

研究报告

泡桐丛枝和枣疯病植原体 *tuf* 基因上游序列结构、功能和遗传变异比较分析

于少帅¹ 林彩丽¹ 潘皎² 任争光³ 朴春根¹ 汪来发¹ 郭民伟¹ 田国忠^{1*}

(1. 中国林业科学研究院森林生态环境与保护研究所 国家林业局森林保护学重点实验室 北京 100091)

(2. 南开大学生命科学学院 天津 300071)

(3. 北京农学院植物科学技术学院 北京 102206)

摘要:【目的】探究泡桐丛枝和枣疯病植原体 *tuf* 基因上游序列结构、功能差异及其遗传多样性。【方法】利用热不对称交错式 PCR (TAIL-PCR) 扩增枣疯病植原体 *tuf* 基因上游未知序列, 利用启动子探针载体 pSUPV4 构建了泡桐丛枝和枣疯病植原体 *tuf* 基因上游序列的大肠杆菌异源表达体系, 分析泡桐丛枝、苦楝丛枝、莴苣黄化、桑萎缩、长春花绿变等 16SrI 组和枣疯病、樱桃致死黄化、重阳木丛枝等 16SrV 组株系 *tuf* 基因上游调控序列的遗传变异特征和启动子活性。【结果】泡桐丛枝等 16SrI 组植原体株系 *tuf* 基因和其上游 *fusA* 基因之间的间区序列长 129–130 bp, 预测有完整的启动子保守结构。泡桐丛枝植原体 *tuf* 基因上游 130 bp 片段具有启动子活性, 此间区序列在 5 种 35 株 16SrI 组株系中存在 4 种变异类型; 枣疯病植原体等 16SrV 组株系 *fusA* 和 *tuf* 基因间区长 53–54 bp, 未预测到完整启动子结构。枣疯病植原体 *tuf* 基因上游 144 bp 和 346 bp 片段均未检测到启动子活性, *fusA* 和 *tuf* 基因间区序列在 3 种 20 株 16SrV 组株系中存在 2 种变异类型。*fusA-tuf* 基因间区序列相对保守, 基于此序列构建的进化树可清晰区分不同组别的植原体株系。【结论】研究方法和结果为深入研究植原体基因表达与调控、揭示植原体生长繁殖规律及其致病机理等奠定了良好的基础。

关键词: 植原体, 启动子, TAIL-PCR, 上游序列, 遗传变异

Foundation item: National Natural Science Foundation of China (No. 31370644); National Infrastructure of Microbial Resources of China (No. NIMR2014-7)

***Corresponding author:** Tel: 86-10-62889525; Fax: 86-10-62884972; E-mail: tian3691@163.com

Received: October 19, 2015; **Accepted:** January 14, 2016; **Published online** (www.cnki.net): January 21, 2016

基金项目: 国家自然科学基金项目(No. 31370644); 国家微生物资源平台项目(No. NIMR2014-7)

***通讯作者:** Tel: 86-10-62889525; Fax: 86-10-62884972; E-mail: tian3691@163.com

收稿日期: 2015-10-19; 接受日期: 2016-01-14; 优先数字出版日期(www.cnki.net): 2016-01-21

Comparative analysis of structure, function and genetic variation of upstream sequences adjoining *tuf* gene in paulownia and jujube witches'-broom phytoplasmas

YU Shao-Shuai¹ LIN Cai-Li¹ PAN Jiao² REN Zheng-Guang³ PIAO Chun-Gen¹
WANG Lai-Fa¹ GUO Min-Wei¹ TIAN Guo-Zhong^{1*}

(1. Key Laboratory of Forest Protection of State Forestry Administration, Research Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China)

(2. College of Life Sciences, Nankai University, Tianjin 300071, China)

(3. Plant Science and Technology College, Beijing University of Agriculture, Beijing 102206, China)

Abstract: [Objective] Exploring the difference of structure, function as well as the genetic diversity of upstream sequences adjoining *tuf* gene in paulownia and jujube witches'-broom phytoplasmas. [Methods] Thermal asymmetric interlaced PCR (TAIL-PCR) was used to amplify the upstream unknown sequence adjoining *tuf* gene in jujube witches'-broom phytoplasma. Recombinant expression system was successfully constructed with promoter-probe vector pSUPV4 and *tuf* gene upstream sequences from paulownia and jujube witches'-broom phytoplasma. Genetic variation and promoter activity of *tuf* gene upstream regulatory sequences from 16SrI group phytoplasma strains including paulownia witches'-broom, chinaberry witches'-broom, lettuce yellows, mulberry dwarf and periwinkle virescence phytoplasma strains and 16SrV group strains including jujube witches'-broom, cherry lethal yellow, *Bischofia polycarpa* witches'-broom phytoplasma strains were analyzed and identified. [Results] The sequence length of intergenic region between *tuf* and its upstream gene, *fusA*, in 16SrI group strains was 129 to 130 bp, with putative complete promoter conservative structure. 130 bp upstream sequence of *tuf* gene in paulownia witches'-broom phytoplasma had promoter activity. There were four variation types in the intergenic region sequences from five kinds of 16SrI group phytoplasmas including 35 strains; the sequence length of intergenic region between *fusA* and *tuf* in 16SrV group strains was 53 to 54 bp, with predicted incomplete promoter structure. No promoter activity was found in 144 bp and 346 bp upstream sequence of *tuf* gene in jujube witches'-broom phytoplasma. There were two variation types in the intergenic region sequences from three kinds of 16SrV group phytoplasmas including 20 strains. The *fusA-tuf* intergenic region sequences were comparatively conservative, and the phytoplasma strains belonging to different groups were distinguished clearly by the phylogenetic tree constructed based on the sequences. [Conclusion] The methods and results of the study would lay a beneficial foundation in further investigating phytoplasma gene expression and regulation, revealing phytoplasma regularity of growth and propagation as well as its pathogenicity mechanism.

