Microarrays for Highly Multiplexed Microbial Detection and Resistome Analysis

Tomasz Leski
Center for Biomolecular Science and Engineering (CBMSE)
Naval Research Laboratory, Washington, DC
tomasz.leski@nrl.navy.mil

Learning objectives

After this presentation, you should be able to:

1. Identify different types microarray-based technologies used for detection and characterization of microbial pathogens.
2. Describe the unique advantages of microarray-based detection compared to other highly multiplexed technologies.
3. Define the cost and real-world technical challenges associated with application of microarrays for microbial detection and genomic characterization.

Overview

1. Origins and early history of the microarray technology
2. Types and applications of DNA oligonucleotide microarrays
3. Advanced microarrays for microbial detection and characterization
4. Two microarray platforms developed in NRL
5. Future of diagnostic microarray technology

Early binding assays

- Bio-macromolecules as diagnostic tools
  - Nucleic acid based
    - Southern blot, northern blot
  - Protein/antibody based
    - Western blot, lateral flow assays, ELISA

- Early arrays resulted from attempts to develop multiplexed assays taking advantage of specific binding
Origins of microarrays

- 1975 – Edwin Southern introduces “southern blotting” technique for detection of DNA molecules
- 1979 – Kafatos et al. – dot blot technique
- Late 1980s – Roger Ekins and others develop theoretical background for microarray technologies
- 1989 – Saiki et al. – reverse dot blot technique
- 1991 – Fodor et al. – development of photolithography based combinatorial synthesis of peptides on solid support
- 1995 – Shena et al. – first application of DNA microarray (spotted microarrays)
- 1996 – Lockhart et al. – development of the first commercially available short oligonucleotide GeneChip

Microarray fabrication technologies

- Varying densities, feature sizes and probe types depend on manufacturing technology
- Microarray manufacturing technologies
  - Deposition/printing (spotted arrays)
  - In situ synthesis
  - electrasense (electrochemical synthesis)
  - micro-mirrors (NimbleGene)
  - ink-jet in situ synthesis (Agilent)
  - photolithography (Affymetrix)
- Suspension arrays (Luminex, Illumina’s Bead Array)
  - Mixtures of microscopic beads suspended in fluid
  - Beads of different specificities synthesized separately

(DNA) microarray terminology

- (DNA) microarray terminology:
  - Probe – nucleic acid molecule of known sequence, attached to solid surface (or bead in case of suspension arrays)
  - Target – nucleic acid molecule or collection of molecules to be analyzed, which may or may not be present in analyzed sample
  - Feature – in solid support based microarrays, a spot or area containing probes of the same specificity
  - Microarray – spatially defined array of microscopic spots (features) attached to solid support (in case of suspension bead arrays each bead type corresponds to a single feature).

Microarray technologies

- Long vs. short probes (sensitivity vs. specificity)
  - long probes – PCR products, cDNA, long oligos
  - highly sensitive, problems with purity, nonspecific binding, high setup cost
  - short oligo probes
  - highly specific and can be synthesized in situ, lower sensitivity, varying annealing temperatures
- Signal generation and detection
  - optical scanning (CCD, confocal microscopy)
  - electronic (NanoGen)
  - electrochemical (CustomArray ElectraSense technology)
Microarray applications

- Gene expression profiling – mostly research
- SNP/variability detection arrays (VDA)
  - drug resistance tracking
  - genetic disease diagnostics/characterization
- Broad range target detection
  - biothreat detection
  - drug resistance profiling
  - infection diagnostics

Broad-range pathogen detection and characterization

- Need for rapid broad-range detection of infectious agents
  - for many infectious syndromes, symptom-based ID of pathogens is difficult
  - hypothesis based detection is not cost effective
  - traditional culture based ID methods are slow and miss large groups of pathogens
  - molecular methods sensitive but usually optimized for detection of only one or a few agents
- Appropriate therapeutic regimes often rely upon characteristics of the identified pathogen
  - determination of drug resistance of bacterial, viral and parasitic pathogens is necessary for effective therapy

Technologies for broad-range microbial identification arrays

- Sample preparation strategies
  - whole genome amplification (RT-PCR/Phi29)
  - multiplexed specific PCR
- Data analysis algorithms
  - pattern based (majority of microarray assays)
  - resequencing (RPM, PathogenID, CustomArray)

