

彭 凌, 林锦铨, 李玲慧, 等. 一株猪丹毒丝菌的全基因组测序及 *SpaA* 基因分析[J]. 江苏农业学报, 2021, 37(3): 694-698.
doi: 10.3969/j.issn.1000-4440.2021.03.018

一株猪丹毒丝菌的全基因组测序及 *SpaA* 基因分析

彭 凌, 林锦铨, 李玲慧, 刘博婷, 蔡巩林

(韶关学院英东生物与农业学院, 广东 韶关 512005)

摘要: 本研究测定猪丹毒丝菌临床毒株 SG7 的全基因组序列, 并运用生物信息学方法对测定的全基因组序列及猪丹毒丝菌表面保护性抗原 A 基因(*SpaA*)进行分析。SG7 菌株的基因组全长为 1 834 291.00 bp, G+C 含量为 36.3%, 基因总数 1 846 个。将 SG7 菌株与 GenBank 中 8 条完整的猪丹毒丝菌全基因组序列进行比较, 发现国内外不同菌株间基因组的基本信息存在不同程度差异。基于全基因组单核苷酸多态性(SNPs)的系统发育分析结果表明, 9 株菌株聚为 3 个分支, 国内菌株并不完全处于同一分支, SG7 菌株与国内常见菌株不属于同一进化分支。*SpaA* 基因高变区的分析结果显示, 9 株菌株亦可分为 3 个 *SpaA* 型, SG7 菌株为携带 Met203 的高致病性菌株。除 SG7 菌株外, 其他 8 株菌株的 *SpaA* 基因分型结果与基于全基因组 SNPs 分析的结果一致, 说明 SG7 菌株的遗传背景复杂。本研究结果可为猪丹毒丝菌基因组整体水平研究和疫苗的研发奠定基础。

关键词: 猪丹毒丝菌; 全基因组测序; 表面保护性抗原 A(*SpaA*)

中图分类号: S852.61 **文献标识码:** A **文章编号:** 1000-4440(2021)03-0694-05

Whole-genome sequencing and *SpaA* gene analysis of a *Erysipelothrix rhusiopathiae* strain

PENG Ling, LIN Jin-quan, LI Ling-hui, LIU Bo-ting, CAI Gong-lin

(Henry Fok College of Biology and Agriculture, Shaoguan University, Shaoguan 512005, China)

Abstract: In this study, the whole genome sequence of *Erysipelothrix rhusiopathiae* clinical strain SG7 was determined, and the whole genome sequence and the surface protective antigen A (*SpaA*) gene were analyzed using bioinformatics method. The full length of SG7 genome was 1 834 291.00 bp, and the total quantity of the predicted genes were 1 846 with a G+C content of 36.3%. Different degrees of differences of genomic basic information between different strains were found by comparing the complete genome sequences of SG7 and eight *E. rhusiopathiae* found in GenBank. Phylogenetic analysis based on genome-wide single nucleotide polymorphisms (SNPs) revealed that nine strains could be clustered into three distinct clades, and strains isolated from China were not in the same clade, while SG7 and common strains isolated from China did not belong to the same clade. The analysis results of hypervariable region of *SpaA* showed that nine strains could also be divided into three *SpaA* types, and SG7 was a highly pathogenic strain carrying Met203. Except for SG7, the *SpaA* genotyping results of other eight strains were consistent with the analytic results obtained based on the genome-wide SNPs analysis, which indicated that the genetic background of SG7 was complex. The results can provide data basis for the research at the whole genome level and development of *E. rhusiopathiae* vaccine.

Key words: *Erysipelothrix rhusiopathiae*; whole-genome sequencing; surface protective antigen A (*SpaA*)

收稿日期: 2020-09-07

基金项目: 广东省自然科学基金项目(2017A030307041); 国家级大学生创新创业训练计划项目(201910576002)

作者简介: 彭 凌(1975-), 男, 江西奉新人, 硕士, 副教授, 主要从事动物病原微生物及分子生物学研究。(E-mail) 308668576@qq.com

猪丹毒丝菌(*Erysipelothrix rhusiopathiae*)又称红斑丹毒丝菌, 属于丹毒丝菌属(*Erysipelothrix*), 呈纤细杆状, 是一种革兰氏阳性的人畜