Keywords: Phytoplasma, Promoter, TAIL-PCR, Upstream sequence, Genetic variation

泡桐丛枝病和枣疯病是我国常见的两种由植原体引起的植物病害, 长期以来对我国泡桐和枣树两大经济树种造成严重危害。植原体寄主范围广、导致的植物病害多、影响地理区域大, 除泡桐和枣树外, 对我国樱桃、桑树、苦楝等经济作物、绿化树种也影响较大。植原体通过对某些关键基因的调控代谢控制自身的生长繁殖及其与寄主植物的相互作用和对环境的适应性, 进而干扰寄主植物的正常代谢途径而使其表现出不正常的

生理状态^[1]。随着分子生物学和基因组学的发展, 通过对植原体与生长繁殖相关的关键基因表达调控分析, 从分子水平上明确植原体致病机理, 无论从理论上还是病害防治上都有较为重要的意义。而细菌功能基因表达和代谢调控主要体现在基因的转录和翻译水平上^[2]。

启动子是基因调控在转录水平上的一个关键结构, 含有 RNA 聚合酶特异性结合和转录起始所需的保守序列^[2]。迄今为止, 关于植原体启动子的

研究报道很有限。Palmano 等首次将西方 X 病植原体 16S rDNA 上游的启动子序列构建到含无启动子报道基因 *cat* 的质粒上,发现该启动子能在芽孢杆菌中启动 *cat* 基因的表达^[3]。Ishii 等预测植原体质粒序列上的启动子,推断 ORF3 编码蛋白在变异株系 OY-NIM 中不表达可能是由于启动子的缺失或突变造成的^[4]。胡佳续等预测了桑树萎缩、苦楝丛枝和长春花绿变等植原体质粒上可能的启动子,并定位了可能的核糖体结合位点^[5]。Oshima 等认为虽然植原体和支原体很少编码参与启动子识别的 σ 因子,但从洋葱黄化植原体中至少鉴定出编码两种 σ 因子的基因 *rpoD* 和 *fliA*^[6]。

蛋白延伸因子直接影响功能基因的翻译水平。蛋白延伸因子编码基因 *tuf* 编码延伸因子 EF-Tu, *fusA* 编码延伸因子 EF-G。其中, *tuf* 基因在革兰氏阴性菌中为双拷贝,在革兰氏阳性菌中为单拷贝。1997 年以来,国外陆续报道了翠菊黄化组、僵顶病组以及 X-病组植原体的 *tuf* 基因序列测定结果, *tuf* 基因在植原体中为单拷贝^[7-8],并在蛋白生物合成与植原体进化过程中起着重要的作用,被认为是一种进化标签。*tuf* 和 16S rRNA 基因相比保守性略低,可在 16S rRNA 基因分组的基础上进一步区分亚组^[7-8]。Malembic-Maher 等利用 *tuf* 基因对 16SrV 植原体株系进行分类,使组内进化关系较近的各株系得到清晰区分^[9]。Streten 和 Gibb 利用 *tuf* 基因序列将 16SrXII-B 中的 *Candidatus* *Phytoplasma australiense* 株系细分为 16SrXII-B (*tuf*-Australia I)、16SrXII-B (*tuf*-New Zealand I)、16SrXII-B (*tuf*-New Zealand II)^[10]。Wang 等研究发现泡桐丛枝植原体 *tuf* 基因抗血清与 16SrI 组株系有特异免疫反应,与 16SrV 组株系无免疫反应^[11]。目前,关于植原体 *tuf* 基因的研究主要集中于遗传变异、系统进化、免疫反应等方面,关于该基因表达和调控的相关研究较少^[12]。Miyata 等对洋葱黄化植原体(Onion yellows phytoplasma) *tuf* 基因及其上游 *rps12*、*rps7*、*fus* 等基因结构进行分析,推测它们共同组成一个 *str* 操纵子(Streptomycin operon)^[12]。此外,已有研究表明基因启动子序列

的变异会影响启动子的活性从而改变相应基因的表达及代谢效率^[13-14]。

研究采用热不对称交错式 PCR 技术(Thermal asymmetric interlaced PCR, TAIL-PCR)扩增枣疯病植原体 *tuf* 基因上游未知序列,利用直接 PCR 扩增泡桐丛枝植原体 *tuf* 基因上游序列,推测和比较分析两种植原体 *tuf* 基因上游序列可能存在的启动子结构。启动子探针载体 pSUPV4 由四川大学生命科学学院构建并得到大量应用^[15-17],通过该启动子探针载体构建的融合表达体系是否表现出卡那霉素抗性,验证两种植原体 *tuf* 基因上游序列是否存在启动子活性。为探究不同地区不同寄主植原体 *tuf* 基因上游序列的遗传变异状况,根据已获得序列设计引物扩增了来自我国不同地区泡桐丛枝植原体 (Paulownia witches'-broom phytoplasma, PaWB) 及同为 16SrI 组的苦楝丛枝植原体 (Chinaberry witches'-broom phytoplasma, CWB)、桑萎缩植原体 (Mulberry dwarf phytoplasma, MD)、莴苣黄化植原体 (Lettuce yellows phytoplasma, LY)、长春花绿变植原体 (Periwinkle virescence phytoplasma, PeV) 和枣疯病植原体 (Jujube witches'-broom phytoplasma, JWB) 及同为 16SrV 组的重阳木丛枝植原体 (*Bischofia polycarpa* witches'-broom phytoplasma, BiWB)、樱桃致死黄化植原体 (Cherry lethal yellow phytoplasma, CLY) 的 *tuf* 基因上游序列。本研究对于揭示植原体关键基因结构特征和表达调控机制、阐明其遗传变异规律和系统进化关系、探究其与寄主植物和昆虫介体间的互作关系、生境多样性和适应性、致病性等问题具有较为重要意义。

1 材料与方法

1.1 实验材料

不同病害植原体材料于 2013 至 2015 年采集,每个采样地采集健康泡桐、枣树等植物枝叶作为健康对照。研究收集的不同病害植原体株系的具体信息如表 1 所示。

染色体步移、DNA 纯化、质粒提取等试剂盒,分子克隆所用的限制性内切酶、T4 DNA 连接酶、

表 1 研究收集的不同植原体株系信息
Table 1 Information of the different phytoplasma strains collected in the study