Selected microarray assays for broad range microbial identification

- GreeneChip
  - 60-mer oligos/Agilent inkjet
  - 10 to 30 thousand probes
  - varying designs including panmicrobial assay
- RPM/PathogenID
  - 25-mer oligos/Affymetrics GeneChip
  - >1 million probes
  - resequencing, approx. 200 targets
- ViroChip
  - 70-mer oligos/spotted and Agilent inkjet
  - 1.6 to 60 thousand probes
  - varying designs for viral detection and discovery
- LLMDA
  - 50-65-mer oligos/NimbleGen synthesis
  - 388 thousand probes (LLMDA v.2)
  - panmicrobial assay
NRL developed microarray assays

Resequencing Pathogen Microarray (RPM)
- Millions of probes
- Affymetrix GeneChip platform
- Optical scan based detection
- Sequence based ID

Antimicrobial Resistance Determinant Array (ARDM)
- Thousands of probes
- CustomArray platform
- Electrochemical detection
- Pattern based ID

Photo courtesy of Dr. J. Golden

Resequencing microarrays for broad-range detection
- Diagnostic and surveillance tool that can test for hundreds or thousands of agents in parallel by directly accessing sequence information and finding matches in sequence databases.

Chip
Segment
Tile

Influenza B virus segment 4 hemagglutinin (HA)
RPM workflow

- Pooled Multiplex PCR products
- DNA fragments
- Biotin (3) Labeled DNA fragments
  - Fragmentation of DNA: 15 min
  - Terminal Labeling with Biotin: 35 min
  - Hybridization to Complementary Sequences on Microarray: 4 or 16 hours
  - Washing & Staining of Microarrays with Fluorescent Dye: 1 hour 45 min
  - Analysis

Intensities are converted to sequence for BLAST analysis & subsequent organism identification.

RPM microarrays

<table>
<thead>
<tr>
<th>RPM-Flu 3.1</th>
<th>RPM-TEI 1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 viral pathogens</td>
<td>75 pathogens</td>
</tr>
<tr>
<td>22 bacterial pathogens</td>
<td>13 toxins/resistance genes</td>
</tr>
<tr>
<td>188 targets in total</td>
<td>187 targets in total</td>
</tr>
</tbody>
</table>

RPM – identification and strain level discrimination

- Rapid diagnosis for “flu-like” syndromes
  - pathogen identification using single assay
  - detection of biothreat agents
- Serotype level identification of Influenza (RPM-Flu 3.1)
  - only some serotypes are highly pathogenic
  - detection of drug resistance markers
  - FDA emergency use authorization for 2009 pandemic strain

RPM-Flu 3.1 – Influenza coverage

Capability to identify 16 HA and 9 NA subtypes independently allows for discrimination between 120 different serotypes.*

*Metzgar 2010 PLoS One 5,e8995
Pathogen detection in clinical and environmental samples

**Respiratory pathogens in urban Population**
Assay: RPM-Flu 1.0
>400 nasal wash specimens analyzed using microarray and compared with reference assays. 98% agreement with reference methods. Multiple pathogen detection and serotype level identification of influenza.

**Respiratory pathogens in military recruits**
Assay: RPM-Flu 3.1
High prevalence of adenovirus in febrile respiratory cases.

Negative correlation between adenovirus and rhinovirus.

Multi-organism colonization: H. influenzae, S. pneumoniae, N. meningitidis and HRV (human rhinovirus)

**Environmental samples from Kuwait and Iraq**
Assay: RPM-TEI 1.0
Human pathogens: rapidly growing mycobacteria (RGM), Coxiella burnetii, Clostridium perfringens, Bacillus and Brucella in inhalable fractions of soil and airborne dust.

Respiratory pathogens in military recruits**
Assay: RPM-TEI 1.0
High prevalence of adenovirus in febrile respiratory cases.

Negative correlation between adenovirus and rhinovirus.

Multi-organism colonization: H. influenzae, S. pneumoniae, N. meningitidis and HRV (human rhinovirus)

Environmental samples from Kuwait and Iraq***
Assay: RPM-TEI 1.0
Human pathogens: rapidly growing mycobacteria (RGM), Coxiella burnetii, Clostridium perfringens, Bacillus and Brucella in inhalable fractions of soil and airborne dust.

