Workshop on the Feasibility of Using DNA/RNA Microarrays and Related Technologies for High Through-Put Detection of Waterborne Pathogens

March 22-23, 2005
U.S. Environmental Protection Agency
Andrew W. Breidenbach Research Center
26 W. Martin Luther King Dr.
Cincinnati, OH 45268

Executive Summary

The U.S. Environmental Protection Agency (USEPA) held an expert workshop in Cincinnati, Ohio on March 22-23, 2005 to explore the feasibility of using microarray technology for detecting waterborne pathogens. Based on the National Research Council (NRC 2001) and subsequent National Drinking Water Advisory Council (NDWAC 2004) recommendations to identify waterborne pathogens in new ways such as using virulence factor activity relationships (VFAR) and genomics, this was the second workshop organized by EPA to understand the feasibility of using such approaches. The workshop announcement, prospectus, and agenda were provided to all participants. Invited speakers were asked to present their research on microbial pathogen detection by microarrays and related technologies. Following the presentations, a panel made up of the speakers from the conference answered questions posed by those in attendance.

The goals of the workshop were to:

- Evaluate the feasability of using DNA/RNA microarrays for monitoring the presence of pathogens in water.
- Evaluate the feasability of using DNA/RNA microarrays to predict/identify emerging pathogens related to a virulence factor activity relationship (VFAR) approach.
- Identify areas of research that might lead to development of microarray technology and technologies associated with it for detection of waterborne pathogens (e.g., presence, viability, and infectivity).
- Determine when, and to what extent, the technology might be ready for use.
- Identify new technologies that have similar capabilities to microarrays, specifically the simultaneous detection of multiple pathogens.

Presentations and Speakers

Twelve scientists from academia, government, and the biotechnology industry addressed these goals in twelve presentations, and answered questions posed by those in attendance. The majority of the scientists were nationally recognized in their field, and/or were senior scientists from centers that are well recognized for their research activities. About 80 people registered for the workshop.

Topics covered

- Microarray technology: One talk was on the technology itself where the speaker explained the different types and formats of arrays that are currently available, the ways they are used, and their advantages and disadvantages. Of these, bead arrays offer higher

throughput and less complicated imaging. Many targets are possible, including the use of multiple genes and whole genomes.

- Clinical applications: Two talks were on the development of microarray chips to detect pathogenic bacteria and viruses from food and clinical samples such as nasal lavage. Of these, the development of a Viruschip used for clinical diagnostic was novel and is discussed later. Sensitivity and specificity of detection from clinical samples, where the concentration of pathogens is high, is very good.
- Environmental applications: Four talks were devoted to detection of pathogens in water and microbial communities in soil. Besides water, a couple of speakers presented detection results from other matrices including food. Sensitivity was poor and is discussed in more detail in the next bullet.
- Detection sensitivity and amplification strategies and challenges: One talk was devoted to amplification strategies, although several speakers addressed the challenges of detection sensitivity and specificity, especially from water samples. Because very low levels of pathogens are present in water, an initial amplification is always required to achieve a concentration of 10^5-10^7 copies of gene that are required for subsequent microarray detection. Several amplification procedures were discussed including: specific amplification by Polymerase Chain Reaction (PCR) with specific primers and Nucleic Acid Sequence Based Amplification (NASBA), random amplification by PCR with random primers, Klenow fragment amplification, φ29 polymerase. PCR was the most sensitive amplification procedure but suffered from copy number bias, template and primer bias, loss of primer efficiency in multiplex reactions, and amplicon length bias. The use of 16S genes for PCR amplification is hampered by the formation of secondary and tertiary structures during amplification. Culture amplification can be as specific as PCR, has no attendant PCR bias, and measures only viable organisms. While culture does not suffer from PCR bias, different organisms may grow at different rates and faster growing organisms may overgrow slower ones. Non-culturable organisms and hard to culture organisms will not be represented. Detection of organisms with microarray is very sensitive (100%) and specific (one organism for *C. parvum,* 3-5 colony forming unit (CFU) for bacteria) from pure culture and clinical samples where background interference is much less. However, work from three groups demonstrated that detection limit and specificity of detection of pathogens from water samples, is poor and could be as high as 10^3 organisms. This is because organisms are present in low numbers and high volumes of water need to be concentrated (L to μL). Thus primers and probes that give sensitive signals with pure culture may not work with a field sample, giving rise to false positives, and matrix interference is a major challenge. Furthermore, primers and probes may give ambiguous results when used in a different matrix. Trouble shooting will include choice of well validated probes and these need to be validated in the field. The use of multiple probes for the same gene, the simultaneous measurement of all chips, and increasing hybridization temperature on chip to see which signals disappear, may all help to eliminate false positives.
- Measurement of viability: One talk covered the different ways viability of the pathogens could be measured using molecular methods. For viruses, pre-treatment of the sample

with proteases followed by Reverse Transciptase- PCR (RT-PCR) could indicate whether the sample contains live viruses.

- New technology and novel use of microarray: Three talks were on the development of new technology for different aspects of waterborne pathogen detection. The liposome-biosensor technology is able to detect 10 live *Cryptosporidium parvum* oocysts from environmental water samples. MME-PCR in which primers are fixed in polyacrylamide, allows the physical separation of the different primer sets and thus solves most of multiplex -PCR problems. As many as 100 primer sets can be used in Microarray-Enhanced Multiplexed-PCR(MME-PCR). Looking for host response to pathogen infection using a microarray chip (e.g., chip containing all mouse genes), rather than the pathogen itself, can be an alternative way to test the quality of water.

Novel Research

Some promising ongoing research brought out in this workshop includes:

- Development of the 'Viruschip', which has the potential to identify viruses without preconceptions, as well as predict the presence of previously unknown viruses. The 'Viruschip' was able to classify the then unknown SARS virus as Coronavirus. Although this clinical application still needs some work, it could assist in epidemiological studies, show waterborne transmissions, and aid in the surveillance of pathogens (emerging, zoonotic, or otherwise).
- Microarray-Enhanced Multiplexed-PCR**:** may provide a high order of multiplexed amplifications without bias, which is necessary for subsequent microarray detection. Liposome-biosensor technology, combined with NASBA, has the potential to detect multiple live pathogens in a shorter time than PCR but with equal sensitivity.
- Affymetrix resequencing microarrays that could rapidly identify pathogens as well be able to speciate them.

Conclusions
Microarrays may be ready for simultaneous detection of a small number of pathogens in the next five years, but the prediction of pathogens using microarrays remains a long term goal. Limited validation is possible when a small number of probes for pathogens are present. The cost of a microarray system can vary, and although currently is high can come down significantly once used routinely with multiple samples. Some estimates that were given at the workshop were $60-$150 per PCR run, $15 per array and a set-up cost of $500-$1,000. The cost will also depend on the number of samples tested. The more samples tested together, the lower the cost. Viability assays are a possibility in the near future using mRNA in bacteria and prokaryotes, and protease treatment followed by RT-PCR or PCR in viruses but the methods will have to be validated. The hurdles of sample preparation and amplification by multiplex reactions, or by random means, need to be overcome before microarrays can be used in a realistic way. Thus future research should be directed there.

Workshop Committee
EPA
Keya Sen, OW, TSC
Jim Sinclair, OW, TSC
Paul Berger, OW, SRRB
Sandhya Parshionikar, OW, TSC
Robin Oshiro, OW, OST
Jafrul Hasan, OW, OST
Tom Carpenter, OW, TAB
Sam Hayes, ORD, NERL
Shay Fout, ORD, NERL
Phyllis Branson, OW, TSC
The Cadmus Group
George Hallberg
Susan Bjork
Jim Smith
Nelson Moyer

SUMMARY

Note: These notes were transcribed as the speakers were presenting. Speakers have reviewed their respective sections.

