



NRC · CNRC

From Discovery to Innovation...

CROSS-PLATFORM COMPARISON AND VALIDATION OF THREE MICROARRAY TECHNOLOGIES

Jeffrey D. Pylatuik

Plant Biotechnology Institute

Friday, March 19, 2004



National Research
Council Canada

Conseil national
de recherches Canada

Canada



Justification for the study

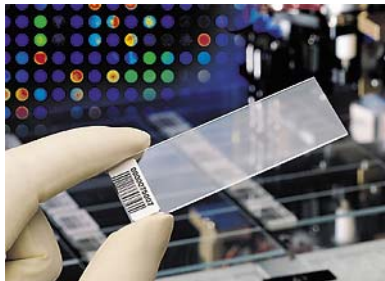
1. Number of microarray studies increasing.
2. Very few studies comparing different microarray platforms.
3. None have examined different Arabidopsis platform technologies.
4. Amplicon platform had not been validated.

Platform Overview

Operon



- Arabidopsis Array-Ready Oligo Set™
- 70mer probes representing 24,921 *A. thaliana* genes
- genes and predicted ORFs based upon the UniGene Database Build At 4 (NCBI)
- T_m normalized to 78°C and have a 5' amino linker.
- Spotted on aminosilane-coated “Superamine” slides



Amplicon



- PBI's annotation project for *Arabidopsis thaliana*.
- Double-stranded amplicons averaging 600bp representing 10739 genes.
- Genes and predicted ORFs based upon the AGI Dataset (TAIR, TIGR).
- Spotted in duplicate on CMT-GAPS amino-silane coated slides (Corning)

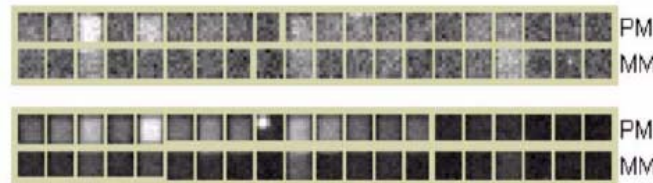


Platform Overview

Affymetrix AG



- 24mer probes representing 7626 genes.
- Genes and predicted ORFs based upon NCBI database.
- 20 PM and MM probes synthesized with photolithography and solid-phase chemistry.

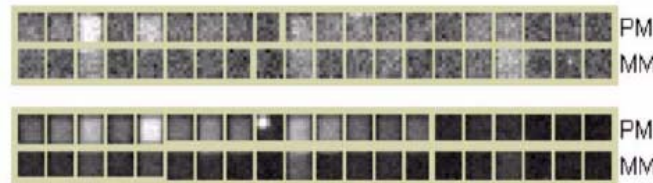


Platform Overview

Affymetrix ATH1



- 25mer probes representing 22810 genes.
- Genes and predicted ORFs based upon TIGR database.
- 11 PM and MM probes synthesized with photolithography and solid-phase chemistry.



Platform Overview

Platform design Summary

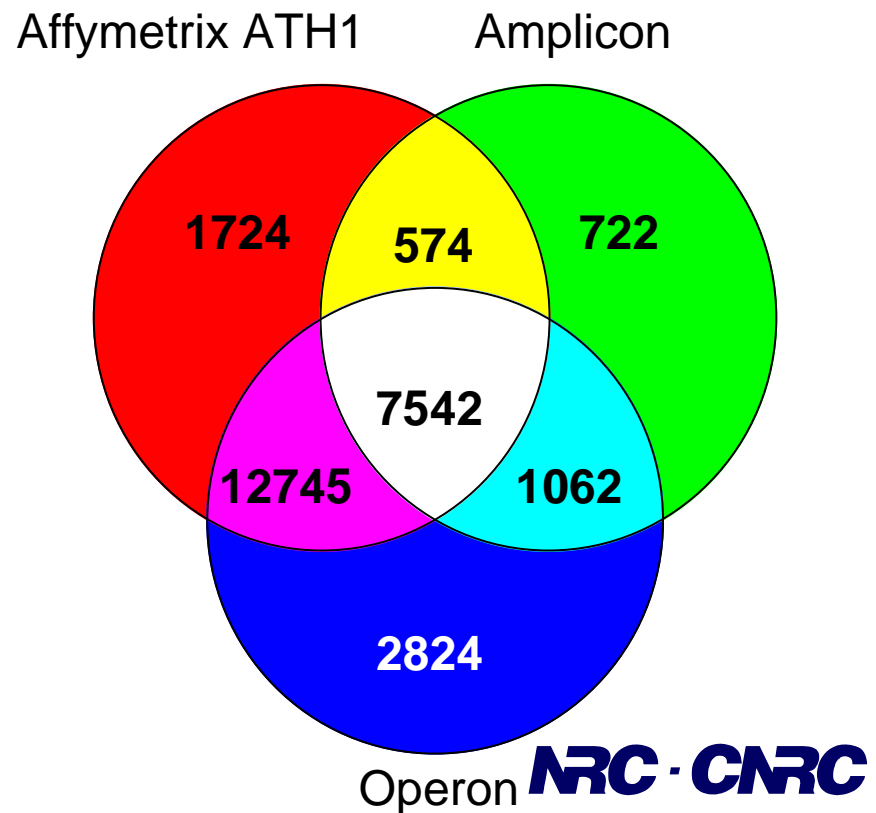
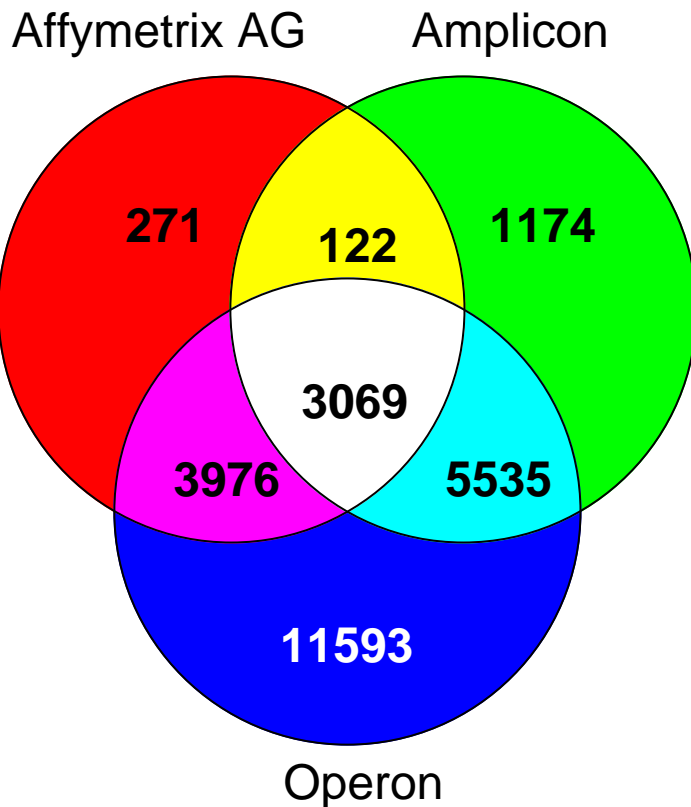
Array type	Total number of elements	Controls	Annotated elements (TIGR)	Non-annotated	Non-redundant gene set (TIGR)
Operon	26791	509	25426	856	24065
Amplicon	12677	15	11822	840	9899
Affymetrix AG	8297	50	8058	189	7437
Affymetrix ATH1	22810	N/A	22591	219	22591
total unique genes					27250