序号 No.	植原体名称 Phytoplasma name	组别 Group	采样地和株系编号 Sampling sites and strain codes	株系数量 Number of strains
1	泡桐丛枝植原体	16SrI	山东兖州(PaWB-sdyz), 江苏南京(-jsnj)、苏州(-jssz1、-jssz2)、北京怀柔(-bjhr), 河南郑州(-hnzz)、濮阳(-hnpv)、洛阳(-hnly)、湖南长沙(-hnsc)、安徽合肥(-ahhf)、江西南昌(-jxnc)、福建福州(-fjz)、河北保定(-hbhd)、辽宁大连(-lnd1、-lnd2)、山西太原(-sxt1、-sxt2)、陕西西安(-sxxa1、-sxxa2)、渭南(-sxwn)	20
2	苦楝丛枝植原体	16SrI	海南三亚(CWB-hnsy1、-hnsy2)、福建福州(-fjz1、-fjz2、-fjz3)、永安(-fjya)、江苏南京(-jsnj)、湖南长沙(-hnsc)、江西南昌(-jxnc)、广东广州(-gdgz)	10
3	莴苣黄化植原体	16SrI	福建永安(LY-fjya1、-fjya2)	2
4	桑萎缩植原体	16SrI	安徽合肥(MD-ahhf)、浙江淳安(-zjca)	2
5	长春花绿变植原体	16SrI	海南海口(PeV-hnhk)	1
6	枣疯病植原体	16SrV	江苏南京(JWB-jsnj)、镇江(-jszj)、北京海淀(-bjhd)、辽宁大连(-lnd1、-lnd2、-lnd3)、葫芦岛(-lnhld1、-lnhld2)、建昌(-lnjc)、河南周口(-hnzk)、濮阳(-hnpv)、洛阳(-hnly)、安徽合肥(-ahhf)、陕西渭南(-sxwn1、-sxwn2、-sxwn3)、重庆武隆(-cqwl1、-cqwl2)	18
7	樱桃致死黄化植原体	16SrV	四川西昌(CLY-sxc)	1
8	重阳木丛枝植原体	16SrV	安徽合肥(BiWB-ahhf)	1

pMD18-T 克隆载体、大肠杆菌感受态细胞 DH5 α 等均购自 TaKaRa 公司 ;PCR 反应所用试剂购自北京博迈德生物技术有限公司 ;DNA 提取试剂购自北京艾德莱生物科技有限公司 ;启动子探针载体 pSUPV4 由四川大学生命科学学院王海燕副教授提供 ;其它试剂为国产分析纯试剂。引物由北京赛百盛基因技术有限公司合成 ;DNA 序列由生工生物工程(上海)股份有限公司测定。

1.2 样品总 DNA 提取和检测

样品总 DNA 提取参照艾德莱植物基因组提取试剂盒提供方法。植原体 16S rRNA 基因用 R16mF2/R16mR1^[18]通用引物 PCR 扩增后测序检测。PCR 反应设阳性对照(实验室组培的部分感染植原体植物)和阴性对照(健康植物)。

1.3 *tuf* 基因上游序列获取

根据已知洋葱黄化植原体(Onion yellows phytoplasma, OY-M, GenBank 登录号 AP006628)全基因组设计引物 Pf1 和 Pr1, 序列及 T_m 值如表 2 所示。通过 PCR 方法扩增泡桐丛枝及其它 16SrI 组植原体株系的 *tuf* 上游序列。为保证所得序列准确性, 每条所得序列均测 3 次。PCR 反应体系及条件

参照博迈德公司 PCR 反应试剂说明。PCR 产物经 1% 琼脂糖凝胶检测后送公司测序。根据 FDTUF-F1/R1^[19]扩增 JWB-jsnj 植原体株系 *tuf* 序列, 所得序列通过 BLAST 验证为 JWB 植原体 *tuf* 序列, 根据所得序列设计 3 条特异性巢式 TAIL-PCR 引物 SP1、SP2、SP3, 如表 2 所示。简并引物 AP 来自 TaKaRa 染色体步移试剂盒, 第一轮 PCR 反应以 SP1 和 AP 为引物, 第二轮 PCR 反应以 SP2 和 AP 为引物, 第三轮 PCR 反应以 SP3 和 AP 为引物, PCR 反应体系和条件参照 TaKaRa 染色体步移试剂盒提供的方法。第三轮 PCR 产物经 1% 琼脂糖凝胶检测后切胶回收, 纯化产物与 pMD18-T 载体连接后转化大肠杆菌感受态细胞 DH5 α , 37 °C 培养 24 h 后挑选单菌落测序。通过引物 Jtf-1 和 SP2 扩增枣疯病及其它 16SrV 组植原体株系的 *tuf* 上游序列, 每条所得序列均测 3 次, 所用引物序列及 T_m 值见表 2。

1.4 序列分析

通过 DNASTar、DNAMAN 5.0 等软件对所得序列进行拼接和多重比对, 通过在线分析软件 iPhyClassifier (<http://plantpathology.ba.ars.usda.gov/>)

表 2 研究所用引物
Table 2 Primers used in the study

引物 Primer	引物序列 Primer sequences (5'→3')	引物大小 Primer size (bp)	T_m (°C)	引入酶切位点 Restriction site added
Pf1	TACCATGATGTCGATTCTTCC	21	55	—
Pr1	ATCCATTGAGCAGCACCAG	20	55	—
FDTUF-F1	ATTGGTCATGTAGACCATGG	20	55	—
FDTUF-R1	GTTCTTCCGCCTTCACGTAC	20	55	—
SP1	GCAACAGTTCTCTTCTTTTACG	24	65	—
SP2	GGCATTACTCCATCTACAGCAGAAAC	26	65	—
SP3	TAATCTGCATGCCAGGACAATC	23	65	—
TPf	cccaagcttACAACCTTACACTAAAAAAC	29	45	Hind III
TPr	cgcggatcCATTTTCAAAGGCCTC	25	45	BamH I
Jtf	cccaagcttACAAGGCAGAGCGAGTT	26	50	Hind III
Jtr	cgcggatcCATTTTAAAAGGCCTCC	25	50	BamH I
Jtf-1	cgaagcttATGGTGGTTAGTTCTTT	26	50	Hind III
Jtr-1	cgcggatcCATTTTAAAAGGCCTC	24	50	BamH I

