



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

孙伟*, 许杰, 周万平, 罗倩, 张薇

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, 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 Wei, XU Jie, ZHOU Wanping, LUO Qian, ZHANG Wei. Rapid identification of *Lentzea* members: based on ribosomal protein biomarkers and MALDI-TOF MS[J]. Acta Microbiologica Sinica, 2023, 63(1): 430-445.

摘要:【目的】伦茨菌属(*Lentzea*)放线菌(*Actinobacteria*)代谢产物具有广泛的生物活性,在医药领域展现出潜在的应用价值。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)技术鉴定伦茨菌属放线菌的方法。【方法】检索基因组数据库,提取伦茨菌属菌种模式菌株 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伦茨菌属菌种模式菌株之间及其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 质谱峰与伦茨菌属各菌种模式菌株的 15 种核糖体蛋白质分别匹配,通过最大匹配数及质谱峰强度模式可鉴定至属或种。【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 技术鉴定伦茨菌属放线菌的方法,可为放线菌纲其他类群的快速鉴定提供借鉴。

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; 伦茨菌属; 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@sjtu.edu.cn

Received: 2022-05-12; Accepted: 2022-07-30; Published online: 2022-08-11

Rapid identification of *Lentzea* members: based on ribosomal protein biomarkers and MALDI-TOF MS

SUN Wei*, XU Jie, ZHOU Wanping, LUO Qian, ZHANG Wei

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

Abstract: [Objective] Metabolites from members of *Lentzea* (*Actinobacteria*) have a wide range of bioactivities, thus showing potential application value in biopharmaceutical field. This study aims to establish a method to identify members of *Lentzea* based on ribosomal protein biomarkers and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). [Methods] We searched public genome database to obtain the sequences of 15 ribosomal proteins from the type strains of *Lentzea* and calculated the theoretical molecular weights. The molecular weight comparison was performed to investigate the matching of 15 ribosomal proteins between the type strains of *Lentzea* and that between the type strains of *Lentzea* and its closely related genera (*Actinosynnema* and *Saccharothrix*). Identification criteria for *Lentzea* at the species and genus level were proposed based on the number of ribosomal protein matches. Subsequently, the identification criteria were tested 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 *Lentzea* type strain and the number of peaks that matched each strain's biomarkers was obtained. Based on the maximum of matches and peak intensity mode, the organism can be identified at the genus or species level. [Conclusion] This study established a method for identifying *Lentzea* members based on 15 ribosomal protein biomarkers and MALDI-TOF MS, which can provide a reference for the rapid identification of other *Actinobacteria* groups.

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

伦茨菌属(*Lentzea*)建立于 1995 年^[1], 是一类丝状稀有放线菌, 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*)。原核生物标准命名列表(list of prokaryotic names with standing in nomenclature, <http://lpsn.dsmz.de/>)显示伦茨菌属现有 23 个有效描述种和 2 个非有效描述种。近年来, 随着越来越多具有特殊生物活性(如 I 型人体免疫缺损病毒整合酶抑制活性、抗肿

瘤活性和抗结核活性等)的天然产物从伦茨菌中分离获得^[2-6], 人们开始重视这类放线菌在新药创制中的应用价值。截至目前有效发表的伦茨菌属菌种仍然较少, 而宏基因组学研究表明该属许多新成员尚未获得分离培养, 存在较大的新物种资源挖掘空间^[7]。目前细菌鉴定方法的“金标准”是 16S rRNA 基因序列比对法, 该方法鉴定准确性高, 但操作较复杂, 鉴定周期较长。因此, 开发一种简单快速的伦茨菌鉴定方法有助于促进该类放线菌资源的开发利用。

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)技术采集待测菌株的蛋白指纹图谱,与数据库中菌种的指纹图谱比对,可用于菌种鉴定^[8]。该方法与 16S rRNA 基因序列比对法相比,具有简单、快速和高通量等优势,近年来在临床诊断、环境监测和食品安全领域得到广泛应用^[9-13]。Demirev 等提出利用 MALDI-TOF MS 技术及蛋白质数据库搜索鉴定菌种的设想,并通过枯草芽孢杆菌(*Bacillus subtilis*)和大肠埃希氏菌(*Escherichia coli*)的研究证实了其可行性^[14]。Pineda 等建立统计模型预测 MALDI 质谱峰与蛋白质组数据库中蛋白质的错误匹配,提出精简蛋白质组(排除低丰度和不表达的蛋白质)减少错误匹配的思路^[15]。在此基础上,Pineda 等利用核糖体蛋白质标志物匹配法分析 7 种细菌发现有 20 个或更多标志物序列信息的菌种可准确鉴定^[16]。随后,Sun 等选取 31 种核糖体蛋白质和 3 种核糖体相关蛋白质作为标志物用于植物乳杆菌(*Lactobacillus plantarum*)的鉴定^[17]。Ziegler 等使用 13 种核糖体蛋白质标志物将种类多样的根瘤菌准确鉴定至菌种水平^[18]。Cheng 等选取 10 种蛋白质(包括 8 种核糖体蛋白质)作为标志物鉴定了 8 个属的菌种,发现属水平的鉴定准确率达 84%^[19]。Kostas 和 Parker 选取 65 种蛋白质(包括 52 种核糖体蛋白质)作为标志物准确鉴定了弧菌目和环境中的某些菌种^[20]。以上研究表明不借助图谱数据库,通过选取合适的生物标志物,利用 MALDI-TOF MS 技术也可实现特定微生物类群的鉴定。然而研究不同菌种使用的标志物各不相同,至今仍缺乏统一标准,限制了这类方法在环境微生物鉴定中的推广。

我们前期利用基因组学技术注释核糖体蛋白质的 MALDI-TOF 质谱峰,找到放线菌纲细菌 MALDI-TOF MS 图谱中常见的 15 种核糖

体蛋白质(L36、L35、L34、L33、L32、L31、L30、L29、L28、L27、S20、S19、S18、S15 和 S14-Z)信号,对测试菌株的分析提示这些核糖体蛋白质标志物可用于放线菌鉴定^[21]。在此基础上,本研究以伦茨菌属放线菌为对象,尝试建立以上述 15 种核糖体蛋白质为标志物,利用 MALDI-TOF MS 技术鉴定伦茨菌的方法,以期放线菌纲其他类群的鉴定提供借鉴。

1 材料与方 法

1.1 伦茨菌属菌种核糖体蛋白质序列提取、分析及分子量计算

以“*Lentzea*”作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(<http://www.ncbi.nlm.nih.gov/genome/>),获取伦茨菌属测序菌种信息,选取模式菌株进行分析。以“ribosomal”作为检索词搜索菌株基因组组装与注释报告的蛋白质列表,提取 15 种核糖体蛋白质(L36、L35、L34、L33、L32、L31、L30、L29、L28、L27、S20、S19、S18、S15 和 S14-Z)的序列。

按我们前期研究所用方法分析核糖体蛋白质序列^[21]。将同种核糖体蛋白质的序列整合为一个 fasta 文件,导入 ClustalX 1.81 软件进行多序列比对^[22],找出氨基酸发生突变的位点。基于 MEGA 6.06 软件的最大似然法(maximum likelihood, ML)构建系统进化树^[23-24]。

应用蛋白质基本性质分析工具 ProtParam(<http://web.expasy.org/protparam/>)^[25]计算核糖体蛋白质的平均分子量作为理论分子量,获得伦茨菌属菌种模式菌株 15 种核糖体蛋白质的理论分子量,统计各种分子量的出现次数。

1.2 伦茨菌属菌种模式菌株间及菌种内不同菌株间的核糖体蛋白质匹配分析

将各菌种模式菌株 15 种核糖体蛋白质的

分子量进行两两比对, 统计模式菌株之间的核糖体蛋白质匹配数, 评估以 15 种核糖体蛋白质标志物的组合区分菌种的可行性。基于各菌种模式菌株与其他菌种模式菌株之间的核糖体蛋白质最大匹配数, 建立鉴定至菌种的标准。

选取有多个测序菌株的菌种, 提取各菌株 15 种核糖体蛋白质的序列并计算分子量, 将同一菌种内不同菌株间 15 种核糖体蛋白质的分子量进行两两比对, 分析菌种内不同菌株间 15 种核糖体蛋白质组合的差异度, 评估模式菌株的代表性。

