Use of microarrays to study gene expression

Part II: Using microarrays to identify differentially expressed genes during defense to a bacterial plant pathogen

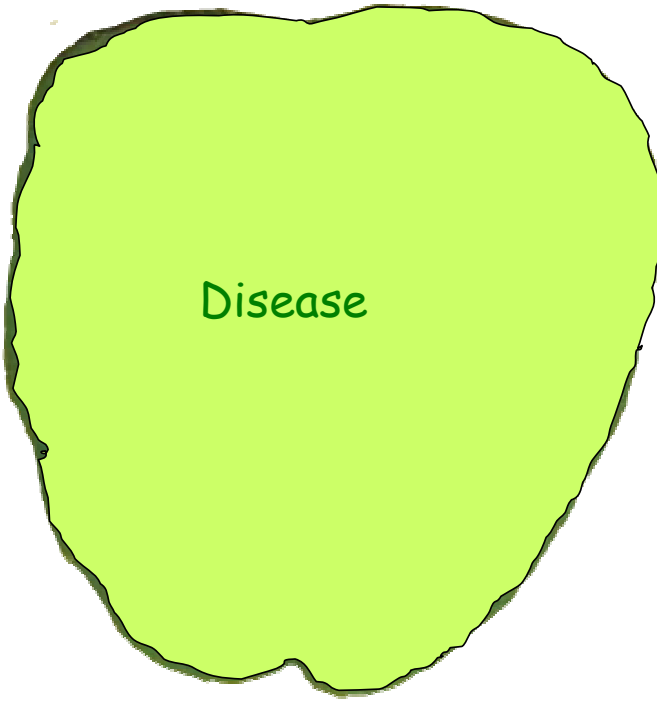
-Steve Clough
sjclough@uiuc.edu

Bacterial Blight (*Pseudomonas syringae*)

Steve Clough, USDA-ARS
University of Illinois, Urbana

Hypersensitive Response (HR)

Susceptible (S)



Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

How do we minimize variation not related to specific treatment?

Can the experiment be statistically validated?

What can we afford?

How do we verify the results?

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

- Best way to inoculate bacteria into leaves is probably vacuum infiltration.

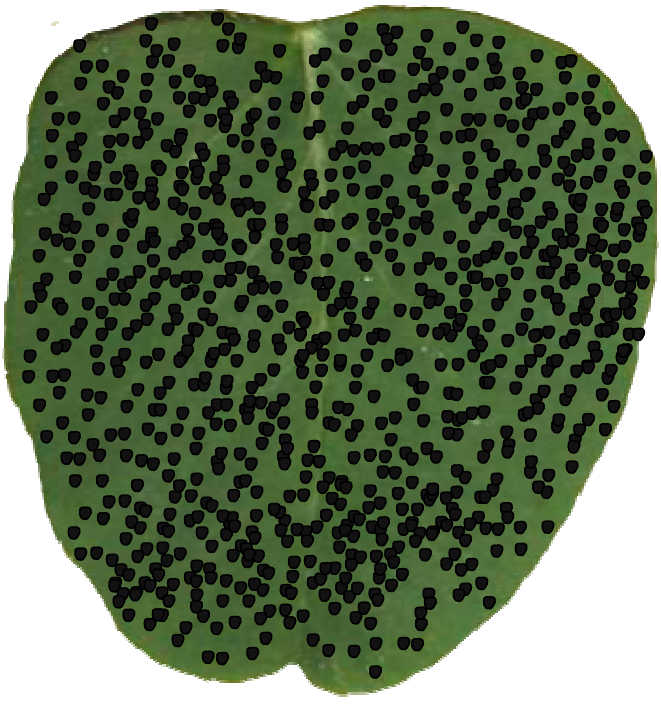
  Little / no wounding
  Even inoculation
  Can do multiple plants at same time

- Use of growth chambers allow controlled environment that is relatively easy to reproduce.

- True biological replicates are best obtained by completely reproducing the experiment for each rep.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Vacuum infiltrate inoculum to maximize number of cells in contact with bacteria for maximal signal



Steve Clough, USDA-ARS
University of Illinois, Urbana

Submerge plants into inoculum within a bell jar

Steve Clough, USDA-ARS
University of Illinois, Urbana

Apply strong vacuum to release air from interior of leaves

Steve Clough, USDA-ARS
University of Illinois, Urbana

After about 60-90 seconds of 'leaf boiling', quick release vacuum

Steve Clough, USDA-ARS
University of Illinois, Urbana

Infiltration is obvious when viewed from underside of leaves.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

Genes that are important in resistance response.
Genes that are important in susceptible response.

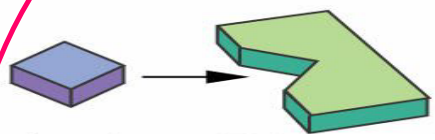
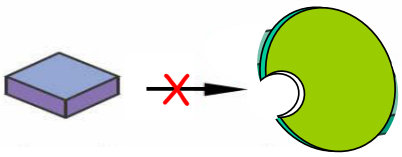
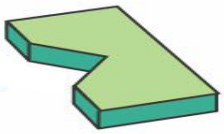
So we need to know something about timing of defense and we need to know what host/pathogen combination to use.

We can use same host, but use different pathogen:
one induces HR resistance
one induces disease

We also need to add an infiltration control and null control

Steve Clough, USDA-ARS
University of Illinois, Urbana

Flor's gene-for-gene model

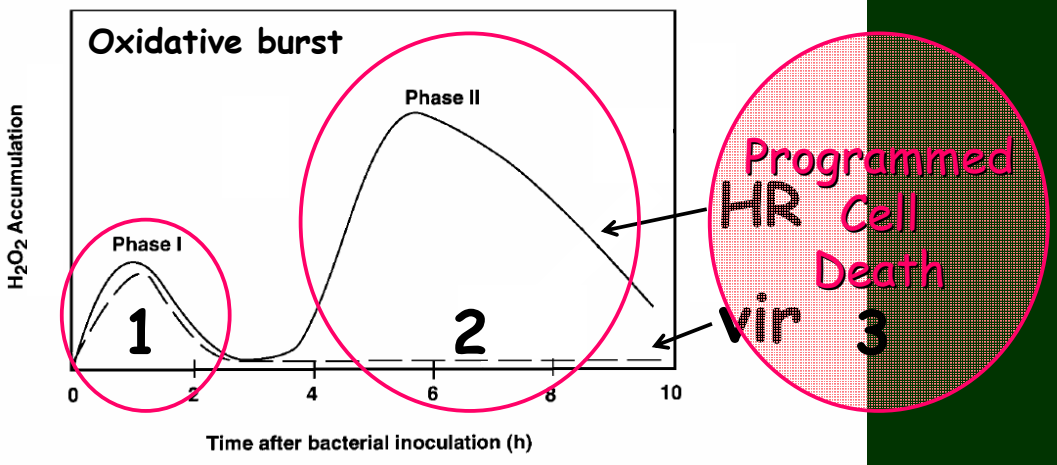
Pathogen genotype	Host plant genotype	
	Dominant R gene	*Recessive r gene*
avr	 AVR R protein Resistance (incompatible)	 AVR r protein Disease (compatible)
avr	 R protein Disease (compatible)	 r protein Disease (compatible)

Steve Clough, USDA-ARS
University of Illinois, Urbana

Bacterial Blight (*Pseudomonas syringae*)

Three Phases associated with HR resistance

Adapted from: Lamb and Dixon 1997 Annu. Rev. Plant Physiol. Plant Mol. Biol. 48 : 251-275



Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

How do we minimize variation not related to specific treatment?

