11-2007

Array-Based Genomic Comparative Hybridization Analysis of Field Strains of Mycoplasma hyopneumoniae

Melissa L. Madsen
Iowa State University, mlmadsen@iastate.edu

Michael J. Oneal
Iowa State University

Stuart W. Gardner
Iowa State University

Erin L. Strait
Iowa State University

Dan Nettleton
Iowa State University, dnett@iastate.edu

See next page for additional authors

Array-Based Genomic Comparative Hybridization Analysis of Field Strains of Mycoplasma hyopneumoniae

Abstract

Mycoplasma hyopneumoniae is the causative agent of porcine enzootic pneumonia and a major factor in the porcine respiratory disease complex. A clear understanding of the mechanisms of pathogenesis does not exist, although it is clear that *M. hyopneumoniae*adheres to porcine ciliated epithelium by action of a protein called P97. Previous studies have shown variation in the gene encoding the P97 cilium adhesin in different strains of *M. hyopneumoniae*, but the extent of genetic variation among field strains across the genome is not known. Since *M. hyopneumoniae* is a worldwide problem, it is reasonable to expect that a wide range of genetic variability may exist given all of the different breeds and housing conditions. This variation may impact the overall virulence of a single strain. Using microarray technology, this study examined the potential variation of 14 field strains compared to strain 232, on which the array was based. Genomic DNA was obtained, amplified with TempliPhi, and labeled indirectly with Alexa dyes. After genomic hybridization, the arrays were scanned and data were analyzed using a linear statistical model. The results indicated that genetic variation could be detected in all 14 field strains but across different loci, suggesting that variation occurs throughout the genome. Fifty-nine percent of the variable loci were hypothetical genes. Twenty-two percent of the lipoprotein genes showed variation in at least one field strain. A permutation test identified a location in the *M. hyopneumoniae* genome where there is spatial clustering of variability between the field strains and strain 232.

Disciplines

Genetics and Genomics | Microarrays | Statistical Models | Veterinary Microbiology and Immunobiology | Veterinary Preventive Medicine, Epidemiology, and Public Health

Comments

This is a manuscript of an article published as Madsen, Melissa L., Michael J. Oneal, Stuart W. Gardner, Erin L. Strait, Dan Nettleton, Eileen L. Thacker, and F. Chris Minion. "Array-based genomic comparative hybridization analysis of field strains of Mycoplasma hyopneumoniae." *Journal of bacteriology* 189, no. 22 (2007): 7977-7982. doi: 10.1128/JB.01068-07. Posted with permission.

Authors

Melissa L. Madsen, Michael J. Oneal, Stuart W. Gardner, Erin L. Strait, Dan Nettleton, Eileen L. Thacker, and F. Chris Minion

JB Accepts, published online ahead of print on 14 September 2007
J. Bacteriol. doi:10.1128/JB.01068-07

Array-Based Genomic Comparative Hybridization Analysis of Field Strains of *Mycoplasma hyopneumoniae*

Melissa L. Madsen,[1] Michael J. Oneal,[1] Stuart W. Gardner,[2] Erin L. Strait,[1] Dan Nettleton,[2] Eileen L. Thacker,[1] and F. Chris Minion[1]*

Departments of Veterinary Microbiology and Preventive Medicine[1] and Statistics[2], Iowa State University, Ames, IA

*Corresponding author. Mailing address: Department of Veterinary Microbiology and Preventive Medicine, Iowa State University, Ames, IA 50011. Phone: 515-294-6347. Fax: 515-294-1401. E-mail: fcminion@iastate.edu.

Abstract

Mycoplasma hyopneumoniae is the causative agent of porcine enzootic pneumonia and a major factor in the porcine respiratory disease complex. A clear understanding of the mechanisms of pathogenesis does not exist although it is clear that *M. hyopneumoniae* adheres to porcine ciliated epithelium by action of a protein called P97. Previous studies have shown variation in the gene encoding the P97cilium adhesin within different strains of *M. hyopneumoniae*, but the extent of genetic variation among field strains across the genome is not known. Since *M. hyopneumoniae* is a worldwide problem, it is reasonable to expect that a wide range of genetic variability may exist given all of the different breed and housing conditions. This variation may impact the overall virulence of a single strain. Using microarray technology, this study examined potential variation of fourteen field strains in comparison to strain 232 on which the array was based. Genomic DNA was obtained, amplified with TempliPhi™, and labeled indirectly with Alexa dyes. Post genomic hybridization, the arrays were scanned and data analyzed using a linear statistical model. Results indicate that genetic variation could be detected in all fourteen field strains but across different loci, suggesting that variation occurs throughout the genome. Fifty-nine percent of the variable loci were hypothetical genes. Twenty-two percent of the lipoprotein genes showed variation in at least one field strain. A permutation test identified a location in *M. hyopneumoniae* genome where spatial clustering of variability between the field strains and strain 232 exists.

Introduction

Genetic variation is thought to occur among bacterial species as a survival mechanism in both adverse environmental and host niches. A natural consequence of evolutionary and environmental pressures, genetic variation in pathogenic species results in changing phenotypes. Genetic variation can occur by three general mechanisms, local nucleotide sequence changes, intragenomic recombination resulting in reshuffling of genome sequences, and acquisition of foreign DNA (1). It can also occur vertically and horizontally. Vertical transmission refers to passage of genetic material to siblings through cell division and its accompanying replication mistakes (point mutations, inversions and spontaneous deletions). Horizontal transmission involves the acquisition of new genetic material. This could occur among closely related species by transformation of native DNA, transduction by phages or by conjugation mechanisms, giving rise to organisms with subtle changes in phenotype. Alternatively, it could occur between dissimilar species by similar mechanisms that result in dramatic changes in phenotype. For example, pathogenicity islands are thought to arise by uptake and insertion of large DNA segments that encode large blocks of genes related to a virulence phenotype (8, 23). While it is clear that numerous species of bacteria have acquired large segments of DNA (24), there is no evidence for gene acquisition in mycoplasmas.

Mycoplasmas are cell wall-less bacteria that are thought to be the smallest organisms capable of self-replication. Their genome sizes range from 580 kilobases to over 1,700 kilobases (20). Their small genomes do not restrict their ability to generate high rates of diversity, however. This type of variation is not a consequence of environmental signals, but rather occurs through random events. There are numerous examples of both small sequence changes and recombination to introduce genetic variation in mycoplasmas (33). This variation is usually expressed by the

generation of new chimeric surface molecules with high rates of antigenic diversity. The mechanisms by which this occurs include slipped strand mispairing during DNA replication and recombination between homologous sequences. Genetic variation can result in phase switching when it occurs within homopolymeric tracts of adenine in promoter regions (34) or in structural gene sequences (36), or by DNA inversion (14, 25). The generation of chimeric genes by intragenic recombination also occurs (13). There is no evidence that any of these mechanisms are operative in *Mycoplasma hyopneumoniae*, however. Analysis of the *M. hyopneumoniae* genome sequence failed to identify families of lipoprotein genes that could undergo phase switching through mechanisms employed by other mycoplasmas for surface variation (18).