cgi-bin/resource/iphyclassifier.cgi)和 16S rRNA 基因序列对不同植原体株系进行分析鉴定。通过 Softberry (<http://linux1.softberry.com/berry.phtml>) 在线分析工具并参考相关分子生物学资料^[2]推测不同植原体株系 *tuf* 基因上游序列可能存在的启动子及其结构。用 MEGA 4.0 软件基于邻接法 (Neighbor-Joining, NJ)构建系统进化树, Bootstrap 值设为 1 000。16SrI 组的洋葱黄化植原体 (Onion yellows phytoplasma, OY-M) (GenBank 登录号 AP006628)和翠菊黄化丛枝植原体 (Aster yellows witches'-broom phytoplasma, AYWB) (CP000061), 16SrXII 组的澳大利亚葡萄黄化植原体 (*Candidatus* Phytoplasma australiense, PAa) (AM422018)和草莓致死黄化植原体 (Strawberry lethal yellows phytoplasma, SLY) (CP002548), 16SrX 组的苹果簇生植原体 (*Ca. P. mali*, AT) (CU469464)等 5 个已完成全基因组测序的株系作为参照, 相关序列通过 NCBI (<http://www.ncbi.nlm.nih.gov>)获取。

1.5 *tuf* 基因上游序列启动子活性验证

将启动子探针载体 pSUPV4 用 Hind III 和 BamH I 限制酶酶切后回收大片段。根据已知序列, 设计引物 TPf 和 TPr 扩增 PaWB-sdyz 植原体株系 *tuf* 上游 130 bp 序列, 设计引物 Jtf 和 Jtr、Jtf-1 和 Jtr-1 分别扩增 JWB-jsnj 植原体株系 *tuf* 上游 144 bp 和 346 bp 序列, 引入酶切位点 Hind III 和

BamH I, 将各 PCR 扩增产物用 Hind III 和 BamH I 双酶切后回收片段, 将回收片段与酶切的载体片段用 T4 DNA 连接酶连接, 连接产物通过热激法转化大肠杆菌感受态细胞 DH5 α , 涂布于含有卡那霉素 (Kan) (Kan 浓度分别为 30 mg/L 和 50 mg/L) 的 LB 平板上, 37 °C 培养 24 h, 挑取单菌落通过酶切和测序验证转化子中目的片段。

2 结果与分析

2.1 泡桐丛枝和枣疯病植原体 *tuf* 基因上游序列扩增和结构解析

通过 PCR 方法扩增泡桐丛枝植原体株系 PaWB-sdyz *tuf* 上游片段, 测序获得 782 bp 序列, 经 BLAST 比对分析, 发现 355 bp 片段为 *fusA* 基因的 3'末端, 编码蛋白延伸因子 EF-G 的 C 端; 后 297 bp 为 *tuf* 基因的 5'前端, 编码蛋白延伸因子 EF-Tu 的 N 端, *fusA* 和 *tuf* 两个基因间区序列长度为 130 bp。所得序列已上传至 GenBank 数据库, PaWB-sdyz 株系 *tuf* 基因及其上游部分序列 GenBank 登录号为 KU095823。经 Softberry 软件在线分析工具分析表明: *fusA* 和 *tuf* 两个基因间区序列具有明显完整的启动子结构特征。

通过 TAIL-PCR 方法扩增枣疯病植原体株系 JWB-jsnj *tuf* 基因上游未知序列, 对 TAIL-PCR 第三次 PCR 产物测序获得 561 bp 序列, 将其与引物 FDTUF-F1/R1 扩增获得的 937 bp 片段经 DNAMAN 组装, 共获得 1 414 bp 长度的序列, 其中有 84 bp

的序列完全吻合。经 BLAST 比对分析发现, 前 306 bp 序列为 *fusA* 基因的 3' 末端, 编码蛋白延伸因子 EF-G 的 C 端; 后 1 055 bp 为 *tuf* 基因序列, 编码蛋白延伸因子 EF-Tu 的 N 端; *fusA* 和 *tuf* 两个基因间区序列长度为 53 bp。所得序列已上传至 GenBank 数据库, JWB-jsnj 株系 *tuf* 基因及其上游部分序列 GenBank 登录号为 KU095824。预测分析发现: *tuf* 基因上游序列存在启动子保守结构特征, 部分保守结构位于 *fusA* 基因的 3' 末端, *fusA* 和 *tuf* 两个基因间区未预测到完整的启动子保守的结构区域。

2.2 不同植原体株系 *fusA* 和 *tuf* 基因间区序列多态性和系统进化分析

扩增来自我国不同地区的 16SrI 组和 16SrV 组植原体株系的 *tuf* 基因上游序列, 多重序列比对发现: 研究所采用的 16SrI 组植原体株系 *fusA* 和 *tuf* 基因间区长度为 129–130 bp, 相似性为 97.7%–100%。PaWB-sdyz 株系与 OY-M 株系本段序列相似性为 100%, 与 AYWB 相似性为 94.6%。35 条基因间区序列存在 4 种变异类型。变异类型 1 (TP1) 包括泡桐丛枝植原体株系 PaWB-sdyz、-jsnj、-jssz1、-jssz2、-bjhr、-hnzz、-hnpv、-hnly、-hnsc、-ahhf、-jxnc、-hbbd、-lnd11、

-lnd12、-sxt1、-sxt2、-sxxa1、-sxxa2、-sxwn 和长春花绿变植原体株系 PeV-hnhk 等共 20 个株系; 变异类型 2 (TP2) 包括苦楝丛枝植原体株系 CWB-hnsy1、-hnsy2、-fjz1、-fjz2、-fjz3、-fjya、-jsnj、-hnsc、-jxnc、-gdgz 和桑菱缩植原体株系 MD-ahhf、-zjca 等共 12 个株系; 变异类型 3 (TP3) 包括茛苳黄化植原体株系 LY-fjya1、-fjya2 等 2 个株系; 变异类型 4 (TP4) 包括泡桐丛枝植原体株系 PaWB-fjz 一个株系。

