**Soybean Microarrays**

An Introduction

By Steve Clough

November 2005

---

**Common Microarray platforms**

- **cDNA**: spotted collection of PCR products from different cDNA clones, each representing a different gene
- **Oligo**: spot collections of oligos, usually 50-70 bp long that span known/predicted ORFs. May have one or more oligos representing each gene.
- **Affy**: Affymetrix gene chips, 25 bp oligos 11 per gene predicted to span ORF
**Pros and Cons of cDNA platforms**

- **cDNA**: need to construct cDNA libraries from a variety of tissues and conditions and to sequence to verify lack of duplication.

- Cheapest approach. Do not need to have a sequenced genome.

- Hybridization involves strands of hundreds of bases, therefore less specificity in binding and cannot differentiate multigene family members. Good if your organism is closely related but not identical to one used to make the cDNA libraries used to make arrays.

---

**Pros and Cons of Oligo-based platforms**

- **Oligo**: spot collections of oligos, usually 50-70 bp long that span known/predicted ORFs. Affymetrix chips use 25mers and 11 or so probes per ORF.

- Need lots of sequence information from your organism.

- Works best if your organism is same or very closely related to the one used to obtain the sequence information.

- More costly than cDNA arrays to manufacture.
Soybean cDNA Microarrays

Produced in the lab of Dr. Lila Vodkin, U of Illinois
l-vodkin@uiuc.edu

cDNA Library Synthesis
(representing expressed genes)

Extract RNA from variety of tissues and conditions

Clone cDNA into vector

Steve Clough, USDA-ARS
University of Illinois, Urbana
cDNA clone

Sequence cDNA

GCTCTAAGTCATCGTACTAGATCT
= protein kinase

Compare EST sequence to database to identify

Eliminate duplicates to generate set of unique clones

PCR amplify insert of unique clone set

Pipette PCR products into microtiter plates to print onto slides

Printing microarrays – picking up PCR samples
Printing PCR products on glass slides

Pin Washing Between PCR Samples
Spots of single-stranded DNA adhered to glass surface

Typically 10-25,000 spots are printed on a standard 1" x 3" microscope slide

Note: DNA does not bind well to glass so glass is specially coated to allow ionic binding (poly-lysine slides) or covalent binding (amine or aldehyde slides)

Steve Clough, USD-ARS
University of Illinois, Urbana

Fluorescently label cDNA from tissue of interest to hybridize to spots on the slide

Extract RNA

cDNA synthesis and fluorescent labelling

Steve Clough, USD-ARS
University of Illinois, Urbana
Direct labelling with Reverse Transcriptase

mRNA 5'------------------------AAAAAAA3'

RT

dATP
dGTP

dTTTTTTTT

dTTP
dCTP
ddUTP

mRNA 5’
cDNA

3’

AAAAAAA3’

TTTTTTT 5’

RNAse

cDNA

3’

TTTTTTT 5’

Steve Clough, USDA-ARS
University of Illinois, Urbana

Indirect labelling with aa-dUTP and Reverse Transcriptase

* * * * * *

* * * * * *

Steve Clough, USDA-ARS
University of Illinois, Urbana
**Indirect labelling with Klenow**

mRNA $5'$ \[\text{AAAAAAA}\] $3'$

\[\text{RT} \quad \downarrow \quad \text{dTTP} \quad \text{dATP} \quad \text{dGTP} \quad \text{dCTP} \quad \text{dTTP} \quad \text{T}^\prime\]

\[\text{mRNA} \quad 5' \quad \text{cDNA} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTTTTTT} \quad 5', \quad \text{dTTTTTTT} \quad 3', \quad \text{dTTP} \quad \text{dATP} \quad \text{dGTP} \quad \text{dCTP} \quad \text{dTTP} \quad \text{T}^\prime\]

Random hexamer primers

Klenow

RNase

Boil

Hybridization Chamber

Flourescently labelled sample is pipetted under the coverslip and allowed to hybridize to spots on slide.
Hybridization in Water Bath

Washing After Hybridization

15 minutes with shaking for each

1  
1X SSC  
0.2% SDS

2  
0.2X SSC  
0.2% SDS

3  
0.1X SSC

Spin dry  
2 minute  
500 rpm
Scan on a Fluorescent Scanner

Theory:
Spot A will fluoresce 3 times brighter than Spot B
Blocking slides to reduce background. Example, positively charged amine slides.

Wash with SDS to block charges and to remove excess DNA. Then place in hot water to generate single strands. Repeat SDS wash.