Host is genetically identical for all samples.
Leaves at a developmental stopping stage.
Pool multiple leaflets per sample.
Minimize handling after treatment.
Rapidly freeze tissue in liquid nitrogen immediately after collecting.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

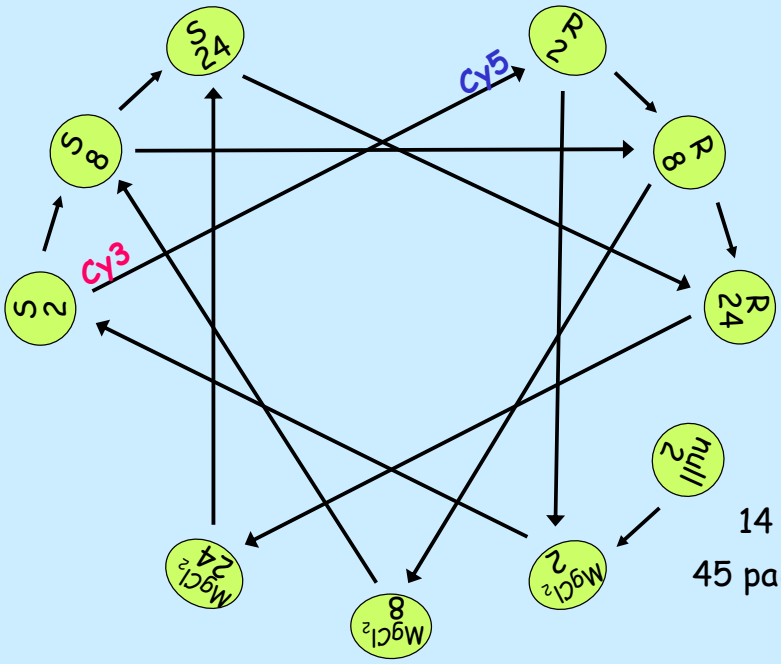
How do we minimize variation not related to specific treatment?

Can the experiment be statistically validated?

Want to determine significance due to treatment.
Want to determine significance due to time.
Want to be able to determine if expression affected by dye.
Want to determine significance due to infiltration.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design



10 samples
14 slides per library
45 pair-wise comparisons

Steve Clough, USDA-ARS
University of Illinois, Urbana

Statistical Analysis

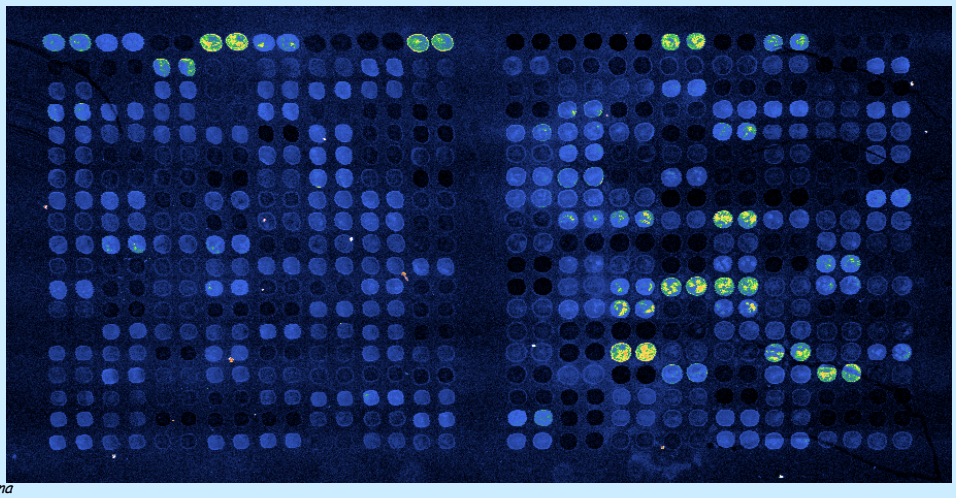
- Used median intensity of pixel values minus background for each spot
- Log 2 transformed
- Adjusted the data from the two scans using glowess normalization:
  smoothes the scatter plot of Ratio (R/G) versus Intensity (R x G)
- Ran linear mixed effects model in SAS
  - Fixed: Dye and sample (time-treatment combination)
  - Random: Array
- P values adjusted for multiple testing using the FDR criterion of Benjamini and Hochberg
- Identified spots whose ratios significantly differed from the mean of the data set for the entire experiment (found 3897 using p-value cutoff of 0.000005)
- Identified genes within the 3897 whose ratios significantly different from the mean of the data set when in direct comparisons (i.e. contrast between 2 specific samples)

Steve Clough, USDA-ARS
University of Illinois, Urbana

Statistical Analysis

- Used median intensity of pixel values minus background for each spot

Currently we do not subtract background, but flag spots if affected by it
Because background not necessarily related to non-specific binding to
Spotted DNA, but to non-specific binding to slide coating

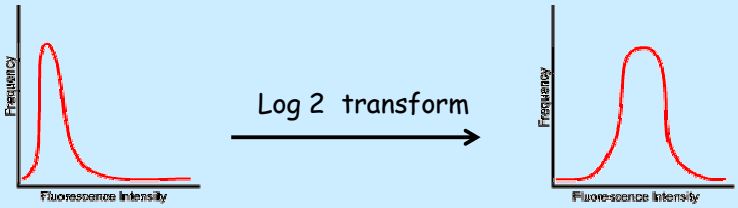


teve Clough, USDA-ARS
niversity of Illinois, Urbana

Statistical Analysis

- Used median intensity of pixel values minus background for each spot
- Log 2 transformed

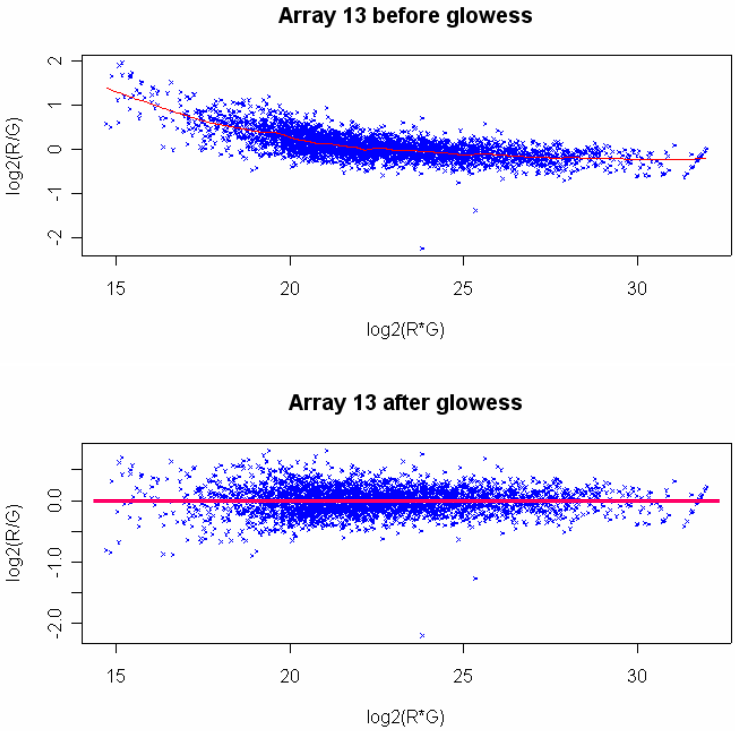
  Removes skew in the data (see plots before and after)



