5-2009

Design and Evaluation of Oligonucleotide Microarrays for the Detection of Bovine Pathogens

Ryan Weldon Black
Utah State University

DESIGN AND EVALUATION OF OLIGONUCLEOTIDE MICROARRAYS FOR THE DETECTION OF BOVINE PATHOGENS

by

Ryan W. Black

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Animal Science

Approved:

Dr. Lee Rickords Major Professor	Dr. Thomas J Baldwin Committee Member
Dr. Ramona T. Skirpstunas Committee Member	Dr. Brian B. Gowen Committee Member
Dr. Kenneth C. White ADVS Department Head	Dr. Byron R. Burnham Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2009

ABSTRACT

Design and Evaluation of Oligonucleotide Microarrays for

the Detection of Bovine Pathogens

by

Ryan W. Black, Master of Science

Utah State University, 2009

Major Professor: Dr. Lee Rickords
Department: Animal, Dairy and Veterinary Sciences

Two microarray designs were developed and produced to screen for multiple bovine pathogens commonly found in the cattle industry today. The first microarray was designed, built, and processed in-house using conventional material and equipment and targeted *Pasteurella multocida*, *Manheimia haemolytica, Histophilus somni,* and *Arcanobacterium pyogenes*. For each pathogen, 12 perfect-match oligonucleotide probes, which were also designed in-house, targeted different sections of the respective 16S ribosomal genes, and were coupled with 12 corresponding mismatched probes for background. These arrays were able to produce distinct hybridization patterns for each pathogen that were easily visible without the need for computer analysis. However, the need for PCR amplification of the 16S gene prior to hybridization motivated us to explore more efficient array options. The second designed microarray, a custom Affymetrix GeneChip, targeted *Escherichia coli, Salmonella typhimurium,* and *Salmonella dublin* in addition to the previously mentioned pathogens and was more successful in overall

performance than the “in-house” arrays. In addition to the 16S gene, oligonucleotide probes targeted other genes (from 2 to >4500, depending on whether the genome was sequenced) that were unique to each pathogen. This array also differed from the “in-house” arrays in that mismatched probes were not designed. The different probe sets performed at different detection limits as *P. multocida, A. pyogenes, S. typhimurium,* and *S. dublin* were detected with as little as 250ng of hybridized genomic DNA (gDNA), while *M. haemolytica, H. somni,* and *E. coli* required as much as 1µg gDNA. These pathogens were also spiked into bovine tissue to simulate multiorgan infections in which they were individually detected with the microarray design.

(76 pages)

ACKNOWLEDGMENTS

I would first like to thank my wife who was my greatest support group throughout this crazy time, and who by marrying me indirectly encouraged me to go back to school, which started all this craziness in the first place. You have my love and appreciation.

I would like to thank my major professor and head of my committee, Dr. Lee Rickords, for taking me in as one of his students, and for his guidance and patience. I need to thank my other committee members, Drs. Tom Baldwin and Ramona Skirpstunas of the Utah Veterinary Diagnostic Laboratory, who provided the project, much of the funding, and the facilities, and Dr. Brian Gowen for his needed support and assistance. Dr. Bart Weimer of the Center for Integrated Biosystems at USU was of especial service in providing the remainder of the funding, necessary expertise, and the means to completing this project.

Additional thanks need to go out to the staff of the Utah Veterinary Diagnostics Laboratory for allowing me to invade their space, sacrificing their time and material for my benefit. I also give thanks to Ninglin Yin for providing me with a flexible job and helping me to take care of my family while completing this project.

Finally, a shout goes out to my labmates, friends, and family for their support and encouragement. Thanks for all your help and good luck to you in your lives.

Ryan Black

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

LITERATURE REVIEW

Introduction

A variety of diagnostic techniques are available to both ranchers and veterinarians for the identification of bovine pathogens. Observing clinical symptoms of a disease such as fever, diarrhea, congestion, and upper respiratory symptoms is often sufficient in making tentative diagnoses. Abnormal behavior, such as depression, lethargy, instability or incoordination can also be the manifestations of various infections. Gross lesions such as hemorrhage, infiltrates, abscesses, scar tissue, and excessive fluid in body cavities, may also be discernible internally in affected animals.[1,2,12,44,71]

Diagnosis of disease using clinical observations requires experience and historical knowledge of herds and individuals within a herd, since symptoms of various diseases can be similar. However, definitive diagnostics are needed to confirm a clinical diagnosis. Veterinary diagnostic laboratories focus on identifying pathogens that cause disease using techniques that include bacterial culture, biochemical testing, serological testing, and molecular diagnostic techniques. Microbiological testing allows a veterinarian to identify specific bacteria based on colony characteristics, biochemical profiles and staining characteristics. Serologic evaluation, Enzyme-Linked ImmunoSorbent Assay (ELISA), serum neutralization (SN), and other techniques are also used to identify specific antibodies or antigens found within the bloodstream or tissues to confirm the presence of or recent exposure to particular pathogens.[1] Additional target factors include specific enzymes, microorganisms or certain components thereof, or

particles that agglutinate when mixed with specific antiserums.[48,65] Molecular techniques, such as polymerase chain reaction (PCR), can identify pathogens by amplification of genetic material of the agent, often targeting specific genes that are unique to an organism. Researchers continue to make improvements on these techniques with the goal of making the tests more accurate, more efficient, less complicated, and most importantly (at least to ranchers), less expensive. A relatively new technology, the microarray, used primarily for gene expression analysis of biological systems, is recently and increasingly finding use as a detection tool for the identification of bacterial and viral genomes within infected tissue and ecological environmental samples.[40] In this report, microarray technology is used to specifically identify pathogens associated with a number of mixed infections that are common to the cattle industry.

Disadvantages of Current Techniques Used to Diagnose Bovine Diseases

The primary methods that veterinary diagnostic laboratories use to diagnose a disease, other than clinical observations, include microbiology, biochemical tests, serological tests, histopathology, and molecular diagnostic techniques. For many diseases, isolation of the infection-causing pathogens is the most direct method of confirming a diagnosis. Materials used for bacterial culture are cost-effective. However, isolating a pathogen from a mixed culture can be difficult and very time consuming, taking anywhere from days to months to cultivate sufficient colonies[27] for definitive diagnosis. Moreover, culture conditions are not known for every pathogen. Finally, culturing zoonotic pathogens poses potential health risks to personnel and requires

Biological Safety Level (BSL) facilities, proper equipment, training and protocols to ensure the safety of those handling the pathogens.[15]

Biochemical and serological tests are designed to be more species and pathogen specific and techniques will vary depending on the type of pathogens involved. The application of specific techniques requires some level of foreknowledge as to the identity of the suspected pathogen, since each test is species and pathogen specific. Individually, these tests may be inexpensive, however, due to the method of testing (serum neutralization assays),[44] multiple tests could be required for confirmation of a disease which would result in higher labor and material costs.

Molecular techniques, like PCR, have made great strides in providing a cost efficient, fast, and safe means for diagnosing pathogens present in samples. However, culture is still sometimes necessary, and many pathogens require different experimental conditions and optimized protocols and reagents, which again requires some suspicion as to the identity of the pathogen.

Microarrays, another molecular technique, provide a means of potentially creating an all-in-one diagnostic assay which provides the ability to screen for multiple pathogens which is often seen in cases of co-infection or identify individual pathogens that present with similar clinical symptoms. Also, all microarray experiments can be run using the same repeatable techniques and experimental conditions. A limitation to many microarray detection assays is the requirement of PCR amplification of the target sequences prior to hybridization.[73] Without this amplification, detection is often limited. This limitation creates a challenge when designing a microarray assay that is not species-

specific. This limitation is usually overcome through the use of non-specific universal primers.

Unfortunately, microarray technology is currently associated with high cost. However, since multiple conventional tests are sometimes required to correctly identify all pathogens in a disease syndrome,[1] the cost of multiple tests could easily be more expensive than the cost of a single microarray test. Ultimately, a single microarray could provide diagnostic answers in a time-efficient and cost-effective manner compared to potential multiple diagnostic tests that could otherwise be required.

Microarrays

Microarray assays are relatively new technology that is finding rapid approval and widespread application. Microarrays are flat solid surfaces, usually glass, that contain numerous sequences of nucleotides (probes) that are physically and systematically bound perpendicularly to the surface in known locations. These probes, varying in length from as small as 18 bases to hundreds of bases long, represent known regions of a genome, and usually correspond to specific genes. Samples of either DNA or ribonucleic acids (RNA) are isolated from biological samples and labeled enzymatically with a fluorescent molecular tag and, in this report, termed as "target." The labeled target is hybridized to the array by binding the target DNA/RNA to a complimentary probe sequence via base-pair binding. Fluorescent molecular tags can then be quantitatively measured at each probe location to determine the relative amount of each target sequence present in an extracted DNA/RNA sample. The power of microarray technology is the ability to analyze tens of thousands of genes (or gene sequences) simultaneously in a single

experiment.[23,47] The primary function of microarray technology is for gene expression profiling, but other applications include identification of potential drug targets, detection of mutations or single nucleotide polymorphisms, detection of short tandom repeats, detection of sequence insertions and deletions, comparative genomic hybridization, and the identification of genomes or parts of genomes (ie. bacterial, viral).[32]

History of Microarray Technology

In the late 1970s, the first parallel hybridization analysis experiments, known as the dot blots, were introduced[43] to the scientific world by taking advantage of an earlier idea to use multiple DNA libraries arrayed on filters to cross-correlate cloned sequences.[68] Creating these dot blots was a manual procedure that was improved in the early 1990s by using robotics to spot probes on filter surfaces which allowed greater spotting density, accuracy, and speed, while at the same time reducing human error.[36] Microarrays became more defined as further improvements were introduced in robotic spotting technique, and in spotting material, such as glass, polypropylene, nylon, and silicon, that enabled an increase in production, spotting density, and in some cases, cost-effectiveness.[68] Microarrays were generally produced in-house for much of the 1990s, but some laboratories were able to commercialize their specialized version of the microarray. In 1996, Affymetrix (Santa Clara, California, USA) began to mass-produce and commercialize their microarray platform, called the GeneChip. Instead of spotting nucleotides, however, they introduced a new technique by synthesizing short oligonucleotides directly onto a glass surface using proprietary photolithographic techniques.[50] Other major companies followed with their own designs. For example,

Agilent Technologies (Santa Clara, California, USA) builds longer oligonucleotides on a glass slide using inkjet printing technology and phosphoramidite chemistry.[38] Over the years, there has been a decrease in the production of in-house spotted arrays as scientists refer to more reliable and consistent commercial arrays, though research labs still commonly produce in-house arrays in an attempt to keep expenses down.

Different Microarray Platforms

As microarray technology has developed, different strategies for microarray design have emerged from both academic and commercial groups. Many of the "homemade" in-house spotted arrays initially employed long cloned cDNA samples as probes, each of which most often represented an entire gene.[66] However, the use of cDNA clones as probes have diminished due to concerns about annotation, clone identity, and probe performance. The use of oligonucleotides as probes has found greater popularity, especially among commercial platforms, as these probes have shown better hybridization characteristics and companies have been able to provide reliable libraries containing the annotation and identity of the oligonucleotides.[77]

Currently, most array platforms utilize either a one-color or two-color dye system. The spotted cDNA arrays regularly use the two-color dye system, in which two samples, each labeled with different fluorescent tags such as cyanine-3 (Cy3) and cyanine-5 (Cy5), are hybridized to a single array.[66] Hybridizing two samples, a control and a treatment, to one array provides an advantage by removing array-to-array variability that would otherwise need to be dealt with when comparing two samples on two arrays. A third fluorescent dye (fluorescein) can also be used to help monitor quality control measures

with prehybridization variables.[33,34]

Affymetrix GeneChips are a one-color platform in which a single sample is hybridized to a single array. These GeneChips are built with a series of short 25-mer oligonucleotide probes that represent the genes within a genome. These oligonucleotides have a greater performance over the much longer cDNA probes, but multiple probes grouped in a probe set (which usually represents one gene) are required to overcome the lack of specificity of a 25-mer sequence and increase the overall confidence level. These probe sets usually contain between 11-20 probes per probe set. Thus, the GeneChips exhibit a much higher density, up to 1,000,000 probes/cm^2,[6] even though the human genome only contains 30,000 genes. Array-to-array variability is reduced through carefully controlled array production and hybridization methods.[6]

Agilent Technologies has more recently produced its own line of oligonucleotide microarrays that use the two-color platform. The Agilent arrays carry longer 60-mer oligonucleotide probes, which they claim allow for higher specificity compared to the smaller (Affymetrix GeneChip) probes.[38] Also, their arrays, which can contain as many as 44,000 features, do not require as high a density coverage as the GeneChips. Longer probes inkjet technology allow for a more cost-effective production of commercial microarrays.

In one study comparing the performance of three commercial microarrays, Affymetrix Genechips (short oligonucleotides, one-color platform), GE Healthcare CodeLink Microarrays (Tempe, Arizona, USA) (long oligonucleotides, one-color platform), and Agilent microarrays (long oligonucleotides, two-color platform)), it was reported that there was a higher level of reproducibility from one-color based arrays

(Affymetrix GeneChips and GE Healthcare CodeLink Microarrays) compared to the two-colored Agilent platform. Furthermore, Affymetrix had higher overall concordance with same-sample qPCR results.[20] Despite this single study, there seems to be no clear consensus as to which platform performs the best or whether the different platforms are even comparable. Various studies claim major differences in the data produced,[46,64,69] while others insist that there is ample levels of concordance among the different platforms.[39,79]

Microarray Applications in Human and Animal Diagnostics

Microarray technology has been applied in both human and animal health studies. For example, microarrays have been used to profile the expression of genes at different stages of cancer,[22,31,41,67] profile gene expression of neurological diseases such as Alzheimer's disease,[4,18,21,35] and identify infectious diseases.[13,16,59] Additional uses of microarray technology in human studies include drug discovery, pharmacogenomics, and toxicogenomics studies,[10,19,76] as well as identifying hypothetical genes that might trigger health-related issues, like obesity, alcoholism, and drug addiction.[24]

Microarrays have been used more in human related clinical applications than animal studies. Nevertheless, microarrays have played a significant role in profiling the gene expression patterns of animal diseases, usually only as starting points for data mining, to be followed later by more specific tests. Beyond gene expression experiments, the most common use of microarrays in animal health has been in the detection and genotyping of animal pathogens.[8,9,11,60,72] A less widely used application is direct-mutation screening for single-gene diseases in dogs, horses, and some livestock.[55] In all

these applications, continued development of genomic maps of individual species will provide the resources necessary to further animal health studies in the future[60] and may even allow microarray technology to assist in breeding decisions.[25]

Hybridization and Washing

Hybridization is the formation of heteroduplex molecules when a target and probe bind together. Under perfect conditions, the target will only hybridize to a probe that is perfectly complimentary. Therefore, it is possible to determine the sequence of a target, since the probe has known sequence information and a known location, both on the array and within a genome. However, microarray experiments are anything but perfect. There are various conditions that will affect hybridization efficiency. These conditions include hybridization temperatures, the concentration of monovalent cations such as sodium, the respective concentrations of the target and probe, GC content, and buffer content.

Optimal hybridization temperature is generally lower than the calculated melting temperature of the probe sequence. However, while the melting temperature can give an estimate, optimal hybridization temperature must be experimentally determined. A lower temperature will result in an increase of non-specific binding, while a higher temperature will reduce true signal.

All nucleotide molecules are composed of a negatively charged phosphate backbone, which can impede heteroduplex formation. To negate this hindrance, monovalent cations, such as sodium, can be added to bind to the negatively charged phosphates. A salt concentration of one molar is generally sufficient for most microarray

experiments.[45] Adjustment of the salt concentration can help to optimize hybridization, but such changes are usually small and insignificant.

The concentrations of both the targets and the probes in a hybridization reaction are important to consider. Hybridization reactions follow either pseudo-first order or second order reaction kinetics.[29] When the target concentration is equal to or less than the probe concentration, second-order kinetics is achieved.[29] Under these conditions, small differences in the concentration of either the target or probe can have a large impact on the reaction rate of hybridization. When the concentration of the target is sufficiently higher than that of the cognate probe, pseudo-first order reaction kinetics is achieved and the target signal will be directly proportional to that of the target concentration.[29] Therefore, a 2-fold increase of target will produce a 2-fold increase in signal. It can be difficult to build microarrays with precise concentrations of probe, as well as adding the exact amount of target to a hybridization reaction due to imprecise measurements and manufacturing techniques. At least a 10-fold greater target concentration than probe is necessary to achieve such a reaction rate that variations due to imprecise microarray manufacturing methods can be minimized.[29]

A post-hybridization wash is necessary to remove unbound target strands from the array. The wash consists of the array being soaked in two stringent salt solutions, the first more stringent than the second, with some type of agitation to help "clean" the array. Optimal stringency of the wash solutions is necessary to not only efficiently remove unbound target strands, but to also unbind random non-complimentary target/probe heteroduplexes. Stringency is adjusted through salt concentrations and wash temperatures. If the stringency of the wash conditions is too low, nonspecific binding

will increase creating false signals. Relatively high-stringency conditions will give the best reproducibility.[45]

Labeling Techniques

Attachment of a fluorescent molecule to a target sample, often termed "labeling," requires the use of various techniques and materials. Company provided protocols are followed when using commercial arrays, but with in-house arrays, a variety of options are available and the best choice to use tends to be determined by the type of experiments being performed. Historically, all target samples were initially labeled in a direct manner, in which nucleotides attached to a fluorescent dye are incorporated into cDNA through reverse transcription. The fluorescent dyes, like the commonly used Cy3 and Cy5 dyes, are attached to one of the four nucleotides, usually dTTPs or dCTPs, and become incorporated into the cDNA at the complimentary dATP (or dGTP) position of the mRNA strands. However, there are some problems associated with direct labeling. First, the number of fluorescent molecules that are incorporated into a cDNA strand are directly proportional to the number of complimentary nucleotides in the mRNA strand. In the case of Cy3-conjugated dTTPs, the number of Cy3 molecules incorporated into a cDNA strand will vary based on the number of dATPs in the mRNA strand, and longer transcripts will invariably have a stronger signal potential than shorter transcripts. Second, the fluorescent molecules are bulky and tend to create interference during reverse transcription.[58] The result is a variation in labeling efficiency among different transcripts. Third, mostly in response to the second issue, direct-labeling protocols can call for up to 50-100μg of starting total RNA. These large amounts of total RNA can be difficult to produce and are not typically ideal for most clinical applications. One method

used to circumnavigate this problem has been to amplify RNA prior to labeling,[53,54,62,74] but the extent of accuracy and precision of that approach is still under question.