DAY 1 - March 22, 2005

Conference Welcome

Eric Bissonette, Deputy Director of Technical Support Center (TSC),OW, EPA welcomed the speakers and audience to the EPA facilities in Cincinnati.

Jerry Stelma, acting Director of the Microbiological and Chemical Exposure Assessment Research Division, National Exposure Research Laboratory, reviewed the history of National Research Council (NRC) suggestions to the Office of Research and Development (ORD) regarding the use of molecular methods in the monitoring of water for pathogens. NRC challenged ORD to look at pathogens in new ways: microarrays and VFAR. This was the 2nd workshop dealing with these suggestions. The first, a VFAR workshop, was in October. The VFAR approach was found to be too limited for the

purpose of water monitoring. In this conference it was suggested that VFAR be used as a screening tool only, rather than a sole method for testing water for pathogens. VFARs, as a means of predicting virulence of emerging pathogens, were found to be at least 5 years away, and interim activity is focused upon development of a database of virulence factors and pilot studies using pathogens from the Contaminant Candidate List (CCL). He spoke about the emerging concern for biosecurity and the development of pilot studies to evaluate vulnerability. Emphasis was placed on needs for improved methods for concentration of viruses and parasites for use in VFAR or microarray. ORD wants the workshop to recommend research needed and an agenda relative to development and introduction of microarrays for various purposes.

Background and Purpose: Pathogen Monitoring, Microarrays and their Applicability to EPA's Drinking Water Programs
James L. Sinclair, TSC, USEPA

Jim Sinclair reviewed the history of EPA'S interest in microarrays, validation requirements for use of a microarray system for water monitoring, and goals for the workshop. Microarrays are used in many disciplines including gene expression studies. It has been suggested that EPA could use microarray for water monitoring. EPA hopes microarrays will overcome limitations of culture for non-culturable pathogens. Microarray allows for the detection of multiple organisms simultaneously, as well as those that are currently non-culturable, and is more rapid than culture techniques. Major issues facing microarray include the very low concentrations of microorganisms in environmental water, the need to validate any microarray system devised, and the fact that molecular methods can detect both viable and non-viable microorganisms. The present EPA regulatory approach has been to use indicator bacteria to monitor treatment efficacy and to promulgate treatment rules (e.g., watershed protection, sanitary survey, CT values, etc.) to provide a margin of safety for drinking water. Additional monitoring for pathogens has been implemented for *Giardia* and *Cryptosporidium*, and occurrence data have been gathered for *Aeromonas*. Monitoring for other pathogens is likely as methods become available. The CCL history was reviewed, beginning with the first CCL that listed 19 pathogens based upon expert judgement, a process that was widely criticized as non-transparent and subjective. EPA requested an NRC review that resulted in recommendation for a CCL classification process that included VFARs for predicting pathogen virulence, and the use of microarrays to detect and identify pathogens in the environment (occurrence), and screen organisms for their ability to cause human disease (virulence). Routine indicator monitoring will continue until microarrays have been shown to have sufficient sensitivity and specificity to detect and predict the virulence of environmental microorganisms. Microarrays can be used in regulatory application only after they have been adequately validated through performance testing and widespread acceptance among laboratories (use history). Finally, molecular methods must determine viability (and perhaps infectivity) before they can be used in regulatory determinations. EPA has not yet used any molecular method for risk determination. The goals of this workshop are to characterize the feasibility of using microarrays, to determine their predictive value, and to compare them with other emerging technologies that may be applied to environmental monitoring and drinking water safety. The workshop provides a

platform for investigator collaboration where the validity of various competing systems can be evaluated, and competing technologies can be matched with their appropriate niche in the continuum from population screening to risk determination.

Nucleic Acid Based Methods for Pathogen Detection in Water: What EPA is Doing
Keya Sen, TSC, USEPA

Keya Sen gave an overview of available nucleic acid methods, to discuss methods and their problems, and to review how EPA is using and plans to use molecular methods in the future. Molecular methods offer an advantage when working with hard to culture pathogens, (e.g. *M. avium*), pathogens for which there are no available culture methods, (e.g. human caliciviruses), and pathogens that enter into non-culturable but viable states under condition of environmental stress. Methods could be either DNA-based methods such as PCR, nested PCR, multiplex PCR, restriction fragment length polymorphisms (RFLP), denaturing gradient gel electrophoresis (DGGE), and pulsed field-gel electrophoresis (PFGE), or RNA-based methods such as reverse transcriptase PCR (RT-PCR), *in situ* hybridization methods such as fluorescent *in situ* hybridization (FISH) and *in situ* hybridization (ISH), and hybrid methods such as integrated cell culture PCR (ICC-PCR). The RNA based methods have proven useful for detection of RNAviruses such as caliciviruses, and in selective detection of viable organisms. Hybrid methods have been used to enrich viral samples for PCR and besides providing information on the viability of the organism, are more rapid than the strictly culture-based methods. Sample sizes range from 0.1 to 1000 liters depending on the pathogen to be detected, and various concentration methods are used (e.g. filtration, centrifugation, immunomagnetic bead capture, etc).

EPA uses several Nucleic acid based methods. One of the *Aeromonas* identification methods is based on a RFLP and a multiplex PCR based method is being developed. A method for virulence determination based on PCR for detection of selected *Aeromonas* virulence genes as well as, a microarray assay based on the detection of host gene expression, is being developed. *Helicobacter pylori* assays include real-time PCR for detection of low numbers of organisms in drinking water. *Mycobacterium avium* assays include real-time PCR and amplified fragment-length polymorphism (AFLP) for differentiation of strains. *Cyanobacteria* and related toxin assays are being developed using gene microarray.

Other projects include characterization of bacteria in distribution systems using rRNA to evaluate response to disinfection and phylogenic analysis based upon 16S rDNA using PCR and sequencing. PCR is also being used to detect nitrifying bacteria in water. EPA is developing databases based upon 16S rDNA profiles for fecal source tracking, studying the environmental genomics of fecal pollution and developing quantitative PCR methods for individual bacterial pathogens. EPA is using methods such as RT-PCR, ICC-PCR and microarrays for detection of enteric viruses. Efforts are underway to incorporate internal controls in RT-PCR assays. For waterborne protozoa, PCR, quantitative PCR, and FISH

methods are used for *Cryptosporidium* genotyping, characterization of microsporidia, and detection of *Cyclospora*.

EPA has not adopted molecular methods because several gaps in these methods need to be filled before they can be used for risk assessment purposes. The detection limits as well as the recovery from actual field samples need to be determined for these methods. A way to determine the viability of the microbes detected is of interest and finally, errors associated with these highly sensitive methods need to be controlled. EPA has published a PCR quality control manual in an effort to reduce or eliminate PCR errors and improve standardization between methods (www.epa.gov/nerlcwww/qa_qc_pcr10_04.pdf). Multiplex PCR is limited to the simultaneous detection of 5-8 pathogens so high throughput methods such as microarray are needed. In conclusion, molecular methods may be the only methods available for some pathogens, and EPA has begun conducting research in nucleic acid based methods. Microarrays offer the advantage of simultaneous detection and amplification of multiple targets and thus may be a powerful tool of the future.

Gary Vora: Doesn't ICC-PCR have the disadvantage of skewing virus population results depending upon the cell line chosen?
Keya Sen: Yes, the results are cell line dependent, we see different results from PCR alone than we do from ICC-PCR due to such biases.

Shay Fout: The bias is not characterized. We need to better understand these biases.

Participant: Which monitoring method will be adopted first?
Keya Sen: We do not know.

Jorge Santo-Domingo: Advancement in methods needs the detection of the entire microbial community in the sample rather than the individual organism.