Mycoplasma hyopneumoniae is the primary agent of porcine enzootic pneumonia (21). There is increasing evidence that *M. hyopneumoniae* has a predisposing influence on other infectious agents (26, 27, 32). Genetic variation is known to occur in *M. hyopneumoniae* (2, 4, 10, 30), but there are few studies examining the extent of variation within field isolates at the molecular level (29). Phenotypic variation does occur within *M. hyopneumoniae* as described by Young et al. within the context of protein immunoblotting (35) and in some cases within specific genomic regions (30), but no studies have been reported that examine genetic differences in field strains of *M. hyopneumoniae* within genes on a global basis. This is due to the difficulty in isolating and cloning *M. hyopneumoniae* from field samples and the fact that adequate tools have not been available until recently (18).

The studies reported here examine genetic variation in *M. hyopneumoniae* on a genome-wide basis using microarray technology. The arrays were based upon the genome sequence of strain 232 (18). Our results with fourteen field strains show that microarrays can be used to examine

genetic diversity and that all of the strains of *M. hyopneumoniae* vary in at least one genetic locus.

Materials and Methods

Mycoplasma strains and culture conditions. Pathogenic *M. hyopneumoniae* strain 232, a derivative of strain 11, was used in this study (17). Fourteen field strains were cultured from case studies from the Iowa State University Veterinary Diagnostic Laboratory (Ames, Iowa) and are from the US Midwest. All *M. hyopneumoniae* strains were grown in Friis media as previously described (7) and are from *in vitro* passage less than 15. Cultures consisted of 125 ml of Friis media in 250 ml Erlenmeyer flasks incubated at 37°C with slow agitation until the culture reached mid log phase as indicated by color change and turbidity. Mycoplasmas were pelleted by centrifugation at 24,000 x *g*, and the cell pellets were stored at -70°C until the chromosomal DNA was isolated.

Microarray. The *M. hyopneumoniae* microarray consists of polymerase chain reaction (PCR) products (probes) spotted to Corning UltraGAPS™ glass substrates (Corning, Inc., Big Flats, N.Y.). Eighty-nine percent (620/698) of the total open reading frames (ORFs) of strain 232 are represented on the array as PCR products of approximately 125-350 base pairs in length. Each product is a unique sequence even within paralogous families as described by Minion et al. (18). No tRNA or ribosomal RNA sequences were included. The primer design, array construction and validation have been described (15, 16). Each slide was divided into two regions (upper and lower), and each region contained the full array of spots, printed in triplicate in a noncontiguous well-spaced format. This design allowed two independent hybridizations simultaneously to reduce variation due to slide interactions.

Experimental design. TempliPhi™ amplified DNA samples from field isolates were compared to control strain 232 using a two-color experimental microarray design. Independent samples from one isolate labeled with one dye were paired with control samples labeled with the alternate dye; the samples were mixed and hybridized to the microarray. For nine of the fourteen isolates (95MP1501, 95MP1502, 95MP1503, 95MP1504, 95MP1508, 95MP1509, 97MP0001, 00MP1301, and 05MP2301), four independent field isolate DNA samples were paired with four independent DNA samples from control 232. In two of the four arrays, the control sample was labeled with Alexa 555 dye and compared to the field isolate sample labeled with Alexa 647 dye (Molecular Probes, Inc., Eugene, Ore.). The dye assignment to control and treated samples was reversed for the other two arrays (dye swap). The arrays were hybridized under identical conditions as described below. This procedure was repeated for isolates 95MP1510 for a total of four arrays, where the control sample was labeled with Alexa 647 dye in three of the arrays and Alexa 555 dye for the fourth array. For isolates 95MP1505, 95MP1506 and 95MP1507, a total of five arrays each, including two dye swaps, were done; and for isolate 00MP1502, a total of six arrays were done with the control labeled with Alexa 555 dye for four of the arrays and Alexa 647 for the other two arrays.

DNA isolation. DNA was isolated from frozen cell pellets as follows. The cells were first resuspended in 1 ml of TNE buffer (10 mM Tris, 140 mM sodium chloride, 1 mM ethylenediamine tetraacetic acid, pH 8.0), and Proteinase K was added to a final concentration of 70 µg/ml. The suspension was incubated at 50°C for 5 min, and then sodium dodecyl sulfate was added to a final concentration of 0.1% and incubation was continued at 50°C for 4 h. The suspension was then extracted with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol three times, and the DNA was precipitated by the addition of one tenth volume of 3 M

sodium acetate and bringing the solution to 70% ethanol as described (22). The DNA pellets were dissolved in nuclease-free water, and samples were quantified and checked for purity using the Nanodrop® ND-1000 Spectrophotometer (Nanodrop, Wilmington, Del.).

TempliPhi™ reactions. Field isolate samples yielded low amounts of genomic DNA compared to strain 232 due to their fastidious growth and lack of adaptation to growth media. To overcome the issue of limited quantities of DNA, genomic samples were amplified using the TempliPhi™ 100 Reaction Kit (Amersham, Biosciences, Piscataway, N.J.) according to the manufacturer's protocol. A total of five reactions were combined for each field isolate and strain 232, yielding approximately 5-8 µg total DNA in each preparation which was subjected to mechanical shearing.

Nebulization. The DNA was mechanically sheared prior to labeling to ensure an optimized fragment size for efficient labeling and hybridization. Each amplified sample was added to the modified nebulizer (product # 4100, MEDEX, Carlsbad, Calif.) containing 2 ml of sterile 50% glycerol. The nebulizer was modified by removing the plastic cuff, trimming the edge and inverting it during reassembly. The samples were sheared using a 10 psi nitrogen stream for 15 min. The fragment size of less than 1,000 base pairs was optimal for efficient labeling and signal strength. This was confirmed by gel electrophoresis on 1.5% agarose gel.

Target generation and hybridization. Targets were generated and purified from mechanically sheared DNA samples using the BioPrime® Plus Array CGH Indirect Genomic Labeling System (Invitrogen Corp., Carlsbad, Calif.). A set of 129 open reading frame-specific hexamer oligonucleotide primers (15) was used to generate amino-allyl modified DNA targets. These targets were then labeled with either Alexa Fluor™ 555 Reactive Dye or Alexa Fluor™ 647 Reactive Dye (Molecular Probes, Inc.) according to the experimental design. Following

purification of the fluorescently labeled cDNA per manufacturer's instructions, samples were dried in a vacuum centrifuge and then resuspended in 10 µl Pronto! cDNA/long oligo hybridization solution (Corning). Targets were denatured at 95°C for 5 min and centrifuged at 13,000 x *g* for 2 min at room temperature. Labeled targets from one 232 control and one field isolate were then combined, pipetted to an array, and covered with a 22 x 22 mm HybriSlip™ (Schleicher & Schuell, Keene, N. H.). Slides were placed in a Corning hybridization chamber and incubated in a 42°C water bath for 12-16 h. Slides were washed according to Corning's UltraGAPS™ protocol and dried by centrifugation.

