APPLYING CLASSICAL MICROBIAL IDENTIFICATION STRATEGIES TO MICROARRAY PLATFORMS.

M. Theodore, N. Boggs, H. Le, H. Ho, W Bethea, and J. Jackman*
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd, Laurel, MD 20723
*corresponding author

ABSTRACT

In the event of a BW attack, rapid and sensitive methods are needed for accurate identification of bacteria to the species level. Current rapid methods for species-specific identification require clean extraction of pure nucleic acid and require amplification of signals/starting material by PCR. Significant microbial expertise may be required to evaluate a sample containing microorganisms so that the correct PCR primers can be selected for positive identification before any tests are run. Although PCR is relatively rapid, the applications are limited to assessment of a small community of microbes and are directed to only one or a few genes at a time. In order to adequately test for all potential bacterial pathogens, many tests are required if no apriori knowledge about the suspected agent is known. This increases the time and cost of analysis results in depletion of available test materials. A better approach would be to reduce the cost of analysis without the concomitant loss of information and sample and without increasing analysis time.

We will describe the adaptation of classical microbiological strategies to molecular methods for rapid and high confidence identification of bacterial pathogens. Classical microbial identification strategies are not typically based on a single test or evaluation but are usually based on a collection of tests and a decision tree for their interpretation that results in a high confidence identification Unfortunately, current classical and gold standard methods for bacterial identification are relatively slow, requiring hours (usually 4 or more) to days for species-specific identification.

While single tests are often broadly applicable to many genus, grouping many tests together provides the specificity and internal consistency for accurate bacterial identification. Using a combination of high density microarrays and hierarchical decision trees, we have successfully generated multiplexed platforms for identification bacterial agents in less than two hours. Due to the large number of probes (<1000), the platform can be broadly and simultaneously applied to all human pathogens. The platform is highly adaptable in that methods have been developed with and without the use of PCR. Assay development time for new targets has been significantly reduced by the capture of all data used in the testing and evaluation into a relational and searchable database. We will describe how retrospective analysis of collected data has permitted us to understand the features of probe development and placement and to down select the best methods, assay components and software construction which provide the greatest reproducibility, sensitivity and specificity for this platform. Finally we will describe "within assay" controls incorporated into the platform to provide end-users with point of use metrics of assay performance (benchmarks) simultaneously with test data.

CLASSICAL MICROBIOLOGICAL AND MOLECULAR METHODS

If there is a possible threat to human lives and safety, quick and accurate identification of these potentially harmful organisms is essential to determine possible post exposure prophylaxis methods. A determination would need to be made as to whether the incident is an actual biological event or hoax. Classical bacterial identification methods are routinely used in the laboratory, and although very accurate, often take several hours or days to complete. PCR, polymerase chain reaction, is a technique that is frequently utilized for bacterial identification. Although PCR can be performed in a relatively short time period, the chief limitation of the technique is that specific primers are needed to increase copies of the target organism, such that enough material is available for a detection device. Without some foreknowledge of the target organism, and associated specific primers needed, detection of unknown organisms cannot be accomplished quickly or reliably. Hence, high confidence bacterial identification strategies generally must pair both these approaches.

Success in using classical identification methods is generally based on use by trained, experienced laboratory staff and technicians in a sterile laboratory setting. Most biological incidents are likely to take place in the field where battle-ready personnel may not have the requisite training to use complicated devices and may not be able to handle the biological material in such a manner as to avoid contamination. The net result can be unreliable assays and low confidence results.

In addition, the logistics associated with a paired

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
00 DEC 2004	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Applying Classical Microbial Identification Strategies to Microarray Platforms	**g**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **The John Hopkins University Applied Physics Laboratory Laurel, MD 29723**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001736, Proceedings for the Army Science Conference (24th) Held on 29 November - 2 December 2005 in Orlando, Florida.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**7**	

approach using classical bacterial identification schema and molecular methods such as PCR can be a significant burden. Multiple, large instruments are sometimes required to conduct biological assays (e.g. incubators, and thermal cyclers). In addition to being bulky and difficult to transport into the field, many of the reagents used with these instruments are perishable or are only stable at certain temperatures and under certain conditions.