Platform Overview

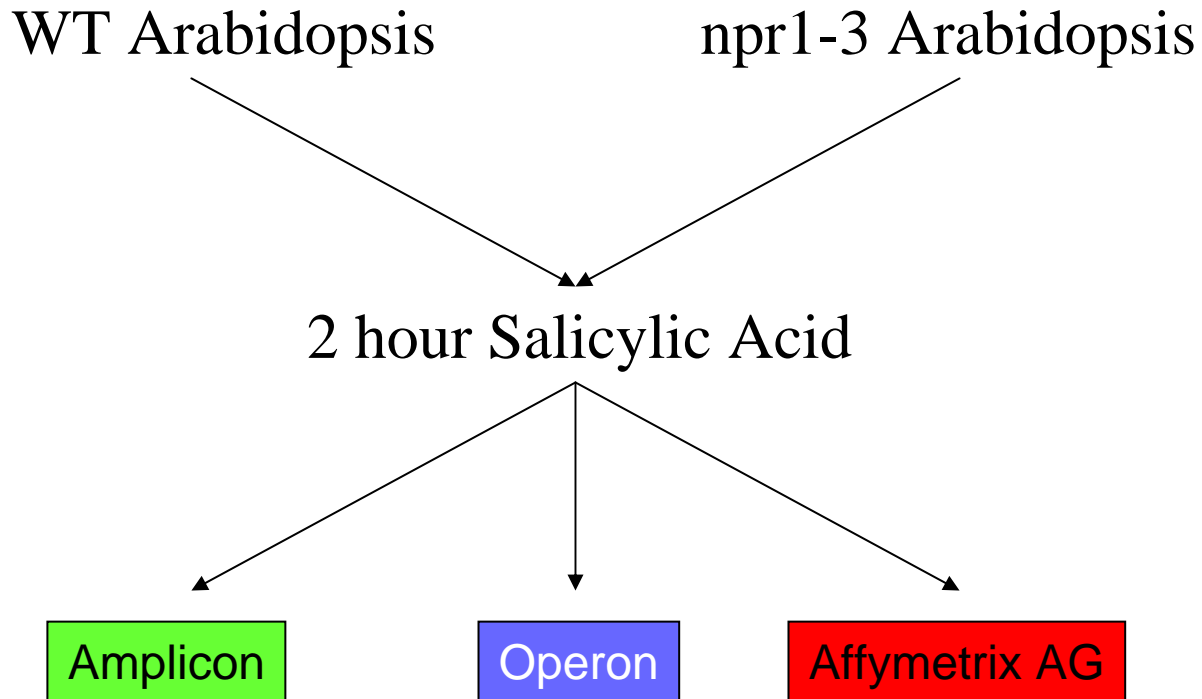
Cross-Platform Comparison of Reporters

2 hour experiment

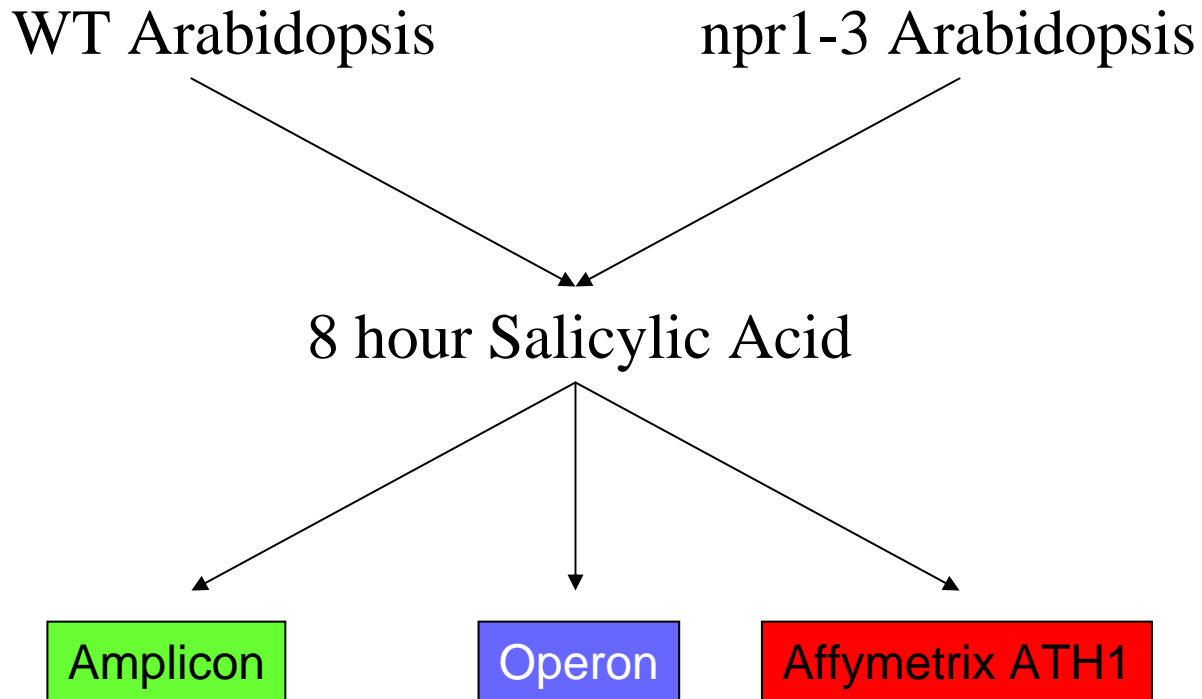
8 hour experiment



Experiment 1



Experiment 2



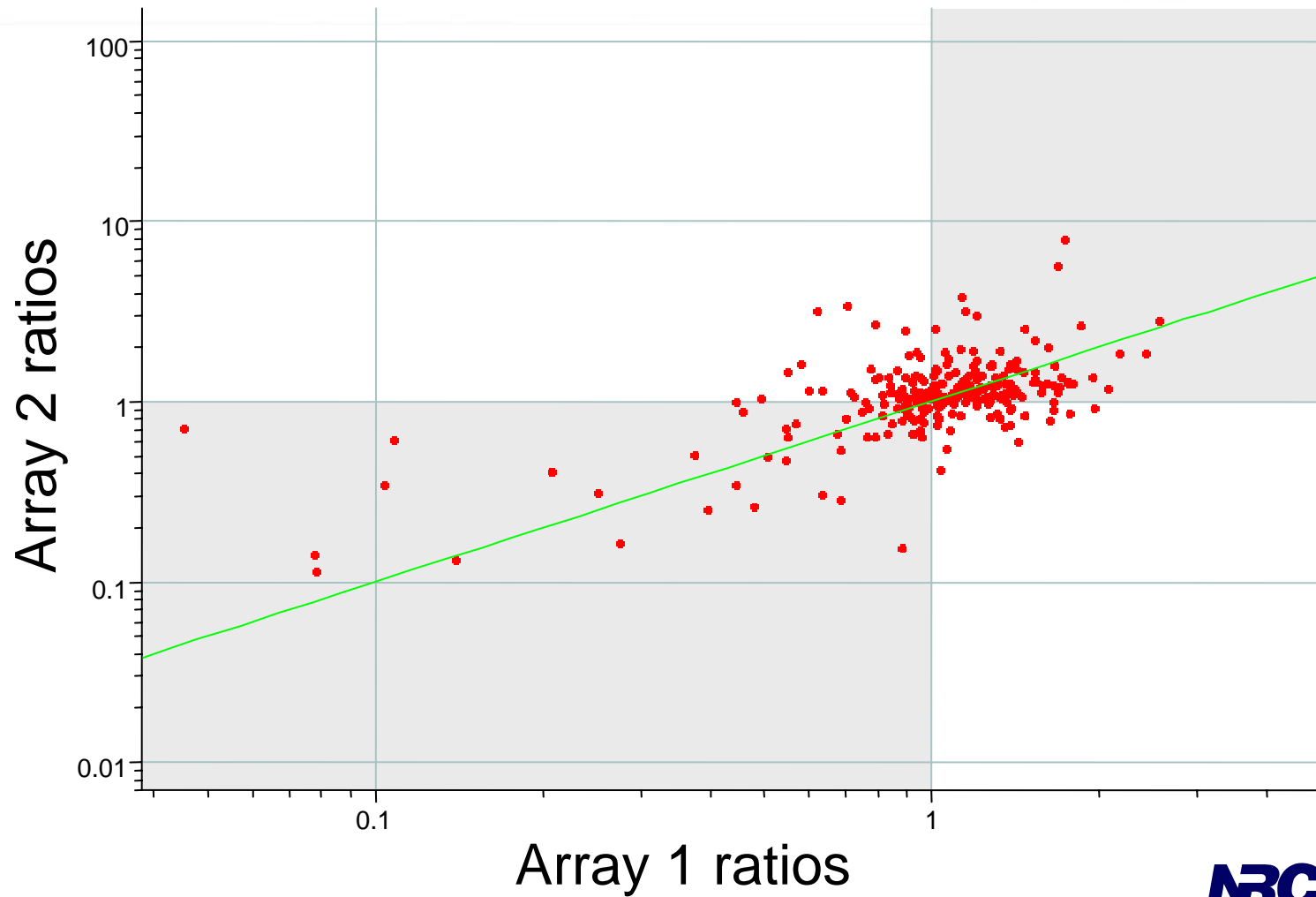
Analysis of Spotted Arrays

1. 4 biological replicates, 2 dye swaps.
2. Localized background subtraction of Median Intensity values. (BASE – Saal et al. (2002) Genome Biology, 3:3.1-3.6)
3. **Shrunk Robust Splines.** (Limma, R statistics, Gordon Smyth, Unpublished)
 - adapts smoothly between print-tip loess and global loess.

Analysis of Affymetrix arrays

- No replicates.
- Analysis by MAS 5.0
- Hybridization and data analysis performed by the Montreal Genome Center.