1.3 伦茨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

束丝放线菌 (*Actinosynnema*) 和糖丝菌 (*Saccharothrix*) 是亲缘关系最接近伦茨菌的放线菌^[7], 区分伦茨菌属与这 2 个邻近属的菌种是准确鉴定至属的前提。提取这 2 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, 与伦茨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伦茨菌属的标准。

1.4 伦茨菌属测试菌株

为测试鉴定标准对目标属菌种的鉴定结果, 收集 7 个伦茨菌属菌株进行分析, 包括: *L. guizhouensis* CGMCC 4.7203^T 和 *L. pudingi* CGMCC 4.7319^T (由江苏师范大学生命科学学院曹成亮高级实验师提供); *L. soli* CCTCC AA 2017027^T、*Lentzea* sp. CCTCC AA 2017024 和 CCTCC AA 2016016 (购自中国典型培养物保藏中心); *Lentzea* sp. 2-8 和 3-5 (从崇明岛东滩湿地土壤中分离获得), 用 ISP2 培养基将菌株活化后进行传代培养。取 30 °C 培养 2-3 d 的测试样本, 采用甲酸提取法处理样本, 提取菌株蛋白^[21]。

1.5 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(型号: ultraflextreme)购自布鲁克公司。按我们前期研究所用方法采集测试样本的质谱图并进行图谱分析^[21]。

蛋白质在基质辅助激光解吸电离下结合一个质子产生 $[M+H]^+$, 因此核糖体蛋白质 MALDI 质谱峰的理论 m/z 为理论分子量 $M+1$ 。质量误差设为小于 300×10^{-6} (线性模式的质量准确度范围)。将测试菌株质谱中的 m/z 与伦茨菌属各模式菌株 15 种核糖体蛋白质的理论 m/z 依次比对, 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, 按鉴定至伦茨菌属和种的标准鉴定并评估准确性。

以测试菌株质谱中信号强度最高的核糖体蛋白质峰为基峰, 计算各种核糖体蛋白质峰的相对信号强度, 分析伦茨菌属菌株核糖体蛋白质的 MALDI 质谱峰强度模式。

1.6 伦茨菌属外菌种的 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(Bruker MBT DB7854, 购自布鲁克公司)选取伦茨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 个菌种进行分析。以图谱数据库收录的唯一一种伦茨菌 *L. flava* IMET 9748T HKJ 作为对照。质量误差设为小于 500×10^{-6} ^[21]。按 1.5 部分所述方法对测试菌株依次比对和统计, 获得最大匹配数, 按鉴定至伦茨菌属的标准鉴定并评估准确性。

2 结果与分析

2.1 伦茨菌属菌种核糖体蛋白质序列提取、分析及分子量计算

从基因组数据库检索到伦茨菌属 19 个种, 包括 18 个有效描述种和 1 个非有效描述种(表 1)。17 个菌种模式菌株的基因组均注释出

表 1 伦茨菌属 19 个模式菌株 15 种核糖体蛋白质的理论分子量表

Table 1 The list of the masses predicted for 15 ribosomal proteins from 19 *Lentzea* type strains

Type strain	L36	L34	L33	L30	L32	S14-Z	L28	L35	L31	S18	L29	L27	S20	S15	S19
01 <i>L. aerocolonigenes</i> NBRC 13195 ^T	4 403.38	5 265.16	6 363.33	6 764.87	6 473.51	6 872.27	6 692.69	6 991.23	7 743.77	8 831.39	8 934.03	8 899.33	9 512.06	10 223.80	10 507.27
02 <i>L. alba</i> NEAU-D13 ^T	/	/	/	6 837.92	/	6 872.27	/	6 961.25	7 821.86	8 831.39	8 948.06	8 869.25	9 455.00	10 039.56	10 507.27
03 <i>L. albidia</i> DSM 44437 ^T	4 403.38	5 265.16	6 349.30	6 723.88	6 514.56	6 872.27	6 692.69	6 991.23	7 823.83	8 804.37	8 948.06	8 869.25	9 456.98	10 223.80	10 506.24
04 <i>L. albidocapillata</i> DSM 44073 ^T	4 403.38	5 265.16	6 381.36	6 798.92	6 517.56	6 872.27	6 692.69	6 966.22	7 764.76	8 804.37	8 962.09	8 891.27	9 471.00	10 209.77	10 506.24
05 <i>L. atacamensis</i> DSM 45479 ^T	4 403.38	5 306.21	6 363.33	6 736.85	6 473.51	6 886.30	6 692.69	6 957.17	7 743.77	8 804.37	8 948.06	8 884.22	9 538.09	10 223.80	10 507.27
06 <i>L. cavernae</i> CGMCC 4.7367 ^T	4 403.38	5 265.16	6 363.33	6 826.93	6 514.56	6 872.27	6 692.69	6 943.18	7 810.83	8 804.37	8 976.11	8 845.18	9 456.98	10 209.77	10 506.24
07 <i>L. flava</i> JCM 3296 ^T	4 403.38	5 320.24	6 363.33	6 794.85	6 457.51	6 886.30	6 692.69	6 957.17	7 743.77	8 831.39	8 962.09	8 873.15	9 538.09	10 237.83	10 507.27
08 <i>L. flaviverrucosa</i> DSM 44664 ^T	4 403.38	5 265.16	6 363.33	6 798.92	6 473.51	6 872.27	6 692.69	6 929.16	7 764.76	8 804.37	8 962.09	8 845.18	9 471.00	10 209.77	10 506.24
09 <i>L. fradiae</i> CGMCC 4.3506 ^T	4 403.38	5 320.24	6 363.33	6 880.00	6 514.56	6 872.27	6 710.72	6 943.18	7 839.83	8 804.37	8 948.06	8 899.33	9 512.10	10 223.80	10 507.27
10 <i>L. guizhouensis</i> CGMCC 4.7203 ^T	4 403.38	5 295.18	6 349.30	6 791.83	6 514.56	6 872.27	6 692.69	6 993.27	7 837.86	8 831.39	8 976.11	8 869.25	9 512.06	10 239.80	10 506.24
11 “ <i>L. indica</i> ” PSKA42 ^{1*}	/	/	/	6 821.96	/	6 872.27	/	6 966.22	7 743.77	8 804.37	8 948.06	8 887.28	9 455.00	10 041.53	10 507.27
12 <i>L. jiangxiensis</i> CGMCC 4.6609 ^T	4 403.38	5 265.16	6 363.33	6 826.93	6 514.56	6 872.27	6 692.69	6 943.18	7 811.77	8 804.37	8 920.01	8 869.25	9 511.07	10 223.80	10 506.24
13 <i>L. kentuckyensis</i> NRRL B-24416 ^T	4 403.38	5 265.16	6 363.33	6 849.97	6 473.51	6 872.27	6 692.69	6 943.18	7 837.86	8 831.39	8 948.06	8 971.36	9 467.10	10 209.77	10 507.27
14 <i>L. nigeriaca</i> DSM 45680 ^T	4 403.38	5 320.24	6 363.33	6 794.85	6 457.51	6 886.30	6 692.69	6 957.17	7 743.77	8 831.39	8 934.03	8 873.15	9 538.09	10 223.80	10 507.27
15 <i>L. pudingi</i> CGMCC 4.7319 ^T	4 403.38	5 309.17	6 363.33	6 826.93	6 514.56	6 872.27	6 692.69	6 929.16	7 837.86	8 804.37	8 962.09	8 845.18	9 455.00	10 209.77	10 506.24
16 <i>L. terrae</i> NEAU-LZS 42 ^T	4 403.38	5 265.16	6 363.33	6 999.24	6 473.51	6 872.27	6 692.69	6 943.18	7 871.92	8 804.37	8 948.06	8 899.33	9 268.84	10 223.80	10 507.27
17 <i>L. tibetensis</i> FXJ1.1311 ^T	4 403.38	5 290.21	6 363.33	6 810.92	6 500.57	6 888.27	6 722.71	6 974.18	7 800.86	8 804.37	8 962.09	9 072.42	9 337.94	10 299.91	10 507.27
18 <i>L. waywayandensis</i> DSM 44232 ^T	4 403.38	5 265.16	6 363.33	6 849.97	6 473.51	6 872.27	6 692.69	6 966.22	7 837.86	8 831.39	8 948.06	8 843.21	9 442.95	10 209.77	10 506.24
19 <i>L. xinjiangensis</i> CGMCC 4.3525 ^T	4 403.38	5 265.16	6 363.33	6 849.97	6 514.56	6 872.27	6 692.69	6 929.16	7 841.80	8 831.39	8 948.06	8 930.30	9 540.11	10 209.77	10 507.27

*：“*L. indica*” is not validly published.