- Adjusted the data from the two scans using glowess normalization:
  smoothes the scatter plot of Ratio (R/G) versus Intensity (R x G)

Steve Clough, USDA-ARS
University of Illinois, Urbana

Normalized data shown as M/A plots



Steve Clough, USDA-ARS
University of Illinois, Urbana

Statistical Analysis

- Used median intensity of pixel values minus background for each spot
- Log 2 transformed
- Adjusted the data from the two scans using glowess normalization:
  smoothes the scatter plot of Ratio (R/G) versus Intensity (R x G)
- Ran linear mixed effects model in SAS
  - Fixed: Dye and sample (time-treatment combination)
  - Random: Array

$$\text{Log}_2\ \text{Intensity}_{ijkl} = \mu + \text{Array}_i + \text{Dye}_j + \text{Sample}_k + \varepsilon_{ijkl}$$

This equation determines the effect of the means of each independent variable (Array, Dye, and Sample) on the dependent variable (log2 Intensity) as well the effect of random noise (ε). The model was written in this manner because we were interested in knowing how genes were changing due to inoculation with the different bacteria at specific time points of the time course experiments, therefore sample is equal to the treatment and time interaction. For example, we wanted to know how the plants were responding during both the resistant and susceptible interactions at 8 hours post inoculation.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Statistical Analysis

- Alternatively, one could use the following model to allow determine significant genes due to treatment effect as well as time effect separately

$$\text{Log}_2\ \text{Intensity}_{ijklm} = \mu + \text{Array}_i + \text{Dye}_j + \text{Treatment}_k + \text{Time}_l + \text{Treatment*Time}_{kl} + \varepsilon_{ijklm}$$

This equation determines the effect of the means of each independent variable as in the previous model but separates the Sample variable into Treatment and Time (Array, Dye, Treatment and Time) on the dependent variable (log2 Intensity) as well the effect of random noise (ε). The model was written in this manner because we were interested in knowing how genes were changing due to inoculation with the different bacteria at specific time points of the time course experiments as well as how they changed independently. For example, we wanted to know how the plants were responding during both the resistant and susceptible interactions at 8 hours post inoculation as well as how this 8 hour time point differed from the 2 hour time point and how treatment A differed from treatment B.

Steve Clough, USDA-ARS
University of Illinois, Urbana

Statistical Analysis

- Used median intensity of pixel values minus background for each spot
- Log 2 transformed
- Adjusted the data from the two scans using glowess normalization:
  smoothes the scatter plot of Ratio (R/G) versus Intensity (R x G)
- Ran linear mixed effects model in SAS
  - Fixed: Dye and sample (time-treatment combination)
  - Random: Array

$$\log_2 Y = \mu + \text{Array} + \text{Dye} + \text{Sample} + \varepsilon$$

- P values adjusted for multiple testing using the FDR criterion of Benjamini and Hochberg
- Identified spots whose ratios significantly differed from the mean of the data set for the entire experiment (found 3897 using p-value cutoff of 0.000005)

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

How do we minimize variation not related to specific treatment?

Can the experiment be statistically validated?

What can we afford?

Soybean cDNA slides ~ $200 for ~ 37,000 genes

Soybean Affy chips ~ $600 for ~ 37,000 genes

Steve Clough, USDA-ARS
University of Illinois, Urbana

Experimental design

How should we set up this experiment?

What do we want to identify?

How do we minimize variation not related to specific treatment?

Can the experiment be statistically validated?

What can we afford?

How do we verify the results?

Quantitative RT-PCR
Northern blots

Steve Clough, USDA-ARS
University of Illinois, Urbana

RNA Quality Control Gel

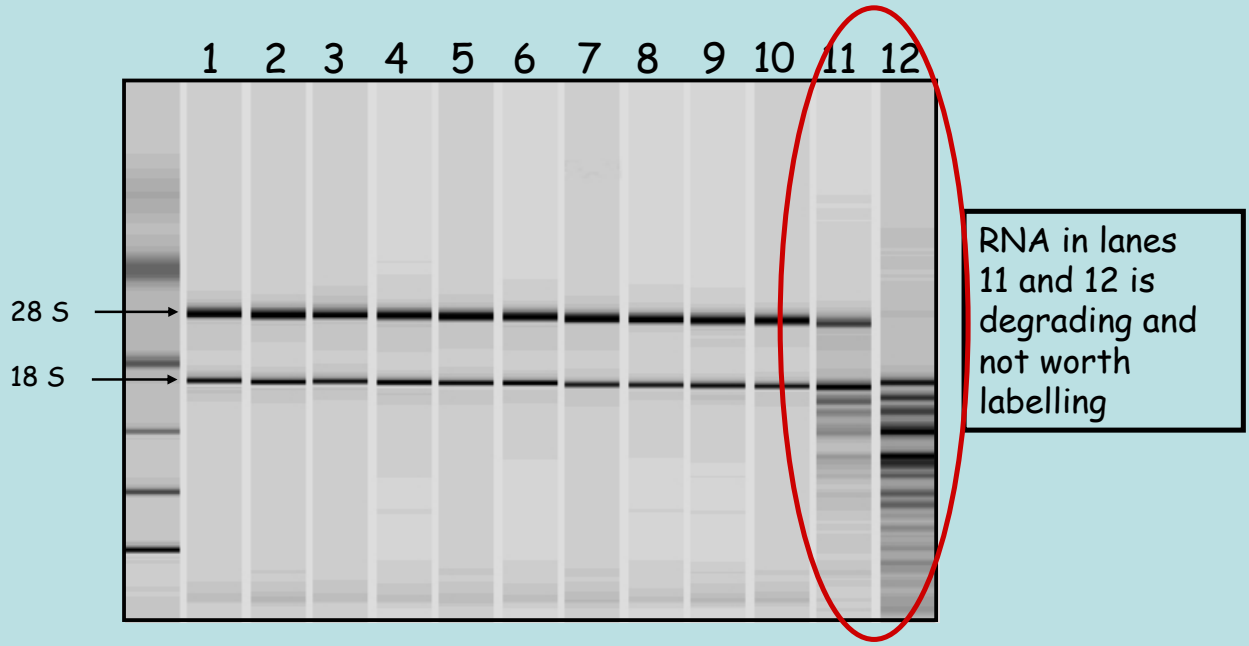
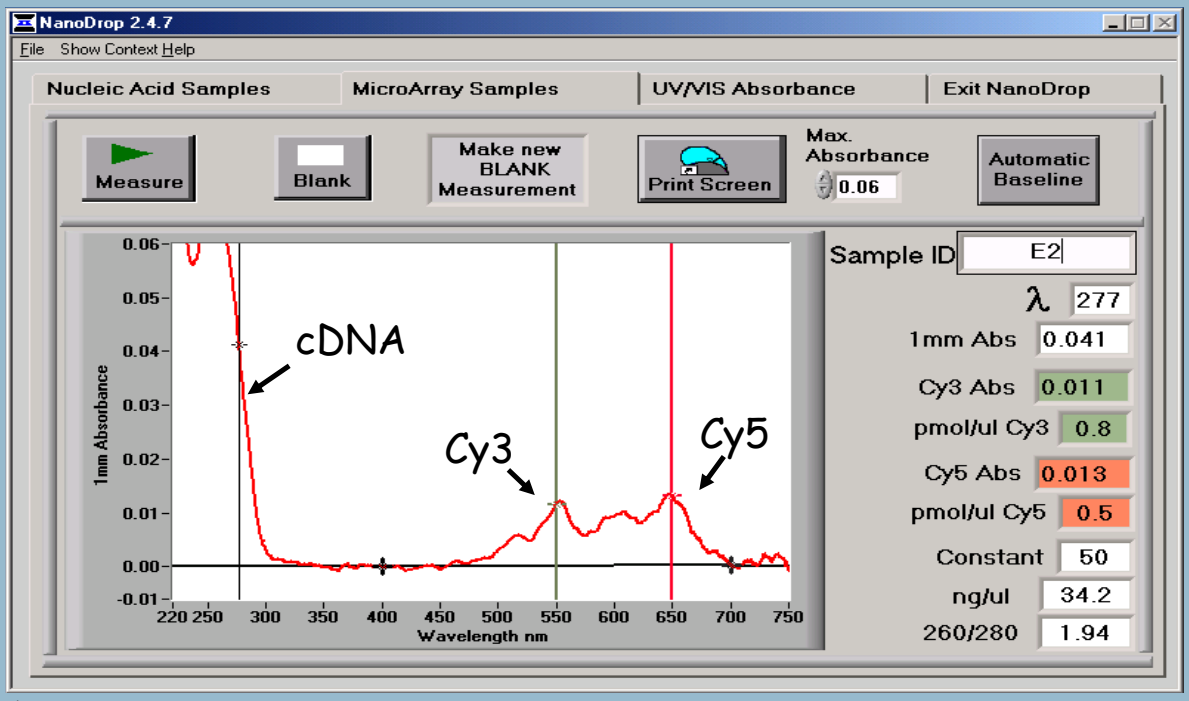