Methods



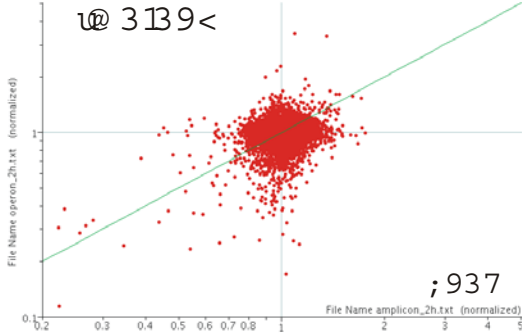
Results

F urvw0Solwirup #F rp sduvrq#0H0#Jhs ruhw

5#rxu

Dp sdf rq#v#R shurq

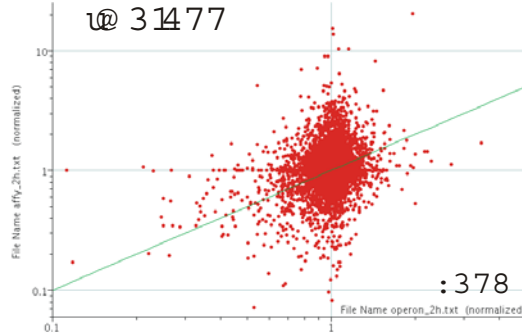
@ 3139<



;937

Dii|#DJ #v#R shurq

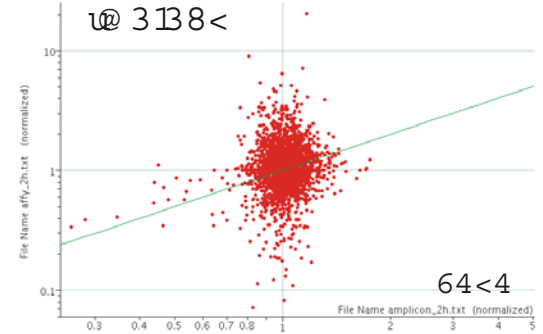
@ 31477



:378

Dii|#DJ #v#Dp sdf rq

@ 3138<

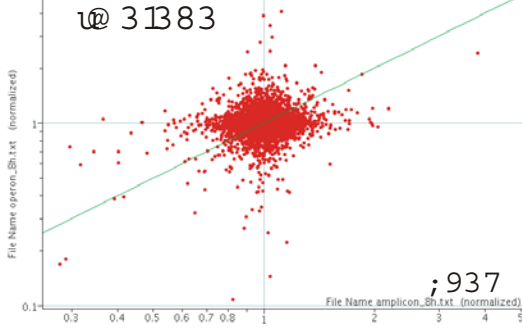


64<4

#;rxu

Dp sdf rq#v#R shurq

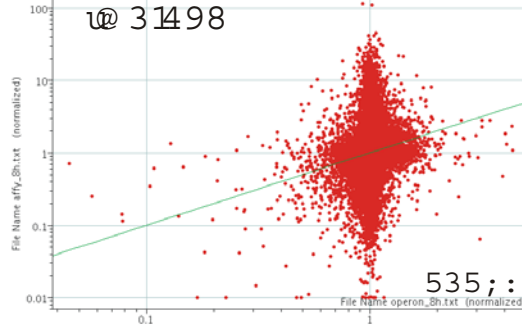
@ 31383



;937

Dii|#DWK4 #v#R shurq

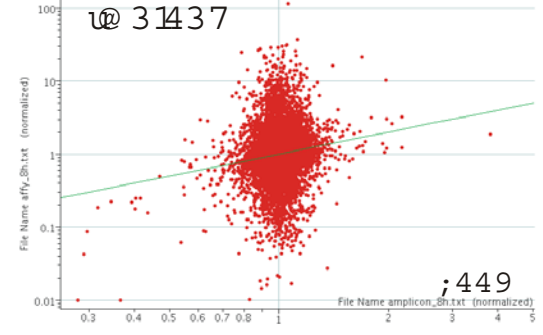
@ 31498



535;:

Dii|#DWK4 #v#Dp sdf rq

@ 31437



;449

Effect of Channel Intensity

1. Percentile ranked average channel intensities.
2. Filtered data for channel intensities that ranked:
 - > 50th percentile
 - > 75th percentile
 - > 90th percentile
3. Had to satisfy these requirements for at least one of the 2 channels for both array platforms.

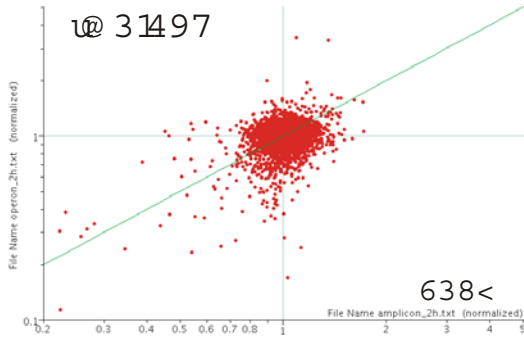
Results

F urw0Solwirup #F rp sduvrq#0# #3wk#Shuf hqwzh

5#krxu

Dp sdf rq#v#R shurq

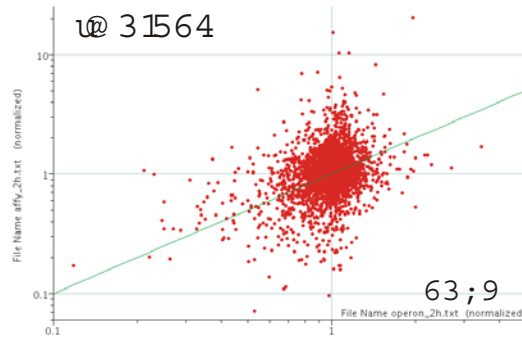
@ 31497



638<

Dii|#DJ #v#R shurq

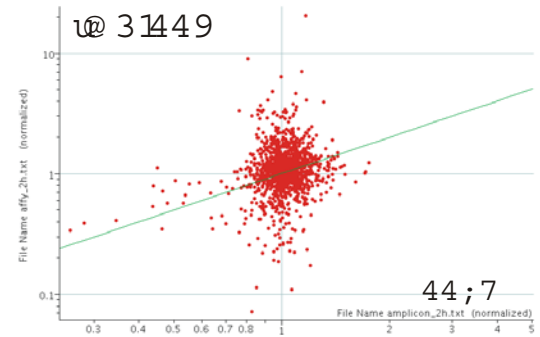
@ 31564



63;9

Dii|#DJ #v#Dp sdf rq

@ 31449

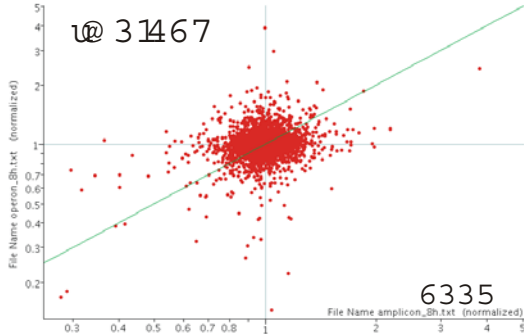


44;7

#krxu

Dp sdf rq#v#R shurq

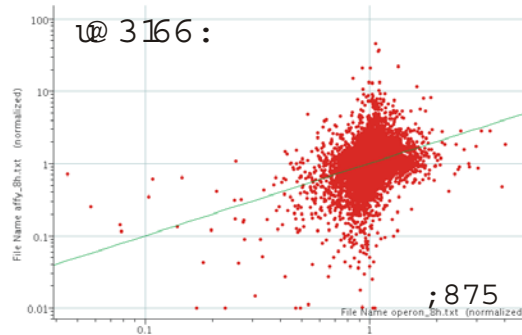
@ 31467



6335

Dii|#DWK4 #v#R shurq

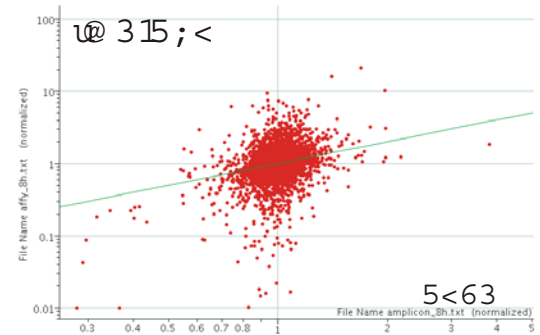
@ 3166:



;875

Dii|#DWK4 #v#Dp sdf rq

@ 315;<



5<63

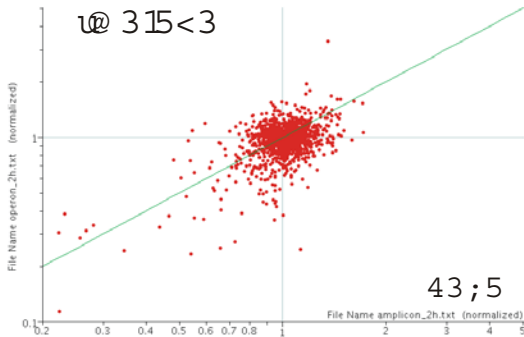
Results

Furw0Solwirup #F rp sdüvrq#0#A #: 8wk#shuf hqwih

5#krxu

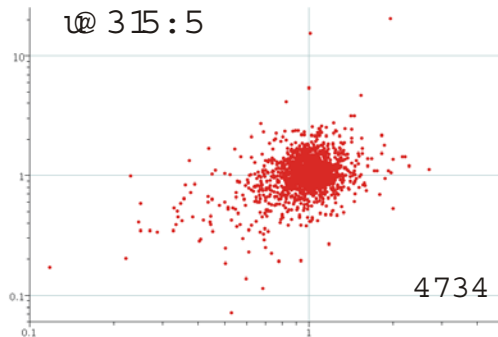
Dp sdf rqr#v#R shurq

@ 315<3



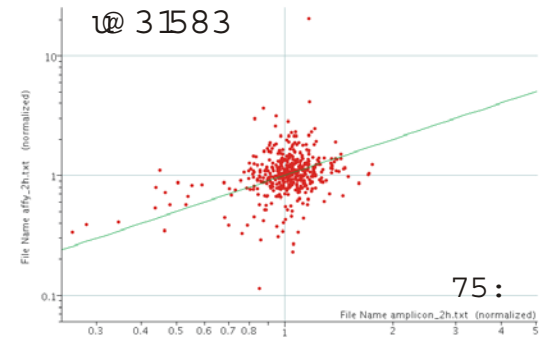
Dii|#DJ #v#R shurq

@ 315:5



Dii|#DJ #v#Dp sdf rqr

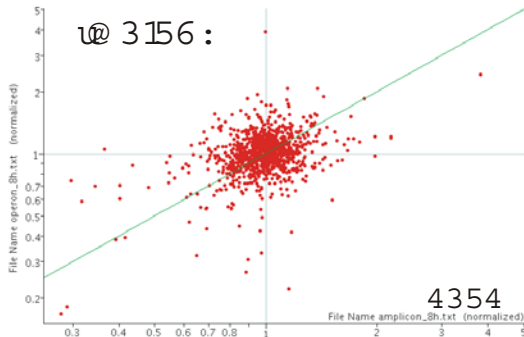
@ 31583



;#krxu

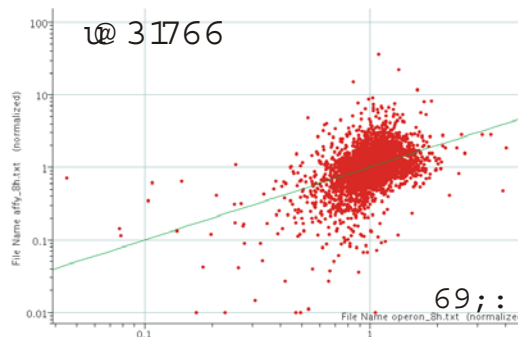
Dp sdf rqr#v#R shurq

@ 3156:



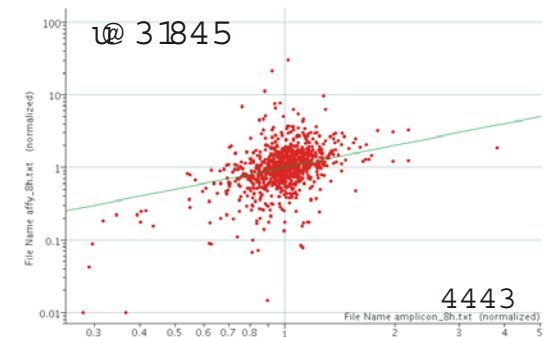
Dii|#DWK4 #v#R shurq

@ 31766



Dii|#DWK4 #v#Dp sdf rqr

@ 31845



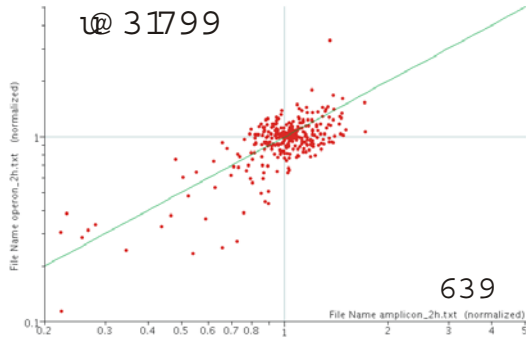
Results

F urw0Sd#wirup #F rp sduvrq#0/A #3wk#shuf#hqw#h

5#krxu

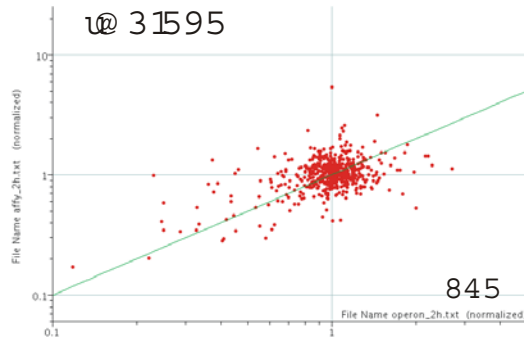
Dp sdf rq#y#R shurq

@ 31799



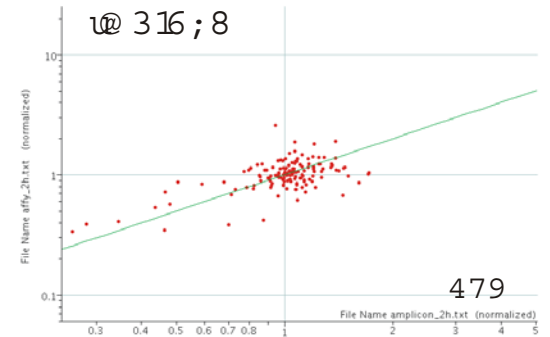
Dii|#DJ #y#R shurq

@ 31595



Dii|#DJ #y#Dp sdf rq

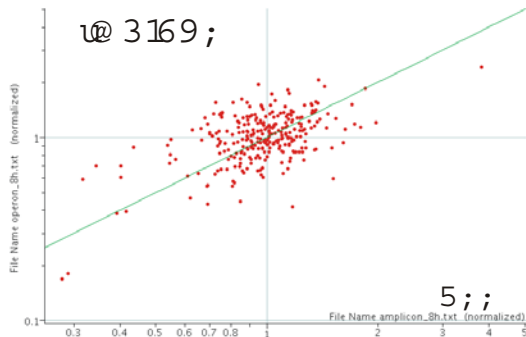
@ 316;8



;#krxu

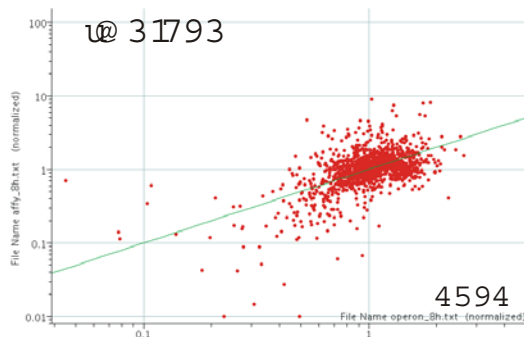
Dp sdf rq#y#R shurq

@ 3169;



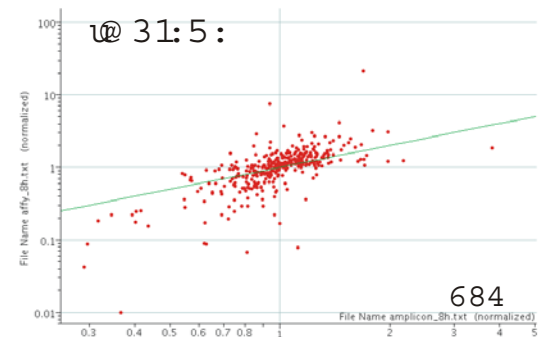
Dii|#DWK4 #y#R shurq

@ 31793

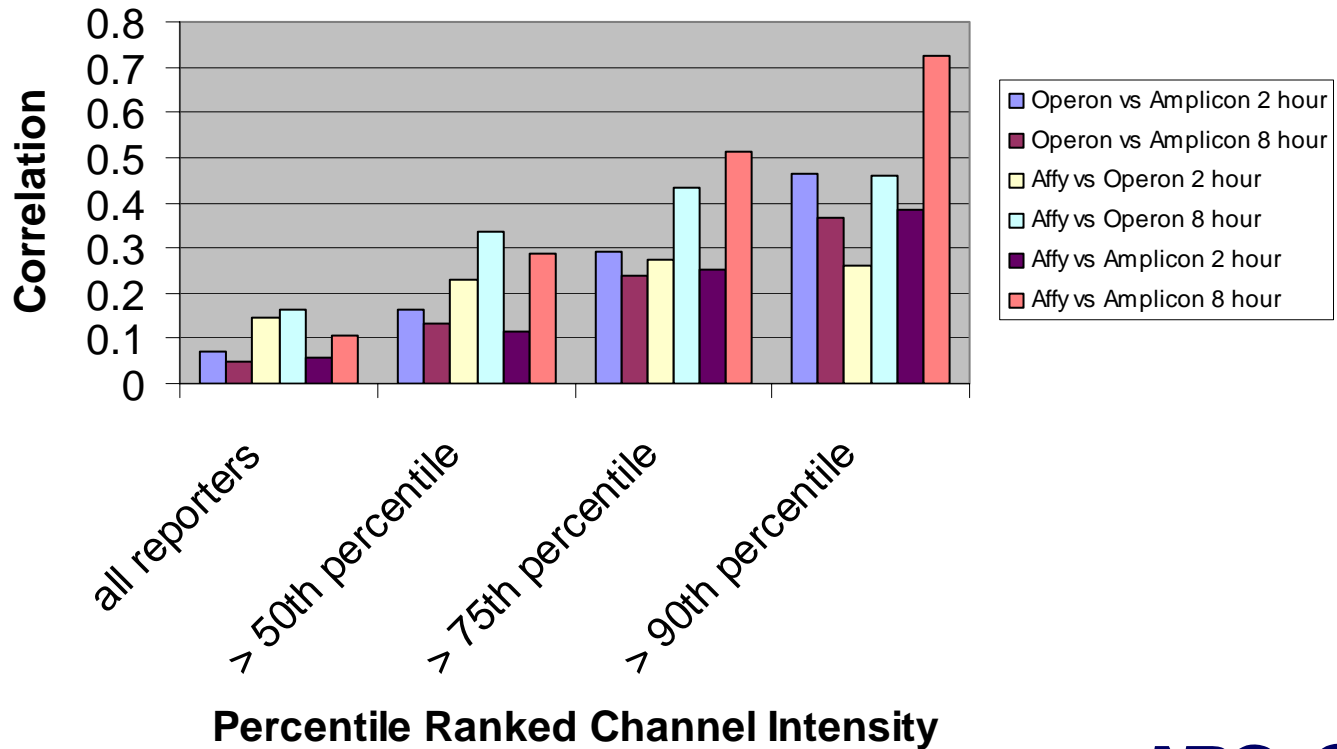


Dii|#DWK4 #y#Dp sdf rq

@ 31:5:



Effect of Channel Intensity on Cross-Platform Correlation



Conclusions

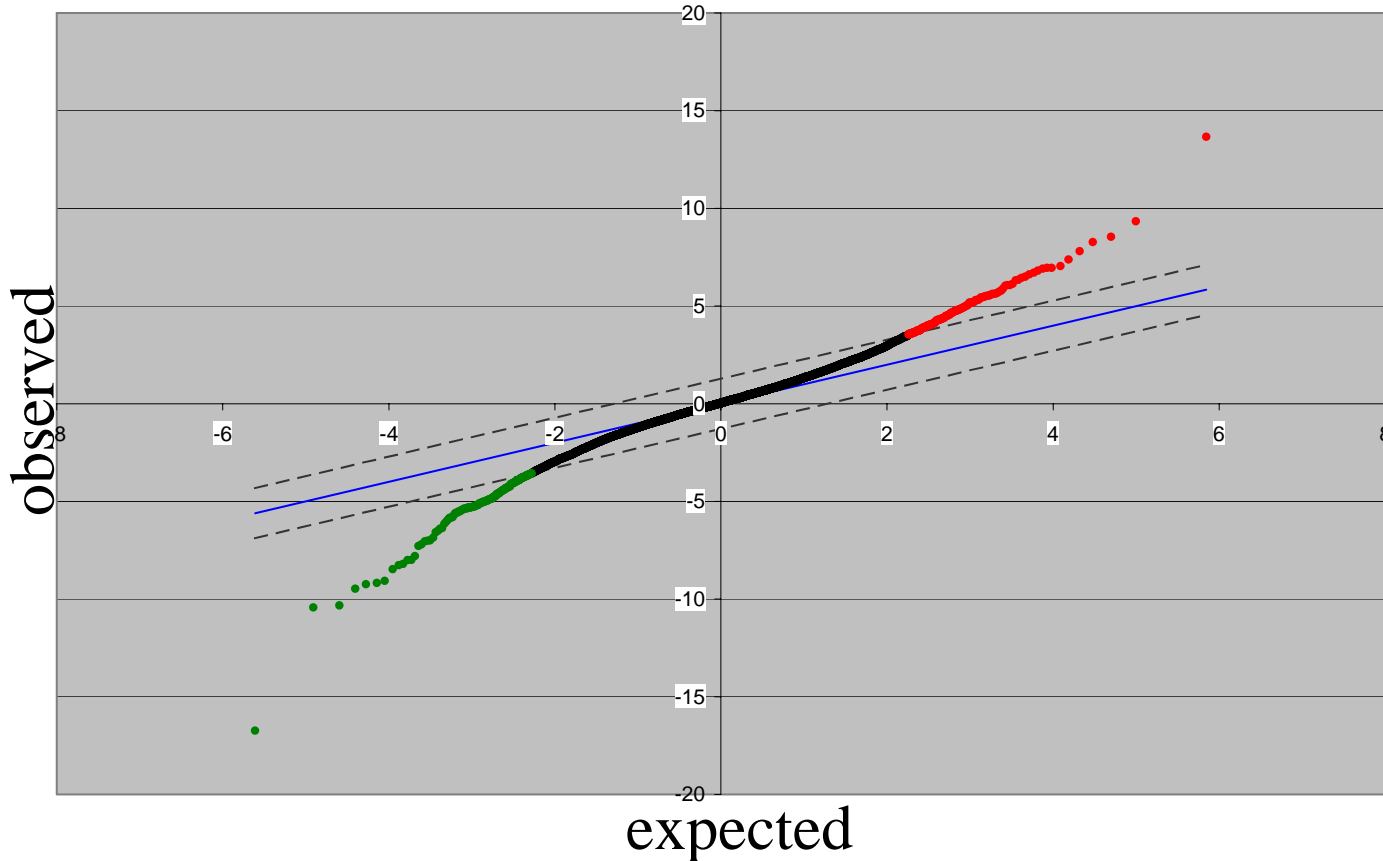
- Similarity of results between platforms is dependent on relative channel intensity:
 - Cross-platform correlation is greater among reporters with relatively higher intensities.

Results

How do significant genes compare across platforms?

Analysis of Spotted Arrays (cont.)

- Significance analysis of Microarrays. (Tusher et al. (2001)
PNAS 98:5116-5121)



Significant = 378

False
Significant = 21.1

Analysis of Affymetrix arrays

- No replicates.
- 2.5 fold threshold criterion
- Minimum of one “present” call between values.

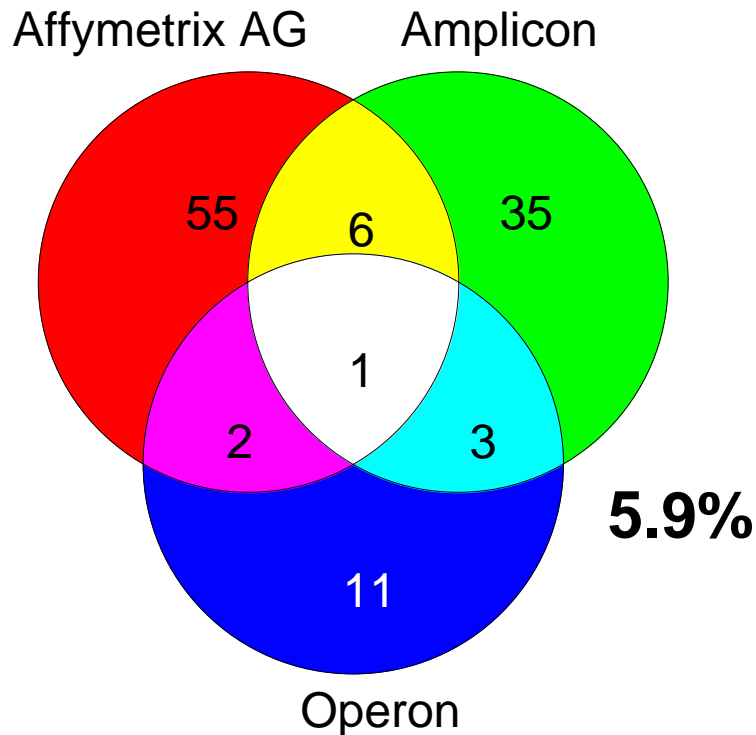
Scoring of candidate genes

2 hour SA data summary						
platform	significant genes	false positives	up-regulated	down-regulated	FDR	delta value
Amplicon	135	6.568	26	109	4.87%	0.15
Operon	179	8.531	3	176	4.77%	0.898
Affymetrix AG	165	N/A	68	97	N/A	N/A
8 hour SA data summary						
platform	significant genes	false positives	up-regulated	down-regulated	FDR	delta value
Amplicon	315	15.691	125	190	4.98%	0.188
Operon	986	47.877	425	561	4.86%	0.642
Affymetrix ATH1	1190	N/A	481	709	N/A	N/A

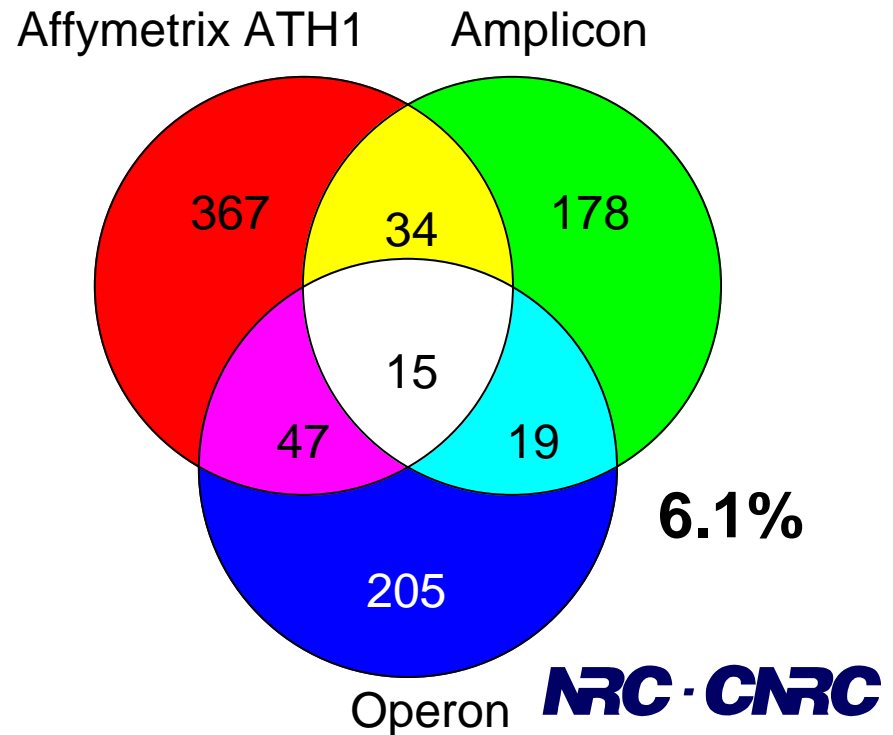
Results

Cross-Platform Comparison of Significant Reporters

2 hour experiment

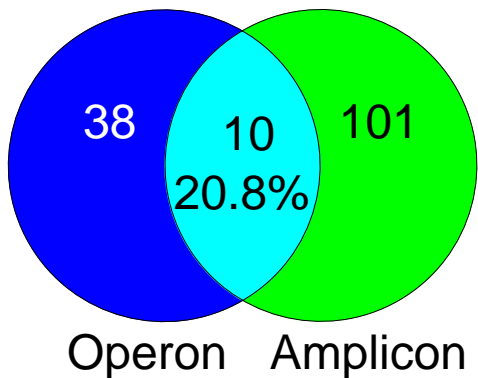
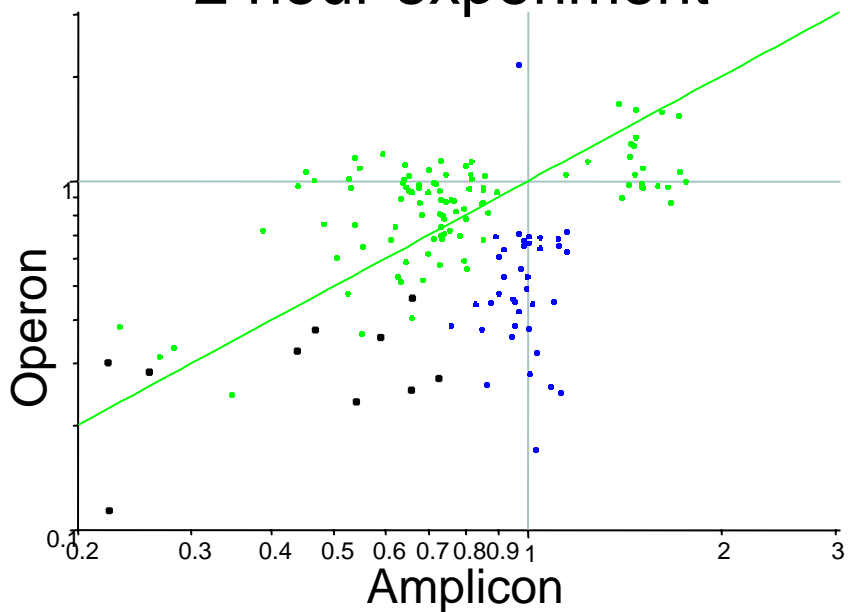


8 hour experiment

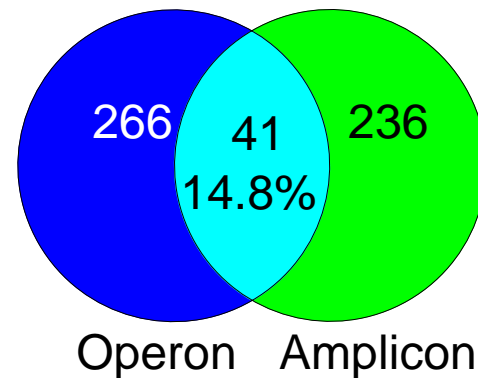
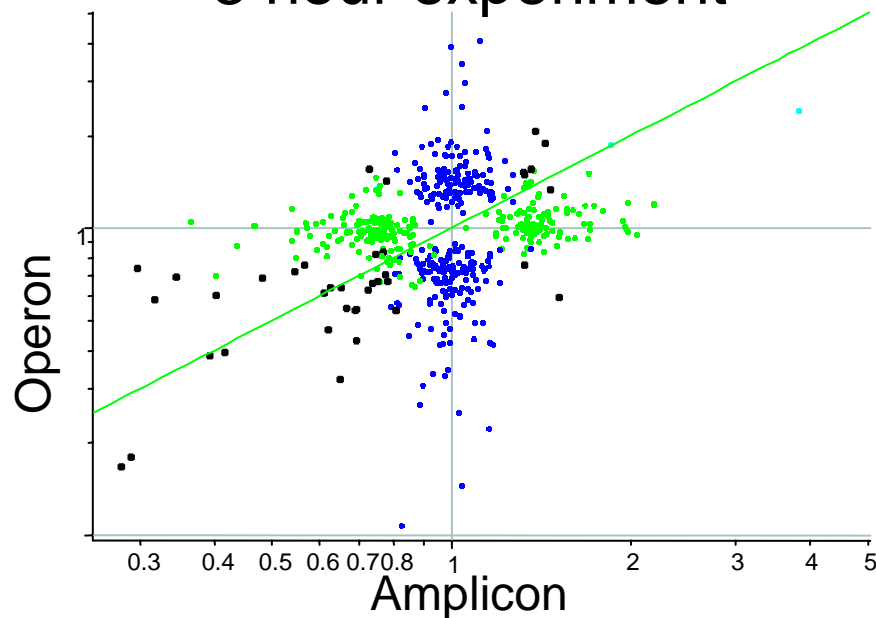


Results

2 hour experiment

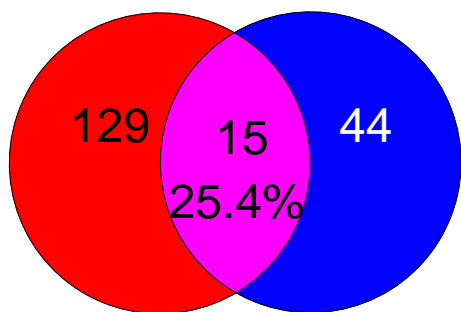
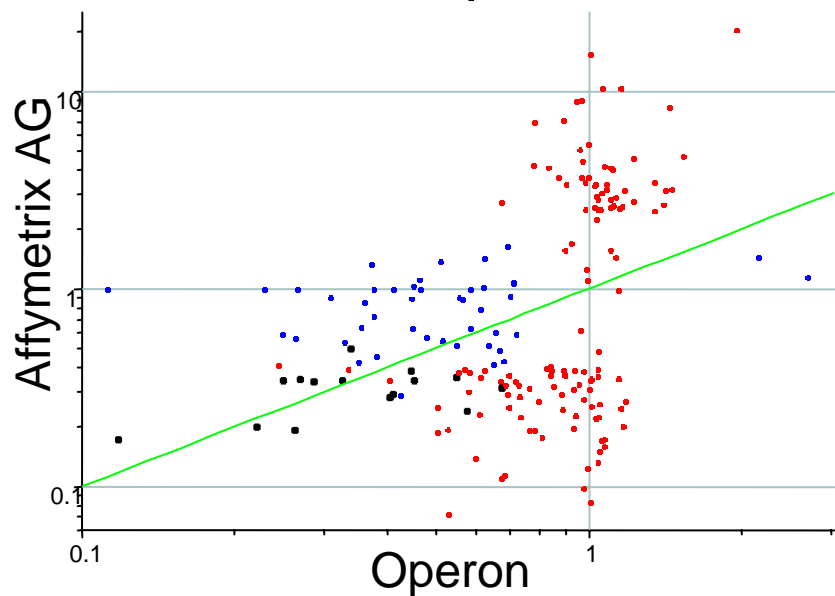


8 hour experiment



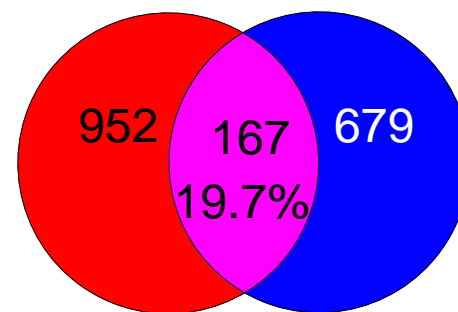
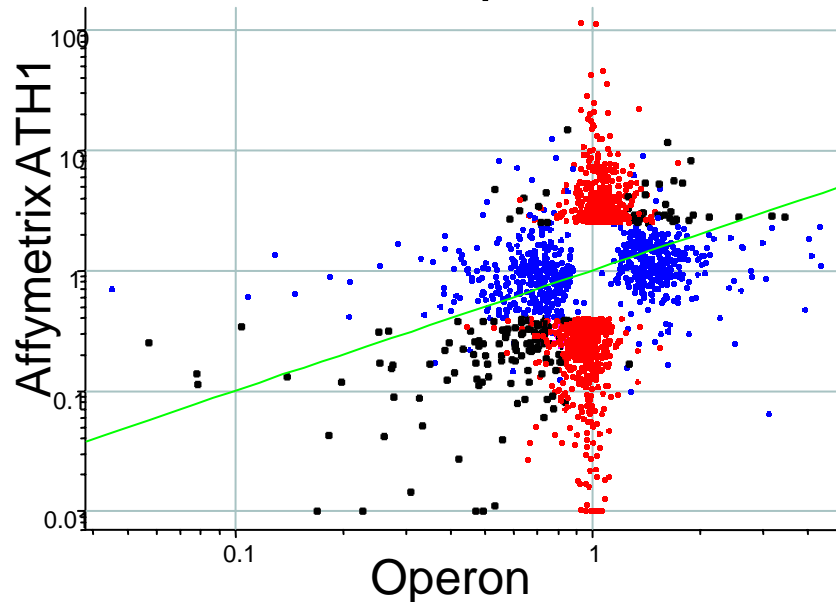
Results

2 hour experiment



Affymetrix AG Operon

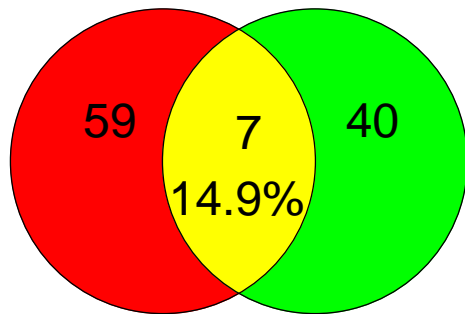
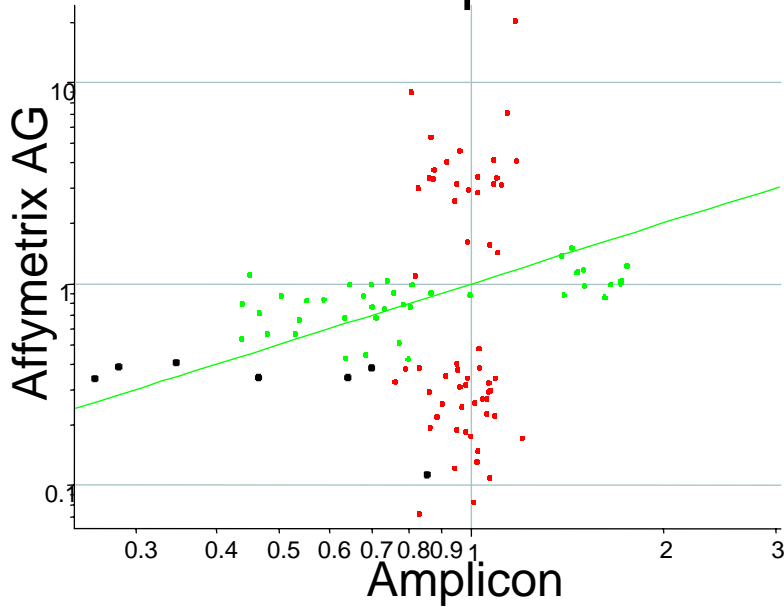
8 hour experiment



Affymetrix ATH1 Operon

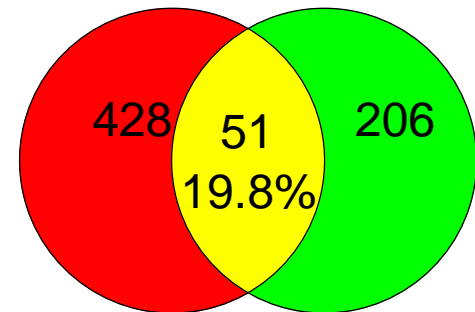
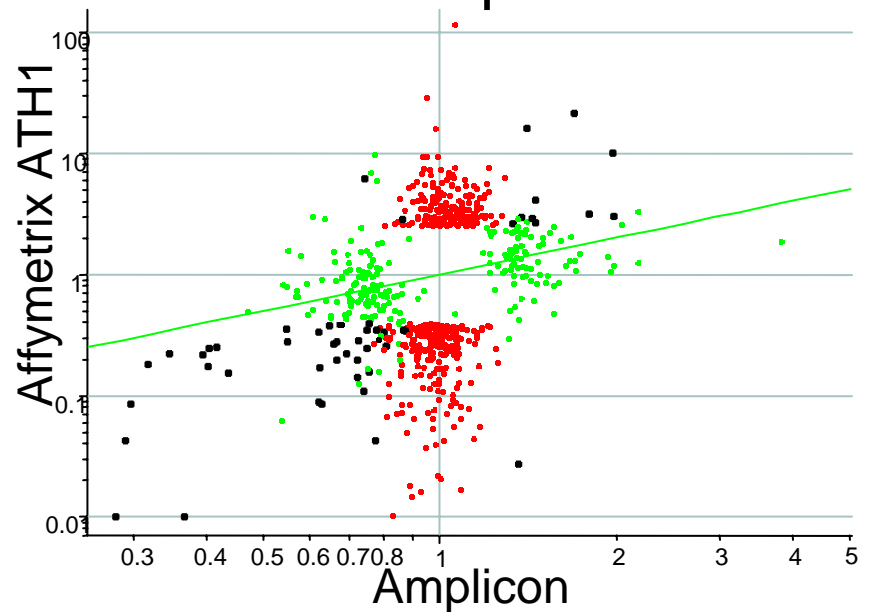
Results

2 hour experiment



Affymetrix AG Amplicon

8 hour experiment



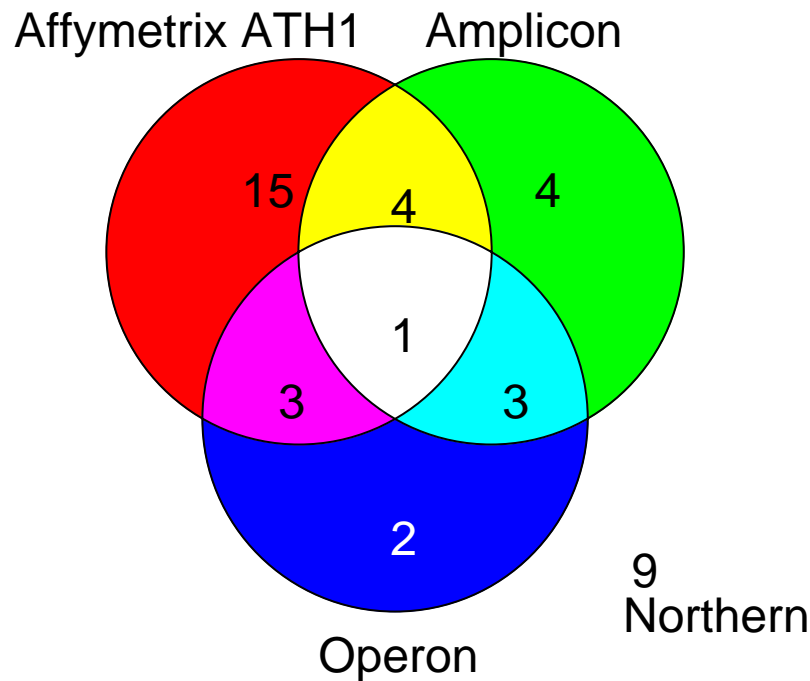
Affymetrix ATH1 Amplicon

Validation by Northern Analysis

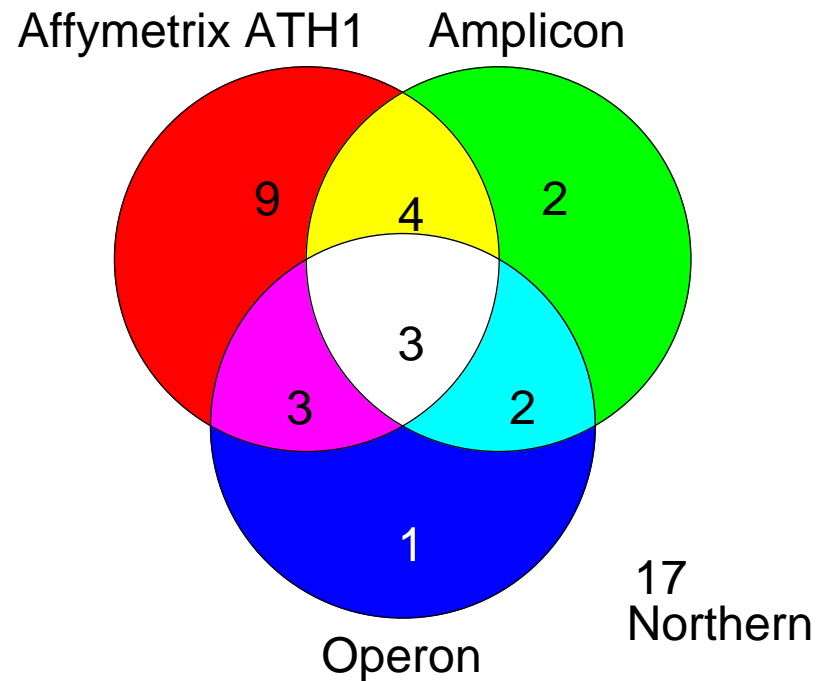
- Chose 41 probes:
 - Significant in 1 platform
 - Significant in 2 platforms
 - Significant in all platforms
- Generate Probes biased to the 3' UTR. Based on Kim et al. (2003) Gen. Res. 13:327-340.
- Hybridization, PhosphorImaging and Volume analysis.
- Normalize to actin.

Representation of Genes Chosen For Northern

2 hour experiment



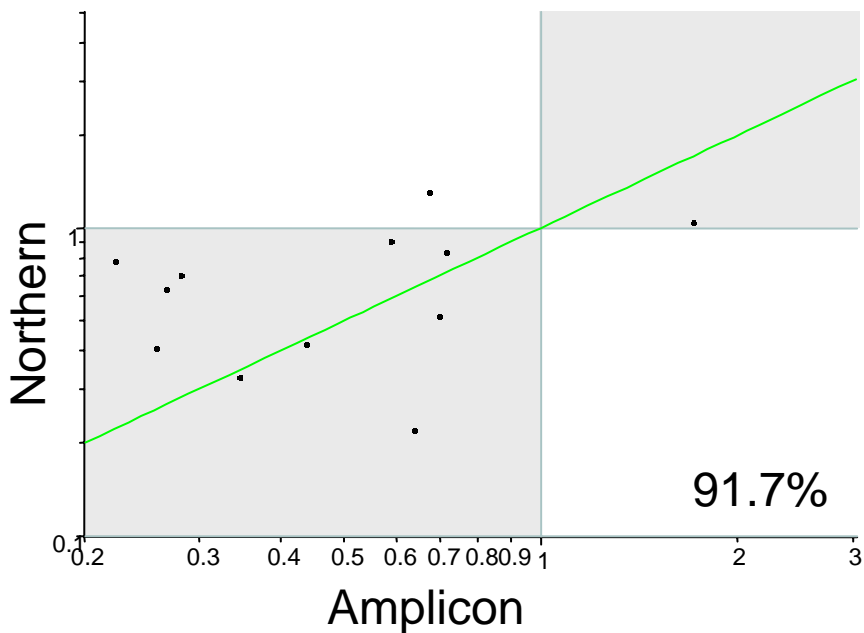
8 hour experiment



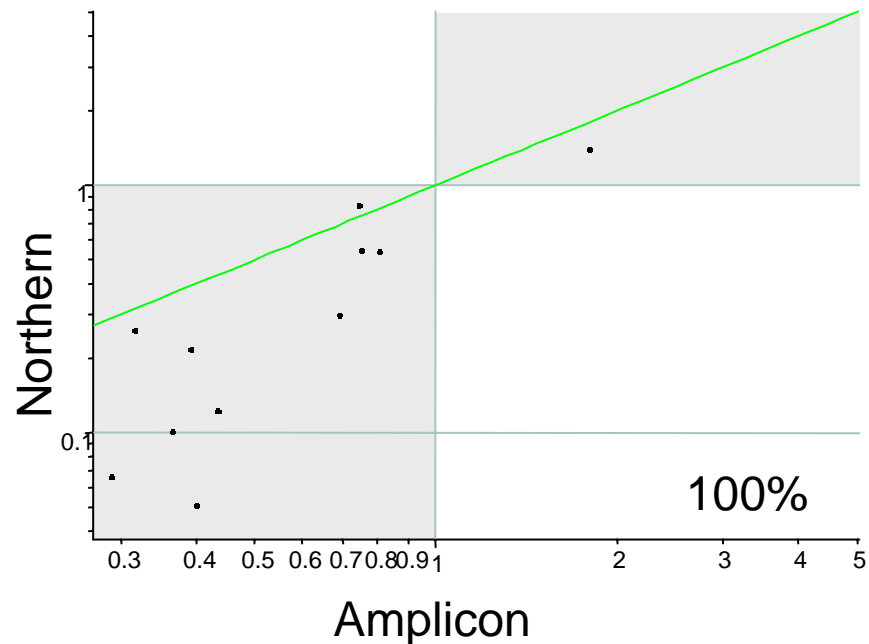
Results

Amplicon.

2 hour experiment



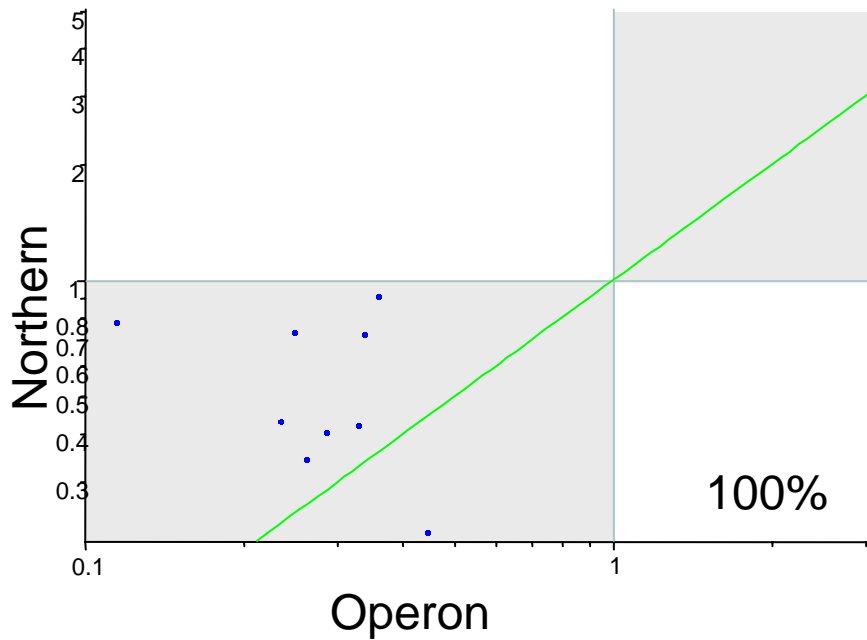
8 hour experiment



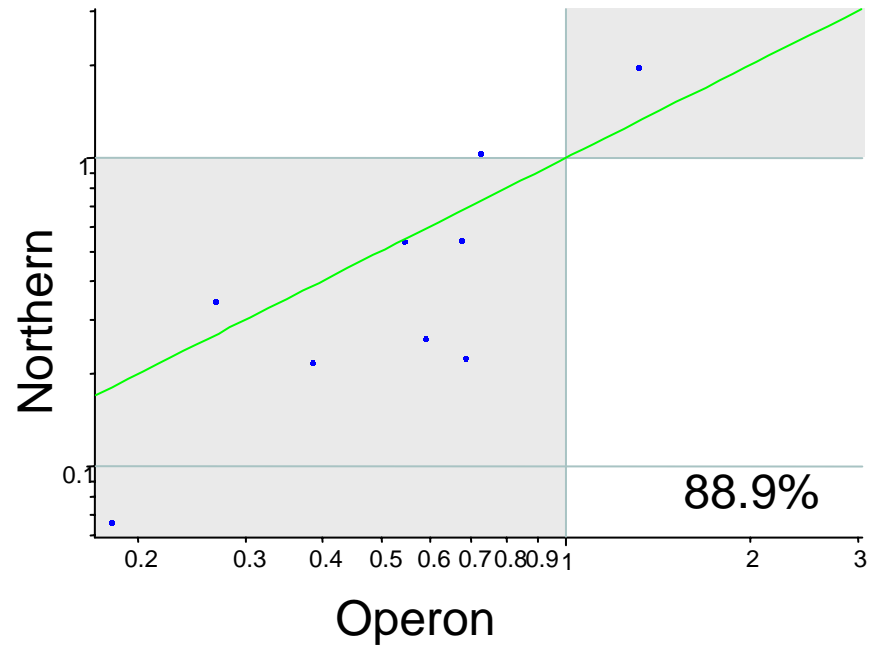
Results

Operon

2 hour experiment