15种目标核糖体蛋白质, 另外2个菌种模式菌株(*L. alba* NEAU-D13^T和“*L. indica*”PSKA42^T)的基因组均注释出10种目标核糖体蛋白质(缺少L36、L34、L33、L32和L28)。15种目标核糖体蛋白质中有6种(L33、L32、S14、L28、L31和S18)注释出长度不同的两个序列。我们前期注释核糖体蛋白质的MALDI-TOF质谱峰时发现伦茨菌属菌株*L. flava* IMET 9748T HKJ实际表达的核糖体蛋白质序列长度为54(L33)、60(L32)、61(S14-Z)、63(L28)、70(L31)和79(S18)个氨基酸^[21], 因此本研究提取这些长度的序列进行分析。

序列分析发现L36和L31的N-端第2个氨基酸是赖氨酸(K), 其前体蛋白的起始甲硫氨酸不切除, 另外13种核糖体蛋白质的N-端第2个氨基酸是丙氨酸(A)、脯氨酸(P)或丝氨酸(S), 起始甲硫氨酸均被切除。进一步分析发现各模式菌株15种核糖体蛋白质的序列保守性差异显著, 大致分三类: 高度保守序列(L36), 该序列在17个模式菌株中完全一致, 可用于伦茨菌属的鉴定, 无法区分菌种; 中度保守序列(S18、S19、L33、S14-Z、L28、L32、L29、L34和S15), 这类序列一般有1-3个位点发生氨基酸替换, 可用于伦茨菌属及少数菌种的鉴定; 高度变化序列(L30、S20、L27、L31和L35), 这类序列一般有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, 可用于菌种鉴定, 数据已提交国家微生物科学数据中心, 编号: NMDCX0000147。这些核糖体蛋白质序列的组合为菌种鉴定奠定了基础。

伦茨菌属19个模式菌株15种核糖体蛋白质的理论分子量见表1, 分子量范围在4 000-11 000 Da。这些核糖体蛋白质产生分子量共计92个, 其中42个分子量在至少2个菌种中出现。出现次数最高的是4 403.38(L36, 17种),

其次是6 692.69(L28, 15种)、6 872.27(S14-Z, 15种)、6 363.33(L33, 14种)、8 804.37(S18, 11种)、10 507.27(S19, 11种)和5 265.16(L34, 10种), 这些分子量在半数以上菌种出现, 可作为鉴定伦茨菌属的指标。另外50个分子量仅在1个菌种中出现, 可作为鉴定菌种的指标。

2.2 伦茨菌属菌种模式菌株间及菌种内不同菌株间的核糖体蛋白质匹配分析

伦茨菌属各菌种模式菌株间一些核糖体蛋白质分子量十分接近, 如6 722.71(L28)和6 723.88(L30); 9 511.07、9 512.06和9 512.10(S20); 10 506.24和10 507.27(S19)等。因此, 模式菌株间的核糖体蛋白质分子量比对时, 需考虑这些分子量引起错误匹配的可能。两种核糖体蛋白质分子量之差在 $\pm 300 \times 10^{-6}$ 内即认为可能引起错误匹配, 计算匹配数时将这些分子量可能引起的错误匹配包括在内。各模式菌株间15种核糖体蛋白质的匹配结果见表2, 详细匹配信息已提交国家微生物科学数据中心, 编号: NMDCX0000147。

表2显示各菌种模式菌株间的核糖体蛋白质最大匹配数在5-13之间, 提示我们伦茨菌属亲缘关系最近的菌种模式菌株间有2种核糖体蛋白质展现质谱仪可辨识的分子量差异, 因此15种核糖体蛋白质的组合理论上可以区分伦茨菌属各菌种的模式菌株。

目前有多个测序菌株的伦茨菌有4种, 可用于分析菌种内不同菌株间核糖体蛋白质的分子量变化。各菌株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: NMDCX0000147。菌种内不同菌株间15种核糖体蛋白质的分子量多数情况下完全一致, 个别情况下1种核糖体蛋白质有分子量差异, 如*L. atacamensis* DSM 45480。但也有例外, 如*L. aerocolonigenes* NRRL B-16140相比模

表 2 伦茨菌属 19 个模式菌株间的核糖体蛋白质匹配数

Table 2 Count of ribosomal proteins matched to 15 ribosomal proteins from 19 *Lentzea* type strains

No.	Organism	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
01	<i>L. aerocolonigenes</i> NBRC 13195 ^T	15	3	7	5	7	6	6	7	7	7	3	8	8	8	5	9	3	8	7
02	<i>L. alba</i> NEAU-D13 ^T	3	10	6	2	2	3	2	2	3	4	5	3	4	2	3	3	1	4	4
03	<i>L. albida</i> DSM 44437 ^T	7	6	15	6	6	8	3	6	7	8	5	9	6	4	7	8	4	6	7
04	<i>L. albidocapillata</i> DSM 44073 ^T	5	2	6	15	4	7	4	11	4	4	4	6	6	3	7	6	4	7	6
05	<i>L. atacamensis</i> DSM 45479 ^T	7	2	6	4	15	5	8	6	6	3	4	6	6	9	5	8	5	6	6
06	<i>L. cavernae</i> CGMCC 4.7367 ^T	6	3	8	7	5	15	4	9	7	6	4	11	8	4	11	8	4	8	8
07	<i>L. flava</i> JCM 3296 ^T	6	2	3	4	8	4	15	5	4	5	2	4	5	13	5	4	5	5	6
08	<i>L. flaviverrucosa</i> DSM 44664 ^T	7	2	6	11	6	9	5	15	5	4	3	7	8	4	10	8	5	9	8
09	<i>L. fradiae</i> CGMCC 4.3506 ^T	7	3	7	4	6	7	4	5	15	6	4	9	7	5	7	9	4	6	7
10	<i>L. guizhouensis</i> CGMCC 4.7203 ^T	7	4	8	4	3	6	5	4	6	15	2	7	6	4	6	4	2	6	6
11	" <i>L. indica</i> " PSKA42 ^T	3	5	5	4	4	4	2	3	4	2	10	3	3	2	4	4	2	4	3
12	<i>L. jiangxiensis</i> CGMCC 4.6609 ^T	8	3	9	6	6	11	4	7	9	7	3	15	7	5	8	9	4	6	7
13	<i>L. kentuckyensis</i> NRRL B-24416 ^T	8	4	6	6	6	8	5	8	7	6	3	7	15	5	7	9	3	12	10
14	<i>L. nigeriaca</i> DSM 45680 ^T	8	2	4	3	9	4	13	4	5	4	2	5	5	15	4	5	4	5	6
15	<i>L. pudingi</i> CGMCC 4.7319 ^T	5	3	7	7	5	11	5	10	7	6	4	8	7	4	15	6	5	8	8
16	<i>L. terrae</i> NEAU-LZS 42 ^T	9	3	8	6	8	8	4	8	9	4	4	9	9	5	6	15	4	8	7
17	<i>L. tibetensis</i> FXJ1.1311 ^T	3	1	4	4	5	4	5	5	4	2	2	4	3	4	5	4	15	3	3
18	<i>L. waywayandensis</i> DSM 44232 ^T	8	4	6	7	6	8	5	9	6	6	4	6	12	5	8	8	3	15	10
19	<i>L. xinjiangensis</i> CGMCC 4.3525 ^T	7	4	7	6	6	8	6	8	7	6	3	7	10	6	8	7	3	10	15

Cells that have the maximum value (self-match) in each row are colored dark gray. Cells that have the second maximum value in each row are colored light gray.