Given the limitations of current biological detection devices, alternative identification methods need to be developed that provide for accurate, timely results, can be easily transported by military personnel to the site of a potential exposure, and are easy to use. The **MAGIChip™** (**M**icro **A**rrays of **G**el-**I**mmobilized **C**ompounds on a **Chip**) is one such technology.

NEW TECHNOLOGY: MICROARRAYS

Microarrays are a technology which is less than a decade old. Microarrays are small platforms that contain many different biological macromolecules logically arranged or "arrayed" (Heller, 2002). The macromolecules used in arrays can be proteins, sugars or nucleic acids (Arenkov, 2000, Eickhoff, 2002, Zlatanova, 2001). Microarrays are analyzed by binding specific portions of target organisms to the array. Target molecules are labeled, , either directly or indirectly and then captured onto the platform surface through a process known as hybridization after which the microarray is imaged using a CCD camera. Preferential binding (adherence of target sequences to capture probes) is realized by site-specific association of labeled molecules to the capture probes (sequences of nucleic acid that exactly match the nucleic acid sequences that will be hybridized onto the microchip).

The use of microarray analysis is a novel approach to bacterial identification that eliminates many of the issues identified above since microarray platforms support multiple assays simultaneously. Once a microarray has been designed for several biological strains, the cost to replicate the array is rather inexpensive. Additionally assay time is short, and the assay itself is relatively uncomplicated allowing for field use with minimal training.

UNIQUE MICROARRAY TECHNOLOGY: MAGICHIP

The MAGIChip (Micro Arrays of Gel Immobilized Compounds on a Chip) is a microarray-based technology developed to specifically address the need for rapid and accurate identification of bacteria. Typical microarrays are printed on the flat surface of a glass slide usually treated with a specific compound upon which the capture probes adhere. However, MAGIChips are unique in that they are 3 dimensional (3-D) DNA microarrays (Proudnikov, 1998). MAGIChip microarrays are formed from oligo impregnated gel pads increasing the surface for hybridization by more than 50 times over that of 2D arrays. The 3-D nature of this microarray allows for increased probe density. Each position on the MAGIChip contains greater than 10^{12} capture molecules.