共患病病原菌。猪丹毒是由猪丹毒丝菌引发的一种急性、热性传染病,能导致家禽、畜类的一系列皮肤感染及败血症。人类因外伤局部感染猪丹毒丝菌,会发生类丹毒,感染时手指出现斑点或肿起^[1-2]。在 20 世纪 80 至 90 年代,猪丹毒与猪肺疫、猪瘟并称中国养猪业三大传染病,给养猪业带来巨大的损失并影响人类健康^[3]。近年来,猪丹毒在中国多个省份均有发生,且发病率呈上升趋势^[4]。

目前预防和控制猪丹毒疫病的主要手段是接种疫苗,传统疫苗虽然对猪丹毒的预防起到一定作用,但仍有保护率低、保护周期短等缺点。细菌的毒力因子是人们研制新型疫苗的标靶^[5]。有研究结果表明,猪丹毒丝菌表面保护性抗原 A(*SpaA*)是其重要的毒力因子^[6-7],可以通过介导猪丹毒丝菌与猪内皮细胞的黏附而侵入宿主发挥作用^[7]。同时亦有很多研究结果证实了 *SpaA* 蛋白的 *N* 端区域为免疫保护区,对猪丹毒丝菌毒株有良好的免疫保护作用^[8-10]。全基因组测序是通过 DNA 测序仪对物种基因进行测序的一种生物信息学手段,全基因组测序分辨率高、分析速度快,为传染性疾病的调查提供了新的方法,对疫病监测具有重要的指导意义^[11]。本研究拟通过测定分离自广东省韶关市的猪丹毒丝菌毒力菌株 SG7^[12]的全基因序列,利用相关软件进行全基因组的注释分析,并分析 *SpaA* 的遗传多样性,以期对猪丹毒丝菌的后续研究提供有利的数据支撑,也为新型亚单位疫苗的开发和改良奠定理论基础。

1 材料与方法

1.1 菌株

从广东省韶关市某猪丹毒发病猪场分离获得猪丹毒丝菌菌株 SG7,通过动物试验鉴定为毒力菌株,半致死量(LD_{50})为 7.3×10^3 CFU^[12]。

1.2 基因组的测序与注释

将得到的猪丹毒丝菌 DNA 样品进行 Illumina HiSeq™ 测序,获得原始图像数据,再经过碱基识别得到原始测序数据。但鉴于 Illumina 测序过程中产生的错误会对最终的分析结果造成影响,为了数据的准确性,通过软件 FastQC 对测序数据质量进行可视化评估,再利用软件 Trimmomatic^[13] 过滤原始数据,以保证信息分析的准确性。使用软件 SPAdes^[14] 对得到

的测序数据进行拼接,然后对拼接得到的重叠群进行填补空隙。在修正拼接过程中,产生剪辑错误及小片段插入缺失会影响分析的准确性,因此需要采用软件 PrInSeS-G^[15] 进行序列矫正,得到最终的结果。拼接结果采用软件 Prokka^[16] 进行基因元件的分析预测,同时利用软件 RepeatModeler 鉴定 SG7 基因组上的重复序列。最后,采用软件 NCBI Blast+ 将蛋白质氨基酸序列与 COG、KEGG、VFDB 等数据库的数据进行比对,获得蛋白质功能注释信息。

1.3 *SpaA* 基因碱基序列分析

从测定的 SG7 菌株以及 GenBank 上公布的猪丹毒丝菌全基因组序列中提取 *SpaA* 基因碱基序列,并翻译为氨基酸序列,根据前人描述的方法^[17-18] 分析 *SpaA* 基因 432 bp 高变区对应蛋白质 *N* 端的氨基酸序列。

1.4 SG7 菌株与其他猪丹毒丝菌的比较基因组学分析

对 SG7 菌株与 GenBank 上公布的猪丹毒丝菌完整的全基因组序列进行比较基因组学分析,并构建基于全基因组单核苷酸多态性(SNPs)系统进化树图。

2 结果与分析

2.1 猪丹毒丝菌菌株 SG7 基因组的基本特征

表 1 显示,猪丹毒丝菌菌株 SG7 的基因组全长为 1 834 291.00 bp, G+C 含量为 36.30%。组分分析发现,基因总数为 1 846 个,碱基数为 1 685 442.00 个,基因平均长度为 913.02 bp,占基因组全长的 91.89%,蛋白质预测数为 1 755 个。采用软件 Aragorn 预测 tRNA,预测数为 55 个,采用软件 RNAmmer 预测 rRNA,预测数为 3 个。利用软件 RepeatModeler 对重复序列进行预测,SG7 菌株共有 107 个平均长度为 170.30 bp 的重复序列,其中 DNA 重复元件有 19 个,低复杂度序列有 12 个,微卫星 DNA 序列有 76 个。