式菌株 NBRC 13195^T 有 7 种核糖体蛋白质(L35、L31、S18、L29、S20、S15 和 S19)发生分子量变化, 只有 8 种核糖体蛋白质匹配模式菌株。以上结果表明本研究选取的模式菌株可代表菌种内多数菌株。

基于对不同菌种模式菌株间及菌种内不同菌株间的核糖体蛋白质匹配分析将鉴定至菌种的标准设为核糖体蛋白质匹配数 ≥ 14 。如测试菌株与表 1 中某模式菌株的核糖体蛋白质匹配数 ≥ 14 , 则鉴定为该模式菌株所属菌种。需注意两点: 一是 *L. alba* 和“*L. indica*”的模式菌株只有 10 种核糖体蛋白质的分子量信息, 无法满足鉴定至菌种的匹配数标准; 二是菌种内某菌株

与模式菌株有 2 种及更多核糖体蛋白质出现差异时(如 *L. aerocolonigenes* NRRL B-16140), 按上述标准无法鉴定至菌种。

2.3 伦茨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

从基因组数据库可检索到束丝放线菌属 2 个种及 2 个亚种和糖丝菌属 13 个种, 数据已提交国家微生物科学数据中心, 编号: NMDCX0000147。分子量比对发现这 2 个属模式菌株的 L36 和 S19 均具有与伦茨菌属模式菌株一致的分子量(4 403.38 和 10 507.27), 糖丝菌属模式菌株的 L33 也具有与伦茨菌属模式菌株一致的分子量(6 363.33), 表明这些分子量并

非伦茨菌属特有。同种核糖体蛋白质分子量接近的情况较普遍, 如 L34 (5 319.25、5 320.23 和 5 320.24)、L30 (6 850.96 和 6 849.97) 和 S14-Z (6 872.23 和 6 872.27) 等。此外, 不同核糖体蛋白质分子量接近的情况偶有出现, 如束丝放线菌的 S18 (8 874.42) 和伦茨菌的 L27 (8 873.15)。因此, 计算匹配数时将相差 $\pm 300 \times 10^{-6}$ 的分子量可能引起的错误匹配包括在内。

束丝放线菌和糖丝菌与伦茨菌属各模式菌株 15 种核糖体蛋白质的匹配结果见表 3。表 3 中束丝放线菌与伦茨菌属各模式菌株的核糖体蛋白质最大匹配数为 3, 糖丝菌与伦茨菌属各模式菌株的核糖体蛋白质最大匹配数为 5。基于表 3 将鉴定至伦茨菌属的标准设为核糖体蛋白质匹配数 ≥ 6 。如测试菌株与表 1 中模式菌株的核糖体蛋白质匹配数 ≥ 6 , 则鉴定为伦茨菌属。

2.4 伦茨菌属测试菌株的 MALDI 质谱峰匹配及菌株鉴定

为测试鉴定标准对伦茨菌属菌种的鉴定结果, 对伦茨菌属的 7 个菌株进行分析, 其 MALDI-TOF MS 图谱见图 1, 与伦茨菌属 19 个模式菌株核糖体蛋白质匹配的峰数量统计见表 4, 详细匹配信息和质谱峰列表已提交国家微生物科学数据中心(编号: NMDCX0000147)。

2.4.1 伦茨菌属菌株鉴定

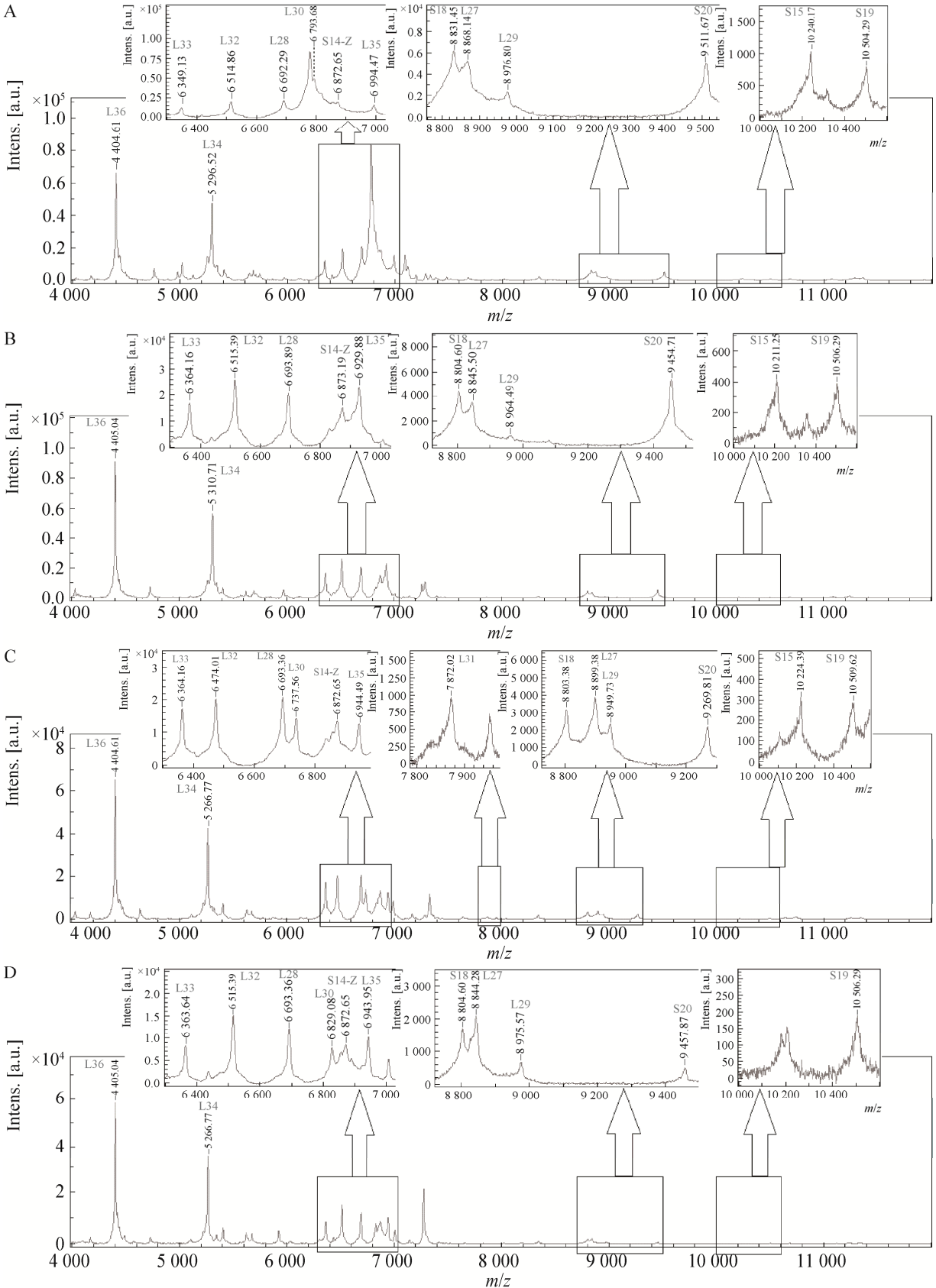
菌株 CGMCC 4.7203^T 与表 1 中 10 号菌种的核糖体蛋白质匹配数最大(14), 按 2.2 部分鉴定至菌种的标准鉴定为 *L. guizhouensis*, 鉴定结果准确。菌株 CGMCC 4.7319^T 与表 1 中 15 号菌种的核糖体蛋白质匹配数最大(13), 但未满足鉴定至菌种的标准。2 个测试菌株均为模式菌株, 核糖体蛋白质理论匹配数为 15, 而 CGMCC

表 3 束丝放线菌和糖丝菌属菌株与伦茨菌属 19 个模式菌株 15 种核糖体蛋白质的匹配数

Table 3 Count of ribosomal proteins from *Actinosynnema* and *Saccharothrix* strains matched to 15 ribosomal proteins from 19 *Lentzea* type strains

