A systems approach to study plant disease resistance mechanisms

Fumiaki Katagiri

Dept. of Plant Biology
Univ. of Minnesota
Microbial and Plant Genomics Institute

Oct. 17, 2005 CBCB seminar Cargill Bldg.

Acknowledgements

Univ. of Minnesota
Masanao Sato
Raka Mitra
Jane Glazebrook
Supercomputing Institute
Zheng Jin Tu
High Throughput Biological Analysis Facility
Nick Hahn

Stanford Functional Genomics Facility
Duke University
Dong Wang
Natalie Weaver
Xinnian Dong
Carnegie Institution
Shauna Somerville
Harvard Medical Sch.
Julia Dewdney
Fred Ausubel

Funding:

CSREES
National Science Foundation
Inducible Defense of Plants

Pathogen

Effectors

Recognition

Signal Transduction

Arabidopsis plants infected with *Pseudomonas syringae* carrying *avrRpt2*

rps2−  RPS2+
How should we study a signaling network?

If we only observe inputs and outputs, we can never specify the mechanism in the black box, however many observations we make.

How should we study a signaling network?

So, we beat it up (perturbations) and try to see what changes occur inside the box (i.e., network).
Our strategy

• T-DNA insertion lines for genetic perturbations.

• A small scale microarray for a wide spectrum, quantitative measurements.

“Mini”-array

• Long oligo (50-70 mers) spotted array
• 464 pathogen-responsive genes – representing diverse expression patterns
• 107 normalization genes for array-to-array normalization – representing a wide range of expression levels
• 5 spiking oligos for quantitation check
• Tracking oligo for pixel-to-pixel calibration
• Locally factorial, globally overlapping design, which allows use of a global linear model.
Locally factorial, globally overlapping design

A single subarray has 144 probes x 2 = 288 spots

In each tile, 72 probes x 2 are common in each of the subarrays.

Technical reproducibility
(Mean, 1 pen x duplicates)

R = 0.945
slope = 0.94
y-intercept = -0.19

R = 0.984
slope = 1.00
y-intercept = -0.02
Technical reproducibility of Miniarray
(Mean, 4 pens x duplicates)

R = 0.990
slope = 1.06
y-intercept = 0.40

R = 0.973
slope = 0.96
y-intercept = -0.32

Statistical model for the miniarray expression value (1)
- linear model

\[ S_{ijr} = Const + A_i + B_j + C_r + E_{jr} + \varepsilon_{ijr} \]

\[ \sum_i A_i = \sum_j B_j = \sum_r C_r = \sum_j \sum_r E_{jr} = 0 \]

Indices: \( i \), probe (gene); \( j \), pen (subarray); \( r \), replicate
\( S_{ijr} \), Log-transformed measured value (median of ratios)
\( A_i \), probe contribution; \( B_j \), pen contribution; \( C_r \), replicate contribution;
\( E_{jr} \), pen-replicate interaction; \( \varepsilon_{ijr} \), error
Technical reproducibility of Miniarray
(linear model)

R = 0.996
slope = 1.04
y-intercept = 0.20

R = 0.978
slope = 0.95
y-intercept = -0.12

Statistical model
for the miniarray expression value

\[ S_{ijr} = Const + A_i + B_j + C_r + E_{jr} + f_j(x_j, y_j) + \varepsilon_{ijr} \]

\[ f_j(x_j, y_j) = \sum_{i=1}^{4} \sum_{w=0}^{v} p_{jvw} x_j^v y_j^w \]

\[ \sum_i A_i = \sum_j B_j = \sum_r C_r = \sum_j \sum_r E_{jr} = 0 \]

Indices: \( i \), probe (gene); \( j \), pen (subarray); \( r \), replicate
\( S_{ijr} \), Log-transformed measured value (median of ratios)
\( A_i \), probe contribution; \( B_j \), pen contribution; \( C_r \), replicate contribution;
\( E_{jr} \), pen-replicate interaction; \( f_j(x_j, y_j) \), spatial correction within \( j \)th subarray; \( \varepsilon_{ijr} \), error
Technical reproducibility of Miniarray
(linear model + spatial smoothing function)

R = 0.982
slope = 0.96
y-intercept = -0.14

R = 0.994
slope = 1.06
y-intercept = 0.37

Distribution of spot-by-spot errors

The mean of measured values as the expression value for each gene (variance = 0.18)

After correction with a linear model only (variance = 0.13)

After correction with a linear model and a spatial smoothing function (variance = 0.062)
Spatial error distribution
(same color scale)

Correlation in the expression ratio between miniarray and Affy data
(linear model + spatial smoothing function)

R = 0.846
slope = 1.08
y-intercept = 0.37
The major cause of the discrepancy is a limited dynamic range of the miniarray at the low end.

Collaboration style in systems biology research
What’s systems biology?  
- My definition

- Study the topology and dynamics of biological networks that underlie biological phenomena.
  - It is not just systematic collection of biological data.

International study of systems biology research

- Panel members: Marvin Cassman (Chair), Adam Arkin (UC Berkeley), Frank Doyle (UCSB), Doug Lauffenburger (MIT), Cindy Stokes (Enteros), Fumi Katagiri (U of Minn)
- Sponsored by NSF, DOE, DARPA, NASA, NCI, NIBIB, NIST, and EPA.
Relatively little instructed global profiling and/or systematic approaches

Other pre-existing data

Validation, Refinement, Parameters, …

Highly instructed, focused experiments

Prediction, Demand, …

Early stage

Later stage

Experimental

Computational/Theoretical

Network inference

Modeling

Old style - Division of labor

Theoreticians/
Computer scientists

Experimentalists

Areas of expertise
Desired style – Highly interactive

An alternative
Welcome to the website of
Plant Biology 5960 sec 002: Introductory Bioinformatics

Meeting time: Monday, Wednesday, Friday, 11:40-12:50 AM
Meeting place: McIlvaine B20
Instructor: Dr. Pamela Komor
Office hours: by appointment
Office address: 316 Old University, B. 206

http://www.cbs.umn.edu/class/fall2005/pbio/5960-2/