Data acquisition and normalization. Eight of the fourteen isolate arrays (95MP1504, 95MP1505, 95MP1506, 95MP1507, 95MP1509, 95MP1510, 00MP1301, and 00MP1502) were scanned with each dye channel using a ScanArray Express laser scanner (Applied BioSystems, Inc., Foster City, Calif.) under varying laser power and PMT gain settings to increase the dynamic range of measurement (5). The other six arrays (95MP1501, 95MP1502, 95MP1503, 95MP1508, 97MP0001, 05MP2301) were scanned with an Applied Precision's ArrayWoRx® Biochip Reader (Applied Precision, Inc., Issaquah, Wash.).

Images were analyzed for spots and signal intensities quantified using the softWorRx Tracker software package (Applied Precision, Inc.). Spot-specific mean signals were corrected for local background by subtracting spot-specific median background intensities. The natural logarithm of the background-corrected signals from a single scan were adjusted by an additive constant so that all scans of the same array-by-dye combination would have a common median. The median of these adjusted-log-background-corrected signals across multiple scans was then computed for each spot to obtain one value for each combination of spot, array, and dye channel. These data for the two dye channels on any given array were normalized using LOWESS normalization to

adjust for intensity-dependent dye bias (6, 31). Following LOWESS adjustment, the data from each channel were adjusted by an additive constant so that the median for any combination of array and dye would be the same for all array-by-dye combinations. The difference in normalized values for each spot was calculated as the signal intensity of Alexa 555 dye minus Alexa 647 dye. The differences for the triplicate spots were then averaged within each array to produce one normalized difference value for each of the 627 probe sequences.

Data analysis. A linear model of the difference in signal intensity for the two dyes was fitted for each probe sequence using the normalized data. The model included an overall mean for the difference in dye effect (Alexa 555 minus 647) and, for each field isolate, a fixed effect for the difference in signal intensity of control minus field isolate. As part of each linear model analysis, a one-sided *t*-test for the difference in signal intensity being greater than zero was conducted for each probe. This test was chosen because in our experimental design, signal intensities can only show a decrease unlike RNA analyses where values can show variation in both directions concomitant with up- or down-regulation. The *p*-values for all the probes and field isolates were then analyzed to obtained false discovery rates (*q*-values) using the method proposed by Benjamini and Hochberg (3).

The analysis of the field isolate data suggests that certain locations of the genome may experience more variation across strains than would be expected by chance. A permutation test was employed to assess spatial clustering of the variation between field strains observed in regions of *M. hyopneumoniae* genome (19). The test consisted of summing the number of field strains with significant variation from strain 232 in consecutively tested genes within a sliding window around the genome. A sliding window size of 10 consecutive tested genes was used.

The data can be access through the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE8306.

Results

TempliPhi™ reactions. Genomic DNA from all mycoplasma isolates were subjected to amplification by TempliPhi™ because of low chromosomal DNA yields in several of the field isolates. The optimal time of nebulization (15 min) and nitrogen stream pressure (10 psi) was determined empirically by taking samples at various time points during the shearing process and analyzing them by electrophoresis. All amplified DNA preparations were then sheared using that time point and nitrogen pressure, and the fragment size was confirmed by electrophoresis prior to labeling.

To determine if there were any signal biases introduced by TempliPhi™ during the amplification reaction, strain 232 chromosomal DNA was subjected to TempliPhi™ amplification, labeled and compared to nonamplified DNA from the same DNA lot in a dye swap experiment. Equal amounts of amplified and nonamplified DNA were sheared and added to labeling reactions. The nonamplified chromosomal DNA sample was mixed with an amplified DNA sample with alternate dye label and hybridized to one array; a dye swap mixture was hybridized to the array at the opposite end of the same substrate. The probe signal intensities were quantified, and these values after background subtraction were compared using correlation analysis. Since TempliPhi™ functions in a rolling circle mode and mycoplasma chromosomes are circular, no differences were expected between amplified and nonamplified samples as confirmed by a Pearson's correlation coefficient of R=0.9803 (Fig. 1).

Microarray studies. Data from each of the field isolate replicates were used in the statistical analysis. Statistical analysis indicated that 123 genes for the combined field isolates had

significant differences from the control strain 232 at $p < 0.004$ and $q < 0.20$. The results are presented in Figure 2 (exact locations are given in the supplemental table). The field strains differed in the number of genes that had significant variation from the control strain 232. The strain with the most variation was strain 95MP1509 with 40 loci differences; strain 95MP1506 had 28 loci differences. One strain, 95MP1507, showed one locus difference at mhp606 while strains 95MP1505 and 97MP0001 showed only two differences. Twenty-two loci showed differences in more than one strain (Fig. 3). Of these twenty-two loci, fifteen loci showed differences in two strains, two loci showed differences in three strains, one locus showed differences in 4 strains, and two loci each showed differences in five and six strains. Fifty-nine percent of the genes showing differences (72/123) were hypothetical with no known function. Twelve of the fifty-one lipoprotein genes showed variation (Fig. 3).

Variation analysis. The identification of variation hot spots among *M. hyopneumoniae* field strains was derived from a permutation test with a sliding window design. Figure 3 shows gene location and number of field strains that have significant variation from 232. The graph in Figure 4 plots the path of this sliding window across the genome with a window size of 10 genes. This path is determined by starting at "gene 1" and adding the number of significantly differing field strains from strain 232 for each gene through "gene 10". This resulted in a total of 3 significant variations. Shift the window one gene, and from "gene 2" to "gene 11" there were also 3 significant variations. This sliding window continues around the genome and since it is circular, the last genes are included in a window with the first genes. If the locations of significant variations were totally random, the sum in the window should vary up and down fairly regularly across the genome. To determine if there were hotspots of variation in the genome, 10,000 random permutations of the observed variation locations were done using the statistical

computing program R 2.4.1 (19). The window with the maximum sum was retained and the 95 percentile of these 10,000 permutations was determined to be 15. Thus, an observed sum of 15 or greater would only be expected to occur 5% of the time by chance. The solid horizontal line in Figure 4 denotes this significance level at the 95% confidence level where only one region of variation is declared significant. The genes in this region of variation are listed in Table 1.