Image Produced on Agilent BioAnalyzer

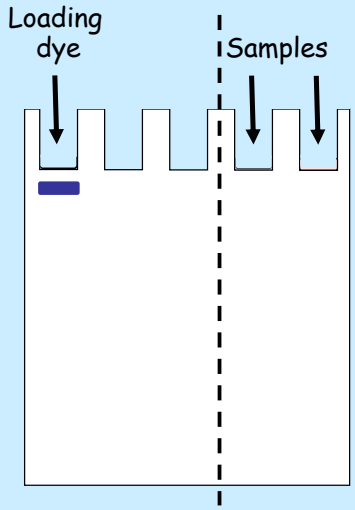
Steve Clough, USDA-ARS
University of Illinois, Urbana

Verifying labelling efficiency using Cy dye absorbance

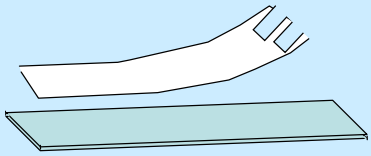


Steve Clough, USDA-ARS
University of Illinois, Urbana

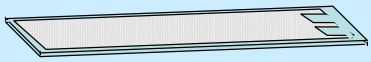
Verifying label efficiency by agarose gel electrophoresis



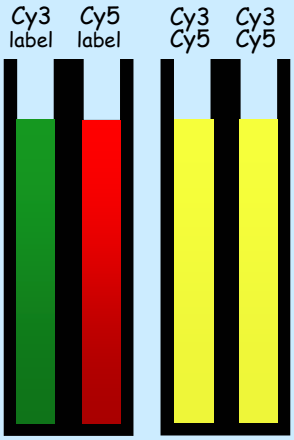
Run on 1% agarose vertical gel



Place on glass slide



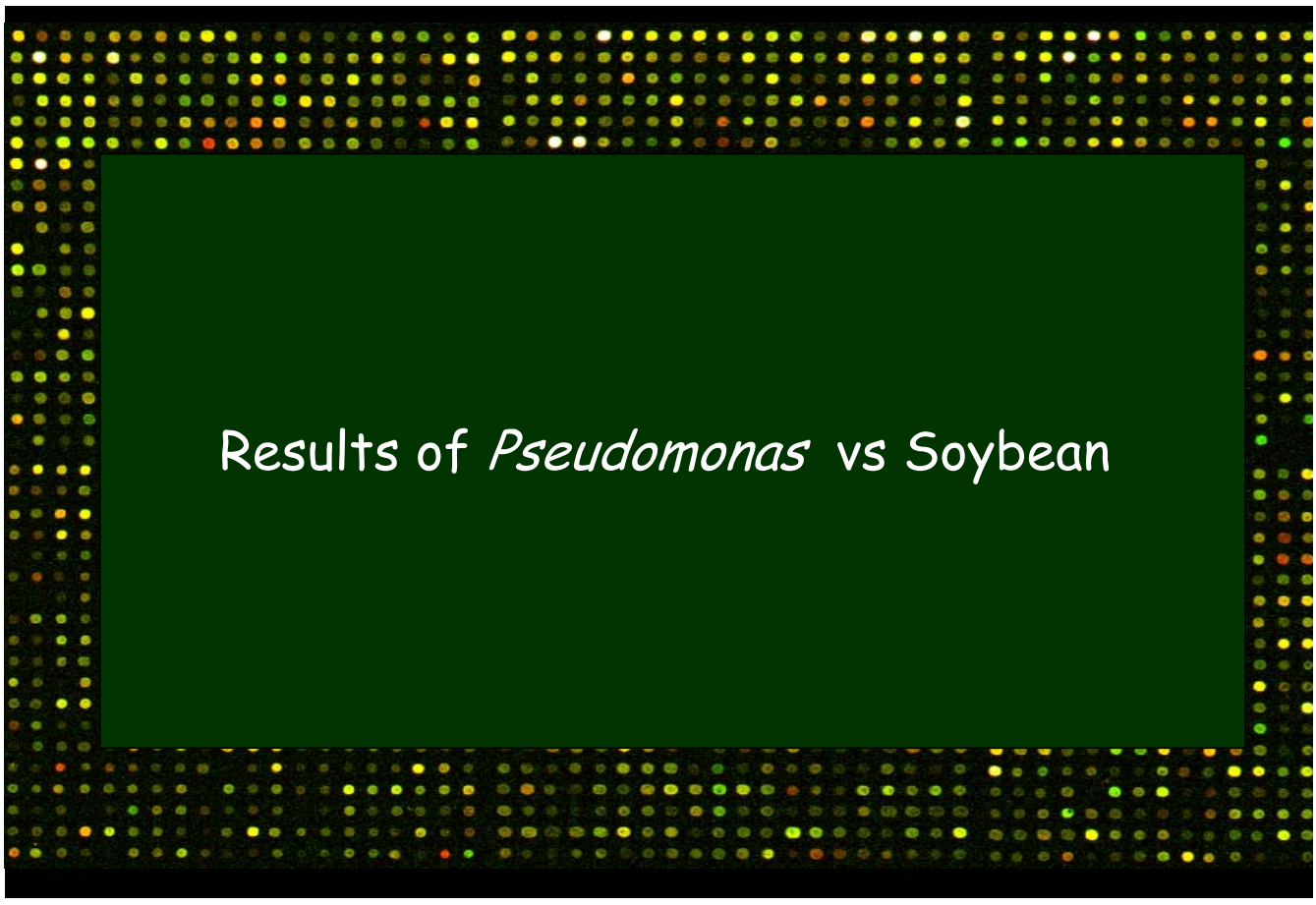
Dry on 70°C heat block



Scan

Steve Clough, USDA-ARS
University of Illinois, Urbana

-Protocol courtesy of Jeff Landgraf, Michigan State University



Expression profiling soybean response to *Pseudomonas syringae* reveals new defense-related genes and rapid HR-specific down regulation of photosynthesis