A Primer on Microarrays and the Prospects and Challenges for Pathogen Detection
Doug Call, Department of Veterinary Microbiology, Washington State University

Doug Call explained that microarrays are like multiple reverse dot blots, and he made the comment that the terminology for target and probe are often interchanged in the literature. During his presentation 'probe' refers to the immobilized nucleic acid on the slide, while 'target' refers to the labeled sample nucleic acid to be interrogated on the microarray. His group frequently uses Teflon coated well slides to produce glass slide microarrays that can accommodate up to 12 microarrays per slide. He presented information on design strategies and cost of arrays. Arrays can provide presence/absence data, subtyping information, and perhaps enumeration, but they are not sufficiently developed and robust to provide all types of data alone. It is questionable that they will be able to determine viability and infectivity. The direct detection limit is currently 6-8 logs of a pathogen, so amplification is required for detection of pathogens in low numbers in

large volume samples. Sample concentration problems and PCR inhibitors are major upstream limitations. In tests to detect *Vibrio* spp. using a multiplex PCR assay, the detection limit was 3 logs without enrichment and 1 cell with 5 hour culture enrichment. For a 16S rRNA-based microarray, he reported PCR contamination problems using universal PCR primers when he exceeded 30 cycles. Reducing the number of cycles to 28 avoided nonspecific amplification in no-template-controls (including no amplification of contaminating DNA present in the Taq enzyme solution), but this severely limited analytic sensitivity. Many challenges need to be overcome. He described a "Polymorphism paradox" whereby we purposely target highly variable regions of a gene to permit discrimination, but find that not all variants within a species may be represented with the array. 16S rDNA sequences a common target for bacterial detection, are also prone to the formation of secondary and tertiary structures which may interfere with the binding of target DNA to the microarray probe. Nick translation can help to bypass the problem of the formation of primary and secondary structures because smaller target sequences are produced and these shorter sequences are less prone to the formation of higher order structure. Some other problems highlighted with the use of PCR for amplification, were copy number bias, PCR template bias and product length bias. Sensitivity and specificity are determined to some extent by the length of the probe, where long probes give better sensitivity at the expense of specificity, and short probes have high specificity but lack sensitivity. Commercial microarrays are expensive, as are labeled nucleotides, and nick translation if it is used. For VFAR applications, planar microarrays are probably too expensive, and bead arrays with flow cytometry may be more cost effective. Bead arrays use beads covered with nucleic acid probes rather than the planar slides that are used in most microarrays. Each bead is color registered and hybridization can be detected using flow cytometry.

Questions:

Roland Brousseau: How long were your short probes?
Doug Call: 25 base pairs

Keya Sen: Were the bacteria detected from fish tissue or spiked samples?
Doug Call: We used fish tissue, no template controls, broth cultures and extractions from agar plates.

Shu Chen: How long did you hybridize?
Doug Call: 4-6 hr or overnight.

Shu Chen: Did you use an automated system?
Doug Call: Yes, we have used an automated system, although this does not work with Teflon masked slides.

Dan Arthur: Stringency?
Doug Call: Stringency is a matter of temperature and salt concentration, but temperature is more important (within certain ranges). With oligonucleotides it is easy to predict the

melting temperature, but it is not clear that these predictions are any more than basic guidelines when we are talking about surface:liquid interfaces.

Jafrul Hasan: Was 5 hr. enrichment minimum?
Doug Call: Yes, 5 hours was required to get that sensitivity for the *Vibrio* work.

Jorge Santo Domingo: Did you concentrate DNA on the beads? Any papers published yet?
Doug Call: Yes, but you can do direct detection on beads as well. No bead papers yet, but some are in the works (one is in press, Journal of Clinical Microbiology).

Development of Microarrays for Pathogens
Vladimir Chizhikov, Laboratory Method Development, Office of Vaccines Research and Review CBER, FDA

Vladimir Chizhikov presented the results of microarray research conducted at the FDA for the detection of bacterial and viral pathogens. The present microarray technology is the result of 5 years of development. He presented the advantages and disadvantages of microarray technology as well as PCR and how the two in combination could be a powerful tool for discrimination of closely related species, multiple pathogen detection, and for genotyping with high sensitivity and specificity. Arrays have already been developed for the detection and genotyping of rotaviruses, detection and analysis of different isolates of vaccine-derived poliovirus recombinants, *Campylobacter* species, *Listeria* species, genotyping of *Bacillus anthracis*, and for detection of mutations in the pncA gene of *M. tuberculosis*. Microarray probe selection was stated to be crucial for reliable detecting, identifying or genotyping of target species. The size of microarray probes can vary from 8-mer to 70-mer oligoprobes depending on application requirements. 'Sliding oligonucleotide' microarray composed of partially overlapping short (14-25 nucleotides) oligoprobes was shown to be a valuable tool for rapid detection of minor genetic changes (e.g., single nucleotide mutations and deletions/insertions) in an antibiotic resistance gene of *M. tuberculosis* (pncA). This approach was also used for detecting recombination events which spontaneously occur during the circulation of the poliovirus vaccine strains in humans. He described some details of developed quality control (QC) system which included the spiking of each microarray oligonucleotide with additional oligonucleotide on unrelated origin provided an opportunity to monitor the quality of all steps of microarray fabrication and analysis. QC and target hybridization images can be taken at different excitation wavelengths.

Questions:

Sandhya Parshionikar: What amplification method was used for the RNA viruses?
Vladimir Chizhikov: RT-PCR.

Gary Vora: Did you fragment after PCR?

Vladimir Chizhikov: No, we found that microarray could be successfully used for analyzing quite long and unfragmented single-stranded DNA up to 7,000 nucleotides in length.

Tim Straub: *B. anthracis* and *B. cereus* can look a lot alike, can you differentiate them?
Vladimir Chizhikov: We didn't do that study, but there is no reason it could not be done.

Keya Sen: With your in house chips what kind of QA/QC can you achieve?
Vladimir Chizhikov: Very good. Developed QC system allowed us not only to control the quality of fabrication but to use signals from QC oligoprobe to normalize signals from specific microarray oligoprobes. It helps to significantly improve the reliability of microarray analysis.

Nelson Moyer: *B. cereus, B. thringiensis* and *B. anthracis* have the same chromosomal genotype, and plasmids confer the differential characteristics of para-crystalline array production or toxin production.

The Virus Chip: Viral Diagnosis and Discovery
Amy Kistler, University of California, San Francisco

Amy Kistler explained that the Viruschip is designed to simultaneously screen for all known virus families whose sequences are available in published databases, allowing sampling without preconceptions. Two bioinformatic tools were created to produce the probes for hybridization on the Viruschip: ViroBLAST and ViroTAX. Both ViroBLAST and ViroTAX select s70 mer sequences, but ViroBLAST searches for highly conserved sequences for detection of viruses at the family level and ViroTAX chooses sequences to fill out the taxonomic tree and allow for differentiation of viruses detected at the family level. ViroBLAST breaks up genomes into 70 mer sequences and blasts each one against the complete genomes of other viruses. It then takes each 70 mer sequence and ranks it by the degree of homology, and finally chooses the top five 70 mers with the most homology with the same viral family. A tool called E-Predict was also created to interpret data from the Viruschip. E-Predict is a virtual library of signature hybridization patterns for all known viruses based on sequences in the Viruschip. Virochip v. 3 is the current production version. For digital readout, the spot data were converted to a strip that resembled a bar code, and bar patterns could be read visually. This system has worked well with the identification of viruses in nasal lavages and exacerbations from patients with upper respiratory tract infections, and has shown very good concordance with data from PCR and culture in the same patients. The system has the potential to recognize emerging viruses. During the SARS outbreak, CDC sent the virus to these investigators in an effort to characterize the then unknown virus. The SARS virus reacted with both rhinovirus and coronavirus oligos. The SARS product was scraped from the slide and sequenced, confirming that SARS was a new coronavirus. The system was modified to permit programable automated elution from specific spots to recover genes using a pH gradient. Sequences of genes of previously unknown viruses can thus be obtained for further characterizations.