Discussion

Previous studies have shown that different strains of *M. hyopneumoniae* vary in their virulence potential (37) and that genetic variation does occur in this species as measured by randomly amplified polymorphic DNA analysis (2, 29). To estimate genetic variation within *M. hyopneumoniae* in a more global, focused fashion, we performed comparative genomic hybridization on microarrays. This study utilized fourteen field strains for comparison with virulent strain 232. Many low-passage field strains are difficult to propagate *in vitro*, a characteristic that has impacted the number of isolates available for analysis. In the United States, the swine serum component of the media is highly variable in its ability to support growth of recent isolates (E. L. Thacker, personal communication), which may explain the difficulty in isolating *M. hyopneumoniae* from clinical samples. The isolation of field strains is also often impeded by the more rapid outgrowth of other mycoplasma species in clinical samples as well as the slow growth rate of *M. hyopneumoniae*. To overcome low yields of chromosomal DNA in the slow growing field strains, other methods of obtaining sufficient quantities of chromosomal DNA were sought for the analysis. Templiphi™ is an enzyme that replicates DNA in a rolling circle replication fashion and was originally developed to amplify plasmid or viral DNA sequencing templates in lieu of culturing and template purification. Our preliminary studies

indicated that Templiphi™ was also capable of amplifying all regions of the AT-rich mycoplasma genome equally well without bias (Fig. 1). Additionally, since the field strains and the control strain 232 DNAs were both amplified by Templiphi™, any specific region bias would be equally reflected in both DNAs and thus not impact the analysis.

An initial analysis of the data showed that the field strains demonstrated significant variation in their reactivities on the microarray (Figs. 2 and 3). Although most of the variation seemed to be randomly spaced around the genome, there seemed to be one hot spot of variation (Figure 2). To test this, the variation data were subjected to a permutation test with a sliding window size of 10 genes. Genes were considered as individual units of equal size to simplify the analysis. The results of that analysis are shown in Figure 4 where one region containing 23 genes was identified at $p < 0.05$. The genes within that region are listed in Table 1. When the sequences of those genes were compared by BLAST analysis with the two other *M. hyopneumoniae* published genome sequences (28), strains J and 7448, it was apparent that the region encompassing genes *mhp522-mhp538* was missing from those genomes. There were an additional five genes in that region that were not part of the array, and three of those five genes, *mhp521, mhp523*, and *mhp534*, were also missing in strains J and 7448. Two genes, *mhp536* and *mhp537*, were present in both strains. Thus, this region of the 232 genome is highly pleomorphic. Interestingly, all of the field strains in this study were from the Midwest and contained at least a portion of these sequences as evidenced by our positive results in the microarray. This highlights one limitation of the analysis. Not all of the *M. hyopneumoniae* genes are represented on the array. The missing genes are listed in the supplemental table. Those missing in the variable region were not included in Table 1 since they were not included in the analysis.

One question of interest was whether lipoproteins showed variation among the different field strains. In other mycoplasma species, lipoproteins generate antigenic diversity as a consequence of phase switching and size variation (33). In *M. hyopneumoniae*, however, similar mechanisms of variation in surface proteins do not exist (18). Our results indicate that variation does exist in lipoprotein genes in *M. hyopneumoniae* field strains since twelve of the fifty-four lipoprotein genes in the genome (18) varied among the field strains examined in this study (Fig. 3). Four of the lipoprotein genes were in the hot spot region of variation, mhp517, mhp532, mhp535 and mhp539 (Table 1). Only two of these showed variation, however; mhp535 varied in two strains and mhp532 varied in 5 strains.

Interestingly, the P97 adhesin (9) varies in one strain (95MP1506), but its companion gene, P102 (11), does not seem to vary among field isolates. Both P97 and P102, however, have multiple paralogs in the chromosome (18). The P97 paralog mhp385 varied in strain 95MP1506 and mhp493 varied in strain 95MP1509. The P102 paralog mhp384 varied in two strains, 95MP1505 and 95MP1506, and mhp683 varied in strain 95MP1509. Although a limited number of field isolates were examined in this study, our data suggests that the cilium adhesin varies little in field isolates because of its critical role in adherence and colonization. One of its paralogs, however, can vary in sequence, possibly as a way to introduce variation in the surface topography.

In unpublished studies, one locus was identified that displayed significant sequence variation in two of the field isolates used in this study, isolates 00MP1502 and 00MP1301 (Strait et al., unpublished). This chromosomal region involved mhp024 and included both a deletion and sequence variation. The region was identified using a nested PCR test that failed to identify these two strains with the inner primer pair (12). It is significant that the present studies confirm the

variation within mhp024 in one of those strains (00MP1502) and are just outside our *q*-value cutoff for 00MP1301 ($p < 0.0053$, $q < 0.234$) since the regions containing this sequence variation are represented on the array. Our analysis also showed genetic variation within mhp024 in strain 95MP1510.

These data indicate that the *M. hyopneumoniae* microarray can identify genetic variability among field isolates across the *M. hyopneumoniae* genome. A potential use of these results is to improve diagnostics by eliminating the variable genes from consideration for PCR targets. Ideally, the PCR target should be homogenous across multiple field strains. In addition, the arrays can be used to screen other mycoplasmal and bacterial species to enhance the specificity of the PCR target sequences for *M. hyopneumoniae* by eliminating those open reading frames that are cross-reactive. One limitation of this approach, however, is the inability to recognize DNA sequences present in field isolates but missing from the microarray. In summary this microarray has proven itself as a powerful tool for genomic analysis.

Acknowledgements

We would like to thank Nancy Upchurch and Barb Erickson for assistance with mycoplasma cultures. Monica Perez contributed to the construction of the microarray. Mike Carruthers assisted with spot finding. Funding for this project was provided in part by the National Pork Board and the Iowa Livestock Health Advisory Council.

References

1. **Arber, W.** 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol. Lett. **24:**1-7.

2. **Artiushin, S., and F. C. Minion.** 1996. Arbitrarily primed PCR analysis of *Mycoplasma hyopneumoniae* field isolates demonstrates genetic heterogeneity. Int. J. Syst. Bacteriol. **46:**324-328.

3. **Benjamini, Y., and Y. Hochberg.** 2000. On the adaptive control of the false discovery rate in multiple testing with independent statistics. J. Edu. Behav. Stat. **25:**60-83.

4. **Dubosson, C. R., C. Conzelmann, R. Miserez, P. Boerlin, J. Frey, W. Zimmermann, H. Hèani, and P. Kuhnert.** 2004. Development of two real-time PCR assays for the detection of *Mycoplasma hyopneumoniae* in clinical samples. Vet. Microbiol. **102:**55-65.

5. **Dudley, A. M., J. Aach, M. A. Steffen, and G. M. Church.** 2002. Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. Proc. Natl. Acad. Sci. USA **99:**7554-7559.