2.2 猪丹毒丝菌菌株 SG7 全基因组的注释

在猪丹毒丝菌菌株 SG7 已经预测的 1 755 个基因中,有 1 316 个基因注释到 COG 数据库中^[19],占预测基因的 74.99%,将这些基因信息依据 COG 分类标准进行分类,可分为 19 类;有 1 079 个基因注释到 KEGG 数据库中^[20],占预测基因的 61.48%,可将这些基因划分为 24 类代谢通路。

表 1 猪丹毒丝菌 SG7 基因预测结果

Table 1 Prediction results of *Erysipelothrix rhusiopathiae* SG7 gene

类别	统计结果
基因组长度 (bp)	1 834 291.00
G+C 含量 (%)	36.30
基因总数	1 846.00
最小基因长度 (bp)	67.00
最大基因长度 (bp)	7 644.00
基因平均长度 (bp)	913.02
碱基数	1 685 442.00
基因比率 (%)	91.89
蛋白质预测数	1 755.00
tRNA 数	55.00
rRNA 数	3.00
重复区数目	107.00
重复区碱基数 (bp)	18 222.00
重复区比率 (%)	0.99
DNA 重复元件	19.00
低复杂度序列	12.00
微卫星 DNA 序列	76.00

VFDB^[21]是专门用于研究致病因子的数据库,包括 SetA 和 SetB 2 部分。将猪丹毒丝菌菌株 SG7 预测得到的基因碱基序列、蛋白质氨基酸序列与 VFDB 数据库中的数据进行比对,把基因和其相对应的毒力因子功能注释信息结合起来,有 108 个基因注释到 SetA 数据库,有 114 个基因注释到 SetB 数据库。

CARD^[22]是耐药基因数据库,能为药物作用研究以及环境治理提供研究依据。将猪丹毒丝菌菌株 SG7 预测得到的基因碱基序列、蛋白质氨基酸序列与 CARD 数据库中的数据进行比对,有 23 个基因注释

到 CARD 数据库。

PHI-base^[23]是病原与宿主互作数据库,用于寻找药物干预的靶基因。将猪丹毒丝菌菌株 SG7 预测得到的基因碱基序列、蛋白质氨基酸序列与 PHI-base 数据库中的数据进行比对,有 34 个基因注释到 PHI-base 数据库。

2.3 猪丹毒丝菌 *SpaA* 基因分析

当前 GenBank 上公布的猪丹毒丝菌完整基因组序列共有 8 条,从 SG7 和 GenBank 中的 8 株猪丹毒丝菌的基因组序列中提取 *SpaA* 基因 432 bp 高变区,并对相应蛋白质的 N 端进行分析,结果表明,9 个菌株中有 5 个菌株为 Met203/Ile257 型,3 个菌株为 Ile203/Ile257 型,1 个菌株为 Ile203/Leu257 型。

2.4 SG7 菌株与其他猪丹毒丝菌全基因组的比较分析

将 SG7 菌株与 GenBank 中 8 株猪丹毒丝菌菌株的全基因组基本信息进行比较,结果(表 2)显示,猪丹毒丝菌全基因组长度为 1 752 910~1 945 690 bp, G+C 含量维持在 36.3%~36.6%,基因数保持在 1 708~1 915 个,而蛋白质数为 1 390~1 801 个,表现出较大差异;tRNA 保持在 55 个左右,但不同菌株中的 rRNA 数量差异较大,SG7 的 rRNA 仅为 3 个,数量远低于其他菌株。9 个菌株的共有基因为 1 234 个,特有基因为 193 个。

将上述 8 株菌株与 SG7 菌株的全基因组进行比较分析,并构建系统进化树图(图 1),该进化树含有 3 个分支,SG7 菌株与英国分离株 NCTC8163、韩国分离株 KC-Sb-R1、中国南京分离株 SY1027 同处一个分支;而其他 4 个国内分离株处于第二分支;日本分离株 Fujisawa 单独为一个分支。

表 2 SG7 菌株与其他猪丹毒丝菌全基因组基本信息比较

Table 2 Comparison of genomic basic information between SG7 and other *Erysipelothrix rhusiopathiae*