Recently, other labeling techniques have become popular including an indirect method of labeling cDNA. In this method, amino-modified (amino-allyl) nucleotides, instead of fluorescent-conjugated nucleotides, are incorporated into the cDNA through reverse transcription. Once more, only one of the four nucleotides is modified. These amino-allyl nucleotides are considerably less bulky and thus are incorporated with greater efficiency. The reactive fluorescent Cy3 and Cy5 dyes are then conjugated to the amino-allyl nucleotides though a chemical coupling step. Labeling efficiency is much higher in an indirect labeling approach; however, this does not resolve the issues of base composition or length of transcripts. Direct labeling still offers one advantage over indirect labeling in that the direct approach is almost always less expensive than the indirect approach.

Genomic DNA (gDNA) can also be labeled with a fluorophore using direct or indirect techniques. The fluorophores are incorporated through PCR instead of reverse transcription. An alternative approach would be to add the fluorophore enzymatically at the end of a DNA strand using terminal transferase. Terminal labeling is useful when labeling small strands of DNA where PCR would be difficult. While this approach would attach only one fluorescent molecule to each strand, the problem of fluorescence intensity due to base composition and transcript length is eliminated.

Cy3 and Cy5 have been commonly used in microarray experiments as the fluorophores of choice, however, there are others on the market that claim a superior performance. The Alexa Fluor family from Invitrogen (Carlsbad, California, USA) is

one example. The claim is that the conjugates of Alexa 546 fluoresce are at least twice as bright as that of the Cy3 conjugates.[63] Affymetrix uses a fluorophore from the streptavidin family, which has a very strong affinity for biotin. Using an entirely different approach, Affymetrix protocols create cDNA from mRNA through reverse transcription, as is standard, but then continue to produce large amounts of cRNA through *in vitro* transcription. During this amplification process, mRNA is amplified 100-fold[7] allowing for large amounts of cRNA to be hybridized to a GeneChip. In addition to amplification, the transcription step also incorporates biotinylated nucleotides. The biotinylated cRNA is first hybridized to the arrays and then coupled with streptavidin phycoerythrin to provide the fluorescent signal. To further increase the signal strength, biotinylated antibodies are attached to the streptavidin phycoerythrin and more streptavidin phycoerythrin is added to the mixture to bind to the biotin on the antibodies.

Bacterial Pathogens

Bovine diseases are usually identified through a combination of clinical observations made by veterinarians and ranchers and diagnostic tests performed in a laboratory. A description of clinical signs and symptoms of the diseases studied in this project, the causative pathogens, and current techniques employed to detect them are provided in order to help convey the impact that the pathogens have on livestock and why further advancement in detection techniques for these pathogens is important. Though the following are descriptions of individual pathogens and the related diseases, it is important to note that, in terms of respiratory diseases, it is common for multiple pathogens to be involved.

The following pathogens were selected for this project because they are common causes of infection afflicting livestock each year in the U.S. These diseases are generally not found in U.S. livestock in epidemic form, but rather as common diseases encountered of importance to U.S. livestock. U.S. ranchers regularly report individual or group cases, and local veterinarians, veterinary diagnostic laboratories and the ranchers themselves usually deal with these cases as they arise. Often, the course of action is to isolate infected cattle to limit the spread of disease and treat animals accordingly.

Pasteurella multocida

Pasteurella multocida subspecies *multocida* is the primary pathogen responsible for hemorrhagic septicemia in cattle, a disease that targets the respiratory system and causes excessive salivation and nasal discharge with swelling in the pharyngeal and ventral cervical regions. Cattle with hemorrhagic septicemia usually die within 24 hours of the first recognizable signs due to respiratory distress. Post-mortem observations include visible congestion of the mucous membranes, widely distributed lesions and hemorrhage and edema of the tissues of the head and neck.[12]

Laboratory diagnosis includes bacterial isolation from heparinized blood or select tissues, biochemical testing, special staining and observing colony characteristics. Moreover, there are immunological tests that include a rapid slide agglutination test for capsular typing, an indirect hemagglutination test for somatic typing, agar gel immunodiffusion tests, and counter immunoelectrophoresis that can identify the pathogenic *P. multocida* serotypes.[1] PCR techniques have also been developed to detect *P. multocida* in tissue.[14,49,56,70]

Histophilus somni (Haemophilus somnus)

Historically thought of primarily as a nervous system pathogen that causes thromboembolic meningoencephalitis (TEME), *Histophilus somni* has the ability to attack numerous different cell types within a cow's body. Thus the disease is now referred to as the "*Haemophilus somnus* complex." The respiratory form of the disease causes severe pneumonia and death. The reproductive and urinary tract form leads to long-term urinary tract infections and can cause abortions. The septicemic form is manifested by many different clinical signs. There are also miscellaneous forms of the disease that can infect such systems as the eyes and ears. Ultimately, if left untreated, one form of the disease can lead to other forms of the disease.[44]

Current techniques for diagnosis greatly rely on clinical signs and bacterial culture. Ancillary serological tests are available including a microagglutination test, tube agglutination tests, a complement fixation test, an ELISA, and a technique that identifies a *H. somni* antigen by using monoclonal antibodies.[44] A PCR test for the detection of *H. somni* in pure and mixed cultures has also been suggested, but is not cost-effective when compared to microbiological diagnostic techniques.[3]

Mannheimia (Pasteurella) haemolytica

Mannheimia haemolytica, along with *P. multocida,* is one of the primary pathogens responsible for pneumonic pasteurellosis, commonly known as "shipping fever." *M. haemolytica* normally resides in the upper respiratory tract of a cow and tends to become pathogenic during suppression of the immune system, often a result of stress (e.g. during shipping) and concurrent viral infections. The disease is typically observed in feeder calves and can range from acute to chronic conditions with morbidity levels

reaching 35% and mortality between 5-10%. Diagnosis primarily relies on bacterial isolation in conjunction with clinical assessment[2] and histopathologic examination. A PCR test that isolates the *lkt* gene in *M. haemolytica* has also been demonstrated, but it lacks species specificity.[26,30]

Arcanobacterium (Actinomyces) pyogenes

Arcanobacterium pyogenes is an opportunistic pathogen responsible for a wide variety of diseases, from arthritis, to mastitis, to pneumonia. Abscesses are often associated with *A. pyogenes* infections, and can target almost every organ and tissue in a cow. Diagnosis is primarily confirmed by isolation of the bacterium from the infected tissue.[17]

Escherichia coli

Escherichia coli is the most predominant species of normal flora within the bovine intestine. Most strains are non-pathogenic, however, pathogenic strains cause severe disease in young animals. Young calves (1-10 weeks) are the most susceptible to pathogenic *E. coli* infections, which commonly include enteritis (white scours, enteric colibacillosis) and septicemia (colisepticaemia). Enteric colibacillosis is primarily caused by Enterotoxigenic *E. coli* (ETEC) infections. Calves are also subject to Enterohaemorrhagic *E. coli* (EHEC) infections, which include the highly pathogenic *E. coli* O157:H7 strain, and Enteropathogenic *E. coli* (EPEC) infections.

Historically, diagnosis of *E. coli* infections required positive culture of the pathogenic species in conjunction with gross clinical evidence such as systemic lesions. While culture is still useful in identifying *E. coli* strains, current techniques like PCR are

widely used for diagnostic detection of pathogenic species[61] using virulence gene identification.

Salmonella typhimurium* and *Salmonella dublin

Bovine salmonellosis is caused by a number of different serovors of *Salmonella* species. However, the most incriminating *Salmonella* by far found in cases of bovine salmonellosis are *typhimurium* and *dublin*. Salmonellosis usually manifests as septicemia or acute or chronic enterocolitis. Salmonellosis is highly contagious as the *Salmonella* species are resilient in the environment. The enteric syndrome of salmonellosis is more common in cattle overall, and the septicemic syndrome is more common in calves. *S. dublin* is observed more often in the septicemic syndrome of the illness. Septicemic salmonellosis has a much higher death rate than the enteric syndrome, however, enteric salmonsellosis still has a fatality rate of about 50% or greater.

Diagnosis is based on a combination of antemortem and postmortem diagnostics including culture of the organism. *Salmonella* sp. can be isolated from tissue or fecal material and from blood in septicemic cases. Serological tests (ELISA, serum agglutination, complement fixation) are also available for antemortem *Salmonella* sp. detection, as well as antibody detection in septicemic animals.[71]

Objectives of This Report

The objective of this project was to correctly identify the above bovine pathogens individually from tissue spiked with a mixture of bacteria using microarray technology. Therefore, two arrays were designed using completely different approaches. A sub-objective to this project was to successfully design the arrays for detection of pathogenic

DNA without a need for PCR amplification. Our first arrays were homemade and designed in-house using equipment and materials that were available in the Center for Integrated BioSystems (CIB) at USU. These arrays had probes that targeted the 16S ribosomal RNA (rRNA) gene for each pathogen. Commercial GeneChips from Affymetrix were utilized for our second array design. Numerous genes, including the 16S rRNA gene, were represented on these arrays for each pathogen. While the genes represented on this array were selected in-house, Affymetrix designed the probes and built the actual arrays. This report covers the research performed using both array platforms.

CHAPTER II

IN-HOUSE ARRAYS

Introduction

Our aim was to design and build an in-house microarray for the detection of common bovine pathogens in a cost-effective manner using equipment and materials that were available at the CIB. In a study that compared a number of different slide types, the nylon-mounted membrane slides seemed to provide the lowest detection limit capabilities while providing the highest signal-to-noise ratios.[28] An additional advantage of nylon-mounted slides is they are generally less expensive among different slide types and are adaptable to multiple protocols that use common reagents found in most laboratories.

The array was designed to detect *Pasteurella multocida, Manheimia haemolytica, Histophilus somni,* and *Arcanobacterium pyogenes*, common pulmonary pathogens that afflict U.S. livestock, using their 16S ribosomal RNA genes for differentiation. These pathogens were individually hybridized to the designed array to determine if the array would be sufficient for further testing with bovine tissue.

Materials and Methods

Designing Probes

The 16S ribosomal RNA gene was used to distinguish between different pathogens. 16S sequence for each pathogen was obtained from GenBank. Numerous 16S gene sequences were obtained for the same pathogen representing different strains available from GenBank (Table 1). Sequences were aligned using Invitrogen's Vector

NTI software. For each pathogen, six regions of the 16S gene were manually selected for use in the design of target probes. The regions selected were homologous with other 16S sequences of the same pathogen species, but unique from the 16S sequences of the other pathogens. For each region, both the sense and anti-sense sequences were used to design complimentary probes (designated A and B). This allowed for a total of twelve different sequences to be used for probe design for each pathogen (Table 2). Sequences were 20-22 bases in length with a calculated melting temperature of 55°C± 1°C according to Vector NTI calculations. Integrated DNA Technology (IDT, Coralville, IA) synthesized the sequences for the probes. Mismatched probes were also designed for each probe sequence to help differentiate positive signal from background noise by changing two nucleotides with their compliment. The positions of these changes were manually selected, based on maintaining a melting temperature of ±+- 1 °C difference from that of the perfect matched sequence. Mismatched probes are designated with a "MM" in table 2.

Table 1. List of GenBank Sequences Aligned to Select Probes for In-house Arrays

A. pyogenes	ACYRR16S	X79225		
H. somni	AF031936	AF549392	AF549396	AF549400
	AF549387	AF549393	AF549397	AF549401
	AF549388	AF549394	AF549398	AF549403
	AF549391	AF549395	AF549399	AF549404
				M75046
M. haemolytica	M75080	PHU57066	PHU57069	PHU57071
	M75063	PHU57068	PHU57070	PHU57072
P. multocida	AY078996	AY299306	AY299316	DQ286927
	AY078997	AY299307	AY299317	DQ286928
	AY078998	AY299308	AY299318	DQ286929
	AY078999	AY299309	AY299319	DQ288145
	AY079000	AY299311	AY324032	DQ288146
	AY299304	AY299313	AY683485	M35018
	AY299305	AY299315	AY999017	

Table 2. Position and Sequence of In-house Array Probes

Well Position	Probe Name	Sequence
A01	M. haemolytica 16S 1A	TGGAGGGGGATAACTACTGG
A02	M. haemolytica 16S 1A MM	TCGAGGGGGATAACTACTCG
A03	M. haemolytica 16S 1B	CCAGTAGTTATCCCCCTCCA
A04	M. haemolytica 16S 1B MM	CGAGTAGTTATCCCCCTCGA
A05	M. haemolytica 16S 2A	TTGGTGAGGTAAAGGCTCAC
A06	M. haemolytica 16S 2A MM	TTGGTCACGTAAAGGCTCAC
A07	M. haemolytica 16S 2B	GTGAGCCTTTACCTCACCAA
A08	M. haemolytica 16S 2B MM	GTGACCCTTTACGTCACCAA
A09	M. haemolytica 16S 3A	CAGTCGATTGACGTTAATCACA
A10	M. haemolytica 16S 3A MM	CAGTCGATTGACGTTAATGAGA
A11	M. haemolytica 16S 3B	TGTGATTAACGTCAATCGACTG
A12	M. haemolytica 16S 3B MM	TGTCATTAAGGTCAATCGACTG
B01	M. haemolytica 16S 4A	GCCATAAGATGAGCCCAAGT
B02	M. haemolytica 16S 4A MM	GCCATAAGATCAGCCCAACT
B03	M. haemolytica 16S 4B	ACTTGGGCTCATCTTATGGC
B04	M. haemolytica 16S 4B MM	ACTTCGGCTCATCTTATGCC
B05	M. haemolytica 16S 5A	AAAGGGTGGGACTTTCGG
B06	M. haemolytica 16S 5A MM	AAAGGGTGGGAGTTTCCG
B07	M. haemolytica 16S 5B	CCGAAAGTCCCACCCTTT
B08	M. haemolytica 16S 5B MM	CCCAAAGTCCCAGCCTTT
B09	M. haemolytica 16S 6A	AGCTGTAAGGTGGAGCGAAT

B10	M. haemolytica 16S 6A MM	AGCTGTAAGCTGGACCGAAT
B11	M. haemolytica 16S 6B	ATTCGCTCCACCTTACAGCT
B12	M. haemolytica 16S 6B MM	ATTCGCTCCAGCTTACACCT
C01	H. somni 16S 1A	GTGATGAGGAAGGCGATTAGT
C02	H. somni 16S 1A MM	GTCATCAGGAAGGCGATTAGT
C03	H. somni 16S 1B	ACTAATCGCCTTCCTCATCAC
C04	H. somni 16S 1B MM	ACTAATCGCCTTCCTGATGAC
C05	H. somni 16S 2A	AGCATGTTAGGGTGGGAACT
C06	H. somni 16S 2A MM	AGCATCTTACGGTGGGAACT
C07	H. somni 16S 2B	AGTTCCCACCCTAACATGCT
C08	H. somni 16S 2B MM	AGTTCCCACCGTAAGATGCT
C09	H. somni 16S 3A	CACGCAGGTGGTGACTTAAG
C10	H. somni 16S 3A MM	CAGGCAGCTGGTGACTTAAG
C11	H. somni 16S 3B	CTTAAGTCACCACCTGCGTG
C12	H. somni 16S 3B MM	CTTAAGTCACCAGCTGCCTG
D01	H. somni 16S 4A	CAGCATTTCAGACTGGGTGA
D02	H. somni 16S 4A MM	CACCATTTCAGAGTGGGTGA
D03	H. somni 16S 4B	TCACCCAGTCTGAAATGCTG
D04	H. somni 16S 4B MM	TCACCCACTCTGAAATGGTG
D05	H. somni 16S 5A	AGATACTGACGCTCGAGTGC
D06	H. somni 16S 5A MM	AGATACTCACGCTGGAGTGC
D07	H. somni 16S 5B	GCACTCGAGCGTCAGTATCT
D08	H. somni 16S 5B MM	GCACTCCAGCGTGAGTATCT
D09	H. somni 16S 6A	CAGAGATGGTGGTGTGCCTA
D10	H. somni 16S 6A MM	CACAGATGGTGGTCTGGCTA

D11	H. somni 16S 6B	TAGGCACACCACCATCTCTG
D12	H. somni 16S 6B MM	TAGCCACACCACCATCTGTG
E01	A. pyogenes 16S 1A	CCTTGTCTTTGGGATAAGCC
E02	A. pyogenes 16S 1A MM	CCTTCTCTTTCGGATAAGCC
E03	A. pyogenes 16S 1B	GGCTTATCCCAAAGACAAGG
E04	A. pyogenes 16S 1B MM	GGGTTATCCCAAACACAAGG
E05	A. pyogenes 16S 2A	GAGTGTGGTAGGGGTAATTGG
E06	A. pyogenes 16S 2A MM	GAGTCTGGTAGCGGTAATTGG
E07	A. pyogenes 16S 2B	CCAATTACCCCTACCACACTC
E08	A. pyogenes 16S 2B MM	CCAATTAGCCCTACGACACTC
E09	A. pyogenes 16S 3A	GGTTACTGGGCCATTACTGAC
E10	A. pyogenes 16S 3A MM	GGTTAGTGGCCCATTACTGAC
E11	A. pyogenes 16S 3B	GTCAGTAATGGCCCAGTAACC
E12	A. pyogenes 16S 3B MM	GTCAGTAATGGCCGACTAACC
F01	A. pyogenes 16S 4A	GGCTTGACATACACTGCGAT
F02	A. pyogenes 16S 4A MM	GGCTTCACATACACTGCCAT
F03	A. pyogenes 16S 4B	ATCGCAGTGTATGTCAAGCC
F04	A. pyogenes 16S 4B MM	ATCGCACTGTATGTGAAGCC
F05	A. pyogenes 16S 5A	GTGGTGTACAGGTGGTGCAT
F06	A. pyogenes 16S 5A MM	GTGGTGTACAGCTGCTGCAT
F07	A. pyogenes 16S 5B	ATGCACCACCTGTACACCAC
F08	A. pyogenes 16S 5B MM	ATGCACCACGTGTAGACCAC
F09	A. pyogenes 16S 6A	GCCTGTGAGGGTGAGCTAAT
F10	A. pyogenes 16S 6A MM	GCCTGTGAGGCTGAGGTAAT
F11	A. pyogenes 16S 6B	ATTAGCTCACCCTCACAGGC