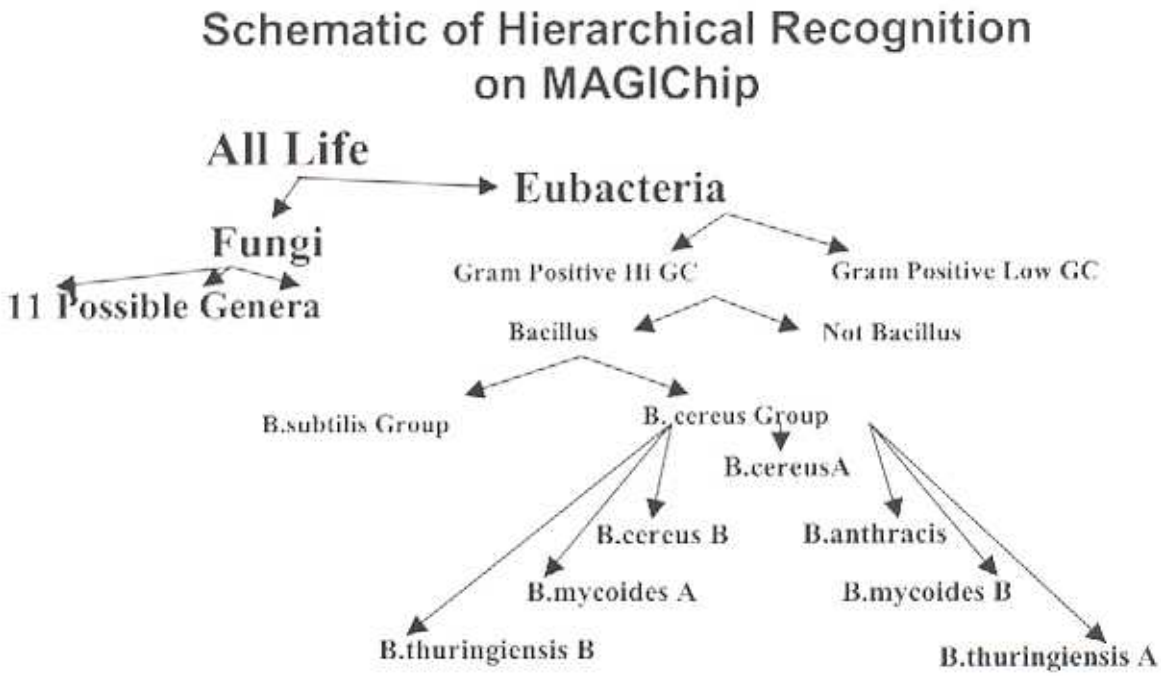
Specific probes are designed to recognize unique ribosomal RNA sequences and are placed into individual pads on the chip. The small size of the 3-D gel pad allows for a high- density array to be deposited onto a single microarray. There is the capacity for over 2700 probes to be loaded onto a single MAGIChip. It is possible to multiplex 2700 separate capture assays simultaneously or to compare fewer different assays with a greater number of replicates for each capture probe, thereby increasing the confidence of each assay.

The MAGIChip technology was initiated as collaboration between Dr. David Stahl now at University of Washington and Dr. Andrei Mirzabekov, formerly of Argonne National Laboratory. The MAGIChip is a portable multiple assay technology which is capable of discriminating all human pathogenic bacteria simultaneously. This is accomplished by using a single set of genes to evaluate all bacteria, namely ribosomal RNA (rRNA). All live bacteria have many copies of rRNA (as many as 1000 copies per bacterium). rRNA molecules are extremely abundant, making up at least 80% of the RNA molecules found in a typical eukaryotic cell (Bavykin, 2001). Since the copy number of rRNA is pre-amplified by bacteria, no further amplification is required. Therefore, the logistics burden of performing PCR and the time required for amplification do not apply to the MAGIChip microarray assay. Woese first suggested that the phylogeny of all life could be traced thorough rRNA(Woese,2002). Stahl and Mirzabeckov expanded on this concept suggesting that comparison of single nucleotide polymorphorisms (SNPs) in rRNA would be informative with regards to genus and species. The MAGIChip exploits these polymorphorisms in bacteria as a mechanism to discriminate between bacterial genus and even to the species level.

Ribosomal RNA (rRNA) is highly conserved within species, using single base pair mismatches in the design of the capture probes allows for discriminating results requiring species-specific identification. Single nucleotide variations present in canonical bacterial rRNA sequences are used to differentiate one bacterial species from another. Probes are loaded onto the chip, which distinguish organisms using hierarchical recognition schema. Capture probes are phylogenetically ordered by domain, kingdom, phylum, class, order, family, genus

FIGURE 1:

Hierarchical recognition uses classical microbiological schema. Probes are phylogenetically ordered by domain, kingdom, phylum, class, order, family, genus and species. Structured multiple probe analysis provides internal consistency check and increases assay confidence.



and species within this schema (Domrachev,2004) (Figure 1). This hierarchical based identification strategy provides an internal consistency check within the automated analysis program. Using the same schema for distinguishing bacteria as performed by classical microbiological assays, the MAGIChip has provided molecular tools which mimic classical microbiological identification methods. Identification of Bacillus species is distinguished by hybridization to less than 100 probes; so, by virtue of available "real estate" on the chip, multiple probes for the same target species can be designed and analyzed simultaneously. Multiple probes for the same target also increases confidence of identification. Thus Bacillus species are not identified merely as Bacillus but also are identified as gram positive bacteria with high GC content.

