SEQUENCE OF THE GENOME OF *XYLELLA FASTIDIOSA* CAUSING PIERCE'S DISEASE IN CALIFORNIA

Project Leader:
Edwin L. Civerolo
USDA-ARS
Parlier, CA 93648

Cooperators:

Marie-Anne Van Sluys	Mariana C. Oliveira	João P. Kitajima	João C. Setubal
Instituto de Biocências	Instituto de Biocências	Instituto Computação	Instituto Computação
USP, Sao Paulo Brazil	USP	Unicamp	Unicamp
Sao Paulo, Brazil	Sao Paulo, Brazil	Sao Paulo, Brazil	Sao Paulo, Brazil

FAPESP, ONSA
Agronomical and Environmental Genomics
Sao Paulo, Brazil

Reporting Period: The results reported are from work conducted from October 1, 2001 to September 30, 2002.

INTRODUCTION

Xylella fastidiosa (*Xf*) causes economically important diseases of agronomic, horticultural and landscape plants (Freitag 1951; Hopkins 1989; Purcell and Hopkins 1996). In addition to a wide diversity of *Xf* -host plant relationships, diseases such as Pierce's disease (PD) of grapevines and citrus variegated chlorosis (CVC) exhibit distinct symptoms and have different geographical distributions. In the previous reporting period, essentially the complete genomic sequence of an *Xf* strain associated with Pierce's disease in California was determined to help elucidate the molecular basis of *Xf* pathogenicity. Here we report the comparative analyses of the complete genome sequences and annotations of *Xf*-PD and *Xf*-CVC to provide further insight into *Xf*-plant host interactions and the relationships among *Xf* strains.

OBJECTIVES

1. Complete the sequencing of the genome of a *Xylella fastidiosa* strain associated with Pierce's disease (PD) in California.
2. Comparatively analyze the genome sequences and annotations of *Xylella fastidiosa* strains associated with PD in California and CVC in Brazil.

RESULTS AND CONCLUSIONS

The *Xf*-PD genome is composed of a single circular chromosome (2,519,802 bp) and a small plasmid (1,345 bp) similar to that reported in other *Xf* strains (Hendson et al. 2001). The major differences between the genomes of *Xf*-PD and *Xf*-CVC strains are the (1) 159,503 bp smaller size of the *Xf*-PD chromosome and (2) absence of the large pXF51 plasmid in the *Xf*-PD strain. Of the 2,066 protein coding genes annotated in *Xf*-PD, 2025 (98%) are also present in the *Xf*-CVC strain. The average amino acid identity of the ORF's in both strains is 95.7%. The most conserved *Xf*-PD genes include those that determine the basic metabolism and cellular functions of the bacterium, and, we conclude, are mostly identical to those of the *Xf*-CVC described previously (Simpson et al. 2000). Genomic structural/organizational differences between these two strains are associated with phage-mediated chromosomal rearrangements and deletions that also account for strain specific genes present in each genome (Figure 1). All of the rearrangements are flanked at one border by a putative phage-related integrase. Two genomic islands (gi), one specific to each genome, are characterized by regions with marked decreases in protein identities, different GC content and codon bias. In *Xf*-PD, giPD1 is 15.7 kb long with 61.2% GC content and harbors an extra copy of a hemagglutinin gene with a phage related integrase at one end. In *Xf*-CVC, giCVC1 is 67 kb long with 63.3% GC content and is inserted with tRNA Gly-2. The presence or absence of giPD1 and giCVC1 was associated with different groups of *Xf* strains. Essentially, all of the differences between the genomes of these two strains can be accounted for by the number and relative position of clusters of phage-related genes and insertion/deletion events, including giPD1 and giCVC1. We propose that the evolutionary divergence of these two *Xf* strains is due mainly to the lateral gene transfer mediated mostly by phage. Despite the genome rearrangements, most of the genes in these two strains are highly conserved including not only those concerned with basic cellular house keeping but also those likely to have a direct role in pathogenicity. This suggests that diseases caused by different *Xf* pathotypes most probably rely on the expression of a common set of bacterial genes to become established *in planta* (i.e., plant colonization, pathogenesis) permitting convergence of functional genomic strategies. Knowledge acquired from the comparison of the complete genomes of both *Xf*-PD and *Xf*-CVC strains has numerous applications, including designing strain specific primers for *Xf* detection and differentiation to screen germplasm and in clinical field samples to control pathogen dissemination.