F12	A. pyogenes 16S 6B MM	ATTAGCTCACCCTGAGAGGC
G01	P. multocida 16S 1A	TAACTGTGGGAAACTGCAGC
G02	P. multocida 16S 1A MM	TAACTGTGGCAAACTGGAGC
G03	P. multocida 16S 1B	GCTGCAGTTTCCCACAGTTA
G04	P. multocida 16S 1B MM	GCTGCACTTTCCCACACTTA
G05	P. multocida 16S 2A	CGGTAATGAGGAAGGGATGT
G06	P. multocida 16S 2A MM	CGGTTATCAGGAAGGGATGT
G07	P. multocida 16S 2B	ACATCCCTTCCTCATTACCG
G08	P. multocida 16S 2B MM	ACATCCCTTCCTGATAACCG
G09	P. multocida 16S 3A	ATTTGGGGATTGGGCTATAT
G10	P. multocida 16S 3A MM	ATTTGGGCATTGGCCTATAT
G11	P. multocida 16S 3B	ATATAGCCCAATCCCCAAAT
G12	P. multocida 16S 3B MM	ATATAGCCCAATGGCCAAAT
H01	P. multocida 16S 4A	GACTGCCAGTGACAAACTGG
H02	P. multocida 16S 4A MM	GACTGCCAGTGAGAAAGTGG
H03	P. multocida 16S 4B	CCAGTTTGTCACTGGCAGTC
H04	P. multocida 16S 4B MM	CCAGTTTGTGACTGGGAGTC
H05	P. multocida 16S 5A	CTCAGAGATGAGCTTGTGCC
H06	P. multocida 16S 5A MM	CTCAGAGATCACCTTGTGCC
H07	P. multocida 16S 5B	GGCACAAGCTCATCTCTGAG
H08	P. multocida 16S 5B MM	GGCAGAAGCTCATGTCTGAG
H09	P. multocida 16S 6A	ATACAGAGGGCAGCGAGAGT
H10	P. multocida 16S 6A MM	ATACAGAGGCCACCGAGAGT
H11	P. multocida 16S 6B	ACTCTCGCTGCCCTCTGTAT
H12	P. multocida 16S 6B MM	ACTGTCGGTGCCCTCTGTAT

Tailing Probes

A polyinosine tail was added to each oligonucleotide probe using terminal transferase.[78] 12nmols of each probe was synthesized by IDT in a deep-well 96-well plate. Probes were diluted with TE buffer (pH 7.0) to a 10μM concentration. The probes were then tailed with dITPs in a 20μl reaction that contained 10μl of 10μM probe oligonucleotides (IDT), 4μl 25mM CoCl (New England Biolabs, Ipswich, MA), 1μl 20U/μl TdT (New England Biolabs), 4μl 5X TdT buffer (New England Biolabs), and 1μl 10μM dITP (Roche Applied Science, Foster City, CA). The reaction was incubated at 37°C for 2 hours and then allowed to freeze at -20°C overnight. After freezing, the reaction was thawed in a biological safety cabinet and allowed to completely evaporate for 24 hours. After evaporation was complete, the oligonucleotide probes were reconstituted in a 1X TE buffer (pH 7.0) with 5% glycerol solution. This TE buffer/glycerol solution helped to prevent evaporation of the probe solution during spotting of the microarrays.

Spotting Microarrays

Vivid microarray slides (Pall, East Hills, NY) were used as the platform for our experiments. The Vivid arrays have a thin uniform nylon membrane that covers the glass slide, which is reported to improve the target/probe binding and reduces background over traditional glass slides.[28] The Vivid slides were spotted with a QArray^{mini} spotter (Genetix, Queensway, UK) following the manufacturer's protocol using solid pins and were spotted at 10°C at 50% humidity.

Culturing Pathogens

Pure cultures of the pathogens were obtained from the American Type Culture Collection (ATCC, Manassas, VA). A list of ATCC numbers is found in Table 3. Bacteria were streaked onto blood agar plates and were incubated under aerobic conditions at 37°C for 24-48 hours depending on the growth rate of each bacterial species. Individual colony forming units (CFU) were selected for DNA extraction as needed. The only exception was that *A. pyogenes* was incubated under anaerobic conditions. All of the pathogens were isolated in pure culture the Utah Veterinary Diagnostics Laboratories (UVDL), Logan, UT. PCR was used to confirm bacterial identity after DNA was extracted.

Table 3. ATCC Numbers of Pathogens for In-house Arrays

Pathogen	ATCC #
P. multocida	12945
M. haemolytica	55518
H. somni	700025
A. pyogenes	49698

Extracting DNA

DNA was extracted from bacterial CFU's by lysing the cells with CTAB and purifying the DNA using a phenol/chloroform separation.[5] DNA concentration and purity ratios were measured using a ND-1000 spectrophotometer (Nanodrop, Wilmington, DE).

PCR

PCR was used to confirm the identity of each isolate by amplifying the 16S ribosomal genes using primers listed on Table 4. Also, as will be discussed later, it was necessary to amplify the 16S ribosomal gene in each group or sample through PCR prior to adding a fluorescent label. The initial objective was to achieve sufficient labeling without prior amplification, but was unsuccessful. The PCR protocol was as follows: 200-300ng of gDNA was added to a 50µl reaction containing 1µl Advantage 2 polymerase (Clontech, Mountain View, CA), 1µl dNTPs (10µM) (New England Biolabs), 1µl 10µM forward and reverse primers (IDT) each, and 5µl 10X buffer (Clontech). Reactions were initially denatured (95°C-1 min), followed by 30 amplification cycles (95°C-30s, 57°C-30s, 68°C-3min), and completed with a final extension step (68°C-3min). PCR products were run on a 1% TAE gel stained with 0.5X SYBR Safe (Invitrogen) to confirm successful amplification.

Table 4. Primer Sequences Used to Amplify the 16S Gene of Bovine Pathogens

Pathogen	Primers
P. multocida	F-5'-AACACATGCAAGTCGAACGG R-5'-TGACGGGCGGTGTGTACAA
M. haemolytica	F-5'-TCAGATTGAACGCTGGCGGC R-5'-CACACCCCAGTCATGAATCATACCG
H. somni	F-5'-GAGTGGCGGACGGGTGAGTAAA R-5'-ACTTCTGGTACAACCCACTCCCATG
A. pyogenes	F-5'-GCGTGCTTAACACATGCAAGTCG R-5'-TCACCGCAGCGTTGCTGATC

Labeling DNA

16S ribosomal amplicons (2 μg) were labeled with Alexa Fluor 546, because of its superior fluorescence compared to Cy3,[63] using Invitrogen's BioPrime Plus Array CGH Indirect Genomic Labeling System. The protocol was followed according to manufactures' instructions except the optional DNA digestion step prior to labeling was skipped and was regarded as unnecessary for this study.

Hybridization, Washing, Scanning

Vivid slides were soaked in a preheated pre-hybridization buffer (100mM NaPO4, 20% SDS, 0.1g Casein) for 30 minutes at 45°C, after which excess buffer was allowed to run off. A total of 1μg of labeled target was added to a hybridization buffer (100mM NaPO4, 20% SDS, 0.1g Casein) for a total volume of 100μl. Hybridization cocktail was heated to 95°C for 5 minutes to denature the DNA. The entire hybridization cocktail, 100μl, was pipetted onto the array, after which a cover slip was laid. Arrays were incubated in a Boekel Scientific Hybridization Oven (model # 241000) at 45°C for 2 hours. Filter paper soaked with pre-hybridization buffer was laid in the oven to retain humidity within the chamber. Arrays were washed in a preheated "Wash A" solution (2X SSC, 0.5% SDS) at a range of temperatures (40°C-55°C) for 30 minutes with agitation. During this step, cover slips were removed if they did not already slide off by themselves as a result of shaking. Arrays were then transferred to a preheated "Wash B" solution (0.5X SSC, 0.5% SDS) and washed at a range of temperatures (40°C-55°C) for 30 minutes with agitation. Arrays were finally transferred to a rinse solution (0.5X SSC) at room temperature for about 10 seconds with agitation. The slides were dried by spinning them in 50mL conical tubes in a centrifuge at 3000 rpm for 5 minutes. Arrays

were immediately scanned with an Axon Genepix 4200A scanner (Molecular Devices Sunnyvale, CA) with a PMT setting of 400.

Data Analysis

Scanned images of arrays were analyzed with GenePix Pro, v5.1 (Molecular Devices). For each hybridization spot, a foreground and background signal was calculated. The foreground signal was then divided by the background signal to get a signal ratio. A signal ratio of 2 or higher was considered a positive signal.

Results

Layout of In-house Array

The organisms, *P. multocida, H. somni, A. pyogenes, and M. haemolytica*, are represented on the in-house array. Each array consisted of 10 repeated regions as shown in figure 1. Each region was divided into four quadrants. The two left quadrants were blanks that acted as negative controls for hybridization. The top right quadrant contained all PM probes for all pathogens, while the bottom right quadrant contained the corresponding MM probes. Probes representing the 16S ribosomal gene for *P. multocida* were located within the 1st and 2nd columns as shown in the red box of figure 1. Probes for *A. pyogenes* are found in the 3rd and 4th columns as shown in the yellow box, *H. somni* in the 5th and 6th columns as shown in the blue box, and *M. haemolytica* in the 7th and 8th columns as shown in the white box. The position of a MM probe matches that of its corresponding PM probe.

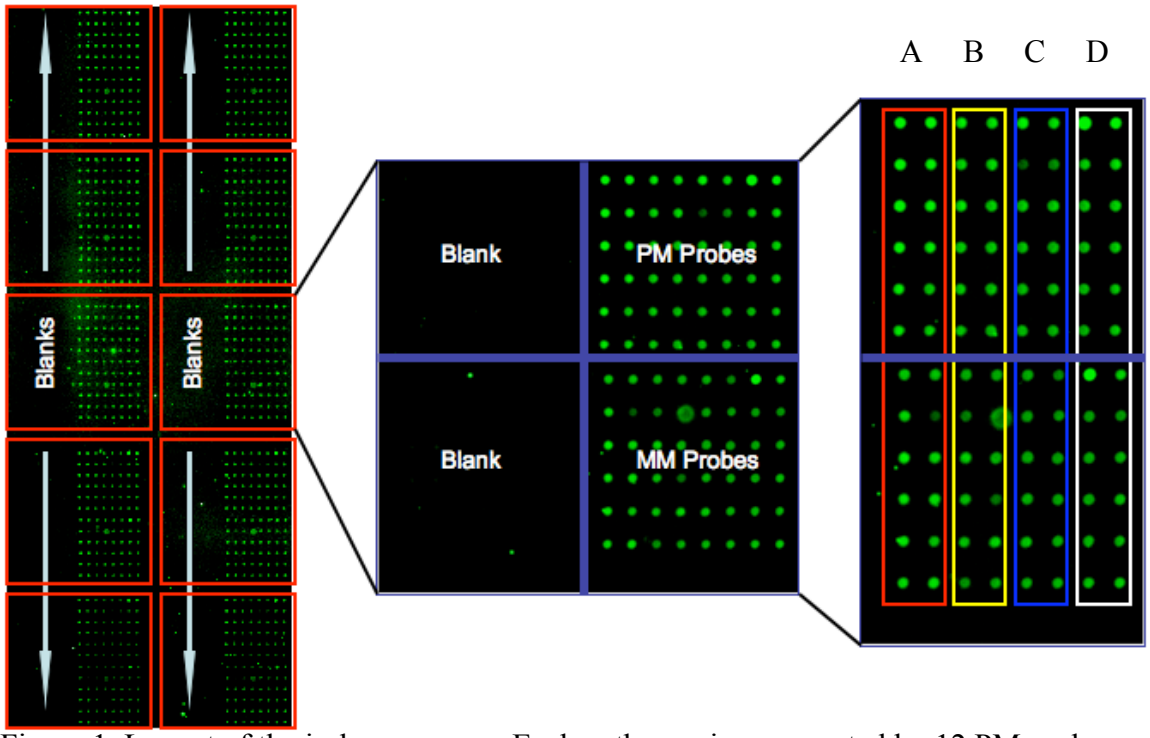

Figure 1. Layout of the in-house array. Each pathogen is represented by 12 PM probes and 12 MM probes. Each block of probes is repeated on the array 10 times.
(A) *P. multocida* (B) *A. pyogenes* (C) *H. somni* (D) *M. haemolytica*

Hybridizing Genomic DNA

The initial aim of this study was to detect proper signals with hybridized gDNA without amplification. This possibility was tested by hybridizing whole gDNA, sheared DNA by use of a sonicator, and enzymatically digested DNA using restriction enzyme RSA. However, this procedure was not able to produce any signal patterns (figure 2), nor were any significant changes seen with the signal as hybridization and wash conditions were changed. Ultimately, gDNA signals were so weak that the strongest signals seen in the images in figure 2 corresponded to the greatest indent on the nylon membrane inadvertently created from spotting the arrays. Due to these findings, it became

necessary to amplify the 16S gene for each pathogen through PCR in order to enhance signal patterns.

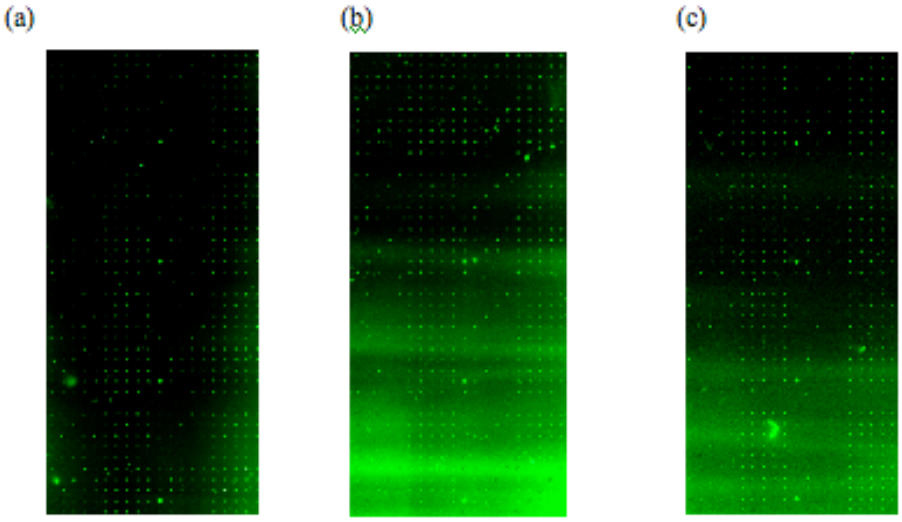

Figure 2. Three arrays hybridized with unamplified *P. multocida* gDNA. (a) Whole gDNA (b) sheared DNA (c) restriction enzyme-digested DNA. No signal pattern was detected.

Optimizing Washing Conditions

P. multocida amplified gDNA was hybridized to a series of arrays, all at 45°C. Wash temperatures were varied from 40°C to 55°C to find an optimal wash temperature. From a series of experiments, the best hybridizations occurred in wash conditions at 50°C where the positive signals were consistently easier to distinguish from the background. As wash temperatures decreased to 45°C and less, an increase in signal was seen for both PM and MM probes indicating the possibility of non-specific binding. At wash temperatures higher then 50°C, most of the probes produced lower signals at or near background levels. Based on these observations and information by Loy et al., 50°C was chosen as optional wash temperature.[51-52]

Hybridization Patterns

Numerous *P. multocida* arrays were initially run to optimize conditions until a strong hybridization pattern for *P. multocida* became apparent (figure 3). With the exception of probes designated *P. multocida 16*-3 A&B , all *P. multocida* probes showed the highest signals, though the analyzed signal values never reached a level of 2, which was the designated positive signal threshold. However, despite this limitation, the green hybridization pattern for *P. multocida*, within the first two columns in the PM quadrant, was easily distinguishable from background noise.

As amplified DNA from the other pathogens was introduced to the array, each produced a hybridization pattern of green spots within the columns of the PM quadrant corresponding to the correct pathogen. In the arrays hybridized with *A. pyogenes* DNA, seven of the twelve probes showed strong signals, most with a value of 2 or greater (figure 4). However, three MM probes consistently produced high signals, while one of the counterpart PM probes did not show any signal. Amplified *H. somni* DNA hybridized well to its probes as all of the PM probes showed strong signals, though four were significantly less than the other eight (figure 5). Arrays hybridized with amplified *M. haemolytica* DNA seemed to be the least successful (figure 6). Only four probes, designated *M. haemolytica* 16-1 A&B and 16-4 A&B, had signal values higher than 2, which were followed distantly by 16-3 A&B.

Though some random signaling not associated with the target DNA in each of the arrays was observed, indicating a need for additional optimization, we were mostly concerned with probes *M. haemolytica* 16-1 A & 16-4 A, which produced strong signals in all arrays.