菌株	来源	长度 (bp)	G+C 含量 (%)	rRNA 数	tRNA 数	基因数	蛋白质数
Fujisawa	日本	1 787 940	36.6	21	55	1 757	1 660
ZJ	中国成都	1 945 690	36.5	18	55	1 915	1 801
ML101	中国长沙	1 854 250	36.4	15	55	1 830	1 711
WH13013	中国武汉	1 778 060	36.5	15	55	1 745	1 654
GXBY-1	中国南宁	1 876 490	36.5	27	57	1 835	1 727
SY1027	中国南京	1 752 910	36.4	10	53	1 719	1 390
NCTC8163	英国	1 770 410	36.6	27	55	1 715	1 617
KC-Sb-R1	韩国	1 771 670	36.6	21	55	1 708	1 606
SG7	中国韶关	1 834 291	36.3	3	55	1 864	1 755

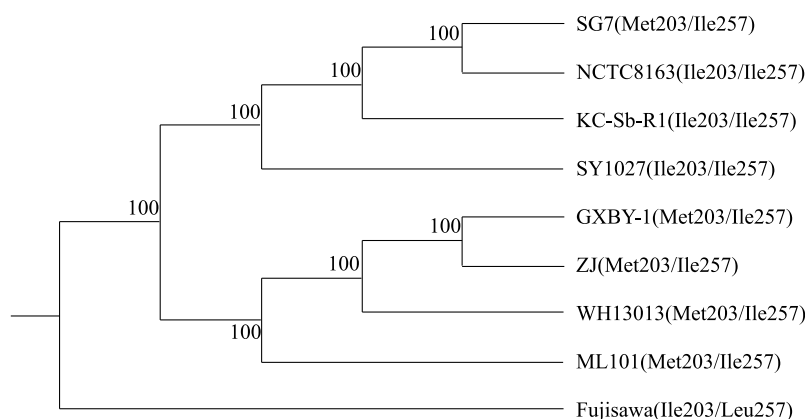


图 1 9 个猪丹毒丝菌菌株全基因组单核苷酸多态性 (SNPs) 的系统进化树

Fig.1 Phylogenetic tree of nine *Erysipelothrix rhusiopathiae* strains based on single nucleotide polymorphisms (SNPs) of whole genome sequence

3 讨论

猪丹毒作为一种人畜共患病,给养猪业带来巨大的经济损失并影响人类健康,该病的主要预防手段是疫苗接种,但传统灭活疫苗存在免疫效果不持久,不能产生细胞免疫等局限性^[24]。猪丹毒的药物治疗大多采用 β -内酰胺类抗生素,但过分依赖抗生素,会导致耐药性的出现以及药物残留的风险^[25-27],有地区就曾分离出猪丹毒丝菌青霉素耐药菌株^[28]。因此,开发更为有效的亚单位疫苗是科学防控猪丹毒的关键。尽管从1876年首次分离该菌至今已经过去了100多年,但我们对猪丹毒丝菌分子机制的了解仍然十分有限^[29-30]。2011年德国北部暴发大肠杆菌疫情,造成数千人感染,数十人死亡,但病原菌株的病原培养特性和细菌多位点序列分型 (MLST) 分析结果均与肠出血性大肠杆菌菌株相差甚远,有学者怀疑是新型的致病性大肠杆菌,经过全基因组测序分析发现,该菌株与肠聚集性大肠杆菌的亲缘关系更为接近,在进化上处于同一分支,但在进化过程中获得了志贺毒素编码基因 *stx2* 及多种耐药基因,使其获得了更强的生存能力,造成了病原菌的广泛传播^[31]。可见,全基因组分析对病原菌株的监测以及遗传变异检测有很大帮助,全基因组测序为我们从分子水平研究猪丹毒丝菌提供了便利。

本研究测定了猪丹毒丝菌菌株 SG7 临床毒株^[12]全基因组序列,并通过使用相关的生物信息学软件及数据库对其进行注释,揭示了该菌株的基因组结构特征,丰富了猪丹毒丝菌的基因组数据库。将 SG7 与 GenBank 上公布的 8 株猪丹毒丝菌完整基因组序列的基本信息进行比较,除 tRNA 数和 G+