Use of broad-range diagnostics in developing world

- Use of broad-range detection methods for identification of pathogens is feasible in low-resource settings.
- Use of these methods leads not only to rapid confirmation of presence and absence of suspected pathogen, but enables detection of pathogens that were not expected in the same step.
- In many situations use of more complex and expensive broad-range detection methods is more cost-effective and less logistically challenging than using a large number of simple assays.

Antibiotic resistance: a threat to modern medicine

- World Health Organization (2013):
  - “Antimicrobial resistance … threatens the success and continuation of clinical medicine as we know it.”
  - “Effective surveillance is the cornerstone of …international efforts to control antimicrobial resistance.”

  - 1st report on global antimicrobial resistance trends
  - High levels of resistance found in all regions
  - 2-fold greater mortality from resistance in select species
  - ~1.5% decrease in gross domestic product

Sierra Leone experience

Mercy Hospital Research Lab (MHRL)
Bo, Sierra Leone*
- operational RPM/GeneChip microarray system deployed in 2010
- Investigation of suspected avian influenza outbreak at a poultry farm in Bo

**Samples**
Outbreak farm, Bo (34)
Reference, Freetown area (68)

**Results**
No influenza detected.
Klebsiella pneumoniae found almost exclusively in samples from outbreak farm in Bo.

*Leski 2012, Health Bus Policy 40, 19-22

<table>
<thead>
<tr>
<th>Farm location</th>
<th>Number of samples</th>
<th>Sample designation</th>
<th>Condition of chickens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bo</td>
<td>34</td>
<td>34</td>
<td>Sick</td>
</tr>
<tr>
<td>Hamilton</td>
<td>20</td>
<td>20</td>
<td>All healthy</td>
</tr>
<tr>
<td>Hastings</td>
<td>20</td>
<td>20</td>
<td>All healthy</td>
</tr>
<tr>
<td>Allen Town</td>
<td>16</td>
<td>16</td>
<td>All healthy</td>
</tr>
<tr>
<td>Joe Town</td>
<td>12</td>
<td>12</td>
<td>Some chickens sick</td>
</tr>
<tr>
<td>Wellington</td>
<td>12</td>
<td>12</td>
<td>Some chickens sick</td>
</tr>
</tbody>
</table>

**Use of broad-range diagnostics in developing world**

**Respiratory pathogens in urban Population**
Assay: RPM-Flu 1.0
>400 nasal wash specimens analyzed using microarray and compared with reference assays. 98% agreement with reference methods. Multiple pathogen detection and serotype level identification of influenza.

**Respiratory pathogens in military recruits**
Assay: RPM-Flu 3.1
High prevalence of adenovirus in febrile respiratory cases.

Negative correlation between adenovirus and rhinovirus.

Multi-organism colonization: H. influenzae, S. pneumoniae, N. meningitidis and HRV (human rhinovirus)

**Environmental samples from Kuwait and Iraq**
Assay: RPM-TEI 1.0
Human pathogens: rapidly growing mycobacteria (RGM), Coxiella burnetii, Clostridium perfringens, Bacillus and Brucella in inhalable fractions of soil and airborne dust.
Antimicrobial Resistance Determinant Microarray (ARDM)

DNA microarray for simultaneous detection of broad range of antibiotic resistance genes and mechanisms versus phenotypic testing.

>200 resistance genes (v.3: >500 genes)

Gram-positives (Streptomyces, Bacillus, Staph, Clostridium)
Gram-negatives (Bacteroides, E. coli, Vibrio, Pseudomonas)
15 classes of antibiotics

ARDM - technology

CustomArray/ElectraSense
- 2240 probes/sub-array
- 4 subarrays/chip
- Electrochemical detection

ARDM v. 2.0
- 238 genes
- 1 array per assay

ARDM v. 3.0
- 540 genes
- 2 arrays per assay

ElectraSense detection

Source: Combimatrix (www.combimatrix.com)

ARDM – benefits vs. limitations

- **Benefits**
  - Isolate ID not required
  - Broad coverage (many species)
  - Detects multiple mechanisms for resistance
  - Detects unexpected genes (e.g., tet(X), qnrVC)
  - Tracks movement of horizontally transferred genes