所测定的 16SrV 组植原体株系 *fusA* 和 *tuf* 基因间区长度在 53–54 bp 之间。以 JWB-jsnj 株系本段基因间区为参照进行多重比对, 结果发现 JWB-hnpv 株系存在一个碱基 A 的插入, CLY-scxc 和 BiWB-ahhf 株系与之完全一致。20 条基因间区序列存在 2 种变异类型。变异类型 1 (TJ1) 包括枣疯病植原体株系 JWB-jsnj、-jszj、-bjhd、-lnd11、-lnd12、-lnd13、-lnhld1、-lnhld2、-lnjc、-hnzk、-hnly、-ahhf、-sxwn1、-sxwn2、-sxwn3、-cqwl1、-cqwl2, 樱桃致死黄化植原体株系 CLY-scxc 和重阳木丛枝植原体株系 BiWB-ahhf; 变异类型 2 (TJ2) 包括 JWB-hnpv 一个株系。16SrI 和 V 组不同植原体株系 *fusA* 和 *tuf* 基因间区的具体变异细节如图 1 所示。

TP1	ACAACCTTAACTAAAAACATTTTATTAACAAAATAGT	40
TP2	ACAACCTTAACTAAAAACATTTTATTAACAAAATAGT	40
TP3	ACAACCTTAACTAAAAACATTTTATTAACAAAATAGT	40
TP4	ACAACCTTAACTAAAAACATTTTATTAACAAAATAGT	40
Consensus	acaaccttaactaaaaacatttttattaacaaaata	
TP1	TGTAAAAAGCCAAAAATAGTTAAATTTAAATGGTAAAT	80
TP2	TGTAAAAAGCCAAAAATAGTTAAATTTAAATGGTAAAT	80
TP3	TGTAAAAAGCCAAAAATAGTTAAATTTAAATGGTAAAT	80
TP4	TGTAAAAAGCCAAAAATAGTTAAATTTAAATGGTAAAT	80
Consensus	tgtaaaaagccaaaaataagttaaaatttaaatggtaat	
TP1	ATATTTATTAACAAACCTTAAACATAATAAAAGGAGG	120
TP2	ATATTTATTAACAAACCTTAAACATAATAAAAGGAGG	119
TP3	ATATTTATTAACAAACCTTAAACATAATAAAAGGAGG	120
TP4	ATATTTATTAACAAACCTTAAACATAATAAAAGGAGG	120
Consensus	atatttatataaacacacctaaaacataataaaaggagg	
TP1	CCTTTGAAAA	130
TP2	CCTTTGAAAA	129
TP3	CCTTTGAAAA	130
TP4	CCTTTGAAAA	130
Consensus	cctttgaaaa	
TJ1	TATATAATATTTATGTAAAAGATAAAAAAATAAAAAAA	39
TJ2	TATATAATATTTATGTAAAAGATAAAAAAATAAAAAAA	40
Consensus	tataataatttatgtaaaagataaaaaataaaaaaa	
TJ1	GGAGGCCTTTTAAA	53
TJ2	GGAGGCCTTTTAAA	54
Consensus	ggaggcctttttaa	

图 1 16SrI 和 V 组植原体株系 *fusA* 和 *tuf* 基因间区序列变异类型分析

Figure 1 Analysis of variation types in *fusA* and *tuf* intergenic sequences from 16SrI and V group phytoplasma strains

注: 不同植原体株系 *fusA* 和 *tuf* 基因间区序列变异位点用竖框标出; 序列中的“-”表示碱基缺失。

Note: Variable sites of *fusA* and *tuf* intergenic sequences of different phytoplasma strains were shown in column; “-” in the sequences represented base deletion.

基于不同植原体株系 *fusA* 和 *tuf* 基因间区序列构建 NJ 树, 系统进化分析表明: 同为 16SrV 组的 JWB、CLY 和 BiWB 植原体株系形成一个独立的分支, 支持率为 86%; 16SrI 组的 PaWB、MD、LY 和 PeV 植原体株系与同组别的参考株系 OY-M、

AYWB 聚在一起, 支持率为 63%; 16SrX 组的 AT 株系与 16SrXII 组的 PAa、SLY 株系分别形成独立的进化分支。不同植原体株系系统进化关系如图 2 所示, 不同组别的植原体株系可以清晰地形成不同的分支。

