋白质匹配数

Table 6 Count of MALDI spectrum peaks from *Pseudonocardiales* test strains matched to ribosomal proteins from 16 *Saccharothrix* type strains

Test strain	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
<i>Lentzea flava</i> IMET 9748T HKJ	1	1	3	1	3	3	1	3	4	2	3	1	1	3	1	2
<i>Pseudonocardia hydrocarbonoxydans</i> DSM 43281 <sup>T</sup>	3	3	2	2	2	2	2	2	3	4	2	3	3	3	4	1
<i>Saccharopolyspora erythraea</i> HKI 184 HKJ	4	3	3	3	2	1	4	3	2	3	2	5	4	1	4	4
<i>Saccharopolyspora hirsuta</i> HKI 9709 HKJ	1	2	1	1	2	1	3	1	2	1	1	1	2	1	2	1
<i>Amycolatopsis alba</i> DSM 44262 <sup>T</sup>	1	2	2	0	2	3	3	3	2	3	2	0	0	1	0	1
<i>Amycolatopsis balhimycina</i> DSM 44591 <sup>T</sup>	0	1	1	1	0	2	0	0	1	0	0	0	0	2	1	2
<i>Amycolatopsis mediterranei</i> DSM 40773	1	1	0	3	1	2	0	0	1	0	1	1	1	3	1	0

Cells that have the maximum value in each row are colored gray.

定为糖丝菌。后续的菌株基因组测序分析也直接证实了  $m/z$  5 352.47 匹配菌株的 L34 (理论  $m/z$  5 350.28),  $m/z$  6 562.18 匹配菌株的 L32 (理论  $m/z$  6 560.60)。此外, 亲缘关系较远的假诺卡氏菌目内的个别菌种也满足匹配数 $\geq 5$ 的标准(表 6), 但这种情况不同于 3 个邻近属的菌种, 主要是随机错配所致, 可结合 2.5.2 部分的峰强度模式予以排除。综上所述, 与表 2 中各模式菌株的最大匹配数为 6 和 5 的待测菌株, 如匹配核糖体蛋白质包括 L34 和 L32, 匹配峰的强度模式符合糖丝菌属特征模式, 也可鉴定为糖丝菌。

本研究通过理论分析提出鉴定标准, 基于测试结果修正鉴定标准并建立糖丝菌属菌种的鉴定方法, 具体步骤如下: (1) 构建糖丝菌属菌种模式菌株 15 种核糖体蛋白质的理论分子量表(表 2); (2) 甲酸提取法处理待测菌株; (3) MALDI-TOF 质谱仪采集菌株蛋白提取液的质谱; (4) 将待测菌株的 MALDI-TOF 质谱峰与表 2 中各菌种模式菌株 15 种核糖体蛋白质分别匹配并获得最大匹配数; (5) 结果判定: 最大匹配数 $\geq 7$ , 匹配峰的强度模式符合糖丝菌属特征模式, 则鉴定为糖丝菌属; 最大匹配数为 6 和 5, 匹配中包括 L34 和 L32, 匹配峰的强度模式符合糖丝菌属特征模式, 则鉴定为糖丝菌属; 鉴定为菌种需满足与该菌种模式菌株的核糖体蛋白质匹配数 $\geq 14$ 。

注意事项包括: (1) 糖丝菌属现有 23 个有效描述种, NCBI 和 JGI 基因组数据库共收录 16 个, 表 2 中缺少 7 种糖丝菌(*S. ghardaiensis*、*S. hoggarensis*、*S. isguenensis*、*S. lopnurensis*、*S. mutabilis*、*S. stipae* 和 *S. xinjiangensis*)的核糖体蛋白质分子量数据, 这些菌种的菌株无法鉴定至种。(2) 按核糖体蛋白质匹配数 $\geq 14$ 的标准鉴定至菌种可确保鉴定准确, 但质谱中有 2 个