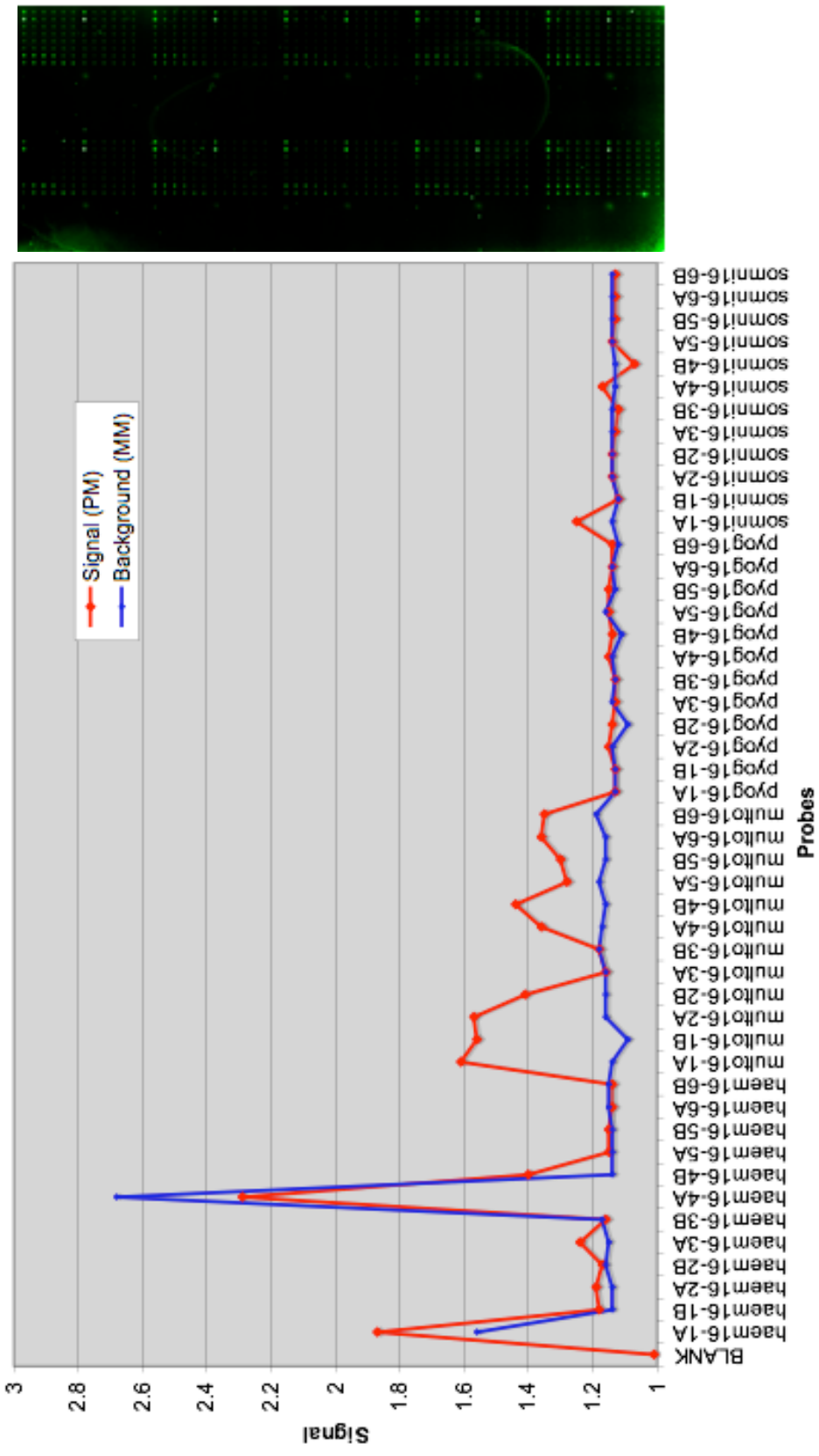

Figure 3. *P. multocida* hybridization pattern and signal intensities.

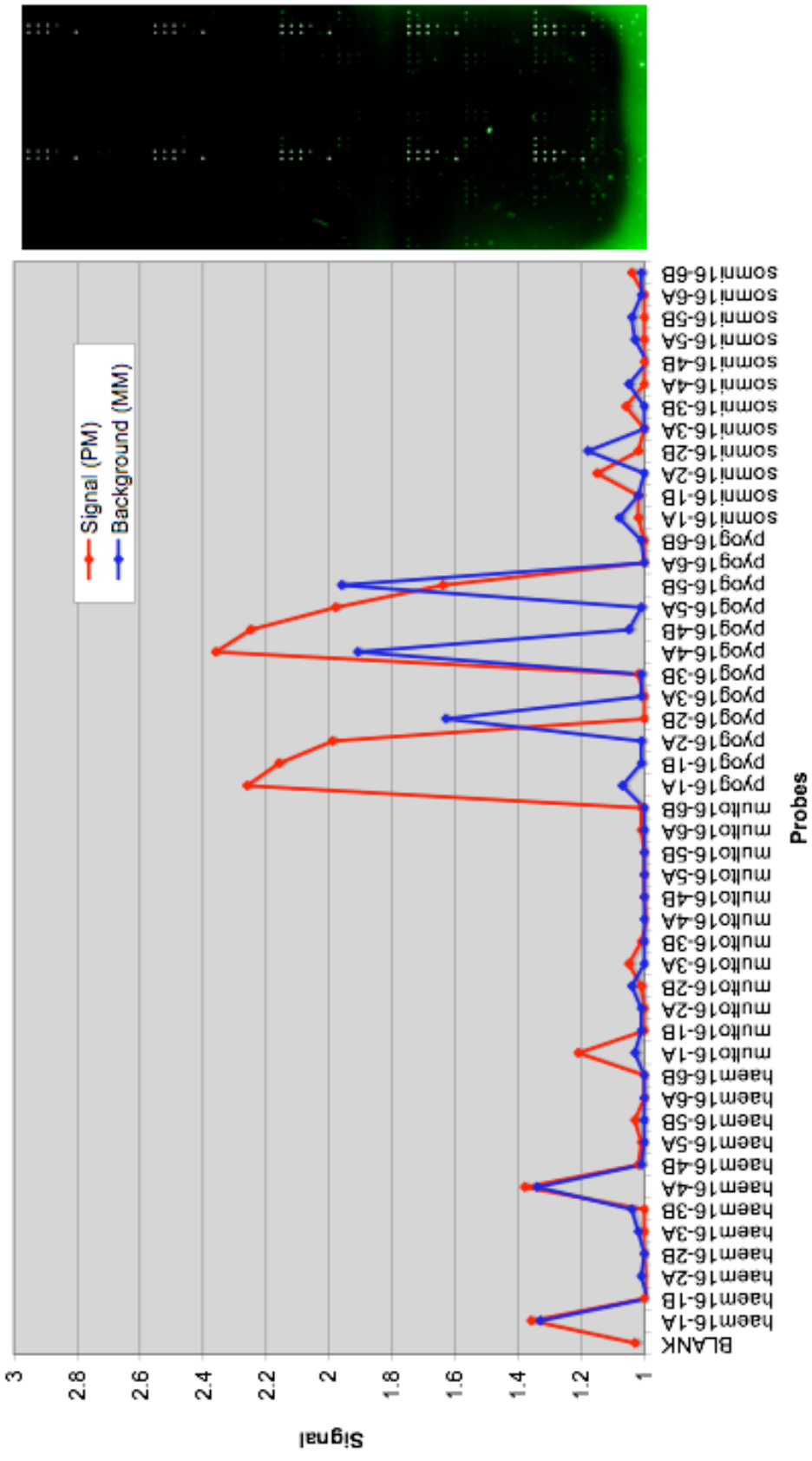

Figure 4. *A. pyogenes* hybridization pattern and signal intensities.

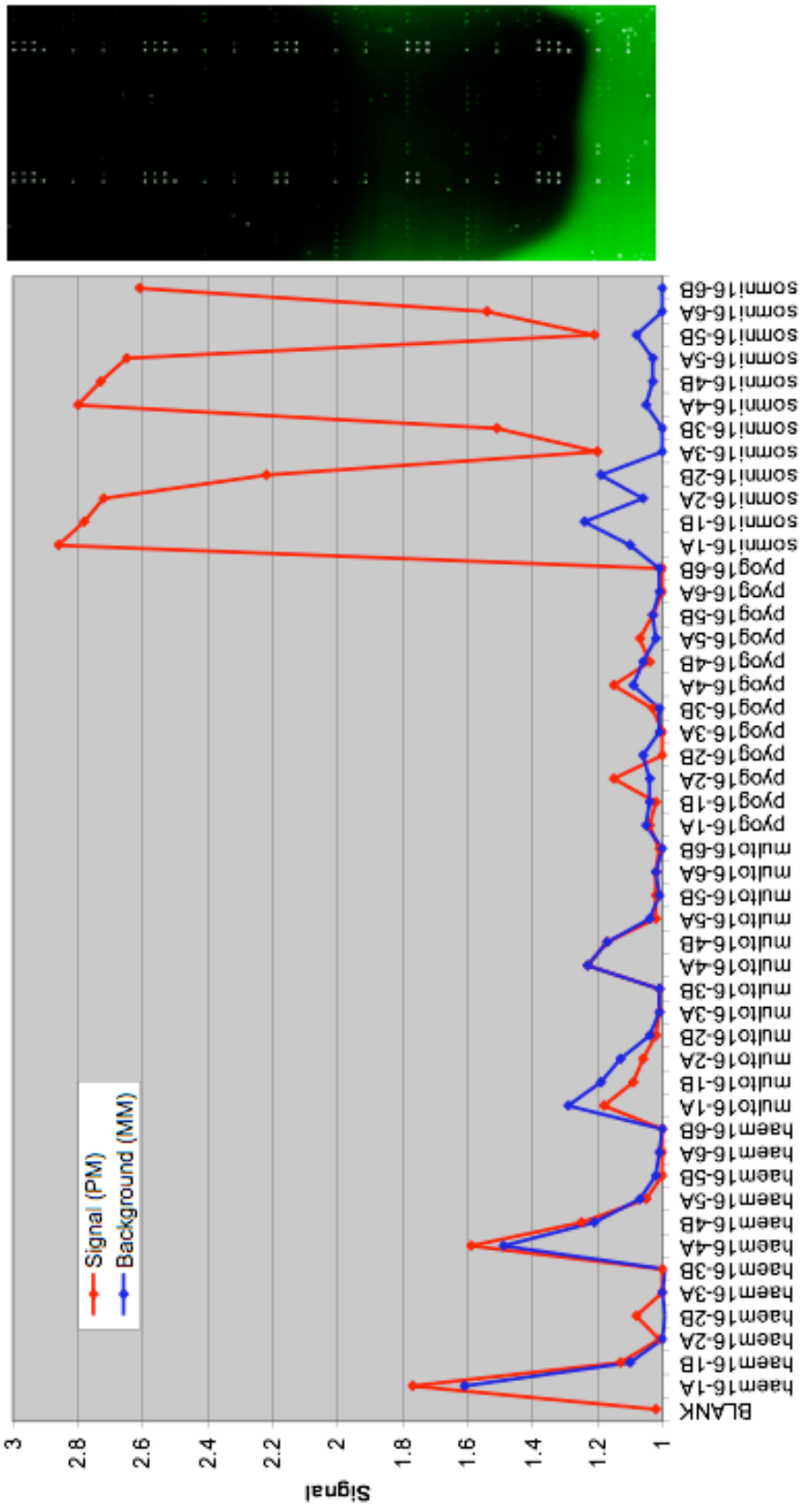

Figure 5. *H. somni* hybridization pattern and signal intensities.

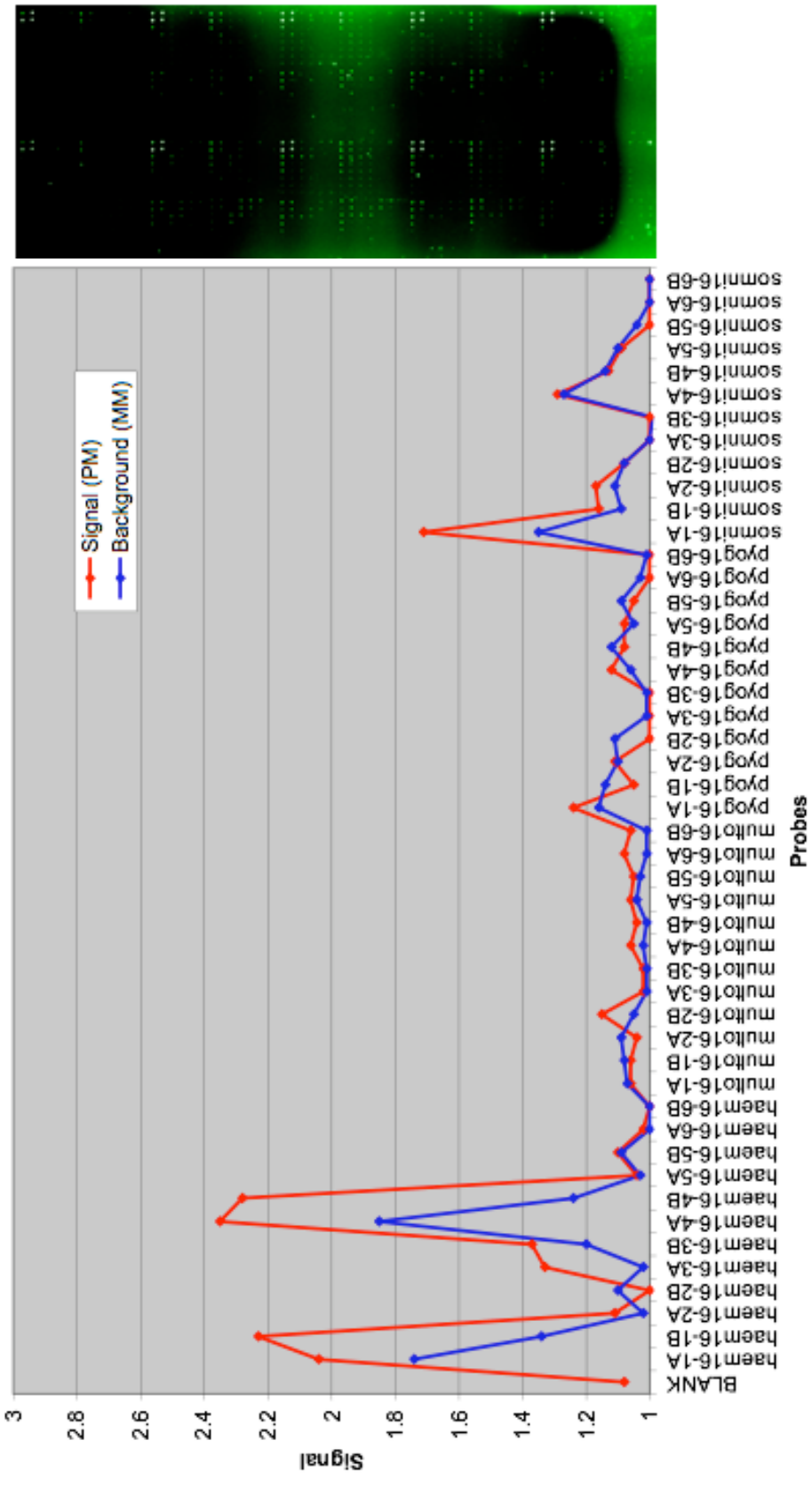

Figure 6. *M. haemolytica* hybridization pattern and signal intensities.

Discussion

Our primary aim in this study was to design a microarray that could correctly detect specific bovine pathogens within infected cattle tissue. Our secondary objective was to design a platform that would be relatively inexpensive compared to current standard microarray costs. The overall aim was to provide an approach that could potentially challenge current veterinary diagnostic techniques in terms of accuracy, ease, safety, and cost-effectiveness.

The Good

After much trial and error in finding optimal assay conditions, we successfully hybridized gDNA from the four pulmonary pathogens, *P. multocida, M. haemolytica, H. somni,* and *A. pyogenes*, to the "in-house" microarray and produce distinct signal patterns for each pathogen. The probes for each pathogen were arranged such that it was possible to identify the pathogen hybridized from the signal pattern alone without the need of additional software analysis.

Not all of the probes designed for specific pathogens demonstrated successful hybridization with the target DNA. Since this was the first trial phase, and thus only the first round of elimination, we were not surprised that there were some probes that refused to hybridize. Overall, hybridization patterns repeated in each of the ten regions and were also consistent among repeated arrays. With the current data being consistent, we feel it would be safe to remove the unsuccessful probes from future arrays and use only the successful probes for diagnostics.

There were a total of twelve probes for each pathogen that targeted the 16S ribosomal gene. Probes for *P. multocida* had the highest success rate with 10 out of 12 showing successful hybridization. Both probe sets for *H. somni* and *A. pyogenes* had an average success rate of 58% (7/12 probes showed successful hybridization), which is sufficient for positive identification. Probes for *M. haemolytica* had the lowest success rate with only 4 out of 12 probes showing successful hybridization. However, two of the probes that target *M. haemolytica* also showed varied levels of signal when hybridized with the other pathogens, making them less useful for diagnosing mixed infections. The corresponding MM probes also produced a signal that was equal to or less than the PM probes, and even on the *M. haemolytica* array, some MM signal was evident for these two probes. This casts doubt as to whether these *M. haemolytica* signals were true signals. We feel that they were false positives created from some unknown hybridization issue, probably a technical issue, since the signals were evident to some degree in all of the hybridized arrays.

The Bad

We were able to produce distinct, visible signal patterns that confidently represented each pathogen, except *M. haemolytica*. However, there are still additional challenges that will be faced in any future versions of these in-house arrays. First, though the probes showed fairly good consistency in terms of whether the probe was successful or not at producing a signal, there was some signal variation that seemed to be dependant on the probe's physical location on the array. The strength of the signal overall tended to be stronger on one side of the array than the other. As you view the arrays presented in this report, from top to bottom, the strength of the signal was typically stronger on the

bottom of the array. We feel this represents a mechanical error with the scanner since this pattern was consistent with each array regardless of which way the array was inserted into the scanner. This could certainly be corrected in the future, but no doubt allowed a bias to enter in the calculations made for the signal strength of each probe during this series of experiments.

Second, these arrays were prone to random spots, smears, and scars that produced extra noise as can be seen on the pictures of each array. These extra blemishes affected the signal of any probes that shared the same physical location with the blemishes on an array. These blemishes were the result of the arrays being exposed to a "dirty" environment during processing. The issue could be corrected by improving the processing protocols and working environment. Taking steps to "clean" the procedure and working environment should lessen the formation of these blemishes.

