Bulletin of Insectology ***64*** *(Supplement): S89-S90, 2011*
ISSN 1721-8861

Rapid and automated analysis of single nucleotide polymorphisms (SNPs) in *secY* gene sequences for finer differentiation and characterization of phytoplasmas

Ing-Ming LEE, Yan ZHAO, Xiaobing SUO, Robert E. DAVIS
Molecular Plant Pathology Laboratory, USDA-Agricultural Research service, Beltsville, MD 20705, USA

Abstract

The number of newly discovered phytoplasma strains has increased dramatically in the last decade. For epidemiological studies and for international quarantine, appropriate molecular tools are essential to identify and distinguish phytoplasmas at taxonomic ranks at and below ribosomal group level. Currently, use of RFLP analysis of the *secY* gene is one of the finest tools available to achieve the resolving power needed for fine differentiation of closely related phytoplasmas. In the present study, we developed a sytem for computer-simulated SNP-based analysis of secY sequences that incorporates SNPs located both within and outside of restriction sites, increasing the number of characters beyond those used in RFLP analyses. With the inclusion of the additional informative characters, the SNP-based system further enhances strain separation and characterization. This new system will provide additional molecular markers and should aid identification and characterization of strains that are ecologically distinct and/or originate from different geographical regions.

Key words: RFLP, SNP, differentiation, phytoplasmas, *secY* gene.

Introduction

Phytoplasmas comprise a group of genetically diverse plant pathogens that are transmitted by insect vectors, and are associated to a variety of economically important diseases (Hogenhout *et al*., 2008). Several systems have been developed for differentiation of phytoplasma strains at different taxonomic levels (Lee *et al*., 2010). RFLP analysis of 16S rRNA gene sequences has yielded a comprehensive classification scheme comprising at present 31 phytoplasma groups and more than 100 subgroups (Lee *et al*., 1998; Zhao *et al*., 2009), each group representing at least one distinct '*Candidatus* species' level taxon.

Because the 16S rRNA gene is highly conserved, phytoplasmal 16S rRNA genes share similarities above 90% (Lee *et al*., 2010). Thus, the relative genetic distances among phytoplasma strains, assessed on the basis of 16S rRNA gene sequence similarities, may not fully reveal the genetic heterogeneity of phytoplasmas, and genetically close, but biologically distinct, strains may remain unresolved. To facilitate the separation of such closely related trains, several less-conserved molecular markers have been employed for phytoplasma classification. For example, RFLP analysis of the *secY* gene for differentiation of strains within a given 16Sr group provides high resolving power, enabling differentiation of closely related strains (Lee *et al*., 2010). SNPs have also been exploited as molecular markers separating phytoplasma lineages (Jomantiene *et al*., 2011). In the present study, we explored analysis of SNPs in *secY* gene sequences for improving strain differentiation. In this approach, we focused on SNPs, within and outside of restriction sites, for comparative analyses among strains within a given group, with the expectation that this approach will further enhance the resolving power of *secY* gene analyses.

Materials and methods

Phytoplasma *secY* gene sequences from 13 strains representative of five subgroups within the 16SrV group (table 1) were aligned using the ClustalW algorithm implemented in the Megalign program of the Lasergene software package.

Table 1. Thirteen 16SrV strains used in this study.

Strain	Geographic origin	16S rDNA RFLP classification	*secY* gene GenBank acc. numbers
EY1	USA	16SrV-A	GU004330
EYEu	Italy	16SrV-A	AY197690
EY626	USA	16SrV-A	AY197691
CLY5	China	16SrV-B	AY197693
PY-In	India	16SrV-B	AY197694
JWB	China	16SrV-B	AY197695
AldY882	Germany	16SrV-C	AY197692
FD70	France	16SrV-C	AY197686
FD-C	Italy	16SrV-C	AY197688
AldY	Italy	16SrV-C	AY197684
SpaWB229	Italy	16SrV-C	AY197689
FD-D	Italy	16SrV-D	AY197685
RuS	Italy	16SrV-E	AY197696

The resulting dendrogram separated these strains into seven distinct branches. Each of these branches contains a single strain or a strain cluster. Nucleotide sequence variations in each alignment position were examined, and subgroup- or strain cluster-specific SNPs were identified computationally (Zhao *et al.*, unpublished).

Results

A total of 256 SNPs were identified in the *secY* sequences (ca. 1,233 bp) of strains classified in the 16SrV group. SNP similarities among the 13 16SrV group strains were calculated; the values ranged from 26.6% to 94.1% (figure 1). Among the seven specific strain or strain clusters, the average similarity ranged from 32.0% to 77.2%. Restriction sites and SNPs in *secY* genes of strains FD-D and FD-C are shown in figure 2.