Zou et al.
Molecular Plant-Microbe Interactions
18 : 1161 - 1174

3897 differentially expressed genes
704 no functional annotation
378 no matches in GenBank

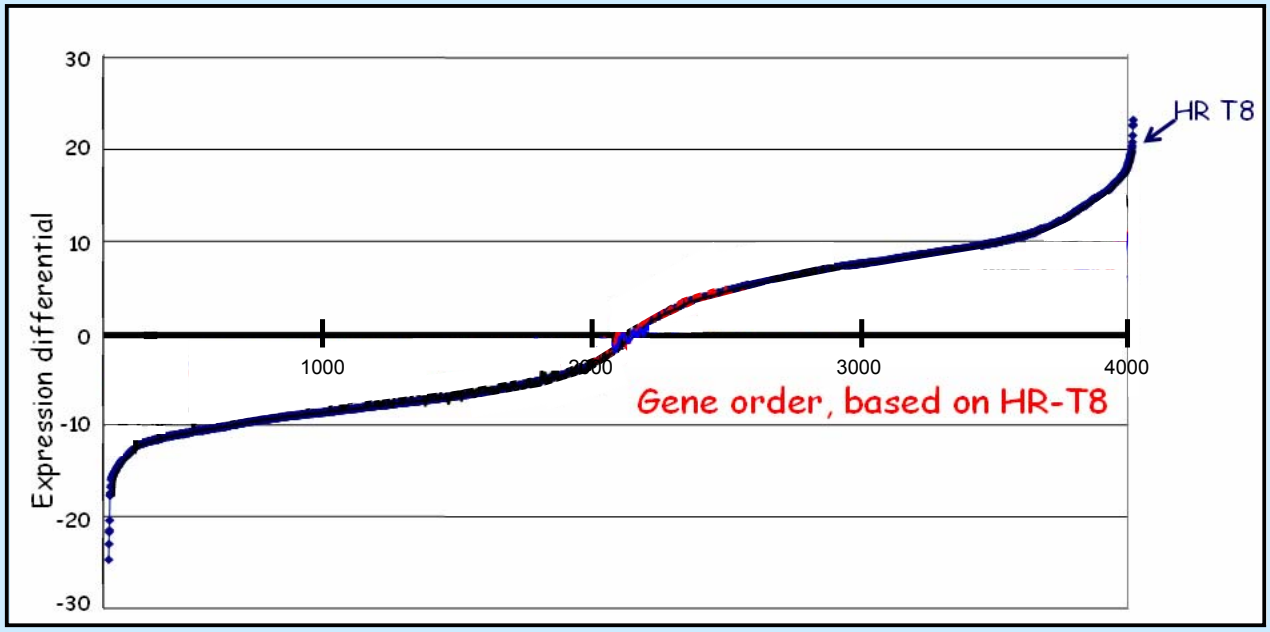
Steve Clough, USDA-ARS
University of Illinois, Urbana

Looking for trends– first sorted data by HR T8

No	Gene ID	Clone ID	R-T8	S-T8	R-T24	S-T24
138	AI960944	Gm-r1088-3	low	-2.03	-4.54	-1.70
364	AW164773	Gm-r1088-573		1.24	3.81	2.31
428	AW184831	Gm-r1088-865		0.83	2.87	0.26
453	AW184933	Gm-r1088-913		-0.64	0.11	-1.26
479	AW184978	Gm-r1088-1017		-1.76	-3.91	-3.01
493	AW185018	Gm-r1088-1057		0.44	1.16	0.89
512	AW185056	Gm-r1088-1127		2.38	4.02	4.76
639	AW185288	Gm-r1088-909		-0.26	-0.51	-0.82
682	AW185387	Gm-r1088-1337		-1.28	-0.44	-1.44
735	AW185514	Gm-r1088-602		0.06	3.55	1.22
747	AW185542	Gm-r1088-622		-1.25	-4.18	-2.57
809	AW185639	Gm-r1088-568		-1.20	-3.00	-1.52
911	AW201421	Gm-r1088-1374		-1.34	-4.23	-2.14
950	AW201518	Gm-r1088-1520		3.08	4.54	2.84
957	AW201528	Gm-r1088-1539		3.66	3.84	3.83
965	AW201550	Gm-r1088-1557		-1.75	-3.36	-1.94
977	AW201578	Gm-r1088-1603		3.76	4.01	4.15
1042	AW201734	Gm-r1088-1843	-7.51	-1.39	-3.25	-1.48
1063	AW201790	Gm-r1088-1592	high	-0.95	-4.38	-2.14
1073	AW201807	Gm-r1088-1604		-0.67	-2.76	-0.37

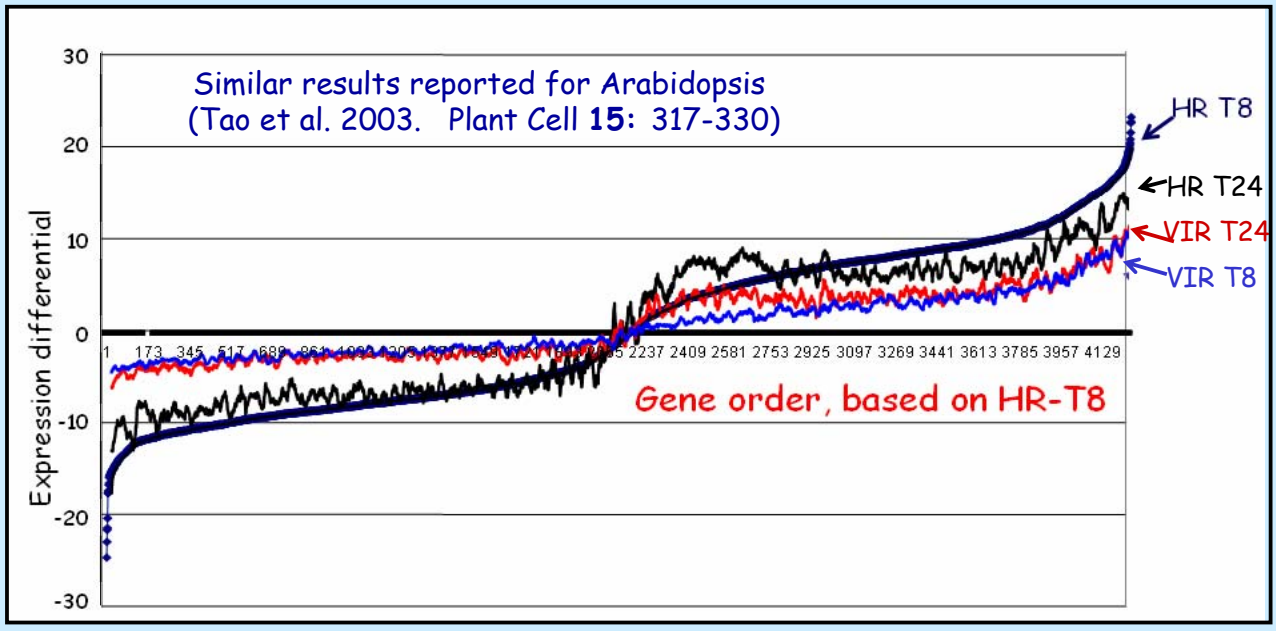
Steve Clough, USDA-ARS
University of Illinois, Urbana