All of the previously mentioned obstacles could be resolved with improved versions of the array through selecting the appropriate probes, optimizing hybridization conditions, including the working environment, and tuning the necessary equipment. The third major issue, however, presented a problem that was not expected. One of our initial aims was to produce positive signal patterns by hybridizing whole gDNA without the need for extra sample preparation (other than labeling). In our initial attempts, we were unable to achieve that objective with these arrays. In order to produce the desired signals, it was necessary to first amplify the 16S gene sequence through PCR using species-specific primers. In our opinion, this was an unacceptable step in preparing target DNA. Using this procedure, one could simply identify the bacterial species

through the PCR amplification step, and further analysis by hybridization of the amplified sequence to a microarray was only redundant.

Conclusion

We designed, built, and processed our arrays in-house, and were encouraged by being able to produce distinct signal patterns for each of the hybridized pathogens. We came across a number of obstacles, both technical and mechanical, but felt that with time most could be resolved. Our most concerning obstacle was that a positive signal was only possible by first PCR amplifying the 16S gene sequence. Due to this unforeseen issue, more efficient array options were explored.

CHAPTER III

AFFYMETRIX ARRAYS

Introduction

For detection of specific bovine pathogens common in the US we designed a custom Affymetrix GeneChip, STYLMONOa520430F. We were interested in detecting pathogens that target various organs; *P. multocida, M. haemolytica, H. somni,* and *A. pyogenes* from the lung and nasal cavities, and *S. typhimurium, S. dublin,* and *E. coli* from the liver, lung, kidney, and spleen. These pathogens were hybridized to the GeneChip both individually and spiked into bovine tissue prior to extraction and hybridization.

The STYLMONOa520430F GeneChip was primarily designed for *S. typhimurium* strain LT2 and *L. monocytogenes* strain EGDe, and their entire genomes are represented on the array, including a series of intergenic regions.[75] Of these two pathogens, only *S. typhimurium* was of interest in this study. *L. monocytogenes* genes were only used to help calculate background signal. For the remaining pathogens, only a select set of genes were represented due to limited space available on the array.

Material and Methods

Array Design

In total 4639 genes were selected for probe design: 56 genes for *P. multocida*, 3 genes for *M. haemolytica*¸ 2 genes for *H. somni*¸ 2 genes for *A. pyogenes*, 4518 genes for *S. typhimurium*, 8 genes for *S. dublin*, and 50 genes for *E. coli* (Appendix A). The

strategy for probe design depended greatly on whether the specific bacterial genome was sequenced. Ultimately, probes were designed to be specific to the target bacterium without significant cross-hybridization. The sequenced genomes of *P. multocida Pm70* and *E. coli 0157:H7* were run through ERGO (a web-based genome-comparison analysis program) to identify a list of possible genes from which to design probes for the array. With that program, the list of potential candidate genes useful for the array was reduced from thousands to only hundreds for each bacterium.

Select candidate gene sequences were analyzed using BLAST from the National Center for Biotechnology Information (NCBI) to search for genes that were the most unique to the target bacterium. Using an E score of $1e^{-5}$ as the cutoff for significance, genes were selected for probe design if they produced a significant hit ($<1e^{-5}$) with only that of its target bacterium. If there was a significant match to another organism other than the target bacterium, the sequence was removed from consideration. According to Affymetrix standard procedure, probes were to be designed only from within the first 600 bases of a selected gene, so any homologous sequences beyond 600 bases, regardless of the E score, were ignored. There were a few cases in which genes were selected for probe design even though the gene sequence had a significant hit ($<1e^{-5}$) with another organism besides that of the target bacterium. In these cases, the genes were admitted only because there was a need for more probes and these other non-target organisms were not closely related to the target bacterium nor were they related to human, bovine, or other farm species, and were thus less of a concern for potential contamination.

For *M. haemolytica*, *A. pyogenes*, *S. dublin*, and *H. somni*, whose genomes are not yet sequenced, we selected genes based on a review of the literature: lktA[26,30] and lpsA

for *M. haemolytica*,[37] plo for *A. pyogenes*,[42] rpoB for *H. somni*.[57] The *H. somni* genome has been sequenced (as of Dec 2006), but was not published at the time the array was designed. There are only seven sequences available in genbank for *S. dublin* of which the functions are unknown. However, since the sequences were specific to *S. dublin*, all seven genes were included. In addition to the gene search and literature review, the 16S and 23S gene sequences were also included for each pathogen.

The *S. typimurium LT2* genome was used to design the probes for *S. typhimurium.*[75] The probe sets designed for *S. typhimurium* differ from the probe sets of the other pathogens of interest in that the entire genome is represented instead of a select number of genes.

Standard Affymetrix protocols for designing GeneChips include the design and synthesis of MM probes for every probe set. Affymetrix utilizes MM probes to increase sensitivity in expression arrays by removing background signal caused by higher levels of non-specific binding as a result of using short oligonucleotide probes. However, since the STYLMONOa520430F GeneChip was used as a diagnostics array in this study, it was opted to exclude the MM probes in order to preferably allow more room for PM probes. Thus this GeneChip contained only PM probes.

Culturing Pathogens

Pure cultures of each bacterium were obtained from either the ATCC or from environmental strains that were isolated from tissues of cattle at the UVDL. A list of the sources is found on Table 5. Unless otherwise stated elsewhere in this report, ATCC strains were hybridized individually to test for cross hybridization, and the field isolates were spiked into tissue homogenates before DNA extraction and hybridization to confirm

that specific detection was possible while mixed with Bovine DNA. With the exception of *A. pyogenes*, the pathogens were streaked onto blood agar plates and were incubated under aerobic conditions at 37°C for 24-48 hours depending on the growth rate of the bacterium. *A. pyogenes* was incubated under anaerobic conditions. Individual CFUs were selected for DNA extraction as needed.

Table 5. List of ATCC Numbers of Pathogens for Affymetrix Arrays

Pathogen	ATCC # or Field Isolate	Source
P. multocida	12945	ATCC
M. haemolytica	55518	ATCC
	Field Isolate	UVDL
H. somni	700025	ATCC
	Field Isolate	UVDL
A. pyogenes	49698	ATCC
	Field Isolate	UVDL
S. typhimurium	35987	ATCC
	Field Isolate	UVDL
S. dublin	Field Isolate 1	UVDL
	Field Isolate 2	UVDL
E. coli	35150 Strain O157:H7	ATCC
	Field Isolate 1	UVDL
	Field Isolate 2 (078:H11)	UVDL
	Field Isolate 3 (Non-pathogenic)	UVDL

DNA Extraction

DNA was extracted from each pathogen by lysing the cells with CTAB and purifying the DNA using a phenol/chloroform separation.[5] DNA was also extracted from bovine lung, kidney, liver, and spleen using QIAGEN's DNeasy Protocol for Animal Tissues.

DNA was also extracted using the QIAGEN's DNeasy Protocol for Animal Tissue from bovine tissue homogenates that were spiked with bacterial cocktails to simulate multi-pathogen infections. Ten to 100 CFUs of each bacterium were added to 25 mg of each bovine tissue (10 mg of spleen) before extraction of DNA.

PCR

PCR was used to confirm that specific bacterial DNA was present in the spiked tissue samples. 200 ng of DNA from the spiked tissue samples was then used as starting material. The 16S gene of each bacterium was targeted with specific primers (Table 6). The GoTaq Green Master Mix protocol (Promega; P/N M7122) was followed using 25ul reactions that were initially denatured (95°C-2 min), followed by 30 amplification cycles (95°C-30s, 63°C-30s, 72°C-2min), and completed with a final extension step (72°C-5min). PCR products were run on a 1% TAE gel stained with 0.5X SYBR Safe (Invitrogen, P/N S33102).

Labeling

Up to 4 µg of extracted DNA was fragmented in a 20 µl reaction containing 0.6U/µg DNaseI (Amersham Biosciences, P/N 27-0514-01) (diluted in 1X One-Phor-All Buffer), 2 µl 10X One-Phor-All Buffer (Amersham Biosciences, P/N 27-0901-02), and

up to 20 µl Nuclease-free water. The reaction was incubated at 37°C for 10 minutes and inactivated at 98°C for 10 minutes. A portion of the extracted DNA was examined on a 1% TBE gel stained with 1X SYBR Gold (Molecular Probes, P/N S-11494).

The entire fragmented DNA sample was labeled on the 3' termini in a 50 µl reaction containing 2 µl of 7.5 mM GeneChip DNA Labeling Reagent (Affymetrix, P/N 900542), 10 µl 5X Reaction Buffer, 2 µl Terminal Deoxynucleotidyl Transferase (Promega, P/N M1875), and up to 50 µl Nuclease-free water. The reaction was incubated at 37°C for 1 hour and was stopped with 2 µl of 0.5 M EDTA.

A gel-shift assay was run to verify labeling efficiency. 5 µl of 2 mg/mL ImmunoPure NeutrAvidin (Pierce Chemical, P/N 31000) was added to a 200 ng aliquot of labeled DNA and incubated at room temperature for 5 minutes. DNA sample was loaded onto a 1% TBE gel, after which the gel was stained in 1X SYBR Gold for 10 minutes.

Hybridization, Washing, Scanning

The labeled DNA was diluted in 6.5 µl of 20X Hybridization controls and Nuclease-free water up to 50 µl. Diluted labeled DNA/Hybridization control mix was added to a hybridization cocktail and hybridized to the STYLMONOa520430F GeneChip according to the Affymetrix Prokaryotic Target Hybridization protocol under the 100 format (Midi). The arrays were incubated at 45°C at 50 rpms for 16 hours in a GeneChip Hybridization Oven 640. Arrays were washed and stained in a GeneChip Fluidics Station 450 and scanned with a GeneChip Scanner 3000. Washing, staining and scanning were done according to the Affymetrix Prokaryotic Arrays: Washing, Staining, and Scanning protocol.

Table 6. Primers Sequences that Target the 16S Gene of Each Pathogen

Pathogen	Primer Sequence
P. multocida	F - 5'-AACACATGCAAGTCGAACGG R - 5'-TGACGGGCGGTGTGTACAA
M. haemolytica	F - 5'-TCAGATTGAACGCTGGCGGC R - 5'-CACACCCCAGTCATGAATCATACCG
H. somni	F - 5'-GAGTGGCGGACGGGTGAGTAAA R - 5'-ACTTCTGGTACAACCCACTCCCATG
A. pyogenes	F - 5'-GCGTGCTTAACACATGCAAGTCG R - 5'-TCACCGCAGCGTTGCTGATC
S. typhimurium	F - 5'-GTTTGATCCTGGCTCAGATTGAACG R – 5'-CGGACTACGACGCACTTTATGAGGT
S. dublin	F - 5'-GTTTGATCCTGGCTCAGATTGAACG R - 5'-TCGCGAGGTCGCTTCTCTTTGT
E. coli	F - 5'-TTGATCATGGCTCAGATTGAACGC R - 5'-CTAGCGATTCCGACTTCATGGAGTC

Data Analysis

The .cel files generated by the GeneChip Operating Software were processed in R. Since the arrays did not have MM probes, background was corrected using the Robust Multi-Array (RMA) average expression method as it was designed to ignore MM signals. For the datasets that required normalization, we used the "constant" normalization method in which all of the arrays in a dataset were scaled to a chosen baseline array. All

datasets were summarized using "pmonly," which calculated a signal value for each probe set by averaging the individual probe signals (Appendix B).

The arrays were analyzed by calculating a single signal value for each pathogen by averaging the probe set signal values. This overall signal value was calculated for each pathogen on all arrays to see if any of the pathogen probe sets were showing positive signals on arrays for which the pathogen was not hybridized. Background was also calculated for each array by averaging the probe set signal values for all non-target probe sets. For the background value, a select number of probe sets were excluded from analysis: the hybridization controls, which produced strong positive signals, and probe sets representing the intergenic regions of *L. monocytogenes* and *S. typhimurium*. These intergenic regions were excluded because they showed significant cross hybridization and produced strong signals that interfered with analysis. Overall signal value for all pathogens was compared on each array with the average background signal to determine whether a pathogen was detected on an array. Ideally, the target pathogen would show a signal higher than that of the calculated background. Also, the signal values for the pathogens not targeted on an array would be equivalent to that of the background signal.

Results

GeneChip Validation

Three STYLMONOa520430F arrays were hybridized with unknown concentrations of *P. multocida* DNA. The signal values for *P. multocida* were 224.39, 2438.39, and 164.21 while the signal for background and other pathogens ranged from 0.7 to 2.31 (figure 7). The hybridized arrays were analyzed further by analyzing the

signal values of the individual probe sets that targeted the different *P. multocida* genes. Of the 56 different probe sets that targeted *P. multocida*, 46 consistently produce positive signals at various strengths. However, 10 probe sets refused to produce any signal with all three hybridized arrays (figure 8, shows data from one array). With PCR and primers designed to amplify those 10 genes represented by the failed probe sets, we determined that those specific gene sequences were not present in our *P. multocida* DNA samples. We also determined that the different probe sets did not respond equally, in that some probe sets produced stronger signals than other probe sets hybridized with the same DNA (figure 8).

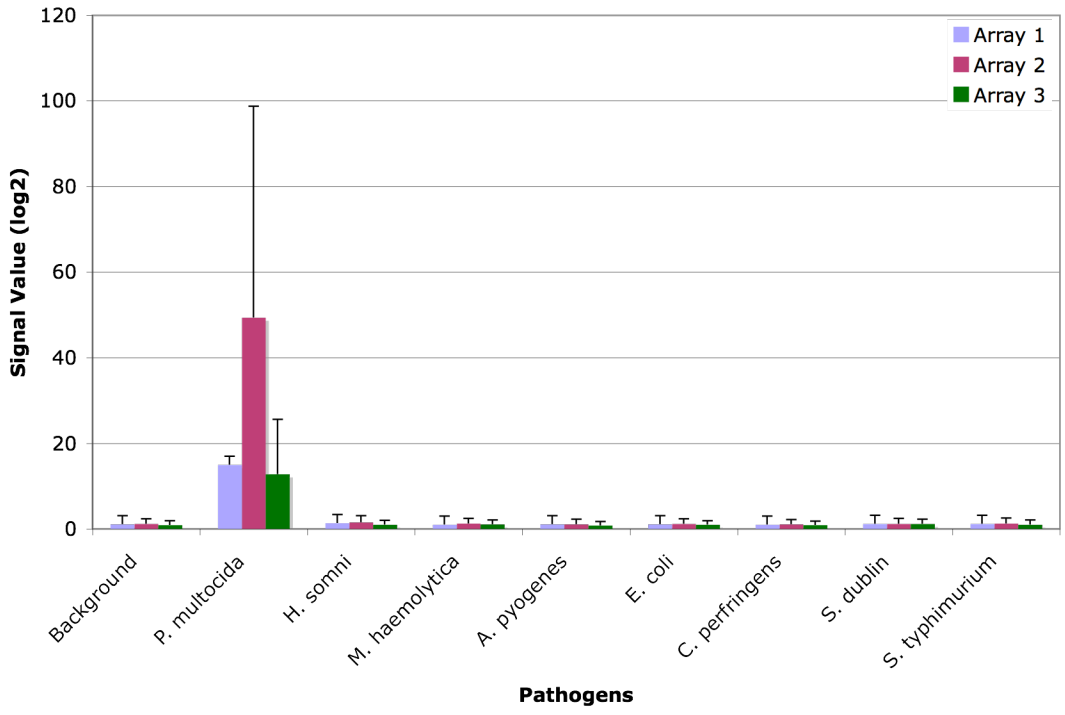

Figure 7. Validation of STYLMONOa520430F GeneChip using *P. multocida*.

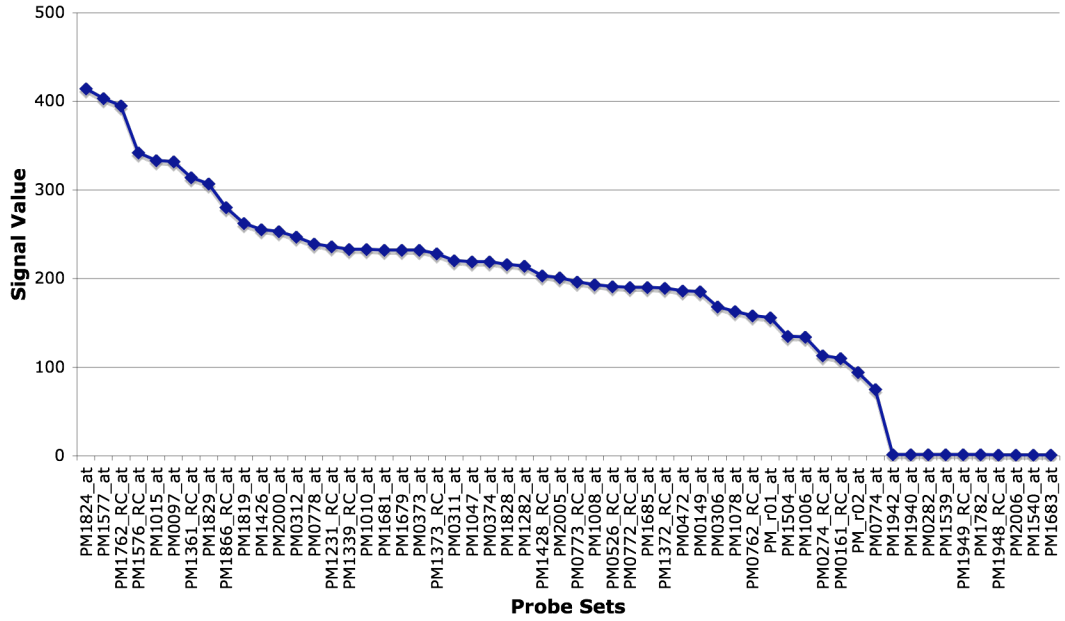

Figure 8. Evaluation of *P. multocida* probe sets.

