DETERMINATION OF GENES CONFERRING HOST SPECIFICITY IN GRAPE STRAINS OF *XYLELLA* FASTIDIOSA USING WHOLE-GENOMIC DNA MICROARRAYS

Project Leaders:

Steven E. Lindow
Department of Plant and Microbial Biology
111 Koshland Hall
University of California
Berkeley, CA 94720

Paul Richardson
Head-Genomic Technologies
DOE, Joint Genome Institute
2800 Mitchell Drive, B400
Walnut Creek, CA 94598

Reporting Period: The results reported here are from research conducted from July 1, 2003 to October 15, 2003.

ABSTRACT

The genomic sequence of several *Xylella fastidiosa* (*Xf*) strains, including one pathogenic to grape (strain 'Temecula1') have been determined. While these strains share most genes, the difference in host range that they exhibit is presumably due at least in part to unique traits that each pathovar possesses. Some of these strains originating from host plants other than grape do not sustain viable populations or are not virulent in grape. Using microarray technology and *Xf* 'Temecula1' as the reference strain, those genes present in the grape strain but not in strains that are unable to grow in grape can be determined through subtractive analysis. A 70-mer-oligonucleotide microarray representing 2543 of the 2574 predicted ORFs of *Xf* 'Temecula1' was designed. *Xf* 'Temecula1' genes identified as putative host-specificity genes of will be subjected to further analysis. Identifying host-specific genes of the grape strain will allow us to target those genes in gene-knockout studies and determine their contribution to virulence in grape.

INTRODUCTION

Some strains of *Xylella fastidiosa* originating from host plants other than grape do not sustain viable populations or are not virulent in grape. In particular, many of the almond strains of *X. fastidiosa* do not infect grape (Almeida and Purcell, *unpublished*). Cross inoculation greenhouse studies showed that the OLS and PD strains of *X. fastidiosa* were not pathogenic to citrus and that the ALS strain was not pathogenic to oleander (Feil et al. *unpublished*). Other cross-inoculation studies provide evidence for host specificity variation among strains of *X. fastidiosa*. Recently, the CVC strain of *X. fastidiosa* was found to be pathogenic to coffee and grape (Li et al. 2001) and these authors are now inoculating citrus with the coffee leaf scorch strain to determine if it is pathogenic to citrus.

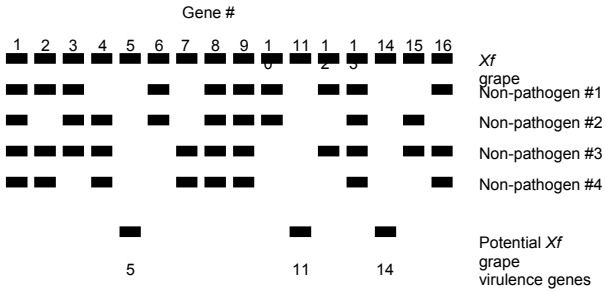
Other studies provide evidence for host specificity among the *X. fastidiosa* strains. Grape strains of *Xylella fastidiosa* were found to cluster together away from oak, plum, mulberry, and periwinkle strains using RFLP data (Chen et al. 1992, Chen et al. 1995). Pooler and Hartung (1995) divided the *X. fastidiosa* into 5 groups (citrus, plum, grape-ragweed, almond, and mulberry) based on RAPD-PCR data. Most almond strains clustered away from the grape strains but a few clustered within the grape strains group whereas oleander, peach, and oak strains were distinct from other strains using RAPD-PCR, CHEF gel electrophoresis, and 16S-23S rRNA sequence analysis (Hendson et al. 2001). In Brazil, CVC strains were found to cluster together away from the coffee and grape strains of *X. fastidiosa* using RAPD (Lacava et al. 2001). At best, the previous studies were able to group the relative strains into groups. The limitations of these earlier studies were that only a small portion of the genome was sampled, and therefore many of the virulence determinants were likely to be missed. The loci they studied may or may not have been related to virulence. In this study we are looking at all of the genes of the genome.

We believe that microarray technology is a good technique to study host specificity between different strains of the same bacteria. The main reason is that the entire genome is represented on a chip to give an immediate global comparison tool. Many studies have shown that microarray technology can differentiate bacterial strains. Microarray analysis of 22 strains of *Salmonella* revealed that approximately 400 genes (out of 4,596) were unique to the reference strain (Porwollick et al. 2002). Another microarray study showed that one serovar of *Salmonella* (serovar Arizona) had a more distant relationship to the subspecies I serovars (Chan et al. 2003). Microarray analysis of 50 strains of *Listeria monocytogenes* relative to a reference array made with sheared DNA from 10 different strains revealed that a number of the polymorphic probes identified in the study are important in the virulence of their respective strain (Call et al. 2003). Strain-specific virulence genes were identified for *Helicobacter pylori* using whole-genome microarray technology (Salama et al. 2000). These authors found that 22% of the *H. pylori* genome was dispensable in other strains and they defined a minimal functional core of 1281 genes but more importantly a differentiating section of the genome referred to as pathogenicity island made of 362 strain-specific genes. Similarly, we will use a *X. fastidiosa* 70-mer-oligonucleotides microarray using the grape strain to do global comparison between the grape strain relative to other strains non-pathogenic to grape.

Comparative analysis of curated pathways and their components present in the finished sequence of the citrus strain of *X. fastidiosa* genome relative to the draft-genomes of the almond and oleander strains of *X. fastidiosa* were described (Bhattacharyya et al. 2002). Based on this analysis, we estimate that ca. 4% of the whole genome of the oleander was unique to that strain. We hypothesized that the grape strain also possesses ca. 4% of unique genes in comparison to other strains that do not infect grape (the CVC strain does infect grape and is thus not helpful in finding host-specificity genes). This means that approximately 100-200 genes are unique to the grape strain. However probably not all 100 to 200 genes will be of

interest and related to virulence, and we can expect that only 20 to 30 will be virulence determinants. To identify these genes, we will use the grape strain microarray as a reference and perform pairwise comparison via hybridization experiments using each *X. fastidiosa* strain that is non-pathogenic to grape.

The cartoon below depicts the procedure:



The three fragments on the bottom of the cartoon represent the unique genes found in the grape strain but not in the other four non-pathogenic strains. The more non-grape pathogens and grape pathogens that can be compared in this process, the more selective such a "filter' will be for those genes that are in common only to grape pathogens. In this process we will eliminate the many housekeeping genes and genes involved in general virulence attributes that all xylem colonists would have, and identify specifically those genes that distinguish grape pathogens from all others. High-density arrays will be generated using 70-mer oligonucleotides that will be designed based on the coding sequence of the >2500 annotated genes from the grape *X. fastidiosa* genome 'Temecula1' (http://integratedgenomics.com). DNA extraction, labeling, hybridization and data analysis will be conducted as described (Salama et al. 2000). Using the subtractive method described above we will identify the unique virulence genes for the *X. fastidiosa* grape strain 'Temecula1' and we will perform site-directed mutagenesis and grape inoculation experiments to investigate their pathogenicity.

OBJECTIVES

1. To identify host-specific virulence determinants of the *X. fastidiosa* grape strain 'Temecula1'.
2. To investigate the role of host specificity genes identified in objective 1 in virulence.

RESULTS

Oligonucleotide and Array Design. 70-mer oligodeoxynucleotides were designed using 'ArrayOligoSelector' software (http://arrayoligosel.sourceforge.net) based on the coding sequence of 2543 of the 2574 predicted ORFs of *Xf* 'Temecula1'. These oligos were generated with a 5' amino linker that will allow for covalent binding to aldehyde or epoxy coated slides. Oligos are in the process of being printed onto glass slides using a GMS 417 Arrayer. The total number of genes represented by gene-specific oligodeoxynucleotides on the arrays will be 2555 including negative and positive controls.

ArrayOligoSelector did not design oligos for the 48 ORFs listed in Table 1. The list includes 19 ORFs which were less than 70 bp in length, and 12 ORFs that were greater than 70 bp but less than 300 bp. The remaining ORFs were from duplicated regions, and the program could not determine a unique sequence for those ORFs. We will manually design 70-mer oligodeoxynucleotides for ORFs greater than 300 bp in length, but only one oligo will be used to represent duplicated ORFs. The remaining oligos that will be included on the array will include 8 negative and 4 positive controls. We are now optimizing the experimental conditions for the printing and hybridization steps of the microarray.

CONCLUSION

We are comparing the genomes of several strains of *X. fastidiosa* that do not sustain viable colonies or are non-pathogenic to grape (i.e several almond, oleander, oak, peach, etc…) to the genome of the grape strain 'Temecula1' by performing hybridization experiments using a microarray representing the whole genome of the grape reference strain 'Temecula1'. In a process of elimination those genes that are in common to all grape strains but missing from other pathovars will be identified as putative host-specificity genes. Their contribution to virulence to grape will be verified by using site-specific gene knockouts of each of the candidate genes in inoculation trials. We are confident that by comparing the genome of the grape strain to many other strains that are non-virulent to grape, we will identify genes unique to the grape strain. The list of putative host-specificity genes will be examined for their contribution to pathogenicity and virulence in subsequent mutagenesis experiments in objective 2. Methodologies for gene knock-outs are working well in our lab, and we expect no problem in being able to disrupt candidate genes as they are identified in objective 1.

Table 1. List of ORFs for which oligonucleotides were not designed by ArrayOligoSelector.

ORF #	Annotation	Length (bp)
RXFZ00021	Hypothetical	192
RXFZ00036	Hypothetical	69
RXFZ00075	Hypothetical membrane spanning protein	840
RXFZ00178	Hypothetical	178
RXFZ00179	Formamidopyrimidine-DNA glycosylase	813
RXFZ00282	Hypothetical	57
RXFZ00404	Hypothetical	60
RXFZ00413	Hypothetical	57
RXFZ00415	Hypothetical	54
RXFZ00595	Hypothetical	282
RXFZ00627	hypothetical	57
RXFZ00643	hypothetical	63
RXFZ00724	hypothetical	60
RXFZ00815	hypothetical	57
RXFZ01099	hypothetical	63
RXFZ01239	hypothetical	222
RXFZ01318	hypothetical	75
RXFZ01429	Tail protein I	567
RXFZ01430	Baseplate assembly protein J	891
RXFZ01438	hypothetical	519
RXFZ01439	hypothetical	324
RXFZ01440	hypothetical	258
RXFZ01476	Phage-related protein	564
RXFZ01480	Phage-related protein	387
RXFZ01532	Phage-related protein	558
RXFZ01537	Phage-related protein	639
RXFZ01538	Phage-related protein	828
RXFZ01539	Hypothetical	315
RXFZ01557	Phage-related DNA helicase	1416
RXFZ01558	Hypothetical	168
RXFZ01590	Hypothetical	201
RXFZ01812	Hypothetical	60
RXFZ01834	Hypothetical	60
RXFZ02159	Hypothetical	57
RXFZ02169	Hypothetical	96
RXFZ02255	Hypothetical	60
RXFZ02281	Hypothetical	195
RXFZ02282	Hypothetical cytosolic protein	258
RXFZ02283	Phage-related protein	1275
RXFZ02285	Antirepressor	1599
RXFZ02286	AhpD protein	2178
RXFZ02287	Phage-related protein	276
RXFZ02289	Phage-related DNA helicase	1416
RXFZ02297	Hypothetical	69
RXFZ02461	Hypothetical	63
RXFZ02612	Hypothetical	60
RXFZ02696	Hypothetical	66
RXFZ02698	Hypothetical	60

REFERENCES

Bhattacharyya, A, S. Stilwagen, N. Ivanova, M. D'Souza, A. Bernal, A. Lykidis, V. Kapatral, I. Anderson, N. Larsen, L. Tamara, G. Reznik, E. Selkov, T. Walunas, H. Feil, W.S. Feil, A.H. Purcell, T. Hawkins, R. Haselkorn, R. Overbeek, P.F. Predki, and N.C. Kyrpides. 2002a. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. Proc. Natl. Acad. Sci. 99: 12403-12408.

Call, D.R., M.K. Borucki, and T.E. Besser. 2003. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes*. J. Clinic. Microbiol. 41: 632-639.

Chan K., S. Baker, C.C. Kim, C.S. Detweiler, G. Dougan, and S. Falkow. 2003. Genomic comparison of *Salmonella enterica* serovars and *Samonella bongori* by use of an *S. enterica* serovar *typhimurium* DNA microarray. J. Bacteriol. 185: 553-563.

Chen, J., C.J. Chang, R.L. Jarret, and N. Gawel. 1992. Genetic variation among *Xylella fastidiosa* strains. Phytopathology 82: 973-977.

Chen, J., O. Lamikanra, C.J. Chang, and D.L. Hopkins. 1995. Randomly amplified polymorphic DNA analysis of *Xylella fastidiosa* Pierce's disease and oak leaf scorch pathotypes. Applied Environ. Microbiol. 61: 1688-1690.

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*. Applied and Environmental Microbiology 67: 895-903.

Lacava, P.T., W.L. Araujo, W. Maccheroni Jr., and J.L. Azevedo. 2001. RAPD profile and antibiotic susceptibility of *Xylella fastidiosa* causal agent of citrus variegated chlorosis. Letters Applied Microbiol. 33: 302-306.

Li, W.-B., W.D. Pria, D.C. Teixeira, V.S. Miranda, A.J. Ayres, C.F. Franco, M.G. Costa, C.-X. He, P.I. Costa, and J.S. Hartung. 2001. Coffee leaf scorch caused by a strain of *Xylella fastidiosa* from citrus. Plant Dis. 85: 501-505.

Pooler, M.R., and J.S. Hartung. 1995. Genetic relationships among strains of *Xylella fastidiosa* from RAPD-PCR data. Current Microbiol. 31: 134-137.

Porwollik, S., R.M.-Y. Wong, and M. McClelland. 2002. Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis. Proc. Natl. Acad. Science 99: 8956-8961.

Salama, N., K. Guillemin, T.K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow. 2000. A whole-genome microarray reveals genetic diversity among Helicobacter pylori strains. Proc. Nat. Acad. Sci. 97: 14668-14673.

FUNDING AGENCIES
Funding for this project was provided by the CDFA Pierce's Disease and Glassy-winged Sharpshooter Board.