Organism	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
<i>A. mirum</i> DSM 43827 ^T	2	1	2	2	2	2	3	2	2	2	1	2	3	3	2	2	2	3	3
<i>A. pretiosum</i> X47 ^T	2	1	2	2	2	2	3	2	2	2	1	2	3	3	2	2	2	3	3
<i>A. pretiosum</i> ssp. <i>auranticum</i> DSM 44131 ^T	1	0	1	1	1	1	2	1	1	1	0	1	2	2	1	1	1	2	2
<i>A. pretiosum</i> ssp. <i>pretiosum</i> ATCC 31280 ^T	2	1	2	2	2	2	3	2	2	2	1	2	3	3	2	2	2	3	3
<i>S. algeriensis</i> DSM 44581 ^T	3	2	4	3	3	3	2	3	3	3	2	3	3	2	3	3	2	3	3
<i>S. australiensis</i> DSM 43800 ^T	2	0	1	1	3	2	2	2	2	1	0	2	3	2	2	2	2	3	3
<i>S. carnea</i> CGMCC 4.7097 ^T	3	1	3	2	3	3	5	3	4	2	1	3	3	5	3	3	3	3	3
<i>S. coeruleofusca</i> DSM 43679 ^T	3	2	4	3	2	3	2	3	3	3	2	3	3	2	3	3	3	3	3
<i>S. deserti</i> BMP B8144 ^T	3	1	2	2	3	3	5	3	4	2	1	3	3	5	3	3	3	3	3
<i>S. ecbatanensis</i> DSM 45486 ^T	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0
<i>S. espanaensis</i> DSM 44229 ^T	3	1	2	3	4	3	3	3	3	2	2	3	3	3	3	3	3	4	3
<i>S. saharensis</i> DSM 45456 ^T	4	2	4	3	3	4	5	4	5	3	2	4	4	5	4	4	4	4	4
<i>S. syringae</i> NRRL B-16468 ^T	3	1	2	2	3	3	3	3	3	2	1	3	3	3	3	3	3	3	3
<i>S. tamanrassetensis</i> CECT 8640 ^T	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1
<i>S. texasensis</i> DSM 44231 ^T	2	0	1	1	2	2	3	2	2	2	0	2	2	3	2	2	2	2	2
<i>S. variisporae</i> DSM 43911 ^T	2	1	2	2	2	2	2	2	2	2	1	2	2	2	2	2	5	2	2
<i>S. violaceirubra</i> DSM 45084 ^T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Cells that have the maximum value are colored gray.



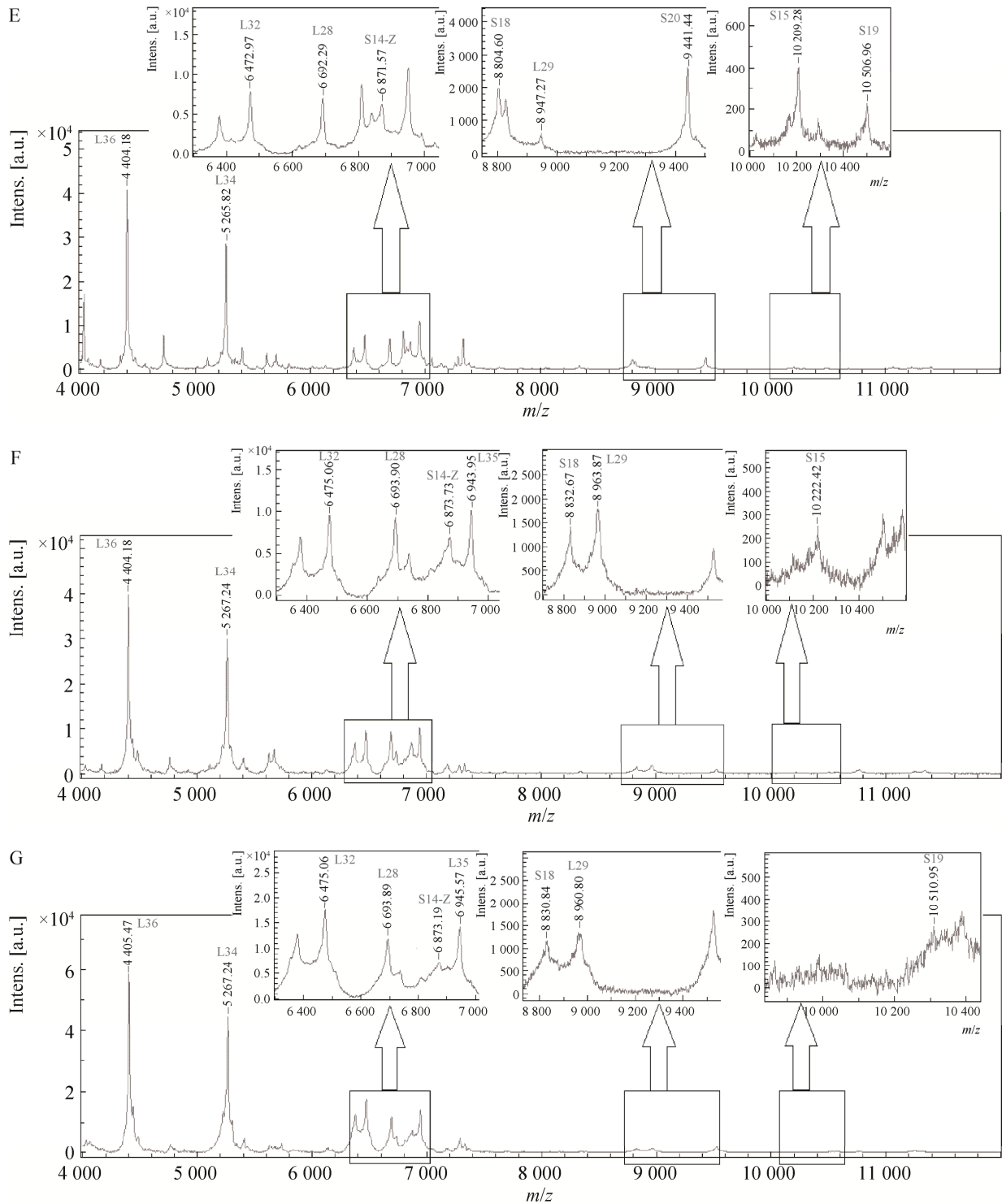


图 1 伦茨菌属测试菌株的 MALDI-TOF MS 图谱

Figure 1 MALDI-TOF mass spectra of *Lentzea* test strains. A: *L. guizhouensis* CGMCC 4.7203^T. B: *L. pudingi* CGMCC 4.7319^T. C: *L. soli* CCTCC AA 2017027^T. D: *Lentzea* sp. CCTCC AA 2016016. E: *Lentzea* sp. CCTCC AA 2017024. F: *Lentzea* sp. 2-8. G: *Lentzea* sp. 3-5. Peaks that match ribosomal biomarkers are marked.

表 4 伦茨菌属测试菌株的 MALDI 质谱峰与伦茨菌属 19 个模式菌株的核糖体蛋白质匹配数

Table 4 Count of MALDI spectrum peaks from *Lentzea* test strains matched to ribosomal proteins from 19 *Lentzea* type strains

Organism	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
<i>L. guizhouensis</i> CGMCC 4.7203 ^T	5	3	7	4	2	6	4	4	4	14	1	7	4	3	5	3	1	5	5
<i>L. pudingi</i> CGMCC 4.7319 ^T	5	3	6	7	5	9	5	10	6	5	4	7	6	4	13	6	5	7	8
<i>L. soli</i> CCTCC AA 2017027 ^T	9	4	8	6	9	8	4	8	9	4	4	9	9	5	6	14	4	8	7
<i>Lentzea</i> sp. CCTCC AA 2016016 ^T	6	3	8	6	5	13	4	8	7	6	4	10	7	4	10	8	5	7	7
<i>Lentzea</i> sp. CCTCC AA 2017024 ^T	7	4	8	7	6	7	3	8	5	4	4	6	8	3	6	8	4	9	7
<i>Lentzea</i> sp. 2-8 ^T	7	2	5	5	5	5	4	6	4	5	1	6	7	4	4	7	4	6	5
<i>Lentzea</i> sp. 3-5 ^T	7	3	4	5	5	5	5	6	4	4	2	5	8	4	4	7	3	6	6

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

4.7203^T的质谱缺失 L31 的信号(图 1A),CGMCC 4.7319^T的质谱缺失 L30 和 L31 的信号(图 1B),提示我们菌株实际测试可能出现少数核糖体蛋白质信号缺失的情况。