Detection Limit

We used *P. multocida* in a series of detection limit experiments. We hybridized *P. multocida* to the STYLMONOa520430F array with 1 µg, 500 ng, 250 ng, and 10 ng of DNA. Each concentration only had 1 replicate except at 250 ng, which had 2 replicates. This dataset was normalized for comparison and as expected, we saw a decrease in signal value as we decreased the amount of DNA added to the array. At DNA concentrations of 1000, 500, and 250 ng, the signal value for the *P. multocida* genes (86.76, 49.75, and 19.97, respectively) was much higher than the background signal (1.39, 1.38, and 1.53, respectively) (figure 9). The 10 ng DNA concentration did not show a significantly higher signal at 1.68 than the background at 1.46. To ensure good detection, remaining experiments were run at a DNA concentration of 250 ng or higher.

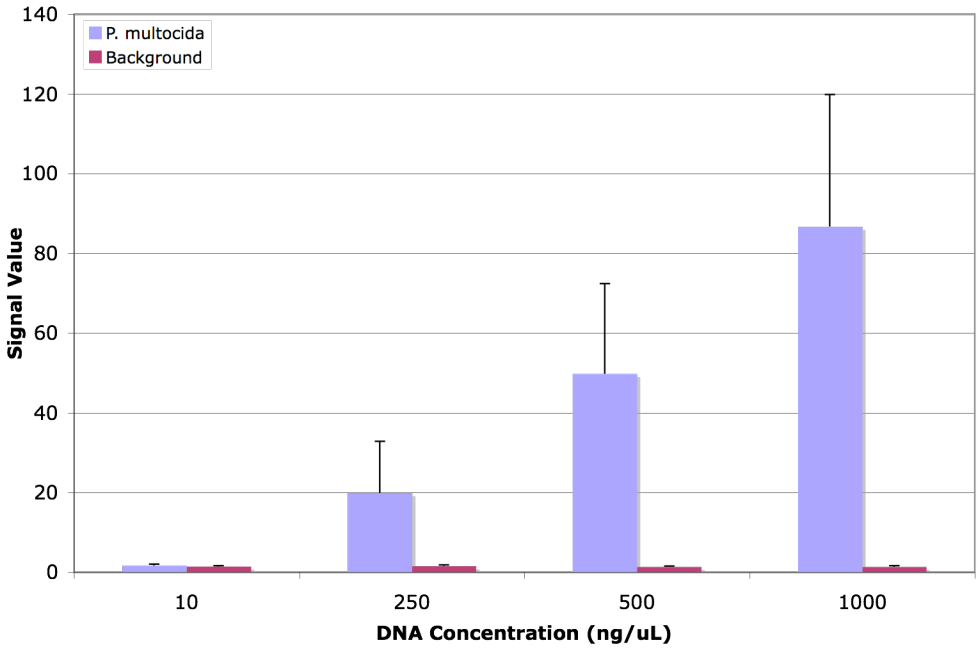

Figure 9. Detection limit of *P. multocida.*

Probe Set and Cross Hybridization Evaluation

To determine which genes gave positive signals for each bacterium and evaluate cross hybridization issues, *M. haemolytica*, *H. somni*, *A. pyogenes*, *S. dublin*, and *E. coli* were hybridized to individual STYLMONOa520430F arrays at 250 ng each. Only field isolates of *S. dublin* were hybridized. The *P. multocida* (250 ng) data from the previous experiment was added to this dataset. DNA from bovine lung, liver, kidney, and spleen was also hybridized at 250 ng each on a single STYLMONOa520430F array, as a negative control and to confirm that DNA from bovine tissue would not cross hybridize to any of the genes on the array. DNA from all bacterial species were hybridized in duplicates on separate days except *S. dublin*, which had only one replicate.

All arrays hybridized with *P. multocida, A. pyogenes,* and *S. dublin* showed positive signals for their target bacterial DNA when hybridized with 250 ng of DNA. For

P. multocida, 46 of 56 genes continued to produce a positive signal. In the two 250 ng arrays, the target probe sets produced signal values of 26.71 and 13.23 with a background of 1.60 and 1.47, respectively (figure 10). Signal values for the other pathogens ranged from 1.20-2.39 on both arrays. Genes designated PM_r01_at and PM_r02_at, which represent 16S and 23S genes, were among the positive signaling genes.

The arrays hybridized with *A. pyogenes* DNA had 1 of 2 genes that showed a positive signal. The two *A. pyogenes* arrays had target signal values of 7.75 and 4.6 with backgrounds of 1.28 and 1.48, respectively (figure 11). The 16S gene designated AP_16S showed strong signals at 14.15 and 7.4 while the published hypothetical genes designated AP_smc_plo_ftsyY showed no signal at 1.36 and 1.8. Signal values for the other bacteria ranged from 1.04 - 1.8 on both arrays.

The single *S. dublin*-hybridized array had a target signal value of 9.39 with a background of 1.17 (figure 12). *S. dublin* genes were the most successful of all the bacterial probe sets in terms of percentage. Only one gene, SD_273, failed to produce a signal, while the remaining genes produced positive signals of various strengths. However, this was the only array to show significant cross hybridization with another pathogen, *S. typhimurium,* with a signal value of 5.55, which will be discussed later. All other pathogen signal values ranged from 0.92-1.19.

S. typhimurium DNA was hybridized to the arrays in a previous project. No cross hybridization to any other bacterial DNA represented on the array was evident.[75]

M. haemolytica and *H. somni* did not demonstrate any significant signal above the background at 250 ng. The two *M. haemolytica* arrays had target signal values of 1.44 and 1.31 with backgrounds of 1.39 and 1.32 (figure 13). The other pathogenic signal

values ranged from 1.29-1.65. *H. somni* signal values were barely detectable at 2.65 and 2.17 with background values of 1.4 and 1.48. Signal values for the other pathogens ranged from 1.23 to 1.56 (figure 14). To determine if these genes were failing to hybridize or were simply below their detection limits, we also hybridized 1μg of *M. haemolytica* and *H. somni* DNA. With these increased concentrations, we saw a higher signal of 34.49 in *M. haemolytica* (figure 13). However, this signal was only produced by one of the three genes designated MH_lpsA_at. The remaining two genes, MH_16S_at and MH_lktA_at, still did not show positive signals with the increased DNA concentration. *H. somni* hybridization responded in a similar fashion with the increase in hybridized DNA. Gene designation HS_rpoB_at produced a stronger signal of 19.37 and gene designation HS_16S_at produced a moderate signal of 5.5 creating an average of 12.44 (figure 14). Ultimately, 250 ng of DNA was below the detection limit for these probe sets.

We hybridized three different strains of *E. coli* to the array (O157 and field isolates 1 & 3). Two *E. coli* O157 (1μg)-hybridized arrays produced overall signal values of 40.28 and 42.88 with backgrounds of 1.14 and 1.11 (figure 15). Neither of the two field isolates (250 ng) showed strong average signals with values of 2.81 and 1.98 with backgrounds of 1.38 and 1.79 (figure 15). However, both isolates did hybridize to different individual probes sets. Isolate 1 hybridized strongly to 4 out of 50 *E. coli* probe sets, and isolate 3, which was considered non-pathogenic and part of the normal flora of a cow's rumen, showed some minor signal with approximately 5 out of 50 *E. coli* probe sets (data not shown). If the 45-46 non-hybridized genes are not considered, and only the positive signals are averaged, signal values of 18.96 and 5.04 for isolates 1 and 3,

respectively, with backgrounds of 1.38 and 1.79 (figure 15) are recognized. All other pathogenic signal values ranged from 1.08-2.3. Furthermore, when *E. coli* probe set performance was evaluated with 1 μg of DNA from *E.* coli field isolates 1 & 2, we saw different hybridization patterns between to the two isolates. We observed the same four genes showing strong signals with *E. coli* field isolate 1 DNA as was in previous experiments, only at higher levels with the increase in target DNA concentration, and five genes showing signal with field isolate 2. Of the nine probe sets that successfully hybridized to target DNA, only two of probes sets hybridized with both field isolate 1 and 2 (figure 16). This demonstrates the limits of the designed microarray when hybridizing different strains of the same bacterial genus.

DNA from bovine tissue (negative control) (lung, liver, spleen, and kidney) hybridized to two arrays did not show any significant binding with any of the bacterial genes (1.41-2.98 with backgrounds of 2.42 and 1.53) (figure 17).

Almost all cross hybridization occurred within the probe sets designed from the intergenic regions of *S. typhimurium* and *L. monocytogenes*. Therefore, these regions were removed from data analysis. Furthermore, cross hybridization was limited to random false positives that were not consistent in any of the repeated experiments. For example, the *S. typhimurium* gene designated STM0417 showed a medium signal on one of the *P. multocida* arrays at 6.23, but only produced signal values that ranged from 1.26-2.95 on the other seven *P. multocida* arrays. One major exception to this pattern was found in the *S. dublin* array in which gDNA cross hybridized to a significant portion of the *S. typhimurium* genes of which less than 1% had a higher signal than the highest *S.*

dublin gene (23.44) and about 40% had a higher signal than the lowest *S. dublin* gene signal considered positive (4.46).