Questions:

Participant: Is the system amenable to all viruses?
Amy Kistler: The system was designed for discovery and diagnosis. We wanted something to allow us to look at diseases with some evidence of viral etiology and see if we can isolate the virus.

Shay Fout: Sensitivity?
Amy Kistler: We have mostly looked at RNA viruses, with these we can detect 3,000 virus copies without competing viruses, with clinical samples sensitivity reduced to 300,000 virus copies.

Vladimir Chizhikov: Sample preparation?
Amy Kistler: Background elimination by homologous tissue adsorption. We have just started to look at DNA viruses, with RNA viruses we digested most of the background by getting rid of the DNA.

Keya Sen: Sample source?
Amy Kistler: Nasal lavage.

Shu Chen: What % hybridization?
Amy Kistler: We are still working on stringency and looking at cross hybridization and what conditions are best for hybridization.

Shu Chen: Any Toronto collaborators?
Amy Kistler: No, just CDC.

Daniel Oerther: With the ViroBLAST tool, did you look at the biological relevancy of the 70 mers that were picked for the Viruschip?
Amy Kistler: Don't know yet, there are lots of things that could be data mined out of this but we have not done this yet.

Vladimir Chizhikov: What is the chip set?
Amy Kistler: Glass slide, array maker, 48,000 spaces, 21,000 spaces in use.

Development and Evaluation of a Microarray for Detection of Waterborne Pathogens
Shu Chen, Laboratory Services, a Division of the University of Guelph, Canada

Shu Chen presented data from a microarray system for testing food and water samples for pathogens. Their oligonucleotides were synthesized and chips were printed in-house, and multiple probes were used for amplified genes. They found that software did not reliably predict good probes. Their printed slides were found to be stable for less than 3 months. They evaluated their own primers as well as those from the literature and found that

primer efficiency needed to be very high since they were amplifying using multiplex PCR. She could achieve a multiplex reaction with 15 primer sets, 12 of which were for bacteria and three for parasites. A new data analysis tool was developed which created a table of outcomes for interpretation of signal intensity of probable, weak, and negative evidence of gene presence and gave each of these categories numerical values. The data analysis tool also incorporates a QA/QC index using positive controls and alerts if QC fails on a particular microarray. Dr. Chen presented a microarray for detection of bacterial and parasitic pathogens in water using 15 genes, including a16S rDNA generic bacterial sequence, for detection of *Salmonella typhimurium* (including DT104), *Listeria monocytogenes*, *Campylobacter*, *E. coli* (STEC and O157:H7), *Cryptosporidium parvum*, *Giardia intestinalis* and *Cyclospora cayetanensis*. The detection limit was approximately 50 cells for parasites and < 10 cells for bacteria. Water samples of 1,000 L were concentrated to 10 mL, which would require 1,000 cells/mL in the original sample to have detectable parasites in the finished sample. For bacteria, the original sample would have to contain 5 logs of cells to be detectable. Detection was better for food samples, where the rates of culture and array detection were typically above 85%. Sample preparation and concentration has a profound effect on detection.

Questions:

Jafrul Hasan: Did you find culture negative, array positive samples?
Shu Chen: Yes.

Amy Kistler: How did you select the target organisms?
Shu Chen: They are all food pathogens.

Keya Sen: What is the sensitivity for water samples?
Shu Chen: 3 logs of parasites and 5 logs of bacteria, but spiked samples did not extend below 3 logs, so the detection limit for parasites could be lower.

Development of a Microarray for Detection of Key Microbial Contaminants in Drinking Water
Tim Straub, Pacific Northwest National Laboratory, Washington

Tim Straub talked about several issues surrounding successful application of microarrays to detection of pathogens in drinking water. The first one was the upstream sample processing. Sample processing is fraught with problems since a 100 to 1,000 L sample must be concentrated to 0.1 to 1 mL before subsequent processing may occur. Furthermore, in order for microarrays to be used to its potential there is great need for unified concentration methods so that co-concentration of parasites, viruses and bacteria may be achieved for subsequent sampling. Use of hollow fiber micro-filtration may achieve this. Secondary concentration to remove debris and artifacts has depended upon gradient centrifugation or immunomagnetic bead separation. However, the latter has been used for single pathogen isolation. Microarrays require 1-10 μg of nucleic acid for detection, which is equivalent to 6-8 logs of cells in the original sample. Thus the

concentrated sample must be amplified. Several issues appear at this stage. Split sampling of the concentrated sample introduces the statistical probability that samples may be diluted to extinction thereby producing false negative results on subsequent PCR amplification of aliquots. Multiplex PCR can be very labor intensive and PCR is sensitive to carryover contamination and inhibitors in environmental samples. The microarray probes that work well at the bench top may not work well with true samples and the matrix may present major challenges. One solution would be to use multiple probes to target different genes for the same pathogen and validate then with different water samples. Another method for enrichment of organisms from samples is by cell culture which can give sensitivities comparable to PCR. But this may change the composition of the community as competition for growth occurs in culture media by the different organisms present and the end result may not always be the true representation of the community. He presented data where 10 cells of *E. coli* 0157: H7 or *Salmonella* could be detected when spiked into 100 ml of secondary sewage, by microarrays that targeted different genes from these bacteria. Another consideration in the use of microarray technology is what constitutes real time. Current processes for concentration, amplification and assay requires at least 6-8 hours, and convenience is served by overnight processes giving results in less than 24 hours. While most of the attention has been focused on the pathogens, perhaps a more practical use of microarrays would be for indicator detection or the use of a probe sequence that is common to both pathogens and indicator organisms. Other issues are the semi-quantitative nature of array results, and the use of arrays for biomarkers in addition to virulence factors. Some of his suggestions were that the investments must be made on the upstream sample processing part of the method so that the μl assay is representative of the 100 ml sample, use of random priming approaches for amplification rather than sequence specific priming and to achieve PCR on a chip so that amplification, hybridization and detection are all achieved in single step on an array.

Questions:

Jorge Santo Domingo: Is there enough sequencing data available?
Tim Straub: Sequence selection is a key event in successful array design. 16S sequences are problematic since they are highly conserved among the *Enterobacteriaceae* and have the potential to form higher order structures. We need more informatics data to better inform sequence selection and oligonucleotide design. More direct sequencing of genetic material from environmental samples needs to be done.

Jafrul Hasan: How do you deal with potential misses associated with aliquots?
Tim Straub: That's a statistical quandary. Random priming may be helpful but random primers may produce random results.

Jerry Stelma: Which genes do you like to get sequence information from?
Tim Straub: We need consensus on targets but there's been no discussion or debate to date.

Challenges in the Application of Microarray Technology
Roland Brousseau, Biotechnology Research Institute, NRC, Canada

Roland Brousseau presented the results of four microarray experiments. The first two experiments were bacterial source tracking in municipal waste water and runoff from swine feed lots using 16S rDNA, chaperonin 60 (*cpn60*), and *wecE* genes. The *wecE* gene was found to be more sensitive than *cpn60* or 16S rDNA. In the urban wastewater experiment they had limited success and were able to detect a few human signals but no animal signals. In swine-lot runoff, no animal-specific probes were positive and again a few human signals were obtained. An improved and enlarged microarray was designed that included a significant number of 16S rDNA sequences from uncultured as well as to known bacterial species. The chip included altogether 210 oligonucleotides directed to indicators of animal fecal pollution, human fecal pollution and to intestinal flora found in humans and animals. This chip worked better than the previous chip and will be used this summer for more sampling. The third experiments involved differentiation of *Helicobacter* species from pure culture based on variation in the *cpn60* gene, and the fourth was characterization of *E. coli* pathotypes in surface water using virulence genes and genes for antibiotic resistance. Preliminary data from the *E. coli* array from an urban beach look good and more sampling is planned. In the *Helicobacter* experiment universal primers for *cpn60* were unable to detect *Helicobacter* even when 5% *Helicobacter* DNA was spiked in samples, redesign of primers for *cpn60* more specific to *Helicobacter* allowed detection down to 0.05% DNA in spiked sample. Specific PCR data were good for *Helicobacter*, but random amplification did not produce enough signal for detection. There are definitely some challenges to direct detection when 50-100 ng is equivalent to 7 logs of cells in a sample and some form of amplification is necessary. Spiking experiments show differences in detection when different levels of pathogen are introduced. This variability is not promising for methods validation studies. Wastewater is not a promising matrix for multiple pathogen arrays. Attempts to use universal primers based on the16S rDNA works well when bacteria are present in pure cultures but not when present in complex mixtures. Thus microarrays could be used for pathotype and antibiotic determination of bacteria, where pure culture is used. Microarrays become practical and cost effective when the number of targets exceeds 20, otherwise, it's more practical to use conventional PCR.