- **Limitations**
  - Cannot detect point mutations
  - Quinolone resistance
  - ESBLs vs non-ESBLs
  - Genotype ≠ phenotype
  - Current status: still time- and labor-intensive
  - 5h hands-on time
  - 20h time-to-result (incl. hybridization)
Portable microarray platform

- ElectraSense scanner
  - small size
  - no separate power supply needed
  - no requirement for alignment/calibration
- Sample and microarray processing (in progress)
  - fluids station for microarray processing
  - simplification and automation of sample processing

ARDM application - resistance genes in unexpected contexts

Cambodia*
- Ralstonia pickettii (wound isolate)
  - opportunistic pathogen
  - respiratory infections
  - immunocompromised
  - CF patients
- 17 resistance genes detected (ARDM v.3.0)
- 3 genes detected for the first time in *R. pickettii*
  - *rrmB* (pan-resistance to aminoglycosides)
  - *qnrVC* (FQ resistance, typical for *Vibrio*)
  - *bla*<sub>IMP-2</sub> (carbapenem resistance)

Sierra Leone**
- Multiple species (urine isolates)
  - Enterobacteriaceae
  - Pseudomonadaceae
- *tetX* gene
  - broad range tetracycline
  - inactivates tigecycline
- First report of *tetX* in human pathogens from clinical samples

ARDM application – diarrheal pathogen surveillance

Trimethoprim resistance genes (*Enterobacteriaceae*)

Microarrays as in vitro diagnostic platform

- PCR – most common molecular diagnostic technology
  - rapid, sensitive, easy to perform, inexpensive
  - limited multiplexing potential
  - ideal for direct detection of small number of targets
- Microarrays – highly parallel detection platform
  - broad-range target detection and near-neighbor discrimination
  - complex processing, pricey
  - syndromic diagnostics, bioterror monitoring
- Majority of FDA approved in vitro diagnostic (IVD) assays based on PCR with increasing number of microarray assays.

*Heang 2014 Int J Antimicrob Agents 44, 81
**Leski 2013 Int J Antimicrob Agents 42, 83
Microarrays vs. HTS

- High Throughput Sequencing for pathogen detection
  - background/non-relevant genes account for >99% of sequences in many sample types (looking for needle in a haystack)
  - HTS run "sees" only a small fraction of the whole sample

Example: whole blood specimen containing *B. anthracis* at 10^3 cells/µL

DNA content of typical blood sample ~50ng/µL = 5 x 10^13 bp human DNA/µL

*B. anthracis* @ 10^3 cells/µL = 5.2 x 10^9 bp/µL bacterial DNA = (0.01% total bp)

Typical output of HiSeq 2500 sequencer = ~ 100 Gb (1.0 x 10^11 bp)

This is enough to sample = 0.22 % of total DNA content of this specimen

Future of diagnostic microarrays in pathogen detection

Enthusiasm of late 1990's

"Because of the predicted central role of microarrays in biomedical research, some experts believe that the biochip revenues will eventually eclipse the sales of computer chips."

Introduction of HTS technologies in ~2005 – significant drop of interest in microarray technologies

Microarrays still method of choice in many areas including pathogen identification and characterization

Emergence of technologies combining microarrays HTS and other technologies

Acknowledgments

**RPM technology**

- David Stenger
- Baohuan Lin
- Anthony Malanoski
- Zheng Wang
- Joel Schnur
- Nina Long
- Kate Bianey
- Carolyn Meador
- Clark Tabetts (TessArae)
- Klaus Schafer (TessArae)
- Brian Weslowski (TessArae)

**ARDM technology**

- Chris R. Taitt
- Gary Vora
- Brian Barrows
- CDR Michael Stockelman

**Mercy Hospital Res Lab (Sierra Leone)**

- Rashid Ansumana
- David Jimmy
- Umaru Bangura

**Environmental samples (Iraq/Kuwait)**

- LCDR Michael Gregory

**NAMRU-2 (Cambodia)**

- LCDR Michael Proudy
- LT Gavin Ford
- Vireak Hsirang
- CDR Steven Newell

*Schena, 1998 Trends Biotechnol 16, 361*