GENOME-WIDE IDENTIFICATION OF RAPIDLY EVOLVING GENES IN *XYLELLA FASTIDIOSA*: KEY ELEMENTS IN THE SYSTEMATIC IDENTIFICATION OF HOST STRAINS, AND IN THE SEARCH FOR PLANT-HOST PATHOGENICITY CANDIDATE GENES

Project Leaders:
Leonard Nunney
Department of Biology
University of California
Riverside, CA 92521

Richard Stouthamer and Robert F. Luck
Department of Entomology
University of California
Riverside, CA

Cooperator:
Donald A. Cooksey
Department of Plant Pathology
University of California
Riverside, CA

Reporting Period: The results reported here are from work conducted from September 1, 2002 (start of grant) through November 1, 2002.

INTRODUCTION

Xylella fastidiosa is not only a bacterium that causes Pierce's disease (PD) of grapevines. It also has a number of genetically distinct host strains (see Hopkins 1989). These strains show varying levels of cross-host pathogenicity. For example, PD strains do not infect peach, and phony peach disease strains do not infect grapevines. Similarly, the strains causing oleander leaf scorch (OLS) do not cause PD in grapevines or cause infectious symptoms in a number of other species (Purcell et al. 1999). The lack of cross-host infection means that genetic differences among the strains must encode the causes of host-specific pathogenicity. Identifying the genes responsible for host-specific effects is an important step in understanding how infection might be controlled.

The genetic differences determining host-specific adaptations are probably only a small fraction of the total genetic differences between the strains. To facilitate identifying host-specificity candidate genes, we need some initial filter that selects those genes most likely to be involved in host adaptation from among the approximately 2700 genes carried by *X. fastidiosa*. To this end, genomic research provides us with some extraordinarily powerful new tools for solving this kind of applied problem.

It is self-evident that adaptive evolution depends upon changes in specific genes. In some cases, a single base substitution in a gene may be sufficient; however, such simple changes recur repeatedly in bacterial populations. The apparent separation of *X. fastidiosa* into stable host strains suggests host adaptations involve more complex changes. For this reason, we believe that the genes involved in host adaptation will be among those exhibiting the most rapid evolutionary change.

OBJECTIVES

The identification of the rapidly evolving genes in the *Xylella fastidiosa* genome. This is the first step towards achieving our four primary objectives. These are:

1. Develop a systematic multigenic method for identifying host strains of *X. fastidiosa.* Our objective is to develop a method that unambiguously identifies the known host strains, and that allows researchers to efficiently recognize the invasion of new strains.
2. Identify plant-host specificity candidate genes. We will use our database of rapidly evolving proteins to test for evidence of strong natural selection and for statistical links between the rapid genetic divergence of host strains and specific biochemical functions.
3. Measurement of clonal variation within host strains. Our objective is to assess within-strain genetic variability at rapidly evolving gene loci and to use these results to assess the evidence that all members of a given host strain share common ancestry.
4. Estimate the frequency of recombination. Our objective is to look for evidence of both within- and between-strain genetic transfer. Genetic transfer can dramatically increase the rate of evolution, and potentially can increase the rate at which new host strains arise.

RESULTS AND CONCLUSIONS

There are estimated to be about 2700 genes in the genome of the CVC (citrus) strain of *X. fastidiosa*; however, many of these genes are of unknown function. In the first phase of this project, we have screened just over 1000 genes of known function from the genomes of the OLS (oleander), ALS (almond) and CVC (citrus) strains. Using a modified relative rate test (with CVC as the outgroup), we have examined the evolutionary pattern of these genes in the OLS and ALS lineage.

Results so far suggest that the rate of gene evolution is generally higher in the OLS strain, with 15% of the genes showing significantly higher rates of change in OLS and 5% in the ALS strain. Since this includes synonymous (silent) and non-synonymous (replacement) base pair substitutions, it is possible (and probable) that the bias is due to a shorter generation time in the OLS strain. On the other hand, we have found a slight trend for faster protein evolution in the ALS strain. After correcting for the underlying rate differences, we find that almost exactly 10% of the 1070 genes screened show unequal rates of protein evolution between OLS and ALS. Of these 106 genes, 75% show faster evolution in ALS, and 25% in OLS. These genes are the initial candidates for being involved in host adaptation. However, since we have done 1070 tests, we can expect about 53 type 1 statistical errors (i.e. 5% of 1070), so that a conservative estimate is that only about 50% of our candidate genes are "real". However, even with this correction, it appears that more than 50 genes in our sample are likely to be involved in adaptive differences between the OLS and ALS strains.

REFERENCES

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

Purcell, A.H., S.R. Saunders, M. Hendson, M.E. Grebus, and M.J. Henry. 1999. Causal role of *Xylella fastidiosa* in oleander leaf scorch disease. Phytopathology 89:53-58.

FUNDING AGENCIES

Funding for this project was provided by the University of California Pierce's Disease Grant Program.