及以上核糖体蛋白质峰未检出的菌株无法鉴定至菌种。上述两点导致本方法鉴定菌种存在应用上的局限性。(3) 按核糖体蛋白质匹配数 $\geq 7$ 的标准鉴定糖丝菌属可确保鉴定准确, 但与表 2 中的糖丝菌亲缘关系较远的糖丝菌属菌种无法得到鉴定。筛选糖丝菌属菌种时按核糖体蛋白质匹配数 $\geq 5$ 的标准能够提高筛选率, 有助于筛选出糖丝菌属新菌种。

### 3 讨论与结论

最初利用基因组学技术注释放线菌纲细菌 MALDI-TOF MS 指纹图谱中的核糖体蛋白质信号, 找到可用于表征放线菌纲菌种的 15 种核糖体蛋白质标志物组合<sup>[14]</sup>, 之后基于 15 种核糖体蛋白质标志物及 MALDI-TOF MS 技术建立了鉴定伦茨菌属放线菌的方法<sup>[15]</sup>。本研究在前期基础上优化方法。首先, 扩大搜索数据库的范围, 通过 JGI 数据库检索, 补充了 NCBI 数据库未收录的模式菌株, 表 2 涵盖了公开基因组序列的全部糖丝菌属模式菌株。其次, 前期研究中伦茨菌属个别菌种模式菌株的基因组仅注释出 10 种核糖体蛋白质, 本研究发现糖丝菌属部分菌种出现同样情况。然而进一步分析发现这并非基因组测序不完整引起, 而是注释信息缺失所致, 通过核糖体蛋白质序列的同源比对可从高度同源的蛋白质序列来源信息中获得这些缺失序列的信息。因此, 表 2 中的菌株注释出全部 15 种核糖体蛋白质, 确保了分子量数据的完整性。上述方法改进扩大了糖丝菌属菌种的鉴定范围。

本研究从数据库检索到糖丝菌属菌种的模式菌株之外, 还检索到鉴定到属的菌株, 基因组比对结果提示一些菌株可能是未发表的糖丝菌属新菌种。这些菌株大部分核糖体蛋白质的序列和分子量与模式菌株一致, 同时不乏菌株

特征性的分子量,尤其是保守性低的核糖体蛋白质如 L31 和 S20,这些分子量数据可作为模式菌株核糖体蛋白质分子量数据的补充,为待测菌株的核糖体蛋白质质谱峰匹配提供更多依据。因此,构建核糖体蛋白质分子量表时除了尽可能涵盖目标属菌种的模式菌株,补充代表潜在新菌种的菌株,可以扩大目标属菌种的鉴定范围。

前期研究中鉴定伦茨菌属放线菌的依据是核糖体蛋白质匹配数和质谱峰强度模式。本研究发现依据核糖体蛋白质匹配数和质谱峰强度模式无法区分糖丝菌与束丝放线菌、梅泽宾夫氏菌及伦茨菌属的一些菌种,需要在鉴定标准中补充糖丝菌属菌种的核糖体蛋白质特征峰,比如 L34 (区分糖丝菌与束丝放线菌及梅泽宾夫氏菌)和 L32 (区分糖丝菌与伦茨菌)的特征峰。建立放线菌纲其他类群的鉴定方法时可以在鉴定标准中设立相应的属水平的核糖体蛋白质特征峰,以便区分邻近属菌种,提高鉴定准确度。

本研究建立的方法中,步骤(4)将待测菌株的 MALDI-TOF 质谱峰与表 2 中各菌种模式菌株 15 种核糖体蛋白质分别匹配的操作比较费时。鉴于 L36 和 L34 在测试菌株的质谱图中均以高信号强度稳定出现,可通过寻找匹配 L36 和 L34 的  $m/z$  对待测菌株进行初筛。分析 57 个属共 80 个非糖丝菌属菌种的质谱峰列表发现,质量误差小于  $500 \times 10^{-6}$  的条件下,仅有 4 个菌种(*Lentzea flava*、*Pseudonocardia hydrocarbonoxydans*、*Saccharopolyspora erythraea* 和 *Kytococcus schroeteri*)的峰列表中同时出现匹配糖丝菌属菌种 L36 和 L34 的  $m/z$ ,表明依据 L36 和 L34 这 2 种核糖体蛋白质的质谱峰多达 95% 的非糖丝菌属菌种可被快速排除。将上述方法用于定向筛选糖丝菌属菌种,可以快速缩小筛选范围,