6. **Dudoit, S., Y. H. Yang, M. J. Callow, and T. P. Speed.** 2000. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. http://www.stat.Berkeley.edu/users/terry/zarray/Html/papersindex.html.

7. **Friis, N. F.** 1975. Some recommendations concerning primary isolation of *Mycoplasma hyopneumoniae* and *Mycoplasma flocculare*, a survey. Nord. Veterinaermed. **27:**337-339.

8. **Hacker, J., and J. B. Kaper.** 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. **54:**641-679.

9. **Hsu, T., S. Artiushin, and F. C. Minion.** 1997. Cloning and functional analysis of the P97 swine cilium adhesin gene of *Mycoplasma hyopneumoniae*. J. Bacteriol. **179:**1317-1323.

10. **Hsu, T., and F. C. Minion.** 1998. Identification of the cilium binding epitope of the *Mycoplasma hyopneumoniae* P97 adhesin. Infect. Imun. **66:**4762-4766.

11. **Hsu, T., and F. C. Minion.** 1998. Molecular analysis of the P97 cilium adhesin operon of *Mycoplasma hyopneumoniae*. Gene **214:**13-23.

12. **Kurth, K. T., H. Tsungda, E. Snook, E. L. Thacker, B. J. Thacker, and F. C. Minion.** 2002. Use of a *Mycoplasma hyopneumoniae* nested polymerase chain reaction test to determine the optimal sampling sites in swine. J. Vet. Diagn. Invest. **14:**463-469.

13. **Lysnyansky, I., Y. Ron, K. sachse, and D. Yogev.** 2001. Intrachromosomal recombination within the *vsp* locus of *Mycoplasma bovis* generates a chimeric variable surface lipoprotein antigen. Infect. Immun. **69:**3703-3712.

14. **Lysnyansky, I., Y. Ron, and D. Yogev.** 2001. Juxtaposition of an active promoter to vsp genes via site-specific DNA inversions generates antigenic variation in Mycoplasma bovis. J. Bacteriol. **183:**5698-5708.

15. **Madsen, M. L., D. Nettleton, E. L. Thacker, R. Edwards, and F. C. Minion.** 2006. Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays. Infect. Immun. **74:**160-166.

16. **Madsen, M. L., D. Nettleton, E. L. Thacker, and F. C. Minion.** 2006. Transcriptional profiling of *Mycoplasma hyopneumoniae* during iron depletion using microarrays. Microbiology **152:**937-944.

17. **Mare, C. J., and W. P. Switzer.** 1965. New Species: *Mycoplasma hyopneumoniae*, a causitive agent of virus pig pneumonia. Vet. Med. Small Anim. Clin. **60:**841-846.

18. **Minion, F. C., E. L. Lefkowitz, M. L. Madsen, B. J. Cleary, S. M. Swartzell, and G. G. Mahairas.** 2004. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. J. Bacteriol. **186:**7123-7133.

19. **R Development Core Team.** 2006. R: A language and environment for statistical computing. URL: http://www.R-project.org, Vienna, Austria.

20. **Razin, S., D. Yogev, and Y. Naot.** 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. Mol. Biol. Rev. **62:**1094-1156.

21. **Ross, R. F.** 1992. Mycoplasmal disease, p. 537-551. *In* A. D. Leman, B. E. Straw, W. L. Mengeling, S. D'Allaire, and D. J. Taylor (ed.), Diseases of Swine. Iowa State University Press, Ames.

22. **Sambrook, J., and W. R. David.** 2001. Purification of nucleic acids, p. A8.9-A8.15. *In* J. Argentine (ed.), Molecular Cloning: a laboratory manual, Third ed, vol. 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

23. **Sanderson, K. E., M. McClelland, and S. Liu, -L.** 1999. "Stable" genomes, p. 217-234. *In* R. L. Charlebois (ed.), Organization of the prokaryotic genome. American Society for Microbiology, Washington, D.C.

24. **Schmidt, H., and M. Hensel.** 2004. Pathogenicity islands in bacterial pathogenesis. Clin. Microbiol. Rev. **17:**14-56.

25. **Shen, X., J. Gumulak, H. Yu, C. T. French, N. Zou, and K. Dybvig.** 2000. Gene rearrangements in the vsa locus of Mycoplasma pulmonis. J. Bacteriol. **182:**2900-2908.

26. **Thacker, E. L., P. G. Halbur, R. F. Ross, R. Thanawongnuwech, and B. J. Thacker.** 1999. *Mycoplasma hyopneumoniae* potentiation of porcine reproductive and respiratory syndrome virus-induced pneumonia. J. Clin. Microbiol. **37:**620-627.

27. **Thacker, E. L., B. J. Thacker, and B. H. Janke.** 2001. Interaction between *Mycoplasma hyopneumoniae* and swine influenza virus. J. Clin. Microbiol. **39:**2525-2530.

28. **Vasconcelos, A. T., H. B. Ferreira, C. V. Bizarro, S. L. Bonatto, M. O. Carvalho, P. M. Pinto, D. F. Almeida, L. G. Almeida, R. Almeida, L. Alves-Filho, E. N. Assunðcäao, V. A. Azevedo, M. R. Bogo, M. M. Brigido, M. Brocchi, H. A. Burity, A. A. Camargo, S. S. Camargo, M. S. Carepo, D. M. Carraro, J. C. de Mattos Cascardo, L. A. Castro, G. Cavalcanti, G. Chemale, R. G. Collevatti, C. W. Cunha, B. Dallagiovanna, B. P. Dambrâos, O. A. Dellagostin, C. Falcäao, F. Fantinatti-Garboggini, M. S. Felipe, L. Fiorentin, G. R. Franco, N. S. Freitas, D. Frâias, T. B. Grangeiro, E. C. Grisard, C. T. Guimaräaes, M. Hungria, S. N. Jardim, M. A. Krieger, J. P. Laurino, L. F. Lima, M. I. Lopes, E. L. Loreto, H. M. Madeira, G. P. Manfio, A. Q. Maranhäao, C. T. Martinkovics, S. R. Medeiros, M. A. Moreira, M. Neiva, C. E. Ramalho-Neto, M. F. Nicolâas, S. C. Oliveira, R. F. Paixäao, F. O. Pedrosa, S. D. Pena, M. Pereira, L. Pereira-Ferrari, I. Piffer, L. S. Pinto, D. P. Potrich, A. C. Salim, F. R. Santos, R. Schmitt, M. P. Schneider, A. Schrank, I. S. Schrank, A. F. Schuck, H. N. Seuanez, D. W. Silva, R. Silva, S. C. Silva, C. M. Soares, K. R. Souza, R. C. Souza, C. C. Staats, M. B. Steffens, S. M. Teixeira, T. P. Urmenyi, M. H. Vainstein, L. W. Zuccherato, A. J. Simpson, and A. Zaha.** 2005. Swine and poultry pathogens: the complete genome sequence of two strains of

Mycoplasma hyopneumoniae and a strain of *Mycoplasma synoviae*. J. Bacteriol. **187:**5568-5577.