Questions:

Jorge Santo Domingo: How did you select primers with host prevalence?
Roland Brousseau: Blast search for genes from pathogens selected. Design was not a problem.

Potential Methods for Characterization of Microbial Communities
Colin Stine, University of Maryland, Baltimore, Department of Epidemiology

Colin Stine began by asking open questions about what specific things EPA wanted from a molecular monitoring system, and spoke about microarrays, multiplex PCR single

nucleotide polymorphisms (SNPs), and about single molecule fluorescence in antibody based methods.

Stine asked what EPA wanted stating that real time varied depending upon whether analyses was performed on site or sent to laboratories. What level of target detection was desired, e.g., genus, species, or strain? What were acceptable costs of analyses? Which of the competing technologies appeared most promising for EPA's goals?

Detection of species within microbial communities may be accomplished using 16S rDNA sequences (highly conserved and may offer poor discrimination of closely related genera) or single copy genes, which offer higher target specificity for species identification. 16S sequence identification is considered qualitative, not quantitative. An example of PCR bias was discussed using failure to reliably detect *Bifidobacter* in stools by conventional PCR and FISH, despite their presence at 9-10 log CFU/g.

Single nucleotide polymorphism (SNP) typing involves multiplex nested PCR amplification of 12 to 48 SNPs at a time, that uses separate enzymatic reaction combined with-fluorometric detection to provide specificity for each SNP. SNP typing may be used for presence/absence detection of a gene. SNP assays have been used to detect flouroquinolone resistance in *E. coli* and *S. aureus*.

Single molecule fluorescence antibody assays depend upon metal enhanced fluorescence (MEF) to boost weak signals to a detectable level. These assays are compatible with microchannel systems for real time detection of pathogens.

Stine stated that sampling protocols would depend on what EPA wanted to know, with one protocol for rapid or real time assays for pathogens in treated waters, and another protocol for monitoring regulated as well as CCL organisms in a broad range of source water matrices. Multiple protocols will be required to cover the wide range of pathogens of interest, and protocols must be flexible as technology improves.

Questions:

Jafrul Hasan: How do you determine false signals with metal enhanced fluorescence systems?
Colin Stine: Antibody specificity and choice of photofluors determines specificity of the assay.

Roland Brousseau: What affinity constant is required to give a signal?
Colin Stine: That calculation has not been performed.

Participant: How much antibody do you need to detect 1 bacterium in 100 ml of water?
Colin Stine: We don't know right now.

Jorge Santo Domingo: Specificity is a double edged sword.

Colin Stine: You have to get it working first. Single molecule fluorescence is similar to culture, in that it provides enrichment of a single molecule that is analogous to culturing a single bacterium.

Darrell Chandler: Molecular sensor recognition, affinity capture.
Colin Stine: Agree.

Monitoring Pathogens by Measuring Host Responses Using Affymetrix Microarrays
Stephen Vesper, ORD, USEPA

EPA wants to know if a glass of water will make you sick if you drink it. The goal is pathogen detection and determination of virulence potential for organisms in drinking water, and recognition of new pathogens and polymicrobial diseases. Assays using mRNA are being explored to look at pathogen and host response to infection. The process involves validation of an animal model for measuring mRNA, then validating a cell culture model to repeat animal model findings, analysis of mRNAs from human cell cultures, and use of proteomics to characterize virulence based upon host response to infection. *Aeromonas* was selected for proof of concept because its virulence is complex and poorly understood. Microarrays were used to study up-regulation of genes and 31 genes were found to be associated with virulence in mice and cell culture. Chemokines, transcription regulators, cell surface receptors, adhesion molecules and signaling systems were identified among these 31 genes associated with virulence and host response. Identification of proteins associated with virulence is underway using two-dimensional SDS-PAGE and MALDI-TOF analysis.

Questions:

Tim Straub: Extrapolation from animal or cell culture data? Can you actually get back to the pathogen with this system? Can you tell the difference from a general inflammation response and response to infection?
Steve Vesper: Immune response of an individual is indicative of potential risk of developing disease.

Richard Montagna: Positive predictive value issues are frequently ignored. While sensitivity and specificity may look very good, application of a test in low prevalence populations may give a very low predictive value.

Shay Fout: What is the response time for expression of infection response genes in cell culture?
Steve Vesper: 5 hours is a midpoint in the range from 20 minutes to greater than 5 hours.

Doug Call: What kind of a response are you measuring?
Steve Vesper: It's complicated and we're not there yet.

General discussion at the conclusion of Day 1:

Shu Chen: Mysterious spots caused by primer dimerization. Others observe this?
All: Yes

Roland Brousseau: 2% of probes are promiscuous, giving spurious signals of unknown origin. Probe redesign may be required to avoid 'noise'.

Participant: Single molecule antibody?
Colin Stine: Sensitivity is determined by the choice of photofluor.

Participant: What antibody affinity do you need?
Colin Stine: Monoclonal, but research is not done yet.

Sandhya Parshionikar: Has anyone been able to strip target DNA off of microarray slides and reuse them?

Gary Vora: In many cases the light from the scanner cross-links the target and probe on the microarray slide and after this bonding occurs you can't reuse the slide.

Amy Kistler: The ones that are scanned wet can be reused.

Doug Call: We have reused some.

Don Stoeckel: Do undetected pathogens give a false-positive signal?
Roland Brousseau: Yes, new pathogens may be detected as new constructs of virulence factors, so signals must be confirmed.

Amy Kistler: We do follow-up with manual reactions, vary annealing temperatures and stringency, culture, etc.

Darrell Chandler: Yes.

Amy Kistler: Need a mechanism for confirmation of signals.

Darrell Chandler: How do you ever know what's real or artifact? I will show a possible method for determining false positives from legitimate signals tomorrow.

Jorge Santo Domingo: The use of one organism or one approach isn't going to tell us much. We're dealing with biological communities and different host responses. Whatever is eventually adopted must be in the hands of many labs. What control samples will be used? How will QC be incurred?

Roland Brousseau: Uncertainty is a problem. Most of the uncertainty is not at the hybridization level, and a panel of genes isn't the answer. What's important is recognition of mutants.

Darrell Chandler: What is the holy grail? Proteomics suggest that nucleic acids are a bogus premise and proteins mediate infections. A sensor for infectivity is a large problem, not of technology but of the underlying biology.

Tim Straub: Infectivity biomarkers are a systems biology problem. That's a long road. Short term graded approaches will be necessary given the limited funding issues. We're ten years out at least. Concentrate on pathogens, and look for altered infectivity and toxicity in cell culture.

Angela Page: EPA has issued an RFA on pathogens in drinking water that is open until June 2005. Five million dollars is available over three years in grants between $100,000-200,000 per grant with a maximum of $600,000/year. Details are available at: www.epa.gov/ncer.