菌株 CCTCC AA 2017027^T与表 1 中 16 号菌种的核糖体蛋白质匹配数最大(14),鉴定为 *L. terrae*,然而该菌株实际为 *L. soli*,鉴定结果错误。在伦茨菌属菌种 16S rRNA 基因序列的系统进化树中 *L. terrae* NEAU-LZS 42^T(MG271812)和 *L. soli* CCTCC AA 2017027^T(MF784397)位于同一分支^[7],16S rRNA 基因序列比对发现其相似度为 99.14%,表明 *L. terrae* 和 *L. soli* 的亲缘关系很近。目前尚未有 *L. soli* 的基因组测序菌株信息,表 1 中缺少该菌种模式菌株的核糖体蛋白质分子量,因此测试菌株错误鉴定为亲缘关系最近的 *L. terrae*。

菌株 CCTCC AA 2017027^T与 *L. terrae* NEAU-LZS 42^T未匹配的一种核糖体蛋白质是 L30 (6 999.24),但其质谱中出现与 *L. alba* NEAU-D13^T的 L30 (6 837.92)匹配的 *m/z* 6 838.24 以及与 *L. atacamensis* DSM 45479^T的 L30 (6 736.85)匹配的 *m/z* 6 737.56(图 1C)。L30 的系统进化树中分子量 6 736.85 和 6 999.24 的序列位于同一分支且相似度高(数据已提交家微生物科学数据中心,编号:NMDCX0000147),而分子量 6 837.92 和 6 999.24 的序列位于距离远的

不同分支且相似度低,因此推测 *m/z* 6 737.56 为 *L. soli* 的 L30 信号,可以通过 L30 的质谱峰区分 *L. soli* 与 *L. terrae* 两个菌种的模式菌株。

菌株 CCTCC AA 2016016 与表 1 中 06 号菌种的核糖体蛋白质匹配数最大(13),菌株 CCTCC AA 2017024 与表 1 中 18 号菌种的核糖体蛋白质匹配数最大(9),菌株 2-8 与表 1 中 01、13 和 16 号菌种的核糖体蛋白质匹配数最大(7),菌株 3-5 与表 1 中 13 号菌种的核糖体蛋白质匹配数最大(8),均未满足 2.2 部分鉴定至菌种的标准,可按 2.3 部分鉴定至属的标准(匹配数 \geq 6)鉴定为伦茨菌属菌种(*Lentzea* sp.)。

7 个测试菌株全部满足鉴定至伦茨菌属的标准(最大匹配数 \geq 6),鉴定为伦茨菌,表明按该标准鉴定伦茨菌属的可行性较好。实际测试表明伦茨菌属内亲缘关系很近的 *L. soli* 和 *L. terrae* 的模式菌株间核糖体蛋白质匹配数为 14,提示我们鉴定至菌种的标准需修正为核糖体蛋白质匹配数为 15,以确保鉴定结果准确。然而按修正后的标准,7 个测试菌株无法鉴定至菌种,表明按该标准鉴定伦茨菌属菌种的可行性较差。

2.4.2 伦茨菌属菌株核糖体蛋白质的质谱峰强度分析

质谱图中除了 *m/z*,峰强度(intensity)也是一个重要参数。图 1 显示 7 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, 以其为基峰, L34 次之(相对强度 60%–80%), 其次是 m/z 为 6 300–7 000 的 6 种核糖体蛋白质峰(相对强度 10%–30%), 之后是 m/z 为 8 800–9 600 的 4 种核糖体蛋白质峰(相对强度 1%–10%), m/z 为 10 000 以上的峰强度最低(相对强度 <1%), 这种峰强度模式可作为鉴定伦茨菌属菌种的必要条件。此外, 伦茨菌属 L36 的分子量一致且在 MALDI 质谱中以高信号强度呈现, 是特征最明显的核糖体蛋白质标志物。

2.5 伦茨菌属外菌种的 MALDI 质谱峰匹配

为测试鉴定标准对伦茨菌属外的放线菌纲菌种的鉴定结果, 对放线菌纲的 80 个代表菌种进行分析, 假诺卡氏菌目的测试菌株与伦茨菌属 19 个模式菌株核糖体蛋白质匹配的峰数量统计见表 5, 另外 7 个目的测试菌株的峰数量统计数据已提交国家微生物科学数据中心(编号: NMDCX0000147)。

对照菌株 *L. flava* IMET 9748T HKJ 与表 1

中各模式菌株的核糖体蛋白质最大匹配数为 9, 满足鉴定至伦茨菌属的标准(匹配数 ≥ 6), 鉴定为伦茨菌, 属水平鉴定结果准确。

表 5 显示假诺卡氏菌目的 8 个测试菌株中有 7 个与表 1 中各模式菌株的核糖体蛋白质最大匹配数小于 6, 不满足鉴定至伦茨菌属的标准, 但有 1 个菌株(红色糖多孢菌 *Saccharopolyspora erythraea* HKI 184 HKJ)与表 1 中各模式菌株的核糖体蛋白质最大匹配数为 6, 满足鉴定至伦茨菌属的标准, 属水平鉴定结果错误。从 NCBI 基因组数据库选取菌株 *S. erythraea* NRRL 2338^T 分析发现, 其并无与伦茨菌一致的核糖体蛋白质序列。菌株 HKI 184 HKJ 的质谱中与伦茨菌属核糖体蛋白质匹配的 6 个 m/z 中有 3 个为 S18、L27 和 S15 产生的随机错配, 3 个为未知蛋白质产生的随机错配。显然, 随机错配数过多导致了错误鉴定。分析菌株 HKI 184 HKJ 质谱发现 m/z 4 402.49 的相对强度仅为 20%, 且缺少 m/z 5 000–7 000 的核糖体蛋白质峰, 明显不同于伦茨菌属质谱峰强度模式, 以此为依据可排除为伦茨菌的可

表 5 假诺卡氏菌目测试菌株的 MALDI 质谱峰与伦茨菌属 19 个模式菌株的核糖体蛋白质匹配数

Table 5 Count of MALDI spectrum peaks from *Pseudonocardiales* test strains matched to ribosomal proteins from 19 *Lentzea* type strains

Organism	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
<i>Lentzea flava</i> IMET 9748T HKJ ^T	1	2	2	2	4	1	9	2	3	2	0	2	1	8	2	1	2	1	2
<i>Saccharothix mutabilis</i> ssp. <i>mutabilis</i> HKI 9771 HKJ ^T	4	2	3	2	5	2	4	2	4	2	2	3	2	5	3	5	4	2	2
<i>Saccharothix mutabilis</i> ssp. <i>capreolus</i> HKI 264 HKJ ^T	2	0	1	1	2	1	1	2	1	1	1	1	2	1	2	3	2	2	1
<i>Pseudonocardia hydrocarbonoxydans</i> DSM 43281 ^T	2	1	3	1	1	2	1	1	2	3	0	3	1	2	2	1	2	1	2
<i>Saccharopolyspora erythraea</i> HKI 184 HKJ ^T	6	2	3	2	3	3	3	3	4	4	1	3	4	5	3	4	2	3	4
<i>Saccharopolyspora hirsuta</i> HKI 9709 HKJ ^T	1	0	2	1	2	1	2	1	1	3	0	1	1	2	1	2	2	1	1
<i>Amycolatopsis alba</i> DSM 44262 ^T	1	1	2	1	1	2	1	2	1	0	1	1	3	1	1	4	0	4	3
<i>Amycolatopsis balhimycina</i> DSM 44591 ^T	1	0	1	1	0	2	0	1	1	1	0	2	1	0	1	1	2	1	1
<i>Amycolatopsis mediterranei</i> DSM 40773 ^T	1	2	3	1	1	1	0	1	3	1	3	1	2	0	1	2	1	2	2

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

能。综上所述,假诺卡氏菌目内与伦茨菌属亲缘关系较近类群有错误鉴定为伦茨菌的可能,但结合质谱峰强度模式可降低发生错误鉴定的几率。

假诺卡氏菌目外的 7 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:NMDCX0000147)与表 1 中各模式菌株的核糖体蛋白质最大匹配数小于 5,不满足鉴定至伦茨菌属的标准。这一结果表明与伦茨菌属亲缘关系更远的假诺卡氏菌目外的放线菌类群一般不会错误鉴定为伦茨菌。

基于以上分析将鉴定至伦茨菌属的标准修正为与表 1 中各模式菌株的核糖体蛋白质最大匹配数 ≥ 6 ,同时核糖体蛋白质的峰强度模式符合伦茨菌属特征模式。按修正后的标准可降低随机错配过多将非伦茨菌属菌株错误鉴定为伦茨菌的几率。