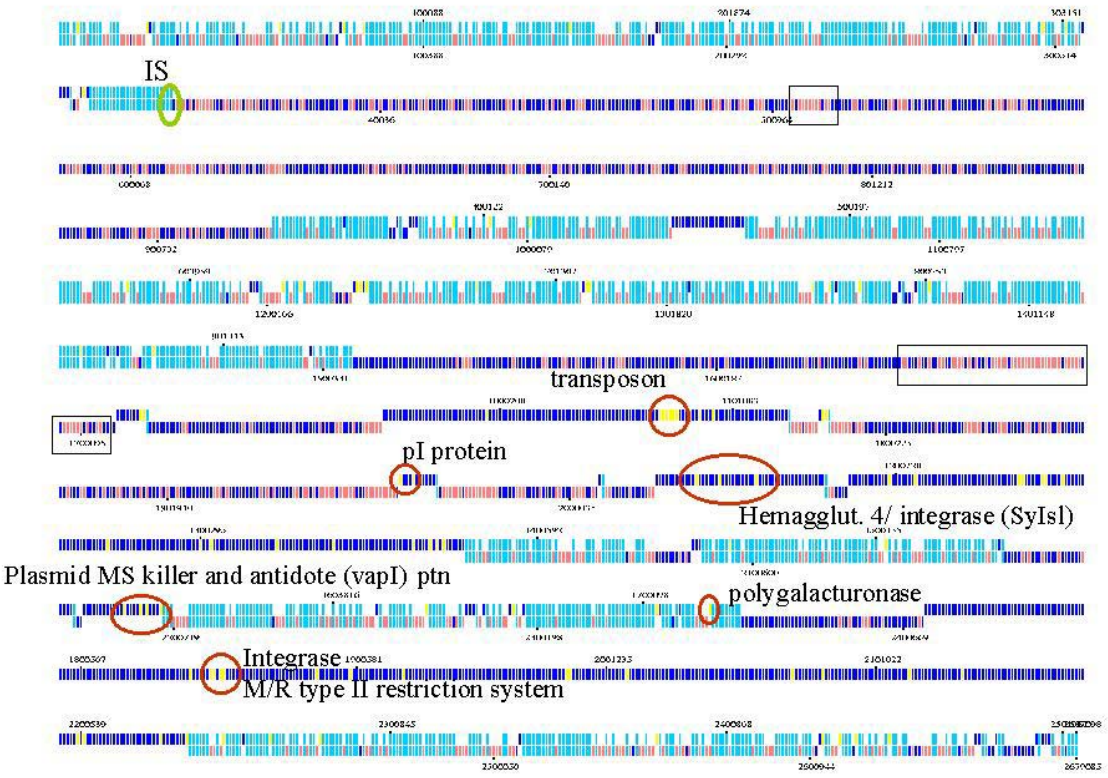

Figure 1. Chromosome alignment between *Xf*-PD Temecula and *Xf*-CVC 9a5c. Light blue dash represents colinear genes and dark blue dash represents rearranged genes in *Xf*-PD and *Xf*-CVC genome while yellow and red dashes represent strain specific genes some of which are highlighted.

REFERENCES

Freitag, H.H. 1951. Host range of Pierce's disease virus of grapes as determined by insect transmission. Phytopathology 41:920-202.

Hendson, M., A. H. Purcell, D. Chen, C. Smart, M. Guilhabert, and B. Kirkpatrick. 2001. Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. Appl. Environ. Microbiol. 67:895-903.

Hopkins, D.L. 1989. *Xylella fastidiosa*: a xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. 27:271-290.

Purcell, A.H. and D.H. Hopkins. 1996. Fastidious xylem-limited bacterial plant pathogens. Annu. Rev. Phytopathol. 34:131-151.

Simpson, A.J.G., *et al*. 2000. The genome sequence of the plant pathogen *Xlyella fastidiosa*. Nature 406:151-157.

FUNDING AGENCIES

Funding for this project was provided by the USDA Agricultural Research Service, the American Vineyard Foundation, the California Department of Food and Agriculture, the State of Sao Paulo Research Foundation (FAPESP), and the Brazilian National Council of Scientific and Technological Development (CNPq).