29. **Vicca, J., T. Stakenborg, D. Maes, P. Butaye, J. Peeters, A. de Kruif, and F. Haesebrouck.** 2003. Evaluation of virulence of *Mycoplasma hyopneumoniae* field isolates. Vet. Microbiol. **97:**177-190.

30. **Wilton, J. L., A. L. Scarman, M. J. Walker, and S. P. Djordjevic.** 1998. Reiterated repeat region variability in the ciliary adhesin gene of Mycoplasma hyopneumoniae. Microbiol. **144:**1931-1943.

31. **Yang, Y. H., S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed.** 2002. Normalization for cDNA microarray data: a robust composite method for adressing single and multiple slide systematic variation. Nucleic Acids Res. **30:**e15.

32. **Yazawa, S., M. Okada, M. Ono, S. Fujii, Y. Okuda, I. Shibata, and H. Kida.** 2004. Experimental dual infection of pigs with an H1N1 swine influenza virus (A/Sw/Hok/2/81) and Mycoplasma hyopneumoniae. Vet. Microbiol. **98:**221-218.

33. **Yogev, D., G. F. Browning, and K. S. Wise.** 2002. Genetic mechanisms of surface variation, p. 417-443. *In* S. Razin and R. Herrmann (ed.), Molecular Biology and Pathogenicity of Mycoplasmas. Kluwer Academic/Plenum Publishers, New York.

34. **Yogev, D., R. Rosengarten, R. Watson-McKown, and K. S. Wise.** 1991. Molecular basis of mycoplasma surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5' regulatory sequences. EMBO J. **10:**4069-4079.

35. **Young, T. F., and R. F. Ross.** 1987. Assessment of antibody response of swine infected with *Mycoplasma hyopneumoniae* by immunoblotting. Am. J. Vet. Res. **48:**651-6.

36. **Zhang, Q. J., and K. S. Wise.** 1996. Molecular basis of size and antigenic variation of a *Mycoplasma hominis* adhesin encoded by divergent *vaa* genes. Infect. Immun. **64:**2737-2744.

37. **Zielinski, G. C., and R. F. Ross.** 1990. Effect of growth in cell cultures and strain on virulence of *Mycoplasma hyopneumoniae* for swine. Am. J. Vet. Res. **51:**344-348.

Figure 1. Comparison of mean log signal intensity values of genomic vs TempliPhi™ amplified chromosomal DNA. Data represent Log mean intensity values following background subtraction. Correlation analysis was performed and Pearson's coefficient was r = 0.9803.

Figure 2. Scatter plot of genetic variation of field isolates. Positions of genetic variation are shown for each *M. hyopneumoniae* field strain.

Figure 3. Locations of loci variation in *M. hyopneumoniae* field strains. The loci that showed variation from strain 232 are shown. Outer circle: the location of variable loci are indicated; small solid circles, variation by a single strain; large open circles, variation among two strains; large open triangles, variation among three strains; solid squares, variation among four or more strains. Inner circle: the locations of putative lipoproteins are shown. Solid circles, loci that show variation; open squares, lipoproteins that did not show variation. The hot spot of genetic variation is shown as a grey bar along the circle.

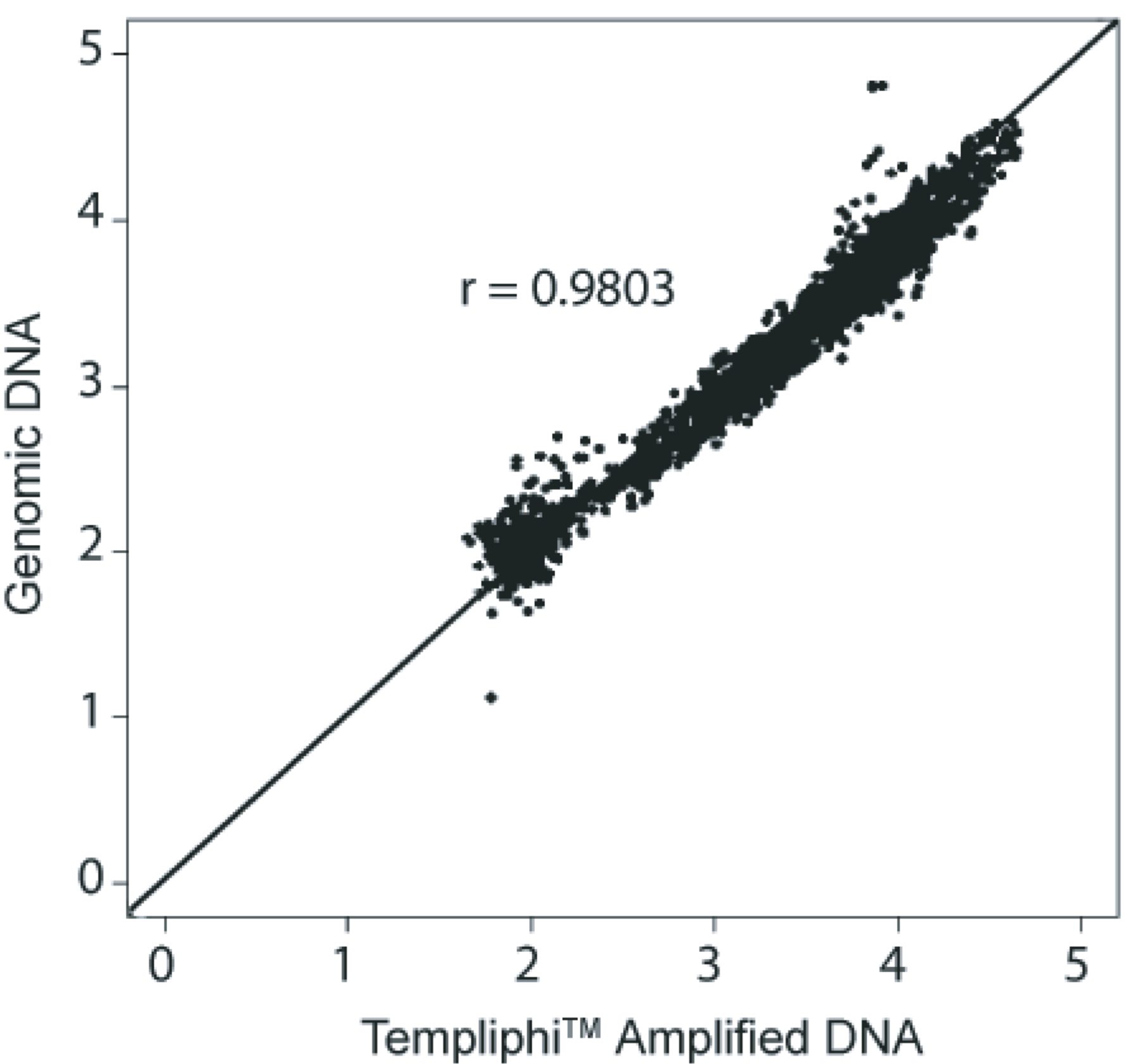
Figure 4. Permutation test of hot spots showing variation around the *M. hyopneumoniae* genome. This graph represents variation around the genome within individual genes in a sliding window of 10 genes. The permutation test indicates significance at $p = 0.05$ (solid horizontal line).