综上所述,对目标属和非目标属测试菌株的分析表明:(1) 伦茨菌属菌株与表 1 中模式菌株的核糖体蛋白质最大匹配数均大于 6;(2) 伦茨菌属外的假诺卡氏菌目菌株与表 1 中模式菌株的核糖体蛋白质最大匹配数均不大于 6;(3) 假诺卡氏菌目外的放线菌纲菌株与表 1 中模式菌株核糖体蛋白质的最大匹配数均小于 5,以上结果为建立核糖体蛋白质最大匹配数法鉴定伦茨菌属奠定了基础。

2.6 伦茨菌属菌种鉴定方法建立

本文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伦茨菌属菌种的鉴定方法,具体步骤如下:(1) 构建伦茨菌属菌种模式菌株 15 种核糖体蛋白质的理论分子量表(表 1), (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 1 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 ≥ 6 且核糖体蛋白质的质谱峰强度模式符合伦茨菌属特征模式,则鉴定为伦茨菌属;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为 15。

注意事项包括:(1) 伦茨菌属现有 23 个有效描述种,NCBI 基因组数据库收录了 18 个,表 1 中缺少 5 种伦茨菌(*L. californiensis*, *L. chajnantorensis*, *L. rhizosphaerae*, *L. roselyniae* 和 *L. soli*)的核糖体蛋白质分子量信息,这些菌种的菌株无法鉴定至种。表 1 中 *L. alba* 和“*L. indica*”只有 10 种核糖体蛋白质的分子量信息,这 2 个菌种的菌株也无法鉴定至种。此外,与表 1 中各菌种模式菌株有核糖体蛋白质差异的菌株及质谱中有核糖体蛋白质信号缺失的菌株无法鉴定至种。上述原因导致本方法鉴定菌种存在应用上的局限性。(2) 按核糖体蛋白质匹配数为 15 的标准鉴定菌种可确保鉴定准确性,但一些菌株无法鉴定至种。降低鉴定标准,将核糖体蛋白质匹配数设为 ≥ 14 ,一些菌株可以鉴定至种,但亲缘关系很近的菌种可能鉴定错误。(3) 本方法鉴定伦茨菌属的效果较好。筛选伦茨菌属菌株时降低匹配数标准,将核糖体蛋白质匹配数设为 ≥ 5 ,可提高伦茨菌的筛选率,但少数非伦茨菌可能错误鉴定为伦茨菌,需通过 16S rRNA 基因序列比对法鉴定确认。

3 讨论与结论

MALDI Biotyper 数据库(Bruker MBT DB7854)仅收录了一种伦茨菌(*L. flava*)的指纹图谱,无法满足对伦茨菌属菌种的鉴定需求。按照布鲁克公司提供的标准方法构建伦茨菌属菌种的图谱数据库,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。本文方法建立在伦

茨菌属菌种基因组数据挖掘基础之上, 操作简单, 对技术要求低, 成本低廉, 可用于鉴定伦茨菌属菌种, 扩大 MALDI-TOF 质谱法菌种鉴定的应用范围。

本文方法鉴定至菌种水平的可行性较差, 主要原因有 3 个: (1) MALDI-TOF MS 图谱中缺少相关核糖体蛋白质信号。比如, L31 的 m/z 范围为 7 700–7 900, 而 5 个测试菌株的质谱中该 m/z 范围内均没有峰[数据已提交国家微生物科学数据中心(编号: NMDCX0000147)]。(2) 伦茨菌属测序菌种不全, 测序菌株数量有限, 一些核糖体蛋白质序列未知, 无法推算分子量, 因此无法识别其信号。(3) MALDI-TOF MS 图谱中存在相关核糖体蛋白质信号, 表 1 中也有与之接近的分子量, 但因实测 m/z 超出理论 m/z 误差范围而被排除。核糖体蛋白质峰的信噪比偏低或其附近有其他蛋白质峰(线性模式分辨率低, 无法实现基线分离)时会出现这种情况。

针对上述制约因素可从 3 个方面进行改进: (1) 在甲酸提取法的基础上优化菌株蛋白提取方法, 尽可能使 15 种目标核糖体蛋白质都出现可检测的质谱信号。(2) 将新测序的伦茨菌属菌种(株)15 种核糖体蛋白质的分子量数据补充至表 1, 为更多菌种的核糖体蛋白质质谱峰匹配提供依据。此外, 序列比对和系统进化分析可为核糖体蛋白质的质谱峰识别提供线索。本研究通过系统进化分析辅助识别了菌株 CCTCC AA 2017027^T L30 的 MALDI 质谱峰, 通过序列比对找出氨基酸突变位点同样能辅助识别核糖体蛋白质的质谱峰。比如, 菌株 CCTCC AA 2017024、2-8 和 3-5 的质谱中均未识别出 L33 的质谱峰, 但分别出现 m/z 6 378.68、6 378.16 和 6 378.68。基于伦茨菌属菌株 L33 的序列比对发现发生突变的氨基酸均为 I (数据已提交国家微生物科学数据中心,

编号: NMDCX0000147)), 若发生 I→K 或 I→Q 突变蛋白质分子量增加约 15 Da, 因此这些 m/z 理论上可能是 L33 的质谱峰。基于序列比对和系统进化分析的预测可通过菌株的基因组测序和分析结果进行验证。(3) 菌株质谱中核糖体蛋白质峰的信噪比偏低或其附近有其他蛋白质峰时, 将质量误差范围由 $\pm 300 \times 10^{-6}$ 提高至 $\pm 500 \times 10^{-6}$, 可识别出对应的核糖体蛋白质信号。因此, 本文方法在菌种水平的鉴定准确度仍然具有提升空间。

本研究发现伦茨菌属菌株 L36 的 MALDI-TOF 质谱峰以高信号强度稳定出现, 因此质谱中出现接近 4 404.38 (L36) 的 m/z 可作为鉴定伦茨菌的必要条件。如测试样本质谱中没有 m/z 位于 4 404.38 误差范围内则判定为非伦茨菌, 有 m/z 位于 4 404.38 误差范围内再按本文方法鉴定。对放线菌纲 57 个属共 80 个菌种的 MALDI-TOF 质谱峰列表分析发现, 其中只有 10 个属(*Saccharothrix*、*Pseudonocardia*、*Saccharopolyspora*、*Kitasatospora*、*Streptomyces*、*Cellulomonas*、*Kytococcus*、*Dermatophilus*、*Cellulosimicrobium* 和 *Oerskovia*) 12 个菌种的质谱中有 m/z 位于 4 404.38 误差范围内, 表明以 L36 一种核糖体蛋白质为标志物筛选伦茨菌可排除 82% (47 个) 的非目标属, 显著提高筛选效率, 提示我们 L36 可在放线菌快速分类中发挥重要作用。

长期以来, 分子探针法一直是定向筛选目标细菌的重要方法, 通过设计目标类群 16S rRNA 基因序列的特异性引物, 从菌株基因组 DNA 中扩增目标基因片段筛选目标细菌。一些放线菌类群的分子探针已有报道^[26-28], 但伦茨菌属放线菌的分子探针迄今未见报道。按本文方法鉴定伦茨菌, 属水平的鉴定准确度高, 因此提供了以核糖体蛋白质为标志物定向筛选目

标细菌的思路。以核酸为靶标的分子探针法提取菌株基因组 DNA, 使用基因扩增仪进行 PCR 扩增, 而以核糖体蛋白质为靶标的 MALDI-TOF 质谱法提取菌株蛋白, 使用 MALDI-TOF 质谱仪采集质谱。相比 DNA 提取常用的酶解法, 蛋白提取使用的甲酸提取法操作更简单; 相比 PCR 扩增, 质谱仪采集数据更快速。目前分析图谱、从图谱中获得结果比较复杂费时, 可尝试开发配套软件用于待测菌株的质谱峰与目标属各模式菌株的核糖体蛋白质分子量比对, 快速获得匹配结果。因此, 以核糖体蛋白质为标志物, 通过 MALDI-TOF 质谱法高通量筛选目标细菌有可能在未来展现出应用价值。