显著提高筛选效率。

先后以伦茨菌和糖丝菌属放线菌为分析案例,对基于核糖体蛋白质标志物和 MALDI-TOF MS 技术鉴定放线菌的方法进行了探索,研究结果可为放线菌纲其他类群的鉴定提供借鉴。随着放线菌纲更多类群 15 种核糖体蛋白质分子量数据库的构建,基于核糖体蛋白质标志物的方法可应用于环境放线菌的分类和鉴定中。相比传统鉴定方法,该方法具有简单、快速、适合高通量鉴定等优势。虽然基质辅助激光解吸电离飞行时间质谱仪的仪器成本高,但其检测成本低,鉴定一个单克隆仅需 1-2 元。在近年来科技部鼓励大型科学仪器对外开放共享的背景下,具备了方法推广的条件。该方法与培养基自动制备分装系统、自动化移液工作站、微生物克隆自动化挑取工作站等仪器设备配套使用,未来有望在工业水平上实现目标放线菌类群的高通量筛选和鉴定。

## 参考文献

- [1] LABEDA DP, TESTA RT, LECHEVALIER MP, LECHEVALIER HA. *Saccharothrix*: a new genus of the Actinomycetales related to *Nocardioopsis*[J]. International Journal of Systematic Bacteriology, 1984, 34(4): 426-431.
- [2] KURANAGA T, TAMURA M, IKEDA H, TERADA S, NAKAGAWA Y, KAKEYA H. Identification and total synthesis of an unstable anticancer macrolide presaccharothriolide Z produced by *Saccharothrix* sp. A1506[J]. Organic Letters, 2021, 23(18): 7106-7111.
- [3] SHEN QY, DAI GZ, LI AY, LIU Y, ZHONG GN, LI XJ, REN XM, SUI HY, FU J, JIAO NZ, ZHANG YM, BIAN XY, ZHOU HB. Genome-guided discovery of highly oxygenated aromatic polyketides, saccharothrixins D-M, from the rare marine actinomycete *Saccharothrix* sp. D09[J]. Journal of Natural Products, 2021, 84(11): 2875-2884.
- [4] SHEN QY, DAI GZ, RAVICHANDRAN V, LIU Y, ZHONG L, SUI HY, REN XM, JIAO NZ, ZHANG YM, ZHOU HB, BIAN XY. Saccharochelins A-H,