The MAGIChip assay is rapid. Identification using the MAGIChip can be completed in less than two hours using only a handful of SNPs on the microarray that can be used to hierarchically discriminate bacteria to the species level. Discrimination to the species level is important because the specific species of an organism is one of the determinants for assessing risk to the warfighter.

CONCEPT OF OPERATIONS: MAGICHIP METHODOLOGY

The process for detecting bacteria using the MAGIChip is relatively straightforward, minimizing the use of sophisticated laboratory techniques that may be more difficult for the naïve user (Figure 2). Bacteria are taken from a sample are washed and pelleted. While the sample is typically a colony from a plate, it does not have to be a pure sample as the predominant species will be the target visualized. The pellet is then lysed using lysozyme, an enzyme that destroys the cell wall, allowing the release of nucleic acid (NA). The lysed bacteria are then added to a preparation column where the NA from the cells remains immobilized while cell debris and soluble contaminating materials are removed. All the processes that take place post lysis (i.e. isolation, purification, fragmentation and labeling of target rRNA) occur by chemical, rather than enzymatic reactions. Use of stable chemicals reduces the logistics burden needed to transport other sorts of enzyme based assays. This purified nucleic acid then undergoes fragmentation and labeling simultaneously (Kelly, 2002) through the use of chemical reactivity. Fragmentation into small, <500 base pair sequences, allows for easier hybridization onto the chip because smaller pieces of NA permeate more easily through the gel matrix. As the NA is cut into smaller pieces free ends are created that are chemically labeled with the fluorescent probe. Once the NA is labeled and fragmented, it is eluted off of the column and ready for hybridization onto the chip. The total amount of all NA (DNA and RNA) eluted off of the chip is determined using ultraviolet spectrophotometry, and then a pre-determined amount of NA is mixed with a hybridization buffer and put onto the chip for one hour. Visualization of the result of hybridization and data interpretation are nearly simultaneous. Automated algorithms interpret the image and determine the identity of the microorganisms in less than two minutes. From start to result, the MAGIChip assay takes less than two hours.

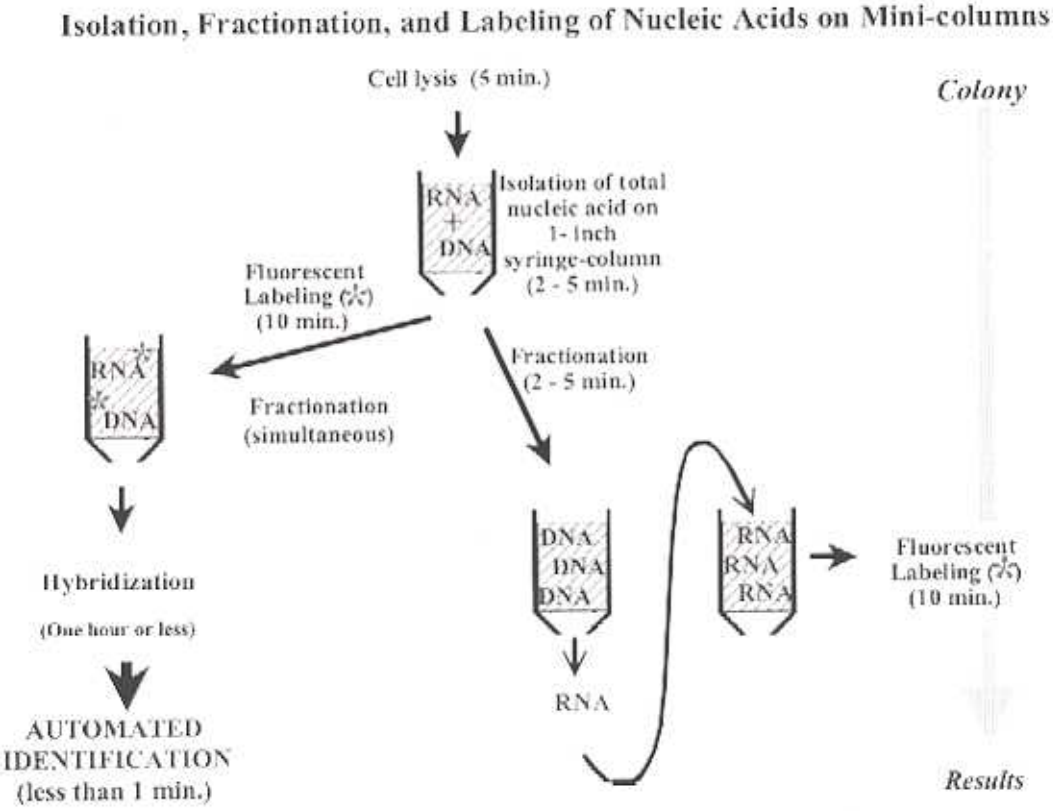
CURRENT APPLICATIONS:

MAGIChip technology harmonizes with available PCR applications by directing users in the selection of PCR primers for specific amplification. While bacterial genus and species identification can be carried out without PCR, confirmation of lethal strains by microarrays requires use of PCR probes. Unlike rRNA, virulence factors which include toxins or proteins which increase the invasive potential of an organism or its toxin, are low copy number transcripts. Also unlike rRNA, they are only produced under specific conditions as found in the host. Hence identification strategies usually focus on the single copy DNA sequence. Current PCR technology is expanding the MAGIChip platform. Generalized PCR reactions can be carried out which amplified genes in several species or genus, the MAGIChip is then used to discriminated to the species/strain level using capture probes from the amplified regions. Once amplicons are generated, many probe validation experiments could be conducted simultaneously. Amplicons are segments of NA in high copy number that can be generated to compliment and probe that is tested on a microarray. Methods being tested are incorporation of fluorescent labels into PCR reactions and generating PCR amplicons using pre-labeled PCR primers, which also generate labeled PCR amplicons, for hybridization onto the chip. Various hybridization methods are also being evaluated for use with PCR amplicons. The use of PCR generated labeled amplicons for hybridization allows for the production of generous amounts of probes specific to those being evaluated on the chip.

While the MAGIChip has a role in detection of virulence, it can also be used to speed the selection of PCR primers and targets. Testing various specific probes using PCR can help determine the optimal size of probes used for hybridization. Probe specification is critical for the determination of target specific hybridizations. Cross reactivity of non-specific probe binding can lead to the mis-identification of an organism. Thus, using PCR for it's specific nature is a useful tool in validating probes while designing a chip. It is, however, not the intention to use PCR for organism detection purposes based on

FIGURE 2

Concept of Operations: Schematic depicts process to generate labeled target which is captured on MAGIChips based on specific sequence recognition. As shown either RNA alone or both RNA and DNA can be simultaneously labeled. The entire procedure takes less than two hours. Automated analysis occurs in near real time (<1 minute) once the hybridized image is obtained.



limitations stated previously.

DATA HANDLING: TOOLS AND ANALYSIS

As part of our testing and evaluation effort at Johns Hopkins Applied Physics Laboratory, we added value to the MAGIChip technology through the creation of a MAGIChip database and analysis tools application that helps provide insight and understanding into performance evaluation, assay design, and hardware/software quality control.

A large portion of the MAGIChip testing and evaluation is captured in a database and analysis application that helps provide performance evaluation, "in-silico" experimentation and hardware and software quality control. The database application was developed to provide the traditional database collection, storage and retrieval tasks for MAGIChip experimental data, and to assist in probe design, assay evaluation and component performance analysis.