	ALY	FD70	FD-D	FD-C	SpaWB	ALY882	RuS	EY1	EYEu	EY626	CLY5	PY-In	JWB
ALY	100	84.0	84.8	77.7	62.5	67.6	64.1	52.0	52.0	52.0	36.3	35.2	35.2
FD70		100	83.2	77.7	63.3	68.0	64.5	52.0	52.0	52.0	36.3	35.2	35.9
FD-D			100	76.2	60.9	66.0	61.7	50.4	50.4	50.4	35.2	34.0	34.0
FD-C				100	64.8	68.0	65.2	51.2	51.2	51.2	34.0	32.8	33.6
SpaWB					100	60.5	57.8	51.6	51.6	51.6	27.7	26.6	27.0
ALY882						100	69.9	52.7	52.7	52.7	33.6	32.4	32.0
RuS							100	51.2	51.2	51.2	32.0	30.9	31.2
EY1								100	93.8	93.8	31.2	30.1	30.5
EYEu									100	94.1	31.2	30.1	30.5
EY626										100	31.2	30.1	30.5
CLY5											100	87.5	87.1
PY-In												100	85.2
JWB													100

Figure 1. Base similarities (%) in 256 SNP positions in the secY genes of 13 phytoplasma strains classified in 16S rDNA RFLP group 16SrV.
(In colour at www.bulletinofinsectology.org)

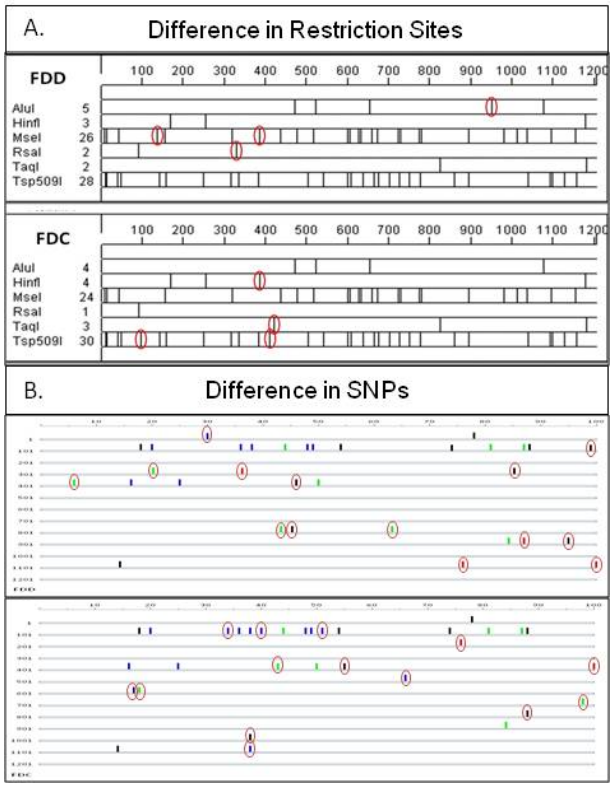

Figure 2. (A) Restriction sites in FD-D and FD-C. (B) SNPs in FD-D and FD-C (Color code for bases: A-green, T-red, C-blue, G-black). Ovals indicate differences between FD-D and FD-C.
(In colour at www.bulletinofinsectology.org)

Discussion

Two decades of research have revealed extensive diversity of phytoplasma ecology and indicated a remarkable complexity of phytoplasma population structure. For epidemiological studies and for international quarantines, appropriate molecular tools are essential to identify and distinguish phytoplasmas at taxonomic ranks at and below '*Candidatus* species' level. Identification and classification of strains may be achieved by RFLP analysis of secY sequences, but bases used for analysis in RFLP-based systems are limited to those within restriction sites. By maximizing inclusion of additional informative characters, the SNP-based analysis further enhances resolving power for strain differentiation and characterization. This new approach will provide additional molecular markers and should aid identification and characterization of strains that are ecologically distinct and/or originate from different geographical regions. A pattern recognition and pattern similarity coefficient calculation program is being developed to analyze query *secY* sequences for identification of SNP patterns.

Acknowledgements

We thank Prachi Bagadia and Kristi Bottner-Parker for technical assistance in preparing this extended abstract.

References

HOGENHOUT S. A., OSHIMA K., AMMAR E-D., KAKIZAWA S., KINGDOM H. N., NAMBA S., 2008.- Phytoplasmas: bacteria that manipulate plants and insects.- *Molecular Plant Pathology*, 9(4): 403-423.

JOMANTIENE R., ZHAO Y., LEE I-M., DAVIS R. E., 2011.- Phytoplasmas infecting sour cherry and lilac represent two distinct lineages having close evolutionary affinities with clover phyllody phytoplasma.- *European Journal of Plant Pathology*, 130: 97-107.

LEE I-M., GUNDERSEN-RINDAL D. E., DAVIS, R. E., BARTOSZK, I. M., 1998.- Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences.- *International Journal of Systematic Bacteriology*, 48: 834-842.

LEE I-M., BOTTNER-PARKER K. D., ZHAO Y., DAVIS R. E., HARRISON N. A., 2010.- Phylogenetic analysis and delineation of phytoplasmas based on *secY* gene sequences.- *International Journal of Systematic and Evolutionary Microbiology*, 60: 2887-2897.

ZHAO Y., WEI W., LEE I-M., SHAO J., SUO D., DAVIS R. E., 2009.- Construction of an interactive online phytoplasma classification tool, iPhyClassifier, and its application in analysis of the peach X-diseasephytoplasma group (16SrIII).- *International Journal of Systematic and Evolutionary Microbiology*, 59: 2582-2593.

Corresponding author: Ing-Ming Lee (e-mail: ingming.lee@ars.usda.gov), Molecular Plant Pathology Laboratory, USDA-ARS, Beltsville, MD 20705, USA.