- cytotoxic amphiphilic siderophores from the rare marine actinomycete *Saccharothrix* sp. D09[J]. *Journal of Natural Products*, 2021, 84(8): 2149-2156.
- [5] GAN ML, LIU B, TAN Y, WANG Q, ZHOU HX, HE HW, PING YH, YANG ZY, WANG YG, XIAO CL. *Saccharothrix*ones A-D, tetracenomycin-type polyketides from the marine-derived actinomycete *Saccharothrix* sp. 10-10[J]. *Journal of Natural Products*, 2015, 78(9): 2260-2265.
- [6] PASIN TM, de OLIVEIRA TB, de ALMEIDA SCARCELLA AS, de LOURDES TEIXEIRA de MORAES POLIZELI M, GUAZZARONI ME. Perspectives on expanding the repertoire of novel microbial chitinases for biological control[J]. *Journal of Agricultural and Food Chemistry*, 2021, 69(11): 3284-3288.
- [7] 程斐, 孙朝晖, 赵玉国, 谢平林, 李式军. 高温分解纤维素的易变糖丝菌菌株筛选及其特性[J]. *南京农业大学学报*, 2002, 25(2): 35-38.  
CHENG F, SUN H, ZHAO YG, XIE PL, LI SJ. Isolation of efficiency decomposing cellulose of *Saccharothrix mutabilis*[J]. *Journal of Nanjing Agricultural University*, 2002, 25(2): 35-38 (in Chinese).
- [8] 曹相生, 孟雪征, 任书魁, 李秀强. 一株 DEHP 降解菌的筛选和分子鉴定[J]. *环境科学与技术*, 2011, 34(12): 217-220.  
CAO XS, MENG XZ, REN SK, LI XQ. Isolation and phylogenetic analysis of a bis(2-ethylhexyl) phthalate degrading bacterial strain[J]. *Environmental Science & Technology*, 2011, 34(12): 217-220 (in Chinese).
- [9] GORNIÁKOVÁ D, PETŘÍČEK M, KAHOUN D, GRABIC R, ZELENKA T, CHROŇÁKOVÁ A, PETŘÍČKOVÁ K. Activation of a cryptic manumycin-type biosynthetic gene cluster of *Saccharothrix espanaensis* DSM44229 by series of genetic manipulations[J]. *Microorganisms*, 2021, 9(3): 559.
- [10] PATEL R. MALDI-TOF MS for the diagnosis of infectious diseases[J]. *Clinical Chemistry*, 2015, 61(1): 100-111.
- [11] ARNOLD RJ, REILLY JP. Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry[J]. *Analytical Biochemistry*, 1999, 269(1): 105-112.
- [12] PINEDA FJ, ANTOINE MD, DEMIREV PA, FELDMAN AB, JACKMAN J, LONGENECKER M, LIN JS. Microorganism identification by matrix-assisted laser/desorption ionization mass spectrometry and model-derived ribosomal protein biomarkers[J]. *Analytical Chemistry*, 2003, 75(15): 3817-3822.
- [13] 周月霞, 胡成进, 陈英剑. MALDI-TOF MS 微生物鉴定数据库应用研究进展[J]. *临床检验杂志*, 2016, 34(6): 447-450.  
ZHOU YX, HU CJ, CHEN YJ. Research progress on application of MALDI-TOF MS microbial identification database[J]. *Chinese Journal of Clinical Laboratory Science*, 2016, 34(6): 447-450 (in Chinese).
- [14] 孙伟, 罗倩, 张薇, 吕霁焯, 祝新德, 许杰, 张风丽. 利用基因组学和 MALDI-TOF MS 技术鉴定放线菌纲细菌的核糖体蛋白质标志物[J]. *微生物学报*, 2022, 62(7): 2671-2685.  
SUN W, LUO Q, ZHANG W, LV JY, ZHU XD, XU J, ZHANG FL. Characterization of the ribosomal protein biomarkers of *Actinobacteria* via genomics and MALDI-TOF MS techniques[J]. *Acta Microbiologica Sinica*, 2022, 62(7): 2671-2685 (in Chinese).
- [15] 孙伟, 许杰, 周万平, 罗倩, 张薇. 基于核糖体蛋白质标志物及 MALDI-TOF MS 技术快速鉴定伦茨菌属放线菌[J]. *微生物学报*, 2023, 63(1): 430-445.  
SUN W, XU J, ZHOU WP, LUO Q, ZHANG W. Rapid identification of *Lentzea* members: based on ribosomal protein biomarkers and MALDI-TOF MS[J]. *Acta Microbiologica Sinica*, 2023, 63(1): 430-445 (in Chinese).
- [16] MEIER-KOLTHOFF JP, GÖKER M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy[J]. *Nature Communications*, 2019, 10: 2182.
- [17] 张娅, 曹成亮, 李荣鹏, 蒋继宏. 糖丝菌属放线菌研究进展[J]. *微生物学报*, 2022, 62(5): 1600-1612.  
ZHANG Y, CAO CL, LI RP, JIANG JH. Recent advance on the genus *Saccharothrix*[J]. *Acta Microbiologica Sinica*, 2022, 62(5): 1600-1612 (in Chinese).
- [18] KIM BY, BROWN R, LABEDA DP, GOODFELLOW M. Reclassification of '*Dactylosporangium variesporum*' as *Saccharothrix variisporae* corr. (ex Tomita et al. 1977) sp. nov., nom. rev.[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2011, 61(2): 310-314.