Trend plot – gene order determined by HR, T8



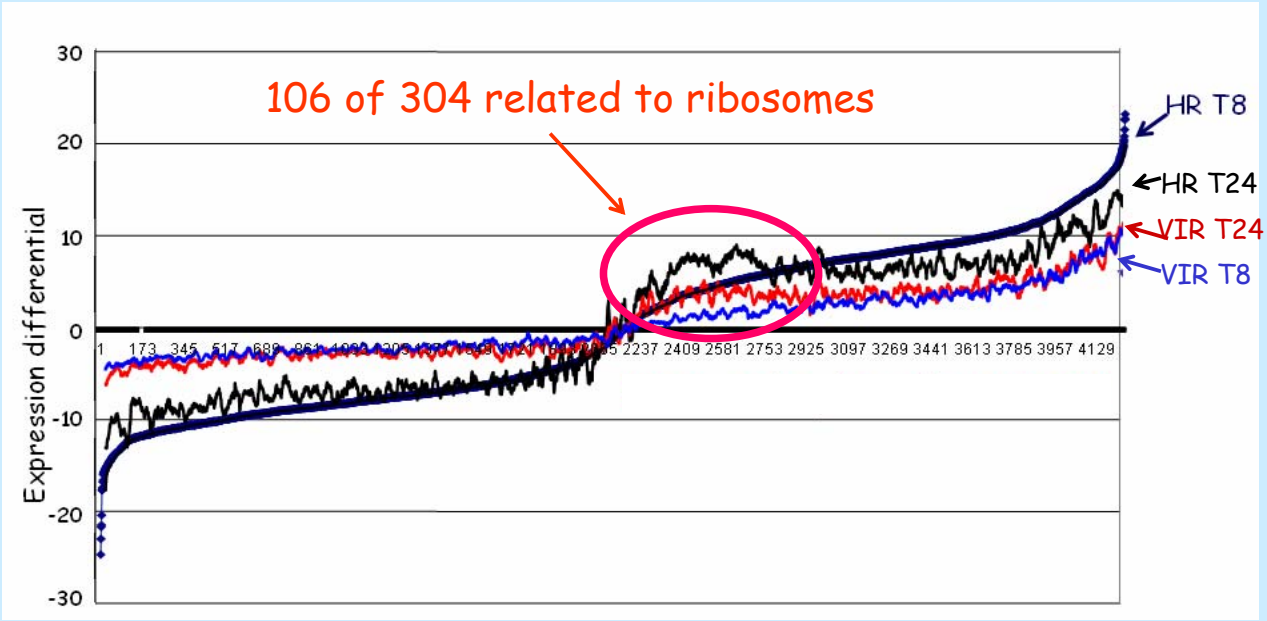
Steve Clough, USDA-ARS
University of Illinois, Urbana

Looking for trends– plot HRT24, VirT8, VirT24 keeping gene order in place as determined by HRT8



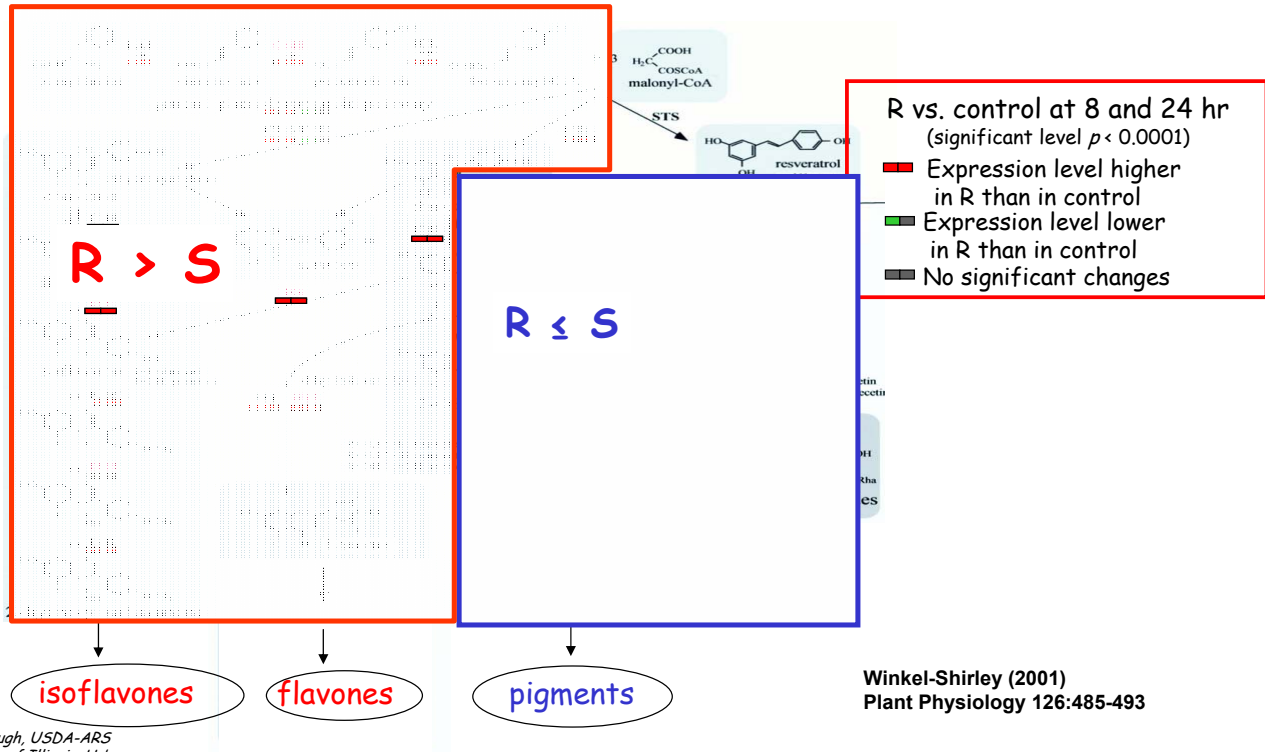
Steve Clough, USDA-ARS
University of Illinois, Urbana

Looking for trends– plot HRT24, VirT8, VirT24 keeping gene order in place as determined by HRT8



Steve Clough, USDA-ARS
University of Illinois, Urbana

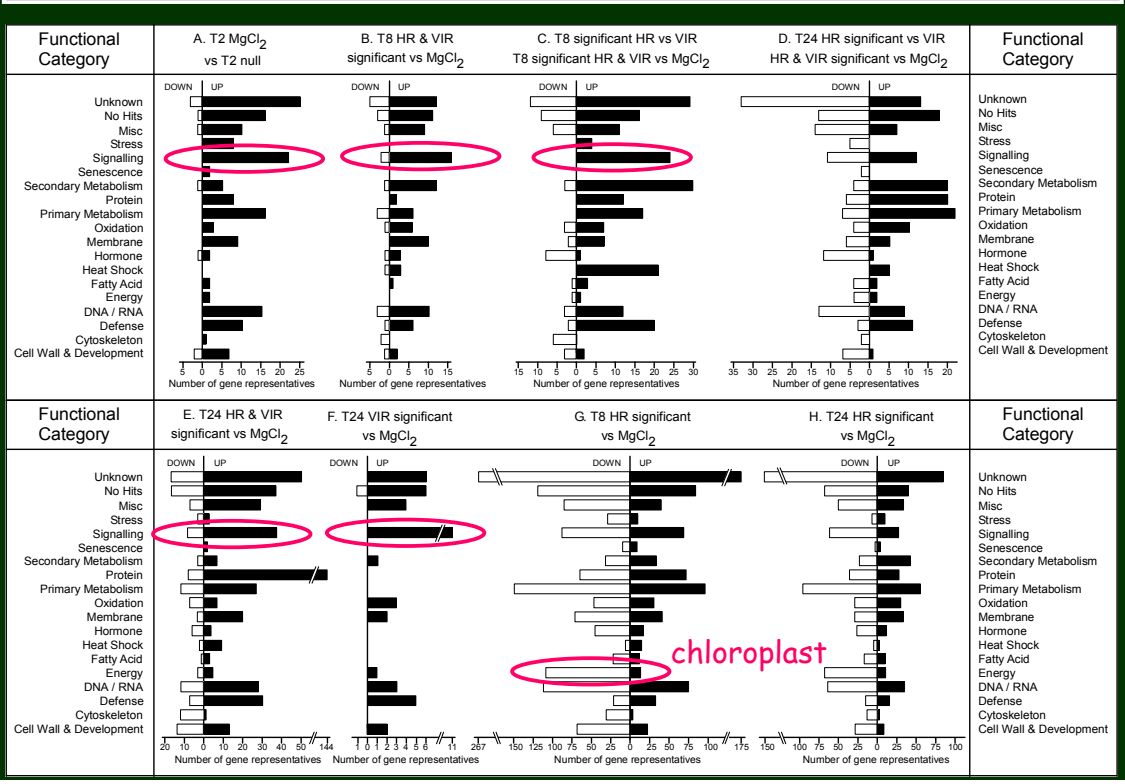
Induction of specific branches of the phenylpropanoid pathway



Winkel-Shirley (2001)
Plant Physiology 126:485-493

Steve Clough, USDA-ARS
University of Illinois, Urbana

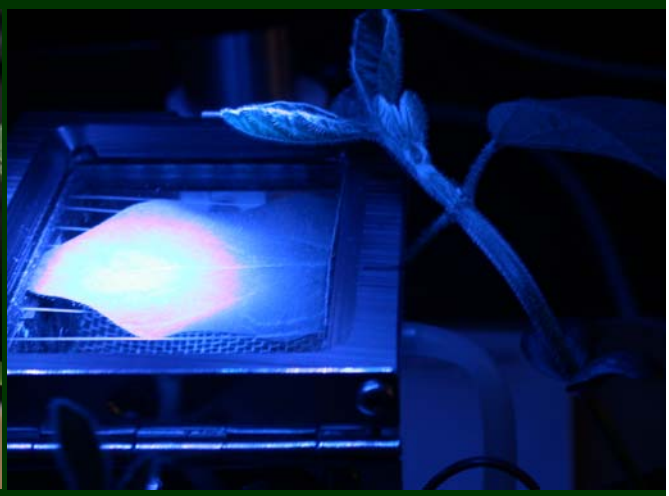
Classifying into functional categories



Steve Clough, USDA-ARS
University of Illinois, Urbana

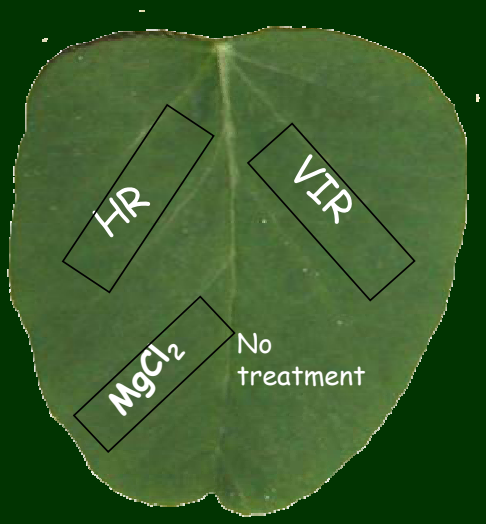
Monitor photosynthesis activity

(in collaboration with Evan DeLucia, U of Illinois)



Steve Clough, USDA-ARS
University of Illinois, Urbana

Monitor photosynthesis activity
(Single leaf assay)



Steve Clough, USDA-ARS
University of Illinois, Urbana

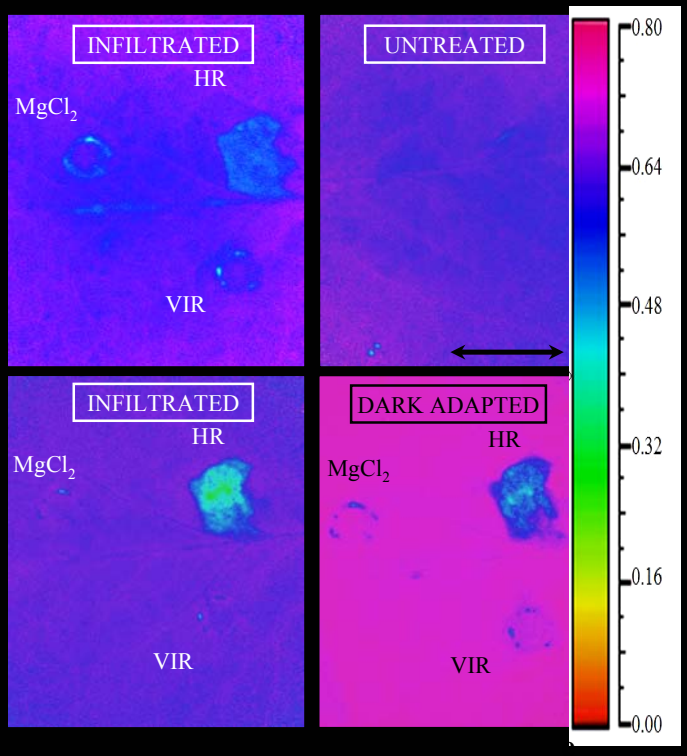
Fluorescent measurement of PS II operating efficiency

8 hpi

18% reduction in PSII efficiency for HR sample

24 hpi

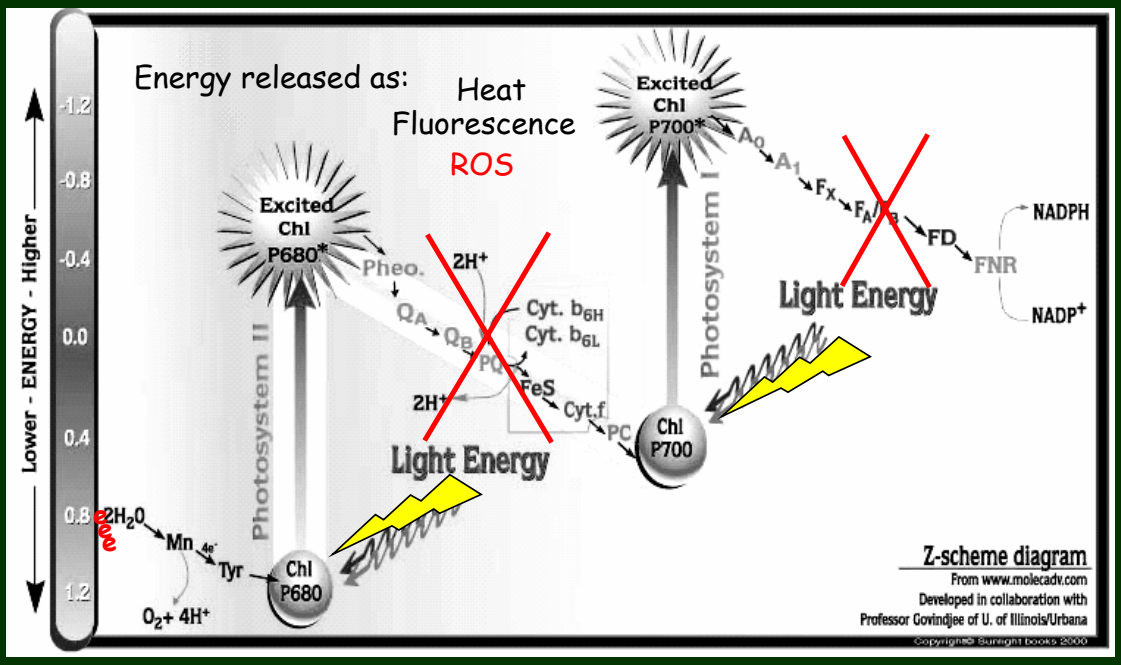
41% reduction in PSII efficiency for HR sample



Steve Clough, USDA-ARS
University of Illinois, Urbana

94 photosynthesis related genes repressed in HR

Reference on HR and photosynthesis:
Allen et al. 1999 Plant Physiol 119: 1233-1241

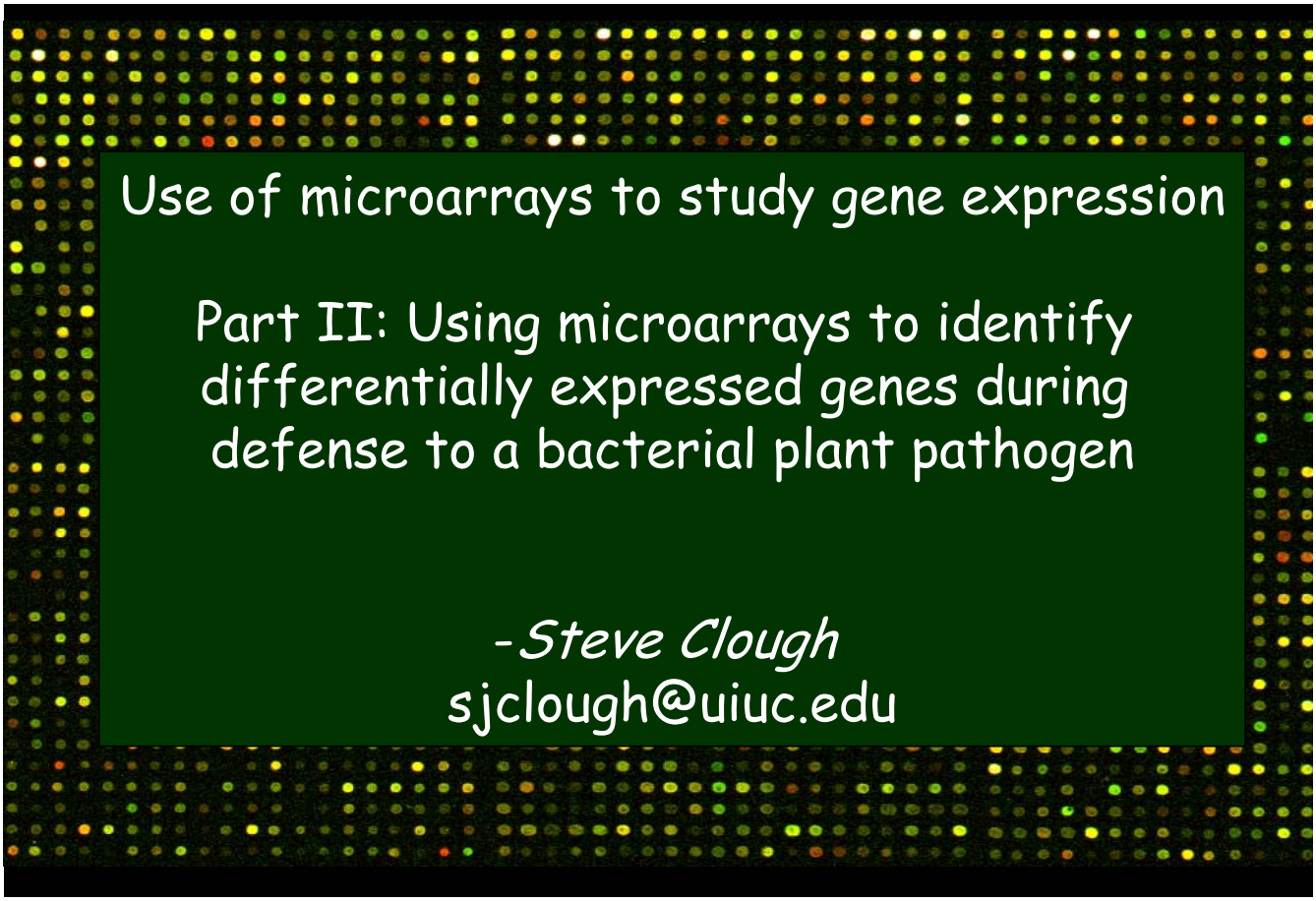


Steve Clough, USDA-ARS
University of Illinois, Urbana

What is the benefit to a plant to stop replacing photosynthetic components if attacked by a pathogen?

- Reduces glucose synthesis
  - Reduces carbon source for pathogen
- Leads to enhanced production of reactive oxygen species
  - Antimicrobial and strong defense signal

Steve Clough, USDA-ARS
University of Illinois, Urbana


Use of microarrays to study gene expression
Part II: Using microarrays to identify differentially expressed genes during defense to a bacterial plant pathogen
-Steve Clough
sjclough@uiuc.edu