Steve Clough, USDA-ARS
University of Illinois, Urbana

False Coloring of Fluorescent Signal

Scale of increasing fluorescent intensities

2^1 2^16
2 65,536

Stronger signal (16 bit image)
Principles behind gene expression analysis

mRNA ➔ Fluorescent cDNA

- Hybridized to array of individual spots of different genes
- Fluorescent intensity of spot is proportional to expression level
- Labelled representation of all recently expressed genes

Fluorescent intensities from quality data (Background ~80)

Steve Clough, USDA-ARS
University of Illinois, Urbana
2 dyes with well separated emission spectra allow direct comparison of two biological samples on same slide.
Ratio of Expression of Genes from Two Sources

Cells from condition A

Label Dye 1

mRNA

cDNA

Cells from condition B

Label Dye 2

- equal
- higher in A
- higher in B

Steve Clough, USDA-ARS
University of Illinois, Urbana

Cy3 Scan

Cy5 Scan

Overlay

Steve Clough, USDA-ARS
University of Illinois, Urbana
Example: Cy3 scan of Uninoculated control

Steve Clough, USDA-ARS
University of Illinois, Urbana

Example: Cy5 scan of Pathogen inoculated sample

Steve Clough, USDA-ARS
University of Illinois, Urbana
Use software such as GenePix to extract data from image

1. Locate spots, define spot area, collect data from pixels within spots
2. Flags bad spots (ex: dust in spot)
3. Calculates ratio Cy5 fluorescent intensity over Cy3 intensity for each spot
4. Produces tab-delineated tables for import to analysis programs
Value of pixels within spot equals the raw data. Software will give pixel value related to fluorescence from both Cy3 and Cy5 scans.

Quick view of expression results per slide can be seen by examining scatter plots of Cy5/3 intensity ratios per spot.
Data analysis

One can identify genes with common expression patterns by hierarchical clustering.

Each horizontal line represents one gene.
**Plant Microarray Controls**

- **Negative**: mammalian genes
- **Positive**: high, medium, low expressers  
  tissue specific  
  ubiquitous
- **Miscellaneous**: transgenes  
  bacterial
- **Labelling efficiency**: spiked control mRNA--  
  genes that are non-homologous to plants.

---

**Negative Plant Controls**

- **Spotting solution**
- **Genes not present in plant**  
  Mammal-specific:  
  antibody / immunoglobulin  
  neuro-related  
  myosin  
  etc.
- **Verify ‘plant negative’ by BLAST against plant databases**
Positive Plant Controls

- Tissue specific, high expressers:
  - ex: cotyledon: conglycinin
  - roots: auxin down regulated gene 12
  - leaves: RUBISCO (small chain)

- Ubiquitous:
  - ex: ubiquitin (med-high)
  - EF1 (med-high)
  - DAD1 (low-med)
  - tubulin (med-high)

Labelling Efficiency Controls

- Spiked mRNAs:
  - Mammalian genes, non-homologous to plant genome.
  - Select several spiking controls (ex: 4).

To use:
- Include them on the array.
- Clone (with polyT tail) into a T7 or T3 expression vector
- Or PCR with T7 or T3 promoter attached to 5’ primer
- and a poly (dT) to the 3’ primer.
- Invitro transcribe with T7 or T3 RNA polymerase.
- Add this ‘mRNA’ to your labelling reactions--
  each one at a different concentration level
  to span the dynamic range of fluorescent intensities
cDNA Arrays vs Oligo Arrays

- **cDNA**: spot a collection of ESTs
- **Oligo**: spot collections of oligos that span known/predicted ORFs
  - need sequence info
  - only option for prokaryotes
  - ‘shagged rug’ spots

---

**Oligo-based Microarrays**

Design specific oligos for every ORF

**Spotted Microarray**
- Synthesize with 5’-amino linker
- Design one to multiple oligos/ORF
- Collect in 384-well plates
- Spot on aldehyde coated slides

**Affymetrix Gene Chips**
- Synthesize oligo directly on chip
- Proprietary photolithography synthesis
- 11 oligo/ORF plus mismatches
- Spotted oligo termed the ‘probe’

---

*Steve Clough, USDA-ARS
University of Illinois, Urbana*
GeneChip® Probe Arrays

GeneChip® Probe Array

Hybridized Probe Cell

Single stranded, labeled RNA target
Oligonucleotide probe

Millions of copies of a specific oligonucleotide probe

> 1,200,000 different complementary probes

Image of Hybridized Probe Array

Glycine max transcripts: 35,611
Phytophthora sojae transcripts: 15,421
Heterodera glycines transcripts: 7,431

Courtesy of Mike Leivelt