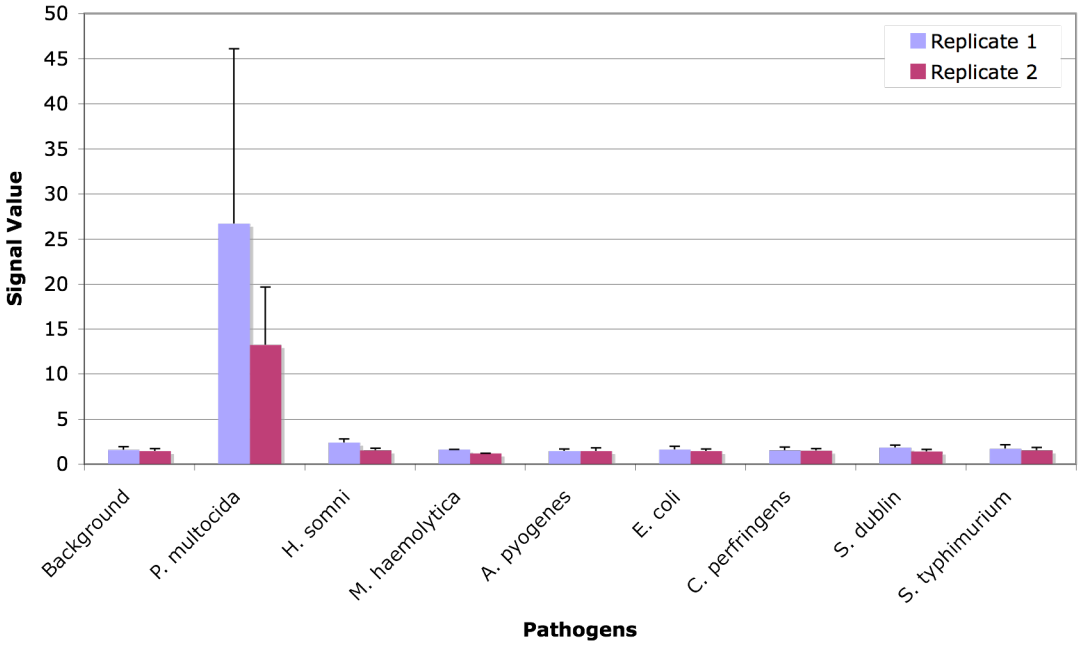

Figure 10. Detection of *P. multocida.*

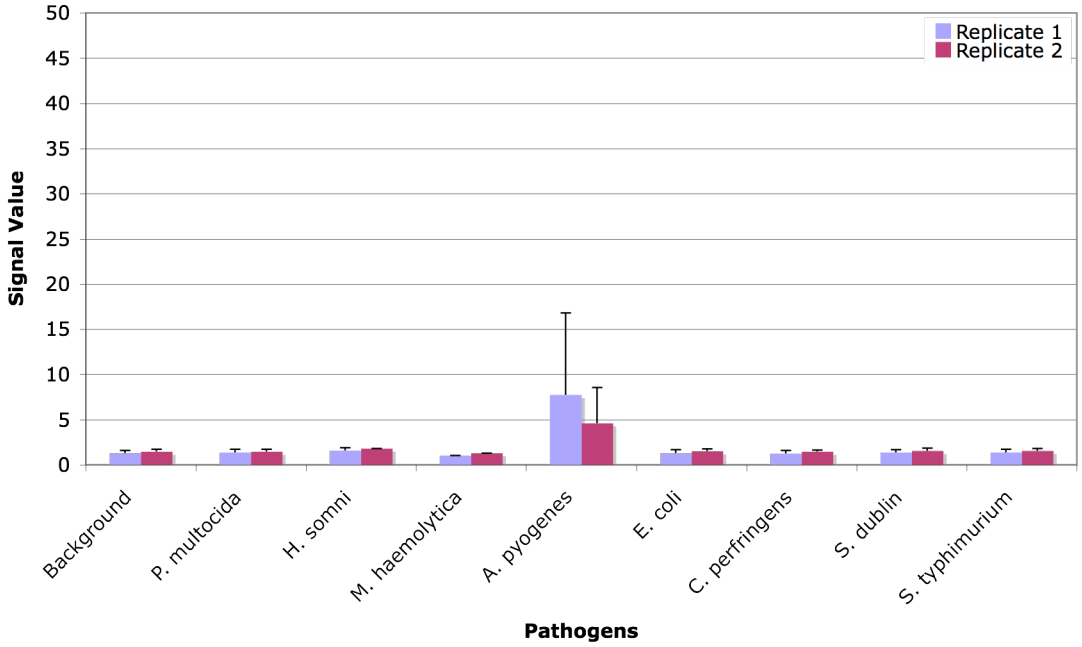

Figure 11. Detection of *A. pyogenes.*

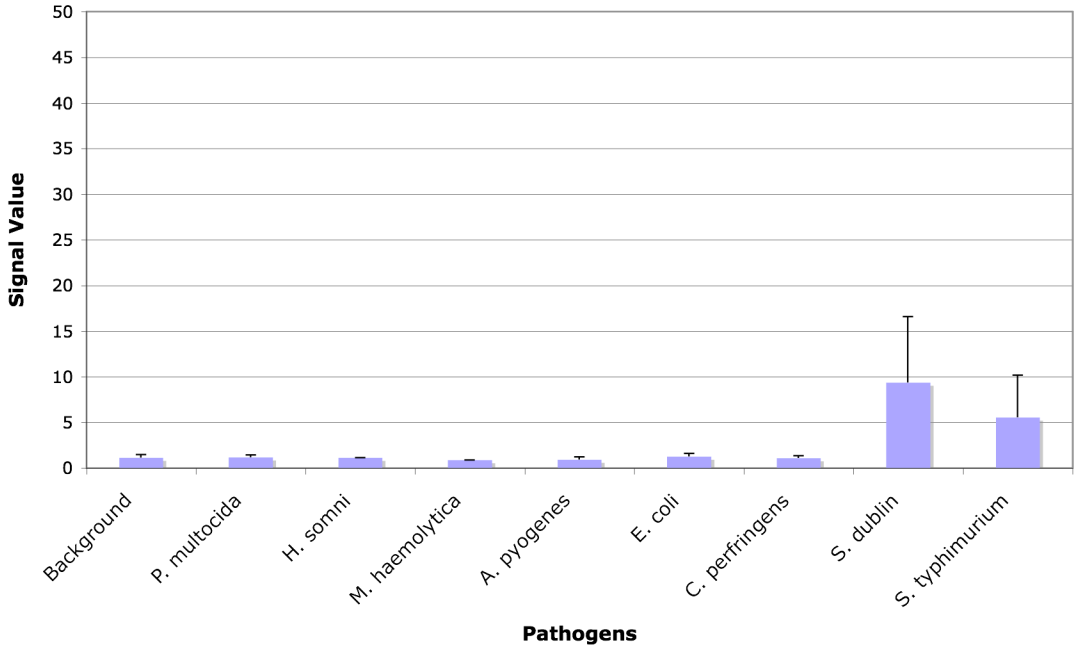

Figure 12. Detection of *S. dublin.*

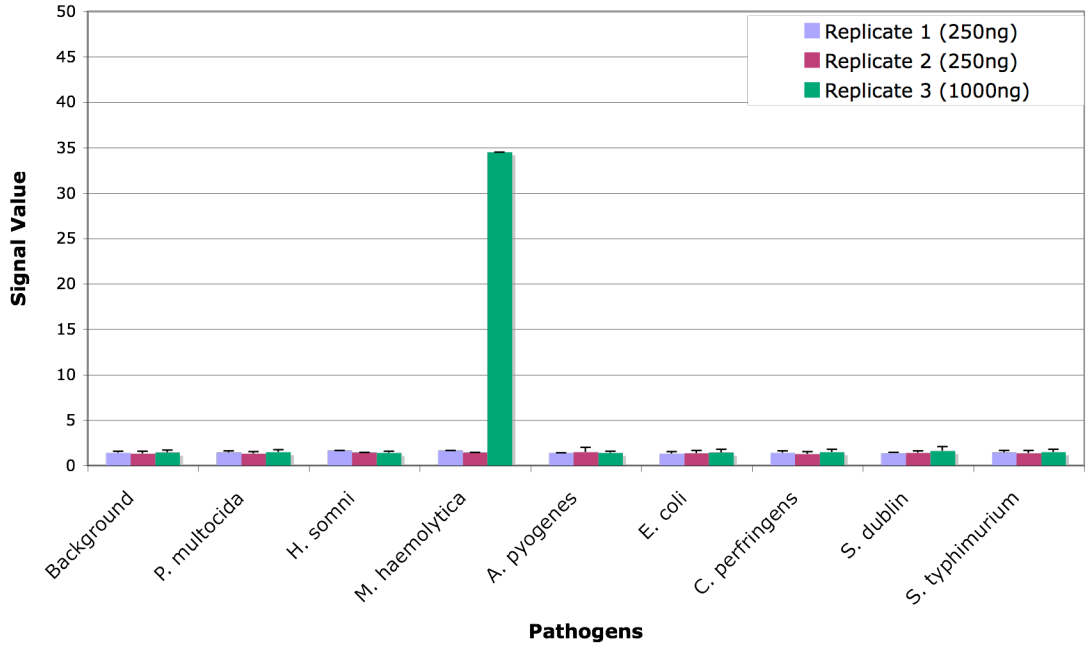

Figure 13. Detection of *M. haemolytica.*

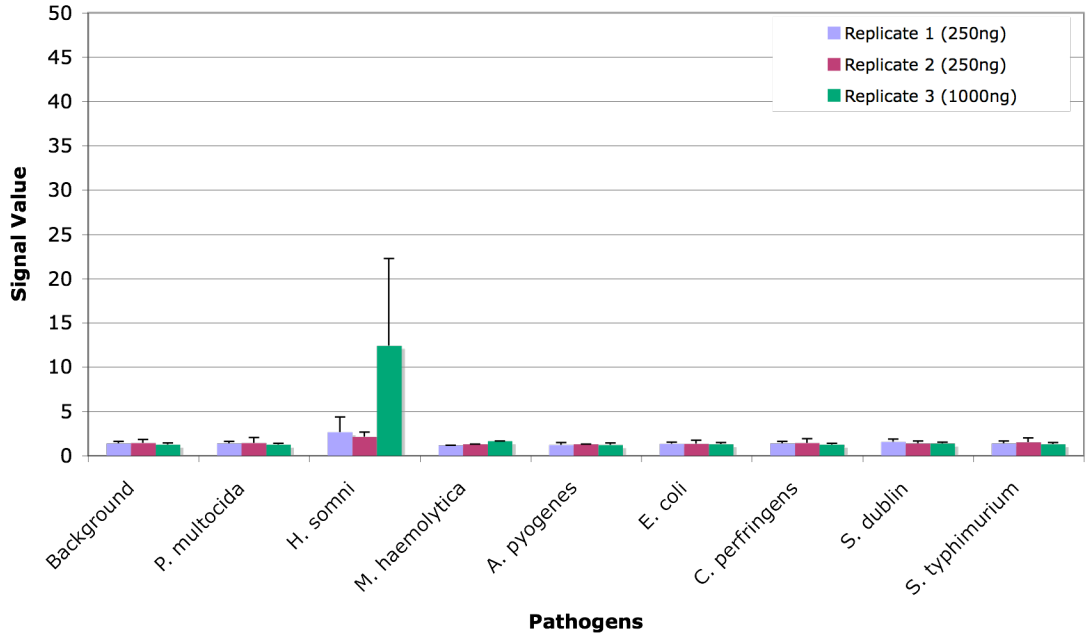

Figure 14. Detection of *H. somni.*

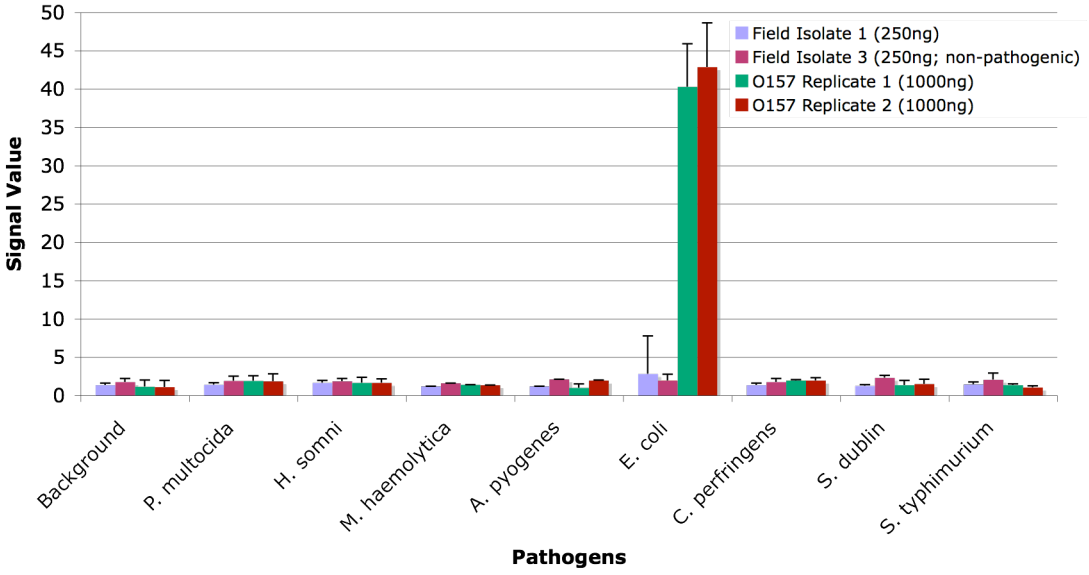

Figure 15. Detection of *E. coli*.

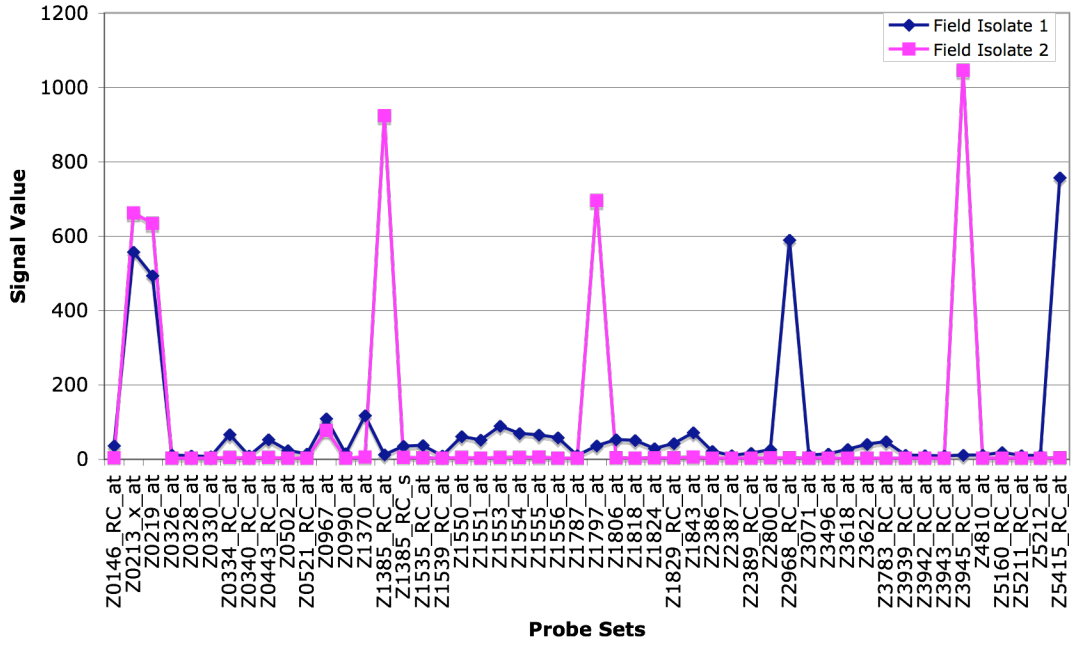

Figure 16. Hybridization Analysis of *E. coli* Probe Sets.

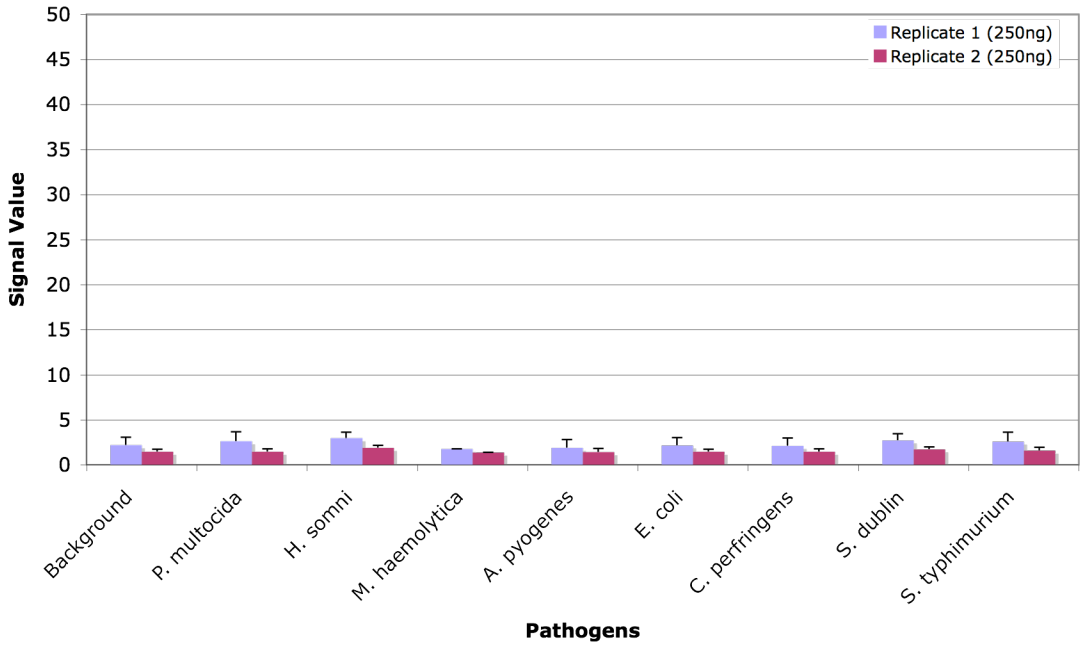

Figure 17. Detection of Bovine DNA extracted from lung, liver, kidney, and spleen.

Detection of Pathogen-Spiked Tissue

DNA extracted from lung, liver, kidney, and spleen, all of which were spiked with multiple bacterial species of confirmed identity (PCR), was hybridized individually to four STYLMONOa520430F arrays. Field isolates of each pathogen were hybridized to the arrays. The two exceptions were P. multocida (ATCC 12945), since we were unable to obtain a field isolate, and E. coli (O157), because we have already shown that different strains of E. coli do not hybridize to most of the E. coli probe sets. The signal value for each pathogen was averaged across the arrays and multiple replicates. The overall background for the arrays for comparison was 12.98. Positive signals were seen with P. multocida (2062.08), H. somni (1527.33), M. haemolytica (2274.99), A. pyogenes (152.78), S. dublin (1688.08), E. coli (2299.20) and S. typhimurium (2040.25) (figure 18, values transformed with squareroot to bring the extreme value differences closer together for easier comparison). With increased concentrations of DNA hybridized to the arrays, some of the genes that were previously not producing signals, or were producing relatively weak signals, were now showing stronger signals. We also observed signals produced by the genes designated HS_16S_at, AP_smc_plo_ftsY_at, and SD_273_at, where previously we saw no signal. However, even though these probe sets were now showing hybridizing signals, they were still lower than their counterparts, further confirming that different probe sets for the same pathogen had different detection limits.

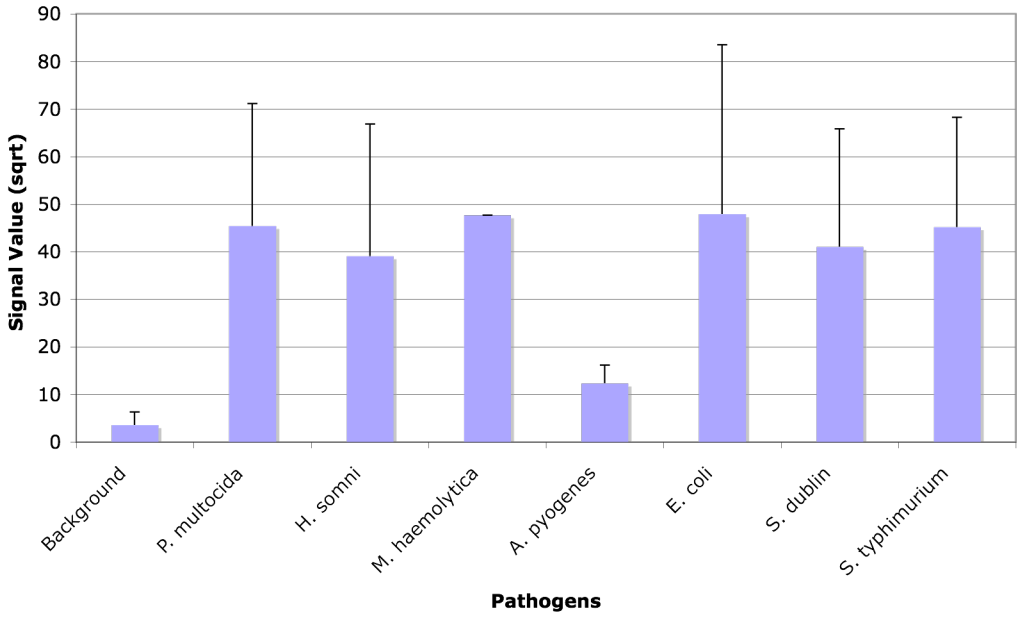

Figure 18. Detection of pathogens spiked into bovine tissue.

Discussion

Microarrays have the potential for detection of multiple bacterial species in cases of mixed infections in a single assay. Here we report on the design and evaluation of a simple DNA microarray used for the detection of pathogenic bacteria that commonly afflict cattle. The prototype array was designed on an Affymetrix platform and contained 4686 probe sets representing 7 bacterial pathogens (*P. multocida, H. somni, M. haemolytica,* and *A. pyogenes*, *E. coli, S. typhimurium,* and *S. dublin*) that commonly are isolated from bovines. All of the genes selected for probe design were either selected from BLAST comparisons or through a review of the literature. As such, each of the bacterial species has a different number of genes represented on the array, from *S. typhimurium,* represented by over 4500 genes, to *H. somni* and *A. pyogenes*, both with only two genes represented.

We have demonstrated the ability to detect *P. multocida*, *A. pyogenes*, *S. dublin*, and *S. typhimurium* with as little as 250 ng of target DNA. We failed to detect *M. haemolytica, H. somni,* and *E. coli* at 250 ng, but were able to detect them at 1 μg and higher.

We also found that different probe sets for each pathogen had their own detection limits. For example, in one of the *P. multocida* microarrays, the signal for each of the positive probe sets ranged from 75.05 to 414.05. Also, some of the probe sets, such as HS_16S_at, SD_273_at, and AP_smc_plo_ftsY_at, failed to produce signal at lower concentrations even though fellow probe sets which target the same bacterial DNA showed a strong signal. However, as increased target DNA concentrations were hybridized, these seemingly "failed" probe sets were now successful in producing a signal, though always lower than their counterparts. This data demonstrates that the probe sets performed at different limits of detection.

Despite increases in hybridized target DNA, some probe sets, notably the 10 *P. multocida* and 2 *M. haemolytica* genes, failed to ever produce a signal. PCR confirmed that most of these exact sequences were not found in the genomes of the strains used in this study. Exceptions include probe sets that targeted the 16S gene, indicating that a few of the probe sets may not have been universally representative for the species of bacterium they targeted.

We demonstrated that the probe sets were robust and able to target multiple strains by successfully hybridizing and detecting at least two different strains for most of the bacterial pathogens. Since two strains was the limit for most of the bacteria, further tests using more strains are needed to confirm this. Exceptions were *P. multocida* (only

one strain available) and *E. coli*. Since *E. coli* probe sets were designed from the O157 genome, we were not surprised when we saw strong signals when hybridized with O157 DNA. However, we were unable to see the same signal strength when we hybridized two other pathogenic *E. coli* strains. Instead, only 4-5 genes sufficiently complimented the probe sets to produce positive signals, and each strain produced a different hybridization pattern.

Finally, we were able to show that we could detect bacterial DNA in tissue spiked with mixed cultures. In experiments in which DNA, extracted from bovine lung, liver, kidney, and spleen spiked with mixed cultures of pathogenic bacteria, was hybridized to individual arrays, the arrays correctly detected and identified each bacterial pathogen. It is interesting to note that the average detection signals for each of the spiked bacterial pathogens were close at values of around 2000. However, *A. pyogenes* produced a value of about 150 which is likely due to the fact that it was the only Gram + species in the group, and may have experienced inefficient cell wall lysis. A change in the DNA extraction protocol that would be more suited towards Gram + species would confirm this.

CHAPTER IV

SUMMARY

Microarray technology has a huge potential in veterinarian diagnostic applications. Currently, there are a multitude of different diagnostic tests used to identify specific pathogens. These tests can be collectively expensive and time consuming to perform. Most of the tests require extensive handling of potentially zoonotic agents requiring the need of BSL facilities, equipment, and proper training for safety reasons. The diagnostics test designed here can detect and correctly identify the presence of pathogenic bacteria using microarray technology by extracting bacterial DNA directly from spiked tissue and hybridizing the mixed sample to our array. Though microarrays can be individually expensive, an all-in-one diagnostic test is potentially less expensive than running multiple tests that may be required for diagnosis of mixed infections.

In the designing of our microarrays, our primary goal was to design an array that could correctly identify bovine pathogens from single and multiple infections. Our secondary objective was to design an array that would not only be relatively inexpensive to current standard microarray costs, but also be comparable to the costs of current veterinarian diagnostic techniques, especially considering multiple tests would be required to diagnose a multi-organ infections.

Our first array, the in-house platform, was designed and built as a cost-effective microarray using in-house equipment and relatively cheap material that could identify four respiratory pathogens (*P. multocida, H. somni, M. haemolytica, A. pyogenes*) that commonly afflict cattle by targeting the 16S gene. The microarray array produced a specific, easy-to-see, hybridization pattern that was unique to each of the four pathogens

without the need for complicated analytical software. However, the biggest flaw in the design was the need to first PCR-amplify the target DNA (16S gene) prior to hybridization. Ultimately, with the technology available to us, we were unable to spot the probes with sufficient concentration to allow detection of the target DNA without prior amplification. One option to overcome this obstacle would be to design a set of primers that would amplify the 16S gene of all the pathogens of interest in the same reaction. However, we were more interested in showing that we could detect the pathogen DNA from an infection without the need for amplification.

The second array was designed by partnering with Affymetrix, who had the ability to create an array with a much higher probe concentration (as much as 1,000,000 probes/cm^2). This custom array, which targeted multiple genes for seven bovine pathogens (the previous four plus *E. coli, S. typhimurium, S. dublin*), was able to detect all of the pathogens by hybridizing unamplified chromosomal DNA. Some of the pathogens were detectable with as little as 250ng of DNA hybridized, while others required at least 1ug. Also, we were able to successfully detect each pathogen from a simulated multi-organ infection. The major drawback to the Affymetrix arrays is the cost associated with each experiment. Indiviually, the GeneChips are more expensive than the conventional in-house arrays. Also, specific and expensive equipment is also necessary to process GeneChip arrays. Nevertheless, despite the added cost, the Affymetrix platform is still be comparable to the cost of multiple diagnostics tests looking for multiple infections.