Gary Vora: Will EPA require nucleic acid microarray? What about flow-through systems with antibody array detection for large volume samples?

Keya Sen: Are antibody arrays sensitive enough?

Gary Vora: Flow through system is useful for concentration. The Navy is already using one.

Shay Fout: Use of antibodies is problematic, antibody methods have not been adequate.

Darrell Chandler: Antibodies are useful for concentration of target organisms from background. Problems occur with the secondary antibodies associated with detection systems, large volumes have to cross array, and background can be complicated.

Jorge Santo Domingo: Stability of epitopes is an issue.

George Hallberg: The focus needs to be on ideas, not requirements.

Shu Chen: In protein array models for *Salmonella* serotyping I have used pure culture, with a detection limit of 8 logs. In a natural sample it may be more difficult.

Keya Sen: Clinical samples have high levels of pathogens compared to water and thus how useful antibody arrays will be with water samples needs to be determined.

Roland Brousseau: Antibody capture with flow through detection is subject to scale. A ratio of 1:100 is easy, but 100 to 1,000 L is not so easy. Antibodies are good for sample cleanup. Antibodies are not available for all pathogens we want to screen. Nucleic acids give cleaner data and serotypes don't tell you anything about virulence.

Tim Straub: Antibody capture in flow through systems is easy to achieve in 100 ml samples, then you can PCR up cells. At 100-1,000 liters everything falls apart and these

systems do not work. We don't have antibodies for all the organisms we are interested in, and you have to know what you are looking for a priori before using them.

Darrell Chandler: Protein-antibody arrays are not going to be easy like nucleic acid arrays.

Keya Sen: What about NASBA?
Richard Montagna: We use it. It is good to look at RNA as a viability marker, not as an infectivity marker. NASBA is powerful but it is subject to problems similar to those found with PCR.

Jorge Santo Domingo: NASBA is promising, but commercial kits are subject to contamination with DNA.

Tim Straub: Isothermal amplification methods are intriguing since you can amplify RNA or DNA. Is there any interest in trying them?

Richard Montagna: Single stranded amplicons are a big advantage.

DAY 2 - March 23, 2005

EPA Perspective on Microarrays and the Goals of the Workshop
Tom Carpenter, OGWDW, USEPA

A toolbox approach is necessary for addressing EPA's varying responsibilities. The CCL classification process from the microbial universe to the PCCL and CCL is less critical, but as EPA considers pathogens for regulatory determination, a more rigorous process with more robust methods is required. Risk determination demands the highest level of rigor in methods used to produce the data considered in risk assessment of individual pathogens. Currently, treatment rules are used to control pathogens where analytical methods are insufficiently developed for use in regulatory monitoring. Homeland security requires methods that can be applied in the field for rapid, specific and sensitive pathogen detection. EPA asks the questions: how ripe is the array technology for the desired applications; how may array methods be applied; what will these methods tell us; how may an iterative approach be used to develop the technology for practical applications in EPA drinking water programs?

Nucleic Acid Amplification Strategies for DNA Microarray-Based Pathogen Detection
Gary Vora, Naval Research Laboratory, Washington D.C.

During methods development, it is desirable to consider the ability to transfer methods between governmental agencies for different applications (e.g., food (FDA), water (EPA), air (homeland security), etc). The military applications of molecular amplification technologies include studies of disease transmission under crowded conditions, e.g.,

aboard ships, and assessments of the level of contamination of the combat environment that may affect troop readiness. The goal is to produce detection technology where one size fits all. The desire is for a black box with broad spectrum application for pathogen detection. In addition to pathogen detection and identification, the perfect system would measure the biological potential of the pathogen to cause disease. The amount and reliability of information from an array is more important than the time of assay. Ideally, arrays could be used to detect contaminants in both food and water.
Vibrio was used as an example, since species specific for disease reflect different sets of virulence factors. *Vibrio* eradication is not feasible because they have an ecological niche promoting their survival and they undergo transformation to a viable but non-culturable state under unfavorable environmental conditions. The pathogenic potential of VBNC *Vibrio* species is under investigation. Toxins of *Vibrio* are mediated by plasmids, hence they have a high potential for horizontal transfer. The ideal array would provide genus, species, serotype, toxin producing potential, and antibiotic resistance profile. *Vibrio* and other pathogens have developed multi-drug resistance and the ability to screen for drug resistance simultaneously with genus/species identification is important to patient diagnosis and treatment. Genes were mapped according to severity of disease using clinical isolates and patient history. Assays require less than 6 hours to perform, have a high degree of specificity for a screening level array. Confirmation and investigation of anomalous results may be necessary using follow-up with other methods. Non-O1 *Vibrio* are pathogenic and some may even produce toxin. Horizontal transfer is dynamic and results in high heterogeneity within the genus *Vibrio*. Toxin genes are expressed in VBNC strains so they are potentially infective. Several strategies were investigated, including random amplification, multiplex PCR, use of Klenow fragment, and φ29 polymerase. Multiplex PCR was 3-4 orders of magnitude over random amplification. Low sensitivity occurred with short oligos, and use of 70 mers improved sensitivity but error range was wider. Tried tandem random amplification by adding Klenow to others, and results approached multiplex PCR. Spiked samples were tested and multiplex PCR gave 3 log sensitivity, while Klenow fragments gave 5 log sensitivity. Field samples and spikes have lower sensitivity by 3 logs.

There is interest in modeling transmission during outbreaks. Evaluated respiratory disease models with random amplification followed by array analysis. Used high density array with one base shift in consecutive oligos to sequence unknown strains, thereby providing strain typing information. This technique is very promising for surveillance and clinical diagnosis.

Questions:

Jafrul Hasan: Antibiotics are an adjunct to rehydration in cholera therapy. Resistance is due to over the counter drugs in developing countries. Are VBNC strains really shed by patients? Are they really infective?
Gary Vora: Both Colwell and Oliver references suggest that VBNC are infective. Use of antibiotics is widespread in aquaculture, which may have a more profound effect on development of resistance than their use in human medicine.

Virus Infectivity and Microarrays: A Promise for the Future?
Sandhya Parshionikar, TSC, USEPA

The issue of virus infectivity is key to assessing risk from viruses in drinking water. Ethidium monoazide staining, mRNA, heat shock proteins, animal infectivity, and cell culture have been used to determine infectivity of various organisms. Current available methods for detection of virus infectivity such as cell culture and integrated cell culture - PCR (ICC-PCR) have disadvantages in that they are slow, expensive, and not applicable to all viruses. Molecular methods such as RT-PCR for detecting viruses although rapid, provide little information about infectivity. At the EPA, a molecular method that can differentiate between viable (infectious) and non-viable (non-infectious) viruses is being developed. Non-infectious viruses often have damaged protein coats, and these coats are more susceptible to protease-mediated degradation. After treatment with protease, samples are treated with RNAse, which degrades the RNA genome from viruses whose capsid coat was degraded by protease treatment. Subsequent RT-PCR yields negative RT-PCR results. Several viruses such as coxsackievirus, echovirus, Norwalk virus and poliovirus were inactivated by various methods and treated with different proteases. Protease E was the most promising of several proteases tested in all of the above viruses. This method has a potential application to detecting infectivity by microarrays in the future, in that, the viral samples can be first treated with Pronase E and RNAse before performing RT-PCR. The PCR products can then be subjected to hybridization on a microarray platform thereby providing information on virus infectivity.

Questions:

Nelson Moyer: There is a presumption at EPA that infectivity is a presence/absence phenomenon, when in fact, it is a continuum between transitory colonization and fulminate disease, including asymptomatic infection and chronic carrier states. Infectivity in cell culture or animal models may not be equivalent to infection in the human host. These considerations are particularly important in use of infectivity data for risk assessments.