4 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 技术鉴定伦茨菌属放线菌的方法, 可为放线菌纲其他类群的鉴定提供借鉴。

致谢

感谢江苏师范大学生命科学学院曹成亮高级实验师无偿提供 2 个模式菌株支持本研究。

参考文献

- [1] YASSIN AF, RAINEY FA, BRZEZINKA H, JAHNKE KD, WEISSBRODT H, BUDZIKIEWICZ H, STACKEBRANDT E, SCHAAL KP. *Lentzea* gen. nov., a new genus of the order *Actinomycetales*[J]. *International Journal of Systematic Bacteriology*, 1995, 45(2): 357-363.
- [2] LI CS, HU YF, WU XH, STUMPF SD, QI YC, D'ALESSANDRO JM, NEPAL KK, SAROTTI AM, CAO SG, BLODGETT JAV. Discovery of unusual dimeric piperazyl cyclopeptides encoded by a *Lentzea flaviverrucosa* DSM 44664 biosynthetic supercluster[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119(17): e2117941119.
- [3] HUSSAIN A, RATHER MA, SHAH AM, BHAT ZS, SHAH A, AHMAD Z, PARVAIZ HASSAN Q. Antituberculous activity of *Actinobacteria* isolated from the rare habitats[J]. *Letters in Applied Microbiology*, 2017, 65(3): 256-264.
- [4] WICHNER D, IDRIS H, HOUSSEN WE, MCEWAN AR, BULL AT, ASENJO JA, GOODFELLOW M, JASPARS M, EBEL R, RATEB ME. Isolation and anti-HIV-1 integrase activity of lentzeosides A-F from extremotolerant *Lentzea* sp. H45, a strain isolated from a high-altitude Atacama Desert soil[J]. *The Journal of Antibiotics*, 2017, 70(4): 448-453.
- [5] YABUTANI T, TSUJIMOTO M, OHIRA S, SHIMIZU S, NAKANO H. Strain improvement of *Lentzea* sp. 7887 for higher yield per unit volume on hydroxylation of cyclosporine derivative FR901459[J]. *Bioscience, Biotechnology, and Biochemistry*, 2017, 81(7): 1456-1459.
- [6] NAKANO H, ŌMURA S. Chemical biology of natural indolocarbazole products: 30 years since the discovery of staurosporine[J]. *The Journal of Antibiotics*, 2009, 62(1): 17-26.
- [7] 高源, 曹成亮, 李荣鹏, 蒋继宏. 伦茨菌属放线菌的研究进展[J]. *微生物学报*, 2021, 61(8): 2294-2305. GAO Y, CAO CL, LI RP, JIANG JH. Recent advance on the genus *Lentzea*[J]. *Acta Microbiologica Sinica*, 2021, 61(8): 2294-2305 (in Chinese).
- [8] SINGHAL N, KUMAR M, KANAUIA PK, VIRDI JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis[J]. *Frontiers in Microbiology*, 2015, 6: 791.
- [9] PATEL R. MALDI-TOF MS for the diagnosis of infectious diseases[J]. *Clinical Chemistry*, 2015, 61(1): 100-111.
- [10] 戴颖欣, 李敏. MALDI-TOF MS 在临床微生物检验中的应用[J]. *检验医学*, 2015, 30(2): 102-107. DAI YX, LI M. Application of MALDI-TOF MS technology in clinical microbial determination[J]. *Laboratory Medicine*, 2015, 30(2): 102-107 (in Chinese).
- [11] 李楠, 刘文森, 刘林娜, 郝镗, 李树民, 孟轲音, 何扬, 万忠海, 王承宇, 李吉平. MALDI-TOF 质谱技术及 Biotyper 数据库在养殖场空气微生物鉴定中的应用[J]. *中国兽医学报*, 2014, 34(12): 2031-2034. LI N, LIU WS, LIU LN, HAO Z, LI SM, MENG KY, HE Y, WAN ZH, WANG CY, LI JP. The application of the MALDI-TOF MS and biotyper database in the

- identification of farm airborne microorganisms[J]. Chinese Journal of Veterinary Science, 2014, 34(12): 2031-2034 (in Chinese).
- [12] 顾春华. MALDI-TOF MS 技术及其在食品微生物检测方面的应用[J]. 中国酿造, 2019, 38(9): 24-27.
- GU CH. MALDI-TOF MS technique and its application challenges in food microbiology detection[J]. China Brewing, 2019, 38(9): 24-27 (in Chinese).
- [13] 杨祖顺, 汤晓召, 邹颜秋硕, 国译丹, 杨庆文. 基质辅助激光解吸电离飞行时间质谱法鉴定食品中金黄色葡萄球菌[J]. 食品安全质量检测学报, 2017, 8(10): 3710-3714.
- YANG ZS, TANG XZ, ZOU YQS, GUO YD, YANG QW. Identification of *Staphylococcus aureus* in food by matrix assisted laser desorption ionization-time of flight mass spectrometry[J]. Journal of Food Safety & Quality, 2017, 8(10): 3710-3714 (in Chinese).
- [14] DEMIREV PA, HO YP, RYZHOV V, FENSELAU C. Microorganism identification by mass spectrometry and protein database searches[J]. Analytical Chemistry, 1999, 71(14): 2732-2738.
- [15] PINEDA FJ, LIN JS, FENSELAU C, DEMIREV PA. Testing the significance of microorganism identification by mass spectrometry and proteome database search[J]. Analytical Chemistry, 2000, 72(16): 3739-3744.
- [16] 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.
- [17] SUN LW, TERAMOTO K, SATO H, TORIMURA M, TAO H, SHINTANI T. Characterization of ribosomal proteins as biomarkers for matrix-assisted laser desorption/ionization mass spectral identification of *Lactobacillus plantarum*[J]. Rapid Communications in Mass Spectrometry: RCM, 2006, 20(24): 3789-3798.
- [18] ZIEGLER D, POTHIER JF, ARDLEY J, FOSSOU RK, PFLUGER V, de MEYER S, VOGEL G, TONOLLA M, HOWIESON J, REEVE W, PERRET X. Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS[J]. Applied Microbiology and Biotechnology, 2015, 99(13): 5547-5562.
- [19] CHENG D, QIAO L, HORVATOVICH P. Toward spectral library-free matrix-assisted laser desorption/ionization time-of-flight mass spectrometry bacterial identification[J]. Journal of Proteome Research, 2018, 17(6): 2124-2130.
- [20] KOSTAS J, PARKER KC. Using matrix-assisted laser desorption/ionization time of flight spectra to elucidate species boundaries by matching to translated DNA databases[J]. Journal of the American Society for Mass Spectrometry, 2020, 31(1): 73-84.
- [21] 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).
- [22] THOMPSON JD, GIBSON TJ, PLEWNIAK F, JEANMOUGIN F, HIGGINS DG. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Research, 1997, 25(24): 4876-4882.
- [23] FELSENSTEIN J. Evolutionary trees from DNA sequences: A maximum likelihood approach[J]. Journal of Molecular Evolution, 1981, 17(6): 368-376.
- [24] TAMURA K, STECHER G, PETERSON D, FILIPSKI A, KUMAR S. MEGA6: Molecular evolutionary genetics analysis version 6.0[J]. Molecular Biology and Evolution, 2013, 30(12): 2725-2729.
- [25] GASTEIGER E, HOOGLAND C, GATTIKER A, DUVAUD S, WILKINS MR, APPEL RD, BAIROCH A. Protein Identification and Analysis Tools on the ExPASy Server[A]//The Proteomics Protocols Handbook[M]. Totowa, NJ: Humana Press, 2005: 571-607.
- [26] QIU DH, RUAN JS, HUANG Y. Selective isolation and rapid identification of members of the genus *Micromonospora*[J]. Applied and Environmental Microbiology, 2008, 74(17): 5593-5597.
- [27] ZHI XY, TANG SK, LI WJ, XU LH, JIANG CL. New genus-specific primers for the PCR identification of novel isolates of the genus *Streptomonospora*[J]. FEMS Microbiology Letters, 2006, 263(1): 48-53.
- [28] MORON R, GONZALEZ I, GENILLOU O. New genus-specific primers for the PCR identification of members of the genera *Pseudonocardia* and *Saccharopolyspora*[J]. International Journal of Systematic Bacteriology, 1999, 49 Pt 1: 149-162.