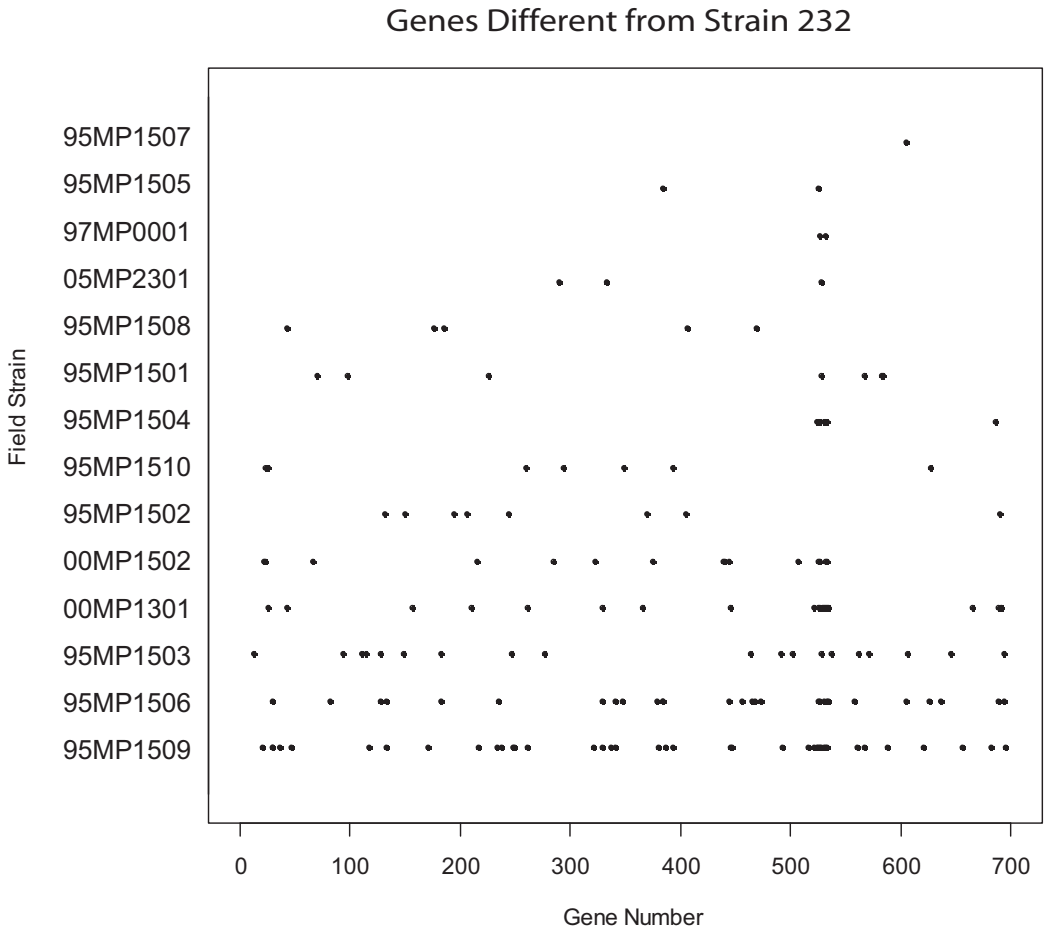
Table 1. Genes identified by hot spot analysis.

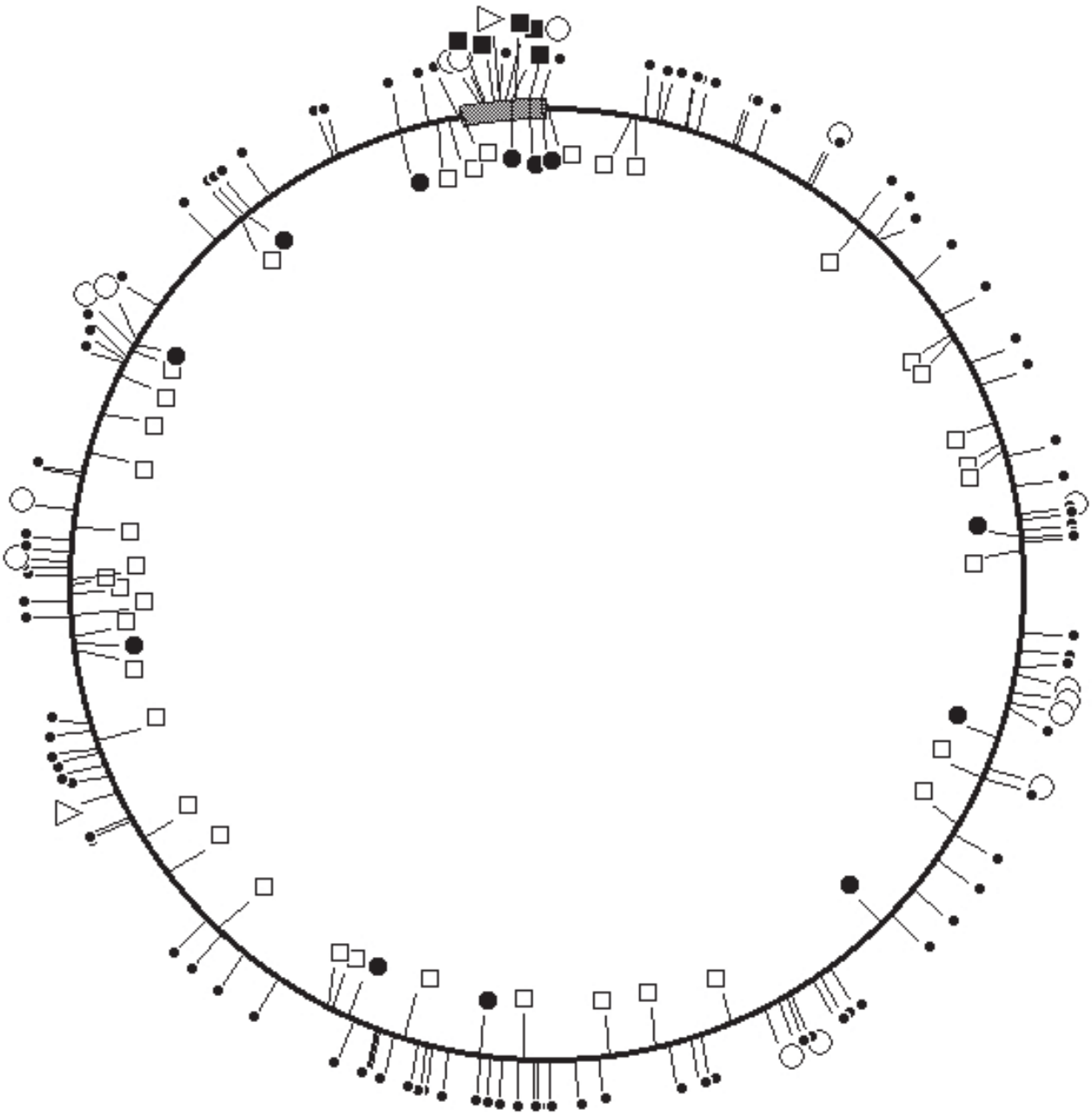
		Region of Variation		
Gene	Description[1]	Instance	J[2]	7448[2]
mhp516	CH	1	+	+
mhp517	UH, Lipoprotein	0	+	+
mhp518	UH	0	+	+
mhp519	UH	0	96 (96)	+
mhp520	*pepF*	0	+	+
mhp522	UH	2	-	-
mhp524	UH	2	-	-
mhp525	UH	0	-	-
mhp526	CH	6	-	76 (62)
mhp527	CH	6	-	-
mhp528	UH	1	-	-
mhp529	UH	3	-	-
mhp530	CH	1	-	-
mhp531	CH	4	-	-
mhp532	*trsE*, Lipoprotein	5	79 (12)	78 (12)
mhp533	CH	5	-	-
mhp535	UH, Lipoprotein	2	-	92 (94)
mhp538	CH	1	97 (10)	-
mhp539	CH, Lipoprotein	0	+	+
mhp540	*tuf*	0	+	+
mhp541	*lon*	0	+	+
mhp542	CH	0	+	+
mhp543	*upp*	0	+	+