Jafrul Hasan: How were the protease pre-treatments performed?
Sandhya Parshionikar: 6 hours for enteroviruses, 1 hour for norovirus.

Jorge Santo Domingo: What were the chlorination values?
Sandhya Parshionikar: 0.6 to 0.8 CT values.

Richard Montagna: Do you need to know the identity of the virus?
Sandhya Parshionikar: yes

Nelson Moyer: We have a tendency to speak of infection in terms of presence or absence, when in fact it can be a continuum, we need to define infectivity for humans.
Sandhya Parshionikar: I agree, but we are looking at the virus and not the host.

Jorge Santo Domingo: What was the time of exposure at 37º C?

Sandhya Parshionikar: 2 weeks.

From Field Portable Microarrays to Microarray-Enhanced Multiplexed PCR
Darrell Chandler, Group Leader, Argonne National Laboratory, Argonne IL.
Darrell Chandler discussed the development of gel element arrays that was capable of a higher order multiplex PCR reaction, without the pitfalls of a regular multiplex PCR. The base technology for three dimensional arrays is to use gel drops or gel pads as an enhancement of two dimensional array technology. Each gel drop represents an independent reaction vessel and involves two rounds of amplification one with specific primers to the gene target, and the second with universal primers. So far the system could achieve a 11-plex reaction but was being extended to perform 98 plexField analysis is possible using portable imaging equipment. The sensitivity of these arrays is $5x10^7$ molecules using Cy3 or Cy5 detection. These systems have been integrated into a micofluidic chip detector.

False positives are a problem that must be dealt with. All signals aren't real. Cross reactions occur with probes having greater than 50% sequence similarity. An imbedded temperature control in the chip allows not only for on chip PCR, but allows a method for determination of false positive signals. Thermal melt is used to confirm amplicon specificity by neutralizing 'noise'. Thermal dissociation curves are used to validate the authenticity of detector signals. T_m alone isn't enough and there is a need for an embedded internal control for each probe on the chip. PCR is the bottleneck, not the array. PCR detection of 100 to 1,000 copies is required for reliable detection. The number of cells (colony forming units) associated with 1,000 copies is 100,000 cells. Splitting samples leads to false negative results as a result of dilution to extinction. Another advantage of using gel drop arrays would be a sample would not have to be split, and hundreds of reactions could be performed simultaneously. Target size also did not seem to affect gel drop or on-chip amplification. While there are many advantages to using the gel drop arrays or on-chip PCR, T. its sensitivity of detection is 4- 6 logs compared to 3 logs by regular multiplex PCR . The goal is to develop an integrated nucleic acid purification, amplification and detection, lab card. The card would be point-of-use. Such developments are already in the pipe line and can handle 10 ml of sample. Commercial intent is primarily for clinical applications because commercial companies doubt that there are sufficient environmental applications to be cost effective. Collaborations between multidisciplinary teams can make it feasible to develop products, but individual investigative efforts are not cost effective.

Questions:

Richard Montagna: Predictive value issue of false positives is a problem. T_m as a built in control is important.
Darrell Chandler: Yes, this is the value added aspect of an integrated chip.

Shu Chen: What volume can the chip take?

Darrell Chandler: 10-25 μL in the current configuration, but 10-20 mL may become feasible.

Keya Sen: Do you plan to push down the detection limit?
Darrell Chandler: Yes, the multiplex PCR power is more important than sensitivity initially.

Doug Call: What is the potential for quantitative results?
Darrell Chandler: None, but we are working on that on paper.

Jorge Santo Domingo: How many genes can you put on a chip?
Darrell Chandler: This is microtechnology, not nanotechnology, but there is no limit on genes. 10,000 to 20,000 features are feasible.

Participant`: Is the on chip PCR process upstream of detection?
Darrell Chandler: The whole process is integrated on the chip and it is field deployable.

Overview of Cornell University/Innovative Biotechnologies Biosensors for Detection of Waterborne Pathogens
Richard Montagna, CEO, Innovative Biotechnologies International, Inc., Grand Island NY

Richard Montagna talked about alternative technologies to the use of microarrays that are available for pathogen detection. Liposome technology has been used in clinical diagnostics for years. Liposomes can have a diameter from 50-800 nm and the surface of the lipid bi-layer is reactive such that detection molecules can be attached to it. Signal molecules can be placed in the lumen of the liposome. Several formats have been developed, including lateral flow, fluorescence-based, and electrochemical detectors.

NASBA amplification of RNA gives 10^9 to 10^{12} target copies. A double sandwich system detection system is used for *C. parvum* heat shock protein mRNA as a viability marker. Reaction time is 4-6 hours from sample collection to result. Using a target concentration of 5 oocysts, 5 of 6 test samples were positive.

A method for detecting *Cryptosporidium parvum* heat shock protein mRNA was developed in which NASBA was used as an amplification method. Oocysts are concentrated using immunomagnetic separation, subsequent mRNA amplification selects viable organisms. The method was a 4-6 hour process and was able to detect as little as 5 oocysts from an original 10 ml concentrated sample. Mixtures of different combinations of organisms showed little interference.

Immunoassay with fluorescence detection has been used for *E. coli* detection with an interdigitated array employing electrochemical detection with the potentiometer on a disposable chip. Theoretically, this technology is capable of measuring nanogram and pictogram quantities, while liposomes can detect femtogram quantities.

Resonant cantilever beam systems have been developed for detection of single cells of pathogens. Use of a vacuum and improved cantilever design has reduced 'noise' and improved the detection limit. Use of a gold spot impregnated with antibodies to localize the binding site has also improved sensitivity. Use of resonance instead of cantilever deflection has improved sensitivity. Data on cantilever nanoelectrical devices has been published.

The use of multiple technologies in an integrated fashion can result in greater specificity and sensitivity than any single test alone.

Questions:

Lu Li: How is cantilever binding accomplished?
Richard Montagna: The antibody is on the cantilever surface.

Shay Fout: How is resonance affected when challenged with an unknown number of viruses?
Richard Montagna: It depends upon the characteristics of the target pathogen. Each one is different. You have to know the characteristics of the organism your looking for.

Jafrul Hasan: What is the efficiency?
Richard Montagna: 5-10% range

Shu Chen: What surfactant do you use?
Richard Montagna: 30mM octoglucol pareneste.

George Hallberg: What is the recycling potential of the chip?
Richard Montagna: None

Keya Sen: What primer optimization do you use?
Richard Montagna: None, not much primer design. Anneals at 60º C to reduce primer-dimer formation. Target is excessive in clinical samples.

Panel Discussion

Question 1 - Paul Berger: EPA is interested in presence of viable and infectious agents. We know that some regions of DNA are very stable in water, but we think that RNA is not very stable, although this is not well characterized in environmental samples. Are heat shock proteins reliable markers of viability and infectivity? What microarray tools might be available in the next few years?

Darrell Chandler: What is your definition of viability and infectivity?

Paul Berger: It takes a higher dose of VBNC organisms to cause illness.

Shay Fout: If a pathogen causes health risk to people, it is infectious.

Darrell Chandler: The demarcation is between nucleic acids vs. proteins. mRNA is a marker of cell activity. An infectivity assay must incorporate host response.

Richard Montagna: Infectivity is a systems biology problem. Inducible RNA is related to infection. No one sensor will answer all questions.

Sandhya Parshionikar: The thinking is pathogen centric, but we need more host information as the next step.

Amy Kistler: Tissue work with high background is similar to water matrix. We want to recover the full genome and use cell lines to test infectivity.

Roland Brousseau: Inducible mRNA is the best proxy for infectivity. For the pathogen side, use of heat shock protein, e.g., release of heat shock proteins indicate response to stimuli hence viability and perhaps infectivity.