C 含量比较一致外,基因组长度、基因数、蛋白质数、rRNA 数 SG7 菌株与国内外其他菌株存在不同程度的差异,这也许与不同地区饲养管理以及环境因素等存在差异有关。基于全基因组 SNPs 的系统发育分析结果表明,9 株菌株聚为 3 个分支,这与 Forde 等^[32]、Ogawa 等^[18]的研究结果一致。国内菌株并不完全处于同一分支,SG7 菌株与南京分离株 SY1027、英国分离株 NCTC8163、韩国分离株 KC-Sb-R1 为同一个分支,而南宁分离株 GXBY-1、成都分离株 ZJ、武汉分离株 WH13013、长沙分离株 ML101 这 4 个国内分离株则为另一分支,说明国内菌株至少有 2 个不同的进化途径,SG7 菌株与英国分离株 NCTC8163 关系最近。*SpaA* 高变区的氨基酸序列分析结果表明,9 株菌株可分为 3 个 *SpaA* 型,SG7 菌株和南宁分离株 GXBY-1、成都分离株 ZJ、武汉分离株 WH13013、长沙分离株 ML101 均为 Met203/Ile257 型。Uchiyama 等^[17]认为携带 Met203 的分离株为高致病性菌株。Ogawa 等^[18]对分离自日本的 34 株菌株进行分析,结果表明这些菌株的 *SpaA* 基因分型结果与基于全基因组 SNPs 分析获得的分类结果是一致的,本研究中除 SG7 菌株外,其他 8 个菌株亦如此。与 SG7 菌株处同一分支的其他菌株均为 Ile203/Ile257 型,SG7 菌株为携带 Met203 的高致病性菌株 (LD_{50} 为 7.3×10^3 CFU),但它与携带 Met203 的其他菌株不处于同一进化分支,因此推测 SG7 菌株在进化过程中也许与携带 Met203 的进化分支的菌株接触后进行了 *SpaA* 基因的重组交换。本研究结果为猪丹毒丝菌基因组整体水平研究和亚单位疫苗的开发提供了数据基础。

参考文献:

- [1] GORBY G L, PEACOCK J E. *Erysipelothrix rhusiopathiae* endocarditis: microbiologic, epidemiologic, and clinical features of an occupational disease[J]. Reviews of Infectious Diseases, 1988, 10(2): 317-325.
- [2] WANG Q, CHANG B J, RILEY T V. *Erysipelothrix rhusiopathiae* [J]. Veterinary Microbiology, 2010, 140(3/4): 405-417.
- [3] 周绪斌, 丹尼, 李聪, 等. 猪丹毒——古老的传染病是否从中国规模化猪场消失了? [J]. 今日养猪业, 2009(5): 22-24.
- [4] 李文春, 杨仕标, 李富祥. 一株猪丹毒杆菌的分离鉴定及 *SpaA* 基因序列分析[J]. 上海畜牧兽医通讯, 2020(3): 9-12.
- [5] 周作勇, 李和贤, 杨浩钺, 等. 伪结核棒状杆菌毒力因子的研究进展[J]. 中国人兽共患病学报, 2017, 33(12): 1115-1119.
- [6] BORRATHYBAY E, GONG F J, ZHANG L, et al. Role of surface protective antigen A in the pathogenesis of *Erysipelothrix rhusiopathiae* strain C43065[J]. Journal of Microbiology and Biotechnology, 2015, 25(2): 206-216.
- [7] HARADA T, OGAWA Y, EGUCHI M, et al. Phosphorylcholine and SpaA, a choline-binding protein, are involved in the adherence of *Erysipelothrix rhusiopathiae* to porcine endothelial cells, but this adherence is not mediated by the PAF receptor[J]. Veterinary Microbiology, 2014, 172(1/2): 216-222.
- [8] 吾鲁木汗·那孜尔别克, 张磊, 何翠, 等. 猪丹毒丝菌天然 *SpaA* 和重组 *SpaA-N* 免疫保护效果的评价[J]. 微生物学报, 2010, 50(3): 367-372.
- [9] 姚焱彬, 陆萍, 杨志鹏, 等. 猪丹毒丝菌 *SpaA* 基因原核表达及表达蛋白质的免疫原性分析[J]. 畜牧兽医学报, 2017, 48(3): 492-500.
- [10] 蒋志琼, 钟泽民, 谭博敏, 等. 丹毒丝菌 *SpaA* 基因免疫保护区的克隆及其在毕赤酵母中的表达[J]. 华南农业大学学报, 2015, 36(3): 20-25.
- [11] 周江林, 彭小川, 胡明达, 等. 基于全基因组测序的细菌进化研究进展[J]. 生物技术通讯, 2018, 29(6): 844-850.
- [12] 周迪, 杨旭夫, 彭凌. 红斑丹毒丝菌的分离鉴定及药敏试验[J]. 动物医学进展, 2020, 41(2): 33-38.
- [13] BOLGER A M, LOHSE M, USADEL B. Trimmomatic: a flexible trimmer for Illumina sequence data[J]. Bioinformatics, 2014, 30(15): 2114-2120.
- [14] BANKEVICH A, NURK S, ANTIPOV D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing[J]. Journal of Computational Biology, 2012, 19(5): 455-477.
- [15] MASSOURAS A, HENS K, GUBELMANN C, et al. Primer-initiated sequence synthesis to detect and assemble structural variants[J]. Nature Methods, 2010, 7(7): 485-486.
- [16] SEEMANN T. Prokka: rapid prokaryotic genome annotation[J]. Bioinformatics, 2014, 30(14): 2068-2069.
- [17] UCHIYAMA M, YAMAMOTO K, OCHIAI M, et al. Prevalence of Met-203 type SpaA variant in *Erysipelothrix rhusiopathiae* isolates and the efficacy of swine erysipelas vaccines in Japan[J]. Biologicals, 2014, 42(2): 109-113.
- [18] OGAWA Y, SHIRAIWA K, OGURA Y, et al. Clonal lineages of *Erysipelothrix rhusiopathiae* responsible for acute swine erysipelas in Japan identified by using genome-wide single nucleotide polymorphism analysis [J]. Applied & Environmental Microbiology, 2017, 83(11): e00130-17.
- [19] TATUSOV R L, GALPERIN M Y, NATALE D A, et al. The COG database: a tool for genome-scale analysis of protein functions and evolution[J]. Nucleic Acids Research, 2000, 28(1): 33-36.
- [20] KANEHISA M, GOTO S. KEGG: kyoto encyclopedia of genes and genomes[J]. Nucleic Acids Research, 2000, 28(1): 27-30.
- [21] CHEN L, ZHENG D, LIU B, et al. VFDB 2016: hierarchical and refined dataset for big data analysis-10 years on[J]. Nucleic Acids Research, 2016, 44: D694-D697.
- [22] MCARTHUR A G, WAGLECHNER N, NIZAM F, et al. The comprehensive antibiotic resistance database[J]. Antimicrobial Agents and Chemotherapy, 2013, 57(7): 3348-3357.
- [23] URBAN M, PANT R, RAGHUNATH A, et al. The pathogen-host interactions database (PHI-base): additions and future developments[J]. Nucleic Acids Research, 2015, 43: D645-D655.
- [24] 姚焱彬. 猪丹毒丝菌灭活疫苗的制备及其对小鼠的免疫效力研究[D]. 合肥: 安徽农业大学, 2018.
- [25] 张珍, 施开创, 王孝德, 等. 2015-2017年广西鸡源沙门氏菌耐药性与致病性的相关性分析[J]. 南方农业学报, 2019, 50(10): 2350-2358.
- [26] 徐重新, 刘敏, 张霄, 等. 农药危害风险及其残留检测用广谱特异性抗体研究进展[J]. 江苏农业学报, 2019, 35(2): 489-496.
- [27] 王志芳, 雷燕, 肖俊, 等. 广西罗非鱼主产区养殖池塘抗生素残留状况分析[J]. 南方农业学报, 2019, 50(4): 891-897.
- [28] 胡晓芬, 王静, 熊剑锋, 等. 一起猪丹毒杆菌病例的临床分离鉴定及诊治[J]. 养猪, 2014(2): 107-109.
- [29] 王力波. 猪丹毒丝菌广西分离株生物学特性研究及其基因组生物信息学分析[D]. 长春: 吉林农业大学, 2015.
- [30] 曹文尧. 丹毒丝菌表面保护抗原 A 作为重组亚单位疫苗和核酸疫苗的初步研究[D]. 乌鲁木齐: 新疆大学, 2007.
- [31] CUI Y, LI D, YANG R. Shiga toxin-producing *Escherichia coli* O104:H4: an emerging important pathogen in food safety[J]. Chinese Science Bulletin, 2013, 58(14): 1625-1631.
- [32] FORDE T, BIEK R, ZADOKS R, et al. Genomic analysis of the multi-host pathogen *Erysipelothrix rhusiopathiae* reveals extensive recombination as well as the existence of three generalist clades with wide geographic distribution[J]. BMC Genomics, 2016, 461(17): 1-15.

(责任编辑: 王妮)