The MAGIChip database and analysis application consists of several components. To handle the traditional task of data collection, a secure, web-based data entry user interface was developed. To handle the traditional data storage and retrieval tasks a secure, relational database was also created. These two components give the system a web enabled, centrally located, secure infrastructure with distributed access to handle data from multiple research activities in different locations. The other components include a user-friendly analysis application that incorporates the analysis by query, and a suite of software tools to accomplish some of the more complex querying of the system. The analysis application has been developed to be flexible and usable by researchers and domain experts and may prove to be quite useful in assessing the capabilities and potential of the MAGIChip.

At the lowest level, analysis is conducted by querying the database. Experienced users of the database application may apply the statistical analysis techniques to stored data or to the results from recently submitted queries. At the highest level, analysis is conducted by encapsulating domain expertise into more complex high-level queries. In some cases these high-level queries are constructed from a sequence of basic queries organized together by the domain expert to simulate how a researcher might investigate the experimental data to draw some conclusions about the performance of the technology. In other cases the high-level queries may be constructed from data processing software that manipulates the stored data. Researchers and domain experts familiar with the types of data captured in the database application can formulate (or select) queries that provide insight into the performance of the assay.

Analysis may also be conducted by using the statistical analysis tools and techniques currently available on the MAGIChip portable reader (developed by Argonne National Laboratory). The portable reader is a fully contained system which can operate from a laptop or mini computer. It is less than 6 lbs and has a foot print less than a briefcase.

CONCLUSIONS

The MAGIChip addresses the needs of the warfighter. It provides a multi assay platform for near real time identification of pathogenic bacteria without complex assays. Use of the MAGIChip requires little training. In addition, the MAGIChip platform can assist in assay development for PCR based assays to assist in the design and validation of new amplicons based assays.

REFERENCES:

Arenkov P, Kukhtin A, Gemmell A, Voloshchuk S, Chupeeva V, Mirzabekov A. (2000). Protein microchips: use for immunoassay and enzymatic reactions. *Anal Biochem* . 278, p123-131

Bavykin SG, Akowski JP, Zakhariev VM, Barsky VE, Perov AN, Mirzabekov AD. (2001). Portable System for Microbial Sample Preparation and Oligonucleotide Microarray Analysis. *Applied and Environmental Microbiology* . 67 (2), p922-928

Domrachev M, Federhen S, Hotton C, Leipe D, Soussov V, Sternberg R, Turner S. *Taxonomy Browser.* NCBI. Available: URL http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=1386. Last accessed 10 February 2004.

Eickhoff H, Konthur Z, Lueking A, Lehrach H, Walter G, Nordhoff E, Nyarsik L, Bussow K. (2002). Protein array technology: the tool to bridge genomics and proteomics. *Adv Biochem Eng Biotechnol* . 77, p103-112

Heller MJ. (2002). DNA microarray technology: devices, systems, and applications. *Annu Rev Biomed Eng* . 4, p129-153

Kelly JJ, Chernov BK, Tovstanovsky I, Mirzabekov AD, Bavykin SG. (2002). Radical-generating coordination complexes as tools for rapid and

effective fragmentation and fluorescent labeling of nucleic acids for microchip hybridization. *Anal Biochem* . 31, p103-118

Proudnikov D, Timofeev E, Mirzabekov A. (1998). Immobilization of DNA in Polyacrylamide Gel for the Manufacture of DNA and DNA- Oligonucleotide Microchips. *Anal Biochem* . 259, p34-41

Weiner MP, Hudson TJ. (2002). Introduction to SNPs: Discovery of Markers for Disease. *Biotechniques* . 32 (S), p4-13

Woese C. (2002). On the Evolution of Cells. *Proc Natl Acad Sci* . 99 (13), p8742-7

Zlatanova J, Mirzabekov A. (2001). Gel-immobilized microarrays of nucleic acids and proteins. Production and application for macromolecular research. *Methods Mol Biol* . 170, p17-38

ACKNOWLEDGEMENTS:

This work was supported by the Defense Advance Research Projects Agency and the Defense Threat Reduction Agency.