Microarrays in Three Easy Steps

Priti Hegde
Microarray Analysis Stages

• Array Fabrication

• Probe Preparation and Hybridization

• Data Collection, Normalization, and Analysis
Array Fabrication

**Microbial**
- ORFs
- Design PCR Primers
- PCR Products

**Eukaryotic**
- THC
- Select cDNA clones
- PCR Products

Microtiter Plate

Microarray Slide
(as many as 10,000 or more spotted genes)

Many different plates containing different genes

For each plate set, many identical replicas

TIGR The Institute for Genomic Research
Development of the TIGR “30k cDNA Gene Set”

**Goal:**
Array 30,000 genes and study gene expression in human cancer to develop stage and tissue specific expression fingerprints.

**Progress:**
- Nearly 48,000 cDNA clones have been selected using the EST Assemblies (THCs) in the TIGR Human Gene Index*; 40,000 have been amplified by PCR and are ready for use in array studies. Funding has been secured to expand to 60,000 clones.
- Priority has been given to arraying known genes and genes with mapping information.
- Additional clones have been chosen representing genes of unknown function.
- Pilot studies are underway with 7,200 and 19,200 clone arrays.

• **PCR Amplify Them**
  - Grow clones overnight
  - Dilute 1:20 (5μl:95μl) in water
  - “Pop” the cells, spin out debris
  - Amplify in 50 μl reaction with Platinum Taq (Life Technologies)

• **Purify the PCR products**
  - 96-well Millipore multiscrreen glass filter plate
  - Bind products in high salt (1:5 5.3M Guanidine-HCl/150 μ M KAc)
  - Elute in water/TE
Microarray PCR Scoring Tool

MAGE PCR Scoring - Plate 25

88% Good
6% Questionable
6% Bad
The Beast: Microarray Robot from Intelligent Automation

- Array The Clones

<http://www.ias.com>
The Beast in Action: #2
• The Glass is Crucial!
Corning CMT-GAPS Slide

Another Slide

TIGR The Institute for Genomic Research
The Effects of Spotting Buffer and PCR Clean-up

• Spotting “Ink” and Clean-up matter

- 50% DMSO
- 3xSSC

- Ethanol Precipitation/DMSO
- Glass Filter Cleanup/DMSO
When Temperature and Humidity Go Bad

• Laboratory Conditions Matter

- 72°F (22.2°C) 40-45% Relative Humidity
- 62°F (16.7°C) 40-45% Relative Humidity
- 80°F (26.7°C) 80-85% Relative Humidity
- 62°F (16.7°C) 40-45% Relative Humidity

TIGR The Institute for Genomic Research
• **Arraying Conditions**

  • **72°F (22.2°C), 40-45% Relative Humidity**
  
  • **50% dimethyl-sulfoxide (DMSO), 20mM Tris HCl, 50mM KCl, pH 6.5**
  
  (Thanks to Robin Cline, Erik Snesrud, Karen Ketchum)

  • **Corning CMT-GAPS silane coated slides**

  • **UV Cross-link at 90 mJ, bake at 80 °C for 2 hours**
Microarray Overview II

Obtain RNA Samples

Prepare Fluorescently Labeled Probes

Control

Test

Hybridize, Wash

Measure Fluorescence in 2 channels red/green

Analyze the data to identify differentially expressed genes
Comparison of Labeling Protocols

• Preparing Labeled Probe

20µg total RNA

4µg total RNA

~1.5µg Poly(A) RNA (Seradyne Beads)
Total and Poly(A) RNA give equivalent results
• Labeling Protocols
  • Oligo(dT) primed labeling using SuperScript II (Life Technologies) and 4-8 μg total RNA
  • Clean Probes using Pharmacia GFX Columns

• Hybridization
  • Prehybridize slides with 5×SSC, 0.1% SDS and 1% BSA to block free amine groups
  • Combine probes with 10×SSC, 0.2% SDS, 50% formamide and hybridize at 42°C O/N
Confocal Laser Scanner: ScanArray 3000
• Data Collection, Normalization, and Analysis
  General Scanning <http://www.genscan.com>
Image Processing Issues

- Spot Finding
- Background Subtraction
- Reproducibility
- Measure - median vs. mean (integrated intensity)
- Quality measures
Data Analysis Issues

- Presentation
- Multiple Views
- Normalization
- Identification of Differentially Expressed Genes
- Multiple Experiments
Microarray Data Display Software

Software displays array data with links to database information about the underlying genes.

Differentially expressed clones can be selectively displayed.
Microarray Data Display Software

Pseudo-false color display allows assessment of hybridization signal strength

TIGR The Institute for Genomic Research
Microarray Data Display Software

Relative Red/Green areas represent relative expression levels

Setting expression ratios show only over/under expressed genes
Data Display/Analysis Software

TIGR The Institute for Genomic Research
Normalization Approaches

• Total Intensity
• Linear Regression
• Ratio statistics described by Chen et al.,


**Any of these using:**

• Entire Data Set
• User-defined Data Set/Controls
Normalization Approaches (II)

Entire Data Set

- Probe Quantification less important
- No assumption on which genes constitute “housekeeping” set
- Uses all the data
- No independent confirmation

User-defined Data Set/Controls

- Requires definition of “housekeeping” set
  - or good added controls
- Requires good RNA quantitation
- Ignores much data
Normalization Approaches (III)

Solution(?)

- Experiment dependent
- Use a combination of techniques
TIGR MultiExperiment Viewer: Data Mining
TIGR MultiExperiment Viewer: Self Organizing Maps
TIGR MultiExperiment Viewer: Self Organizing Map Views

Gene Cluster Analysis Software Package for Microarrays
The Institute for Genomic Research - SOM Result Window
TIGR MultiExpriment Viewer: Principal Component Analysis
Hybridization to a 19,200 Element Human Array
Hybridization to the *Arabidopsis* Chromosome II array
TIGR Human Expression Team
   Kristie Abernathy*
   Sonia Dharap
   Julie Earle-Hughes
   Renee Gaspard
   Cheryl Gay
   Priti Hegde
   John Quackenbush
   Alex Saeed
   Vasily Sharov

H. Lee Moffitt Center/USF
   Tim Yeatman

TIGR Collaborator
   Norman Lee

Funding provided by the National Cancer Institute:
   (R01 CA77049-01; PI: J. Quackenbush)

CMT-GAPS Slides supplied by Corning