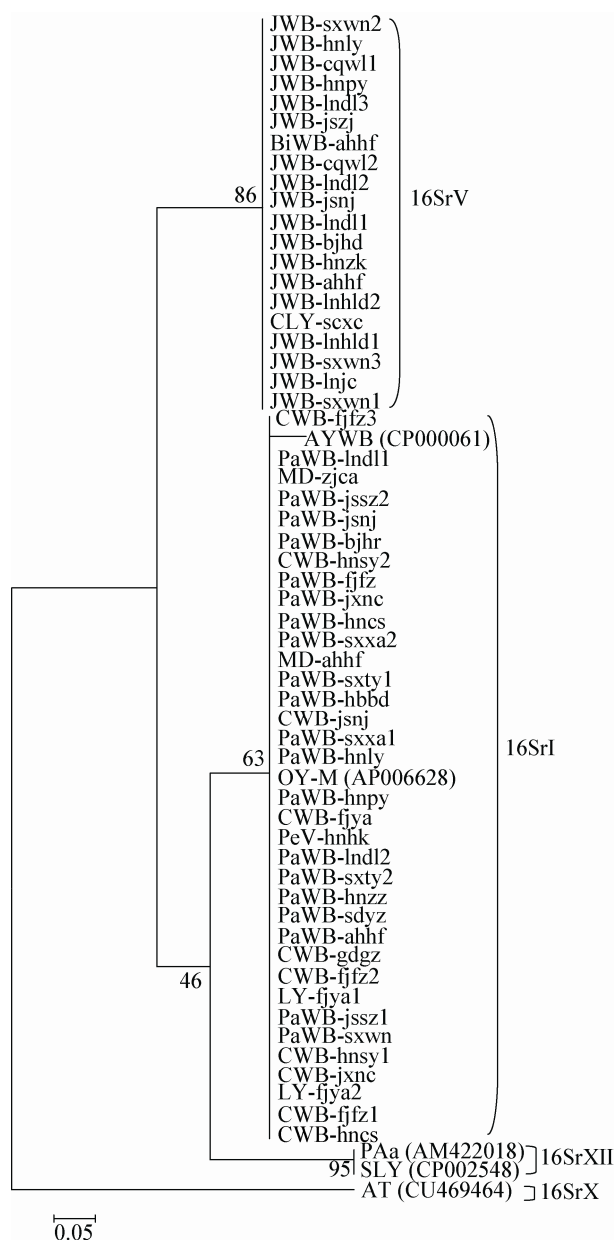


图 2 基于不同植原体株系 *fusA* 和 *tuf* 基因间区序列构建的系统树

Figure 2 Phylogenetic tree constructed based on *fusA* and *tuf* intergenic sequences from different phytoplasma strains

注: 植原体株系编号如表 1 所示; 参考植原体株系序列号在括号中列出; 系统树分支点处的数字代表系统树的自展值 (1 000 次重复); 标尺长度代表遗传距离。

Note: The codes of the phytoplasma strains were shown in Table 1; Accession numbers of the reference phytoplasma strains were listed in parentheses; The numbers on the branch points of the phylogenetic tree represented bootstrap values of the tree (1 000 replicates); Scale length represented genetic distance.

2.3 泡桐丛枝和枣疯病植原体 *tuf* 基因上游序列启动子活性验证

将泡桐丛枝植原体山东兖州株系 PaWB-sdyz *tuf* 上游 130 bp 序列与启动子探针载体 pSUPV4 连接构建融合表达体系, 在含 30 mg/L 和 50 mg/L Kan 的 LB 平板上培养 24 h 后, 均有明显的转化子产生, 挑取单菌落测序发现转化子中含有目的片段。因而判断该区域具有启动子结构和活性。为确定 JWB-jsnj 株系 *tuf* 上游序列是否具有启动子活性及其核心区域, 试验分别截取了枣疯病植原体江苏南京株系 JWB-jsnj *tuf* 上游 144 bp 和 346 bp 序列, 分别与 pSUPV4 启动子探针载体连接构建融合表达体系, 在含 30 mg/L 和 50 mg/L Kan 的 LB 平板上培养 24、48 h 后, 含有 JWB-jsnj *tuf* 上游 144 bp 和 346 bp 的融合表达体系均不能在含有 Kan 的 LB 平板上生长, 因此在 JWB-jsnj 株系 *tuf* 基因上游 144 bp 和 346 bp 序列没有检测到启动子活性 (图 3)。

3 讨论

TAIL-PCR 技术是一种根据已知 DNA 序列扩增相邻未知序列的方法。与其它克隆技术相比, TAIL-PCR 技术具有操作相对简单、特异性较高等特点^[20]。结合枣疯病植原体自身特点, 我们将此技术用于扩增和测定 JWB *tuf* 基因上游未知序列, 并成功获取目的片段。通过 PaWB 等 16SrI 组、JWB 等 16SrV 组以及 OY-M、AYWB、PAa、SLY、AT 等株系比较分析发现, 不同株系间的 *fusA* 和 *tuf* 基因间区存在明显的差异和组间结构特征。PaWB、CWB、LY、MD、PeV、OY-M、AYWB 等 16SrI 组株系 *fusA* 和 *tuf* 基因间区在 129–134 bp 范围, 16SrXII 组的 PAa、SLY 间区序列长为 118 bp, 这些植原体在本段区域序列均能预测到完整的启动子结构。而在 JWB、CLY、BiWB 等 16SrV 组株系 *fusA* 和 *tuf* 基因间区为 53–54 bp, 16SrX 组 AT 株系为 65 bp, 两组植原体的此段序列

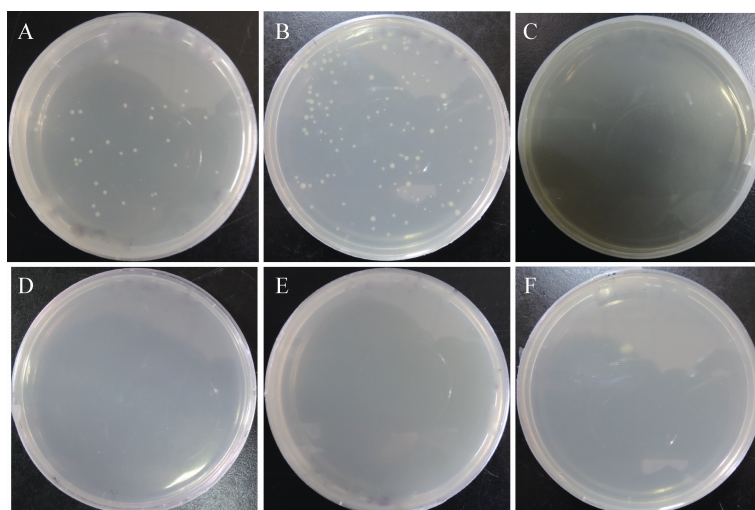


图 3 融合表达体系在含 Kan 的 LB 培养基上的表达状况

Figure 3 Expression of recombinant expression system on LB medium with kanamycin

注: 融合表达体系含有 PaWB-sdyz *tuf* 上游 130 bp 序列, 培养基 Kan 浓度分别为 30 (A)、50 (B) mg/L; 融合表达体系含有 JWB-jsnj *tuf* 上游 144 bp 序列, 培养基 Kan 浓度分别为 30 (C)、50 (D) mg/L; 融合表达体系含有 JWB-jsnj *tuf* 上游 346 bp 序列, 培养基 Kan 浓度分别为 30 (E)、50 (F) mg/L, 所有体系均在 37 °C 条件下培养 24 h。

Note: Recombinant expression system including 130 bp upstream sequence of *tuf* gene in PaWB-sdyz, the concentration of kanamycin in the medium was 30 (A), 50 (B) mg/L respectively; Recombinant expression system including 144 bp upstream sequence of *tuf* gene in JWB-jsnj, the concentration of kanamycin in the medium was 30 (C), 50 (D) mg/L respectively; Recombinant expression system including 346 bp upstream sequence of *tuf* gene in JWB-jsnj, the concentration of kanamycin in the medium was 30 (E), 50 (F) mg/L respectively. All recombinant systems were cultivated under 37 °C condition for 24 h.