Both microarray designs have advantages and disadvantages when compared with each other. The Affymetrix arrays successfully detected each pathogen in simulated

single and multiple infections without first requiring PCR-amplification of the target genes. However, our in-house arrays were able to produce a visible hybridization pattern without the need for software analysis, and the arrays were much cheaper, both in constructing and processing the arrays, than their Affymetrix counterparts.

LITERATURE CITED

1 2004, Hemorrhagic Septicemia. *In*: Manual of standards for diagnostic tests and vaccines for terrestrial animals, 5th ed., World Organization for Animal Health, Paris.

2 2006, Pneumonic Pasteurellosis. *In*: The Merck veterinary manual, ed. Kahn CM, 9th ed., Merck & Co., Inc., Whitehouse Station, NJ.

3 Angen O, Ahrens P, Tegtmeier C: 1998, Development of a PCR test for identification of Haemophilus somnus in pure and mixed cultures. Vet Microbiol 63:39-48.

4 Annibali V, Di Giovanni S, Cannoni S, et al.: 2007, Gene expression profiles reveal homeostatic dynamics during interferon-beta therapy in multiple sclerosis. Autoimmunity 40:16-22.

5 Ausubel M, R. Brent, R.E. Kingston, D.D., et al.: 1988, Current protocols in molecular biology.

6 Barone A, Beecher J, Bury P, et al.: 2001, Photolithographic synthesis of high-density oligonucleotide probe arrays. Nucleosides Nucleotides & Nucleic Acids 20:525-531.

7 Baugh L, Hill A, Brown E, Hunter C: 2001, Quantitative analysis of mRNA amplification by in vitro transcription. Nucleic Acids Res 29:E29.

8 Booth S, Drebot M, Martin I, Ng L: 2003, Design of oligonucleotide arrays to detect point mutations: molecular typing of antibiotic resistant strains of Neisseria gonorrhoeae and hantavirus infected deer mice. Mol Cell Probes 17:77-84.

9 Borucki M, Krug M, Muraoka W, Call D: 2003, Discrimination among Listeria monocytogenes isolates using a mixed genome DNA microarray. Vet Microbiol 92:351-362.

10 Buechler C, Schäffler A: 2007, Does global gene expression analysis in type 2 diabetes provide an opportunity to identify highly promising drug targets? Endocr Metab Immune Disord Drug Targets 7:250-258.

11 Call D: 2005, Challenges and opportunities for pathogen detection using DNA microarrays. Crit Rev Microbiol 31:91-99.

12 Carter RGR: 1998, Hemorrhagic Septicemia. *In*: Foreign Animal Diseases, pp. 265-272. United States Animal Health Association, Richmond, VA.

13 Chen T: 2006, DNA microarrays--an armory for combating infectious diseases in the new century. Infect Disord Drug Targets 6:263-279.

14 Christensen H, Bisgaard M, Larsen J, Olsen J: 2003, PCR-detection of Hemophilus paragallinarum, Hemophilus somnus, Mannheimia (Pasteurella) hemolytica, Mannheimia spp., Pasteurella trehalosi, and Pasteurella multocida. Methods Mol Biol 216:257-274.

15 Clever L, LeGuyader Y: 1995, Infectious risks for health care workers. Annu Rev Public Health 16:141-164.

16 Cobo F, Concha A: 2007, Application of microarray technology for microbial diagnosis in stem cell cultures: a review. Cytotherapy 9:53-59.

17 Collet MG, Bath GF: 1994, Actinomyces pyogenes infections. *In*: Infectious diseases of livestock: with special reference to Southern Africa, eds. Coetzer JAW, Tustin RC, pp. 1406-1415. Oxford University Press, New York, NY.

18 Counts S, He B, Che S, et al.: 2007, Alpha7 nicotinic receptor up-regulation in cholinergic basal forebrain neurons in Alzheimer disease. Arch Neurol 64:1771-1776.

19 Daudén E: 2007, [Pharmacogenetics II. Research molecular methods, bioinformatics and ethical concerns]. Actas Dermosifiliogr 98:3-13.

20 de Reynies A, Geromin D, Cayuela J, et al.: 2006, Comparison of the latest commercial short and long oligonucleotide microarray technologies. BMC Genomics 7:51.

21 Ducray F, Honnorat J, Lachuer J: 2007, [DNA microarray technology: principles and applications to the study of neurological disorders]. Rev Neurol (Paris) 163:409-420.

22 Egervari K, Szollosi Z, Nemes Z: 2007, Tissue microarray technology in breast cancer HER2 diagnostics. Pathol Res Pract 203:169-177.

23 Epstein C, Butow R: 2000, Microarray technology - enhanced versatility, persistent challenge. Curr Opin Biotechnol 11:36-41.

24 Fadiel A, Naftolin F: 2003, Microarray applications and challenges: a vast array of possibilities. Int Arch Biosci, pp. 1111-1121.

25 Feilotter H: 2004, Microarrays in veterinary diagnostics. Anim Health Res Rev 5:249-255.

26 Fisher M, Weiser G, Hunter D, Ward A: 1999, Use of a polymerase chain reaction method to detect the leukotoxin gene lktA in biogroup and biovariant isolates of Pasteurella haemolytica and P trehalosi. Am J Vet Res 60:1402-1406.

27 Freeman H, Noble M: 2005, Lack of evidence for Mycobacterium avium subspecies paratuberculosis in Crohn's disease. Inflamm Bowel Dis 11:782-783.

28 Ganesan B, Chen D, Black R, et al.: 2007, Comparing microarray substrates for optimal hybridization of genome-wide targets to oligomer probes. unpublished.

29 Gold V: 2006, The Gold Book: Compendium of chemical terminology. 2nd ed., International Union of Pure and Applied Chemistry., Research Triangle Park, NC.

30 Green A, DuTeau N, Miller M, et al.: 1999, Polymerase chain reaction techniques for differentiating cytotoxic and noncytotoxic Pasteurella trehalosi from Rocky Mountain bighorn sheep. Am J Vet Res 60:583-588.

31 Haferlach T, Bacher U, Kohlmann A, et al.: 2006, [Microarrays and gene expression profiling for the diagnosis in leukemia. From research studies to routine application]. Med Klin (Munich) 101:908-914.

32 Heller MJ: 2002, DNA microarray technology: Devices, systems, and applications. Annu Rev Biomed Eng 4:129-153.

33 Hessner M, Meyer L, Tackes J, et al.: 2004, Immobilized probe and glass surface chemistry as variables in microarray fabrication. BMC Genomics 5:53.

34 Hessner M, Xiang B, Jia S, et al.: 2006, Three-color cDNA microarrays with prehybridization quality control yield gene expression data comparable to that of commercial platforms. Physiol Genomics 25:166-178.

35 Hoerndli F, Pelech S, Papassotiropoulos A, Götz J: 2007, Abeta treatment and P301L tau expression in an Alzheimer's disease tissue culture model act synergistically to promote aberrant cell cycle re-entry. Eur J Neurosci 26:60-72.

36 Hoheisel J, Ross M, Zehetner G, Lehrach H: 1994, Relational genome analysis using reference libraries and hybridisation fingerprinting. J Biotechnol 35:121-134.

37 Hood D, Deadman M, Cox A, et al.: 2004, Three genes, lgtF, lic2C and lpsA, have a primary role in determining the pattern of oligosaccharide extension from the inner core of Haemophilus influenzae LPS. Microbiology 150:2089-2097.

38 Hughes T, Mao M, Jones A, et al.: 2001, Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nature Biotechnology 19:342-347.

39 Ishii M, Hashimoto S, Tsutsumi S, et al.: 2000, Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. Genomics 68:136-143.

40 Jayapal M, Melendez AJ: 2006, DNA microarray technology for target identification and validation. Clinical and Experimental Pharmacology and Physiology 33:496-503.

41 Jordan B: 2007, DNA microarrays in the clinic: how soon, how extensively? Bioessays 29:699-705.

42 Jost B, Songer J, Billington S: 2002, Identification of a second Arcanobacterium pyogenes neuraminidase and involvement of neuraminidase activity in host cell adhesion. Infect Immun 70:1106-1112.

43 Kafatos F, Jones C, Efstratiadis A: 1979, Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res 7:1541-1552.

44 Kitching JP, Bishop GC: 1994, The Haemophilus somnus disease complex in cattle. *In*: Infectious diseases of livestock: with special reference to Southern Africa, eds. Coetzer JAW, Tustin RC, pp. 1135-1142. Oxford University Press, New York, NY.

45 Korkola J, Estep A, Pejavar S, et al.: 2003, Optimizing stringency for expression microarrays. Biotechniques 35:828-835.

46 Kuo W, Jenssen T, Butte A, et al.: 2002, Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18:405-412.

47 Lennon G: 2000, High-throughput gene expression analysis for drug discovery. Drug Discov Today 5:59-66.

48 Lequin R: 2005, Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clin Chem 51:2415-2418.

49 Liu D, Lawrence M, Austin F: 2004, Specific PCR identification of Pasteurella multocida based on putative transcriptional regulator genes. J Microbiol Methods 58:263-267.

50 Lockhart D, Dong H, Byrne M, et al.: 1996, Expression monitoring by hybridization to high-density oligonucleotide arrays. Nature Biotechnology 14:1675-1680.

51 Loy A, Lehner A, Lee N, et al.: 2002, Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in

the environment. Appl Environ Microbiol 68:5064-5081.

52 Loy A, Schulz C, Lücker S, et al.: 2005, 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order "Rhodocyclales". Appl Environ Microbiol 71:1373-1386.

53 Luo L, Salunga R, Guo H, et al.: 1999, Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat Med 5:117-122.

54 Mahadevappa M, Warrington J: 1999, A high-density probe array sample preparation method using 10- to 100-fold fewer cells. Nat Biotechnol 17:1134-1136.

55 Meyers-Wallen V: 2001, Relavence of the canine genome project to veterinarian medical practice. *In*: Recent advances in small animal reproduction, eds. Concannon PW, England GCW, Verstegen J, Linde-Forsberg C, International Veterinary Information Service, Ithaca, NY.

56 Miflin J, Blackall P: 2001, Development of a 23S rRNA-based PCR assay for the identification of Pasteurella multocida. Lett Appl Microbiol 33:216-221.

57 Mollet C, Drancourt M, Raoult D: 1997, rpoB sequence analysis as a novel basis for bacterial identification. Mol Microbiol 26:1005-1011.

58 Nguyen D, Arpat A, Wang N, Carroll R: 2002, DNA microarray experiments: biological and technological aspects. Biometrics 58:701-717.

59 Odenthal M, Koenig S, Farbrother P, et al.: 2007, Detection of opportunistic infections by low-density microarrays: a diagnostic approach for granulomatous lymphadenitis. Diagn Mol Pathol 16:18-26.

60 Ojha S, Kostrzynska M: 2008, Examination of animal and zoonotic pathogens using microarrays. Vet Res 39:4.

61 Oberst Rd, Hays MP, Bohra LK, et al.: 2003, Detection of Escherichia coli O157:H7 in cattle feces using a polymerase chain reaction-based fluorogenic 5' nuclease (TaqMan) detection assay after secondary enrichment. J Vet Diagn Invest 15:543-52.

62 Pabon C, Modrusan Z, Ruvolo M, et al.: 2001, Optimized T7 amplification system for microarray analysis. Biotechniques 31:874-879.

63 Panchuk-Voloshina N, Haugland R, Bishop-Stewart J, et al.: 1999, Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. J Histochem Cytochem 47:1179-1188.

64 Rogojina A, Orr W, Song B, Geisert E: 2003, Comparing the use of Affymetrix to

spotted oligonucleotide microarrays using two retinal pigment epithelium cell lines. Molecular Vision 9:482-496.

65 Ryan K, Ray C: 2004, Sherris medical microbiology, 4th ed., eds. Ryan KJ, Ray CG, pp. 247-249. McGraw-Hill, New York, NY.

66 Schena M, Shalon D, Davis R, Brown P: 1995, Quantitative Monitoring of Gene-Expression Patterns with a Complementary-DNA Microarray. Science 270:467-470.

67 Song S, Li B, Wang L, et al.: 2007, A cancer protein microarray platform using antibody fragments and its clinical applications. Mol Biosyst 3:151-158.

68 Southern E: 2001, DNA microarrays. History and overview. Methods Mol Biol 170:1-15.

69 Tan P, Downey T, Spitznagel E, et al.: 2003, Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Research 31:5676-5684.

70 Townsend K, Frost A, Lee C, et al.: 1998, Development of PCR assays for species- and type-specific identification of Pasteurella multocida isolates. J Clin Microbiol 36:1096-1100.

71 Venter BJ, Myburgh JG, van der Walt ML: 1994, Bovine Salmonellosis. *In*: Infectious diseases of livestock: with special reference to Southern Africa, eds. Coetzer JAW, Tustin RC, pp. 1104-1112. Oxford University Press, New York, NY.

72 Vernet G: 2007, Use of molecular assays for the diagnosis of influenza. Expert Rev Anti Infect Ther 5:89-104.

73 Volokhov D, Rasooly A, Chumakov K, Chizhikov V: 2002, Identification of Listeria species by microarray-based assay. J Clin Microbiol 40:4720-4728.

74 Wang E, Miller L, Ohnmacht G, et al.: 2000, High-fidelity mRNA amplification for gene profiling. Nat Biotechnol 18:457-459.

75 Weimer, B, Desai P: 2008, personal communication, unpublished data.

76 Wills Q: 2007, SimuGen Ltd: Reliable, early prediction of drug toxicity with toxicogenomics, human cell culture and computational models. Pharmacogenomics 8:1081-1084.

77 Woo Y, Affourtit J, Daigle S, et al.: 2004, A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. J Biomol Tech

15:276-284.

78 Xie Y, Chou L, Cutler A, Weimer B: 2004, DNA Macroarray profiling of Lactococcus lactis subsp. lactis IL1403 gene expression during environmental stresses. Appl Environ Microbiol 70:6738-6747.

79 Yuen T, Wurmbach E, Pfeffer R, et al.: 2002, Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. Nucleic Acids Research 30(10):e48.

APPENDIXES

APPENDIX A - Genes represented on the STYLMONOa520430F GeneChip

Probe Set Designation	**Gene Function**
Controls	
AFFX-BioB-3_at	Affymetrix Hybridization Control
AFFX-BioB-5_at	Affymetrix Hybridization Control
AFFX-BioB-M_at	Affymetrix Hybridization Control
AFFX-BioC-3_at	Affymetrix Hybridization Control
AFFX-BioC-5_at	Affymetrix Hybridization Control
AFFX-BioDn-3_at	Affymetrix Hybridization Control
AFFX-BioDn-5_at	Affymetrix Hybridization Control
AFFX-CreX-3_at	Affymetrix Hybridization Control
AFFX-CreX-5_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioB-3_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioB-5_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioB-M_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioC-3_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioC-5_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioD-3_at	Affymetrix Hybridization Control
AFFX-r2-Ec-bioD-5_at	Affymetrix Hybridization Control
AFFX-r2-P1-cre-3_at	Affymetrix Hybridization Control
AFFX-r2-P1-cre-5_at	Affymetrix Hybridization Control
A. pyogenes	
AP_16S_at	16S ribosomal RNA
AP_smc_plo_ftsY_at	chromosome segregation protein (smc) gene, hypothetical protein, pyolysin (plo), signal recognition particle receptor (ftsY)
H. somni	
HS_16S_at	16S ribosomal RNA
HS_rpoB_at	rpoB gene
M. haemolytica	
MH_16S_at	16S ribosomal RNA
MH_lktA_at	lktA leukotoxin
MH_lpsA_at	glycosyltransferase LpsA
P. multocida	
PM_r01_at	16S ribosomal RNA
PM_r02_at	23S ribosomal RNA
All other probe sets	hypothetical proteins

Probe Set Designation	**Gene Function**
S. dublin	
All probe sets	hypothetical proteins
E. coli	
Z0146_RC_at	yadC putative fimbrial protein
Z0213_x_at	rrsH 16S ribosomal RNA
Z0219_at	rrlH 23S ribosomal RNA
Z0334_RC_at	sidI putative capsid morphogenesis protein encoded in CP-933I
Z0967_at	putative protease encoded in prophage CP-933K
Z1370_at	putative tail component encoded by cryptic prophage CP-933M
Z1385_RC_at	putative secreted protein encoded by cryptic prophage CP-933M
Z1797_at	putative antirepressor of prophage CP-933N
Z1818_at	putative antirepressor protein encoded by prophage CP-933N
Z3622_at	putative resolvase
Z3783_RC_at	putative dimethyl sulfoxide reductase subunit C
All other probe sets	hypothetical proteins
S. typhimurium	
All genes from chromosome and plasmid are represented	

APPENDIX B - R Code Used to Analyze Affymetix Arrays

```
library(affy) # For affy files
#memory.limit(size=4000) # Increase memory use

# Load in library files
library(makecdfenv)

#Choose Mac or Windows, not both.

#Mac
STYLMONOa520430F = make.cdf.env("STYLMONOa520430F.cdf")

#Windows
make.cdf.package("STYLMONOa520430F.cdf", species="Bovine_Pathogens")

# Read in .cel files from laptop
data = ReadAffy(celfile.path="//Users//Ryan//Desktop//RB1-30")

# RMA normalization and avg PM values for each chip
eset = expresso(data[1,30], bgcorrect.method = "rma", normalize.method = "constant",
        pmcorrect.method = "pmonly", summary.method = "avgdiff")

# write the eset file to an excel spreadsheet
library(marray)

write.xls(eset, file = "Data.xls")

# Organize the data; create multiple files; 1 with positive probes; 1 with all background
        noise; save them as .csv files

# Write subset probes to object
PM.sig.values = read.csv("C:\\Documents and Settings\\Ryan\\Desktop\\Data.pm.csv")
bg.sig.values = read.csv("C:\\Documents and Settings\\Ryan\\Desktop\\Data.bg.csv")
```