Tim Straub: What do we want from our assays? There are different camps in EPA. Regarding viability and infectivity, we need better cell culture methods. Microarrays are just part of the answer and new methods other than microarrays may be needed. Maybe apply it to a microarray in 5-10 years, but host factors must be included in infectivity evaluations.

Question 2 - Jafrul Hasan: We need reliable methods from sample collection through concentration, amplification, and detection. How may we ensure a reliable, standardized protocol at reasonable costs?

Richard Montagna: There are some great biosensors moving forward, but the process is dependent on sample processing. We need processing advances, not more sensors!

Darrell Chandler: The steps are development, production, validation, marketing. The technology is not there. EPA needs the end result in mind. There is too much decoupling of science and technology. Glass vs. plastic for example. The end game involves lots of different specialties. Step one is to write a product definition. Step two is to make a cost estimate. For a new device it will take 5 years and $10,000,000. Then for a sample to answer information startup costs may be around $40,000 to $50,000 for individual units during development and testing, and $5,000 when produced in mass. Individual chips may cost $1,000 each during development and may be $2-5 each when produced in high volume.

Richard Montagna: We work with lots of small businesses. It's unrealistic to expect a sample to result in a chip for $2 each with minimal equipment support. Development of chips will be incremental over years.

Roland Brouseau: A typical cost for a microarray experiment would be of the order of $60 for the array and accessory reagents, without including the setup costs. This must be compared with the cost of running perhaps 170 to 200 PCR or biochemical reactions to obtain the same amount of information without using microarrays.

Doug Call: We can produce conventional arrays at the rock bottom price of $15 each because we're located in and supported by an academic setting.

Tim Straub: Realistically, arrays will have a $20-30 million dollar development cost. Method 1623 costs approximately $500 per sample. Using hollow fiber filters for co-concentration of all desired pathogens could bring the cost of array analysis to $500-1,000 per sample.

Shu Chen: Cost is based on number of samples to be tested, which amounts to economy of scale. Therefore the more samples to be tested the cheaper the cost and at a minimum 5 samples should be tested together.

Richard Montagna: If a kit cost $10, would water plants use it?

Question 3 - Jim Sinclair: The purpose of microarrays is to detect multiple targets simultaneously. This is one of the reasons why microarrays are of interest to EPA . We heard a lot of discussion yesterday from the speakers about problems with detection of multiple targets. How many targets can microarrays detect simultaneously if the targets may be diverse and we want good sensitivity in water? When could we expect that we could increase that number?

Tim Straub: The problem isn't how many targets you can put on the array, the problem is the environmental sensitivity of arrays given the desire for a broad range of targets at low levels in water. The real problems are relative concentration issues, sample and concentration scaling problems plus lots of unknowns. We need better multiplex PCR amplification to increase target number.

Jim Sinclair: What number of targets can we detect now?

Darrell Chandler: We might be able to get as many as 1.6 million probes on a chip. We must have a product definition, and you must define the end product. Without a unified effort we will all work on our own parts of a possible system and product specific to your needs will be created.

Gary Vora: There is a risk that the technology will become too sensitive, then it becomes a biological problem. At that point, we have to ask, what does it mean relative to disease?

Shu Chen: We don't have to put all targets on a single chip. We can use more than one.

Richard Montagna: Using the example of HIV in a low incidence population, you must have a high predictive value or you will have lots of false positive tests that will be very

difficult to explain to people. The prevalence of the disease in the community determines the way you design the test and the way you interpret the results.

Question 4 - Alan Lindquist: How you are going to go about validating this new technique?

Participant: It will depend on the answer you are seeking from these techniques. Inter-laboratory validation may have to be done if results from microarrays methods are going to be used for information collection surveys in support of possible future regulations. If the array is going to be used as a research tool for taxonomic or identification purposes, then the validation will be different and will depend on use of methods such as sequencing, qPCR or the use of internal and external controls.

Darrell Chandler: How do you standardize tests relative to public health outcomes? What platform best provides the desired information? We need appropriate testing protocol for arrays. How do you develop a statistically valid testing method?

Tim Straub: Tim Straub stated that EPA needed answers relative to public health outcomes and this would in turn drive the design criteria. Do they need presence/absence or other levels of information? From the perspective of environmental microbiology, build the widgets, while EPA builds the standardization and validation studies, and all is wonderful. When you apply the results to a rule, you have to interpret the outcome. Presence of target alone may not be indicative of a health problem. We need to use data from surveillance sources, environmental, epidemiological, clinical, and public health.

Question 5 - Shay Fout: Array design may be either presence/absence or a fingerprint approach. Would the sequencing/fingerprinting approach be better than straight detection approach for EPA priorities? Are there differences in chip design for either application?

Amy Kistler: We need different levels of monitoring. Different approaches are needed for different questions. We use whole genome detection for obtaining multiple data points, e.g., surveillance, diagnosis, and discovery.

Shu Chen: We can build an array to detect common waterborne pathogens with the goal of replacing traditional methods, but the result may mean more validation and more cost.

Sandya Parshionikar: We can use chips and long oligos for screening, and use other methods for confirmation.

Amy Kistler: Design strategy is very different according to the purpose of the array.

Darrell Chandler: Design of arrays is tied to biology of interest. How will the data be interpreted? Arrays may be constructed to generate data, then interpret it accordingly.

Workshop Summary
Keya Sen, OGWDW, TSC

There are several types of microarrays including planar and bead types. Bead arrays offer higher throughput and less complicated imaging.

There are several possibilities for probe design including short oligonucleotides (8-22 bp) and long oligonucleotides (70 bp). Long oligonucleotides have higher sensitivity at the cost of lower specificity. Amplicons (180 bp -1.5 kb) can also be used.

Many targets are possible including the use of multiple genes and whole genomes. Some promising genes are 16S *rDNA* and *cpn 60.*

Several amplification procedures were discussed including:
Specific amplification: PCR with specific primers and NASBA
Random amplification: PCR with random primers, Klenow fragment amplification, φ29 polymerase. Combinations of these amplification procedures are used.

PCR is the most sensitive amplification procedure but suffers from copy number bias, template and primer bias, loss of primer efficiency in multiplex reactions, and amplicon length bias. The use of 16S genes for PCR amplification is hampered by the formation of secondary and tertiary structures during amplification.

Culture amplification can be as specific as PCR, has no attendant PCR bias, and measures only viable organisms. While culture does not suffer from PCR bias, different organisms may grow at different rates and faster growing organisms may be overgrown relative to the slower ones. Non-culturable organisms and hard to culture organisms will not be represented.

Trouble shooting will include choice of well validated probes and these need to be validated in the field. The use of multiple probes for the same gene, the simultaneous measurement of all chips, and increasing hybridization temperature on chip to see which signals disappear, may all help to eliminate false positives.

Detection of organisms with microarray is very good from pure culture and clinical samples, but not so with water samples, where organisms are much more dilute. Primers that work with pure culture may not work with a field sample and matrix interference is a major challenge

Exciting new work brought out in this workshop include: development of the “virus chip”, MME-PCR in which primers are fixed in polyacrylamide and combined with “nested PCR”, biosensor technology combined with NASBA, Affymetrix resequencing microarrays, and metal enhanced fluorescence technology.

Microarrays may be ready for simultaneous detection of a small number of pathogens in the next five years, but the prediction of pathogens using microarrays combined with a

VFAR approach or based on host genome expression, remains a long term goal. Limited validation is possible when a small number of probes are present. The cost of a microarray system can vary. Viability assays are a possibility in the near future using mRNA in bacteria and prokaryotes, and protease treatment followed by RT-PCR or PCR in viruses.

Future research should be directed toward improvements in sample preparation, greater multiplexing capability, improvement in detection specificity, and shift of focus to analysis of microbial community rather than the analysis of a single pathogen.

George Hallberg, The Cadmus Group, Inc.
The big questions are still on the table, and that is not unexpected given the complexity of the problems.