均无完整启动子结构。我们采用启动子探针载体 pSUPV4 成功构建了植原体调控序列在大肠杆菌中的异源表达检测体系。分析发现：PaWB-sdyz 株系 *fusA* 和 *tuf* 基因间区 130 bp 序列可启动探针载体抗卡那霉素基因的表达，其融合表达体系可在含 Kan 的平板中生长。而通过顺序缺失突变分析发现由 JWB-jsnj 株系 *tuf* 基因上游长 144 bp 和 346 bp 序列构建的融合表达体系均不能在含 Kan 的平板上生长。从而率先揭示了我国重要植原体 *tuf* 基因上游所存在的两大类调控序列的特征及在同种植原体的不同地区来源株系间的遗传变异程度。

已有的研究报道大肠杆菌 *tuf* 基因有两个拷贝，一个拷贝单独表达，另一个 *tuf* 基因与 *fusA* 及其上游的核蛋白编码基因 *rps12*、*rps7* 共同构成一个 *str* 操纵子，其基因结构排列顺序为 5'-*rps12-rps7-fusA-tuf*-3'^[21]。Miyata 等对洋葱黄化植原体株系 *str* 操纵子分析发现其结构与枯草芽孢杆菌更相似，*tuf* 基因只有一个拷贝，基因结构为 5'-*rps12-rps7-fusA-tuf*-3'^[12]。我们对 PaWB、OY-M、AYWB、PAa、SLY、AT 等株系 *rps12*、*rps7*、*fusA*、*tuf* 等基因结构排布进行分析发现它们与大肠杆菌 *str* 操纵子的排列一致，对 JWB 株系已有序列的分析发现，与大肠杆菌 *str* 操纵子的基因结构基本吻合。本研究在泡桐丛枝植原体 *fusA* 和 *tuf* 基因间区序列检测到了启动子活性，因此推测泡桐丛枝植原体 *tuf* 基因可能单独表达，不是 *str* 操纵子的一部分。来自我国不同地区的 5 种 35 株 16SrI 组植原体株系的 *fusA* 和 *tuf* 基因间区序列存在 4 种变异类型，3 种 20 株 16SrV 组植原体株系 *fusA* 和 *tuf* 基因间区存在 2 种变异类型，此段间区序列相对保守。大量研究发现，生物体某些基因启动子序列发生变异将会影响基因的表达及其表达效率^[13-14]。植原体某些关键基因启动子序列的缺失和变异可能会导致其它相关的基因表达及效率发生改变从而表现出不同的寄主(植物和昆虫)特异性和致病性。Ishii 等对洋葱黄化植原体质粒上的编码基因预测分析发现，可虫传植原体株系(OY-M) ORF3 基因由两个启动子调控转录，而非虫传植原

体株系(OY-NIM)仅有一个启动子序列且发生变异，通过免疫反应发现 ORF3 在 OY-NIM 中不表达，由此推测受启动子的缺失或突变影响的 ORF3 基因在植原体与昆虫互作过程起着重要作用^[4]。并推断 PaWB 和 JWB 株系 *tuf* 上游序列结构特点、启动子活性差异及 16SrI 和 V 组株系 *fusA* 和 *tuf* 基因间区序列的变异可能与植原体的生长繁殖能力、环境适应性和与寄主的互作关系相关，从而导致其不同的寄主(植物和昆虫)特异性和致病性状。

植原体的分类鉴定和遗传进化分析主要基于 16S rDNA 序列，这一序列是应用最为广泛的遗传标记^[18]。16S rRNA 基因序列高度保守，不能提供足够的遗传变异信息。因此，一些相对保守的序列如 16-23S rRNA 基因间区、*rp*、*tuf*、*secY* 等被应用于植原体的遗传变异研究中，并揭示了植原体更为丰富的遗传特征^[19]。我们基于 *fusA* 和 *tuf* 基因间区序列多态性和系统进化分析发现：不同组别的植原体株系可以清晰地形成不同的分支。35 株 16SrI 组株系存在 4 个变异位点，PaWB (除福建福州株系外)与 PeV 株系变异类型相同，CWB 与 MD 株系变异类型相同，LY 株系和 PaWB-fjtz 株系分别具有单独的变异类型；20 株 16SrV 组株系存在两种变异类型，JWB (除河南濮阳株系外)、CLY、BiWB 株系此段序列完全一致，JWB-hnpy 株系存在一个碱基位点(A)的插入。*fusA-tuf* 基因间区序列相对保守，可清晰区分不同组别的植原体株系。但与下游编码基因区域相比，因片段较短和变异位点相对较少，因而在同组内植原体株系的鉴定能力较弱。与功能基因序列一样，基因间区也可以反映植原体的遗传变异和分子进化特征，目前在植原体分类鉴定研究中常用的基因间区主要是 16-23S rRNA 基因间区序列。Lai 等利用 16-23S rRNA 基因间区序列结合 16S rRNA、*rp*、*tuf* 等功能基因片段，分析鉴定来自我国的重阳木丛枝植原体株系代表一个新的亚组 16SrV-H^[22]。基于 *fusA-tuf* 基因间区相对保守的特征，可将其与 16-23S rDNA 间区序列一样，或许可以作为一个新的遗传标记用于植原体检测和分类鉴定研究中，

以期揭示植原体更丰富的遗传变异规律和更细腻的系统发育关系。

综上所述, 我们通过比较分析发现: PaWB 等 16SrI 组株系、JWB 等 16SrV 组株系以及 OY-M、AYWB、PAa、SLY、AT 等参考株系的 *tuf* 基因上游序列在结构和功能等方面存在明显差异。55 株 PaWB、JWB 等 16SrI 和 V 组植原体株系的 *fusA-tuf* 基因间区序列存在多种变异类型, 进一步探索这些遗传特征及其对 *tuf* 基因表达可能引起的影响将有助于我们对植原体种群特征、致病机制等的认识和研究。*fusA-tuf* 基因间区相对保守, 可与 16-23S rRNA 基因间区序列一起作为植原体非编码区系统进化分析的分子标记。此外, 我们通过 TAIL-PCR 方法扩增并成功获取枣疯病植原体未知序列, 利用启动子探针载体 pSUPV4 成功构建了植原体调控序列在大肠杆菌中的异源表达检测体系, 为今后植原体启动子和非编码序列鉴定和功能基因的代谢调控分析提供了新的手段。本研究结果为从分子水平揭示植原体遗传变异规律、与寄主植物和介体昆虫间的互作关系等问题奠定了一定的基础, 有利于我们深入了解和探究植原体的致病机制从而有效管理和防治相关病害。

参 考 文 献

- [1] Sugio A, Hogenhout SA. The genome biology of phytoplasma: modulators of plants and insects[J]. Current Opinion in Microbiology, 2012, 15(3): 247-254
- [2] Turner P, McLennan A, Bates A, et al. Instant Notes Molecular Biology[M]. 3rd Edition. Abingdon: Taylor & Francis Group, 2005
- [3] Palmano S, Kirkpatrick BC, Firrao G. Expression of chloramphenicol acetyltransferase in *Bacillus subtilis* under the control of a phytoplasma promoter[J]. FEMS Microbiology Letters, 2001, 199(2): 177-179
- [4] Ishii Y, Kakizawa S, Hoshi A, et al. In the non-insect-transmissible line of onion yellows phytoplasma (OY-NIM), the plasmid-encoded transmembrane protein ORF3 lacks the major promoter region[J]. Microbiology, 2009, 155(6): 2058-2067
- [5] Hu JX, Song CS, Lin CL, et al. Sequencing the full-length DNA and the molecular characterization of four plasmids from plant pathogens of phytoplasma disease[J]. Scientia Silvae Sinicae, 2013, 49(4): 90-97 (in Chinese)
胡佳续, 宋传生, 林彩丽, 等. 4 种植物病害植原体病原质粒全序列测定及分子特征[J]. 林业科学, 2013, 49(4): 90-97
- [6] Oshima K, Kakizawa S, Nishigawa H, et al. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma[J]. Nature Genetics, 2004, 36(1): 27-29
- [7] Schneider B, Gibb KS, Seemüller E. Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas[J]. Microbiology, 1997, 143(10): 3381-3389
- [8] Kouji T, Natsuaki T, Okuda S. Phylogenetic analysis of elongation factor Tu gene of phytoplasmas from Japan[J]. Journal of General Plant Pathology, 2003, 69(5): 316-319
- [9] Malembic-Maher S, Salar P, Filippin L, et al. Genetic diversity of European phytoplasmas of the 16SrV taxonomic group and proposal of '*Candidatus* Phytoplasma rubi'[J]. International Journal of Systematic and Evolutionary Microbiology, 2011, 61(9): 2129-2134
- [10] Streten C, Gibb KS. Genetic variation in *Candidatus* Phytoplasma australiense[J]. Plant Pathology, 2005, 54(1): 8-14
- [11] Wang J, Zhu XP, Gao R, et al. Genetic and serological analyses of elongation factor EF-Tu of paulownia witches'-broom phytoplasma (16SrI-D)[J]. Plant Pathology, 2010, 59(5): 972-981
- [12] Miyata S, Furuki KI, Sawayanagi T, et al. Gene arrangement and sequence of *str* operon of phytoplasma resemble those of *Bacillus* more than those of *Mycoplasma*[J]. Journal of General Plant Pathology, 2002, 68(1): 62-67
- [13] Du HT, Zhu HY, Wang JM, et al. Single-nucleotide polymorphisms and activity analysis of the promoter and enhancer of the pig lactase gene[J]. Gene, 2014, 545(1): 56-60
- [14] Zhang Y, Han QL, Li CY, et al. Genetic analysis of the TBX1 gene promoter in indirect inguinal hernia[J]. Gene, 2014, 535(2): 290-293
- [15] Zhang YZ, Lin F, Xu CH. Construction of gene promoter-probe vectors of *Escherichia coli*[J]. Journal of Sichuan University (Natural Science Edition), 1998, 35(2): 263-267 (in Chinese)
张义正, 林峰, 胥成浩. 大肠杆菌基因启动子探针型载体的构建[J]. 四川大学学报: 自然科学版, 1998, 35(2): 263-267
- [16] Wang HY, Zhang YZ. Cloning of promoters from *Brassica napus* in *Escherichia coli*[J]. Acta Botanica Sinica, 1999, 41(5): 494-497 (in Chinese)
王海燕, 张义正. 利用大肠杆菌克隆在原核生物中有活性的油菜基因启动子[J]. 植物学报, 1999, 41(5): 494-497
- [17] Pan J, Zhang YZ. Cloning and characterization of gene promoters from *Bacillus subtilis*[J]. Acta Microbiologica Sinica, 2004, 44(4): 457-460 (in Chinese)
潘皎, 张义正. 枯草芽孢杆菌基因启动子的分离与鉴定[J]. 微生物学报, 2004, 44(4): 457-460
- [18] Lee IM, Hammond RW, Davis RE, et al. Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasma-like organisms[J]. Phytopathology, 1993, 83(8): 834-842
- [19] Dickinson M, Hodgetts J. Phytoplasma: Methods and Protocols[M]. Totowa: Humana Press, 2013
- [20] Liu YG, Whittier RF. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking[J]. Genomics, 1995, 25(3): 674-681
- [21] Post LE, Nomura M. DNA sequences from the *str* operon of *Escherichia coli*[J]. The Journal of Biological Chemistry, 1980, 255(10): 4660-4666
- [22] Lai F, Song CS, Ren ZG, et al. Molecular characterization of a new member of the 16SrV group of phytoplasma associated with *Bischofia polycarpa* (Levl.) Airy Shaw witches'-broom disease in China by a multiple gene-based analysis[J]. Australasian Plant Pathology, 2014, 43(5): 557-569