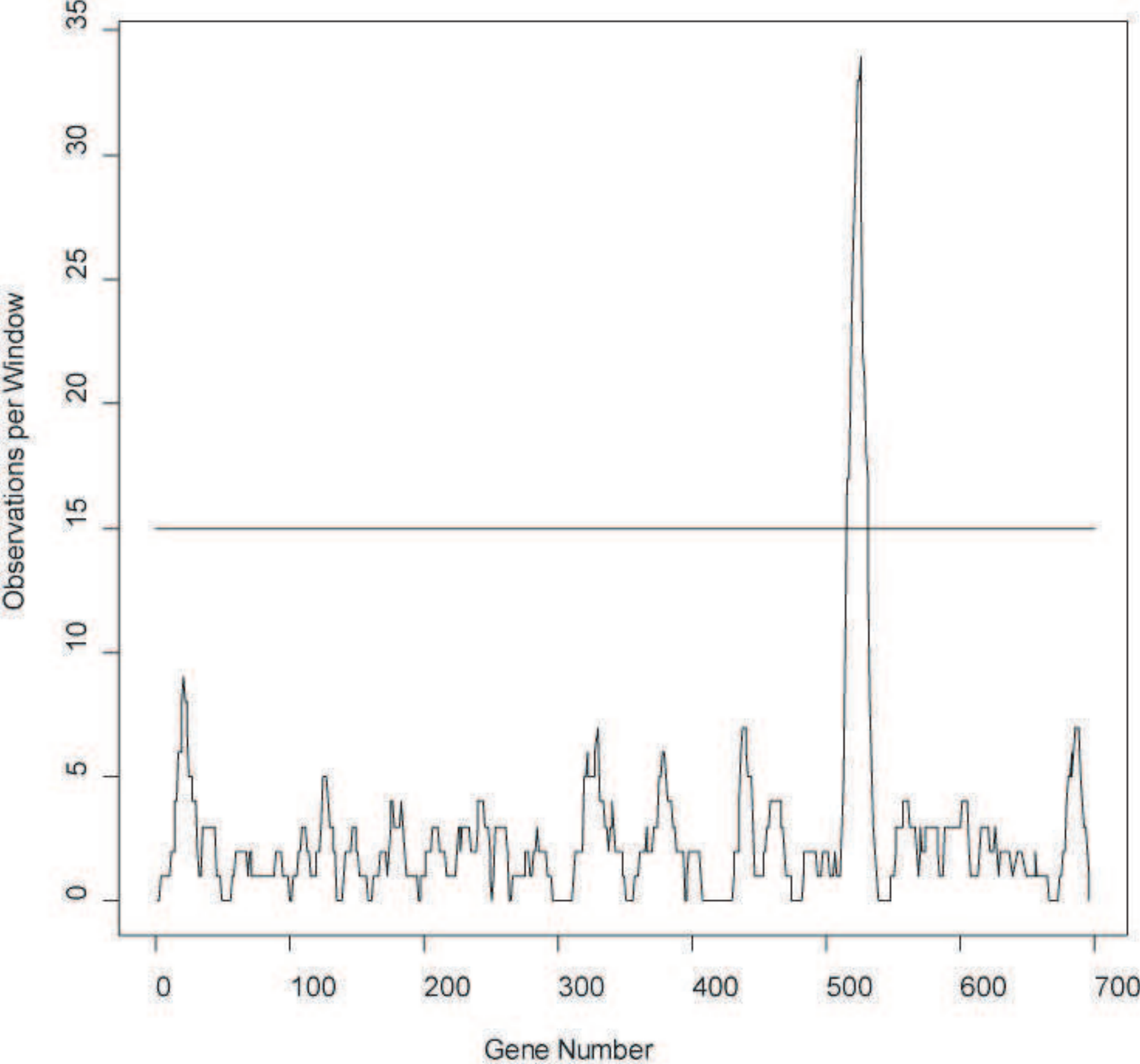
[1] UH= unique hypothetical; CH=conserved hypothetical

[2] BLAST searches were performed with strain 232 sequence against published genome sequences of strains J and 7448 (28). Data represent percent identity with strain 232, parentheses indicate percent of gene sequence present; (-), sequence is missing in published sequence; (+), 98-100% identity with published sequence.


r = 0.9803
Genomic DNA
Templiphi™ Amplified DNA
0
1
2
3
4
5


Genes Different from Strain 232
95MP1507
95MP1505
97MP0001
05MP2301
95MP1508
95MP1501
95MP1504
95MP1510
95MP1502
00MP1502
00MP1301
95MP1503
95MP1506
95MP1509
Field Strain
0
100
200
300
400
500
600
700
Gene Number




Observations per Window
0
5
10
15
20
25
30
35
0
100
200
300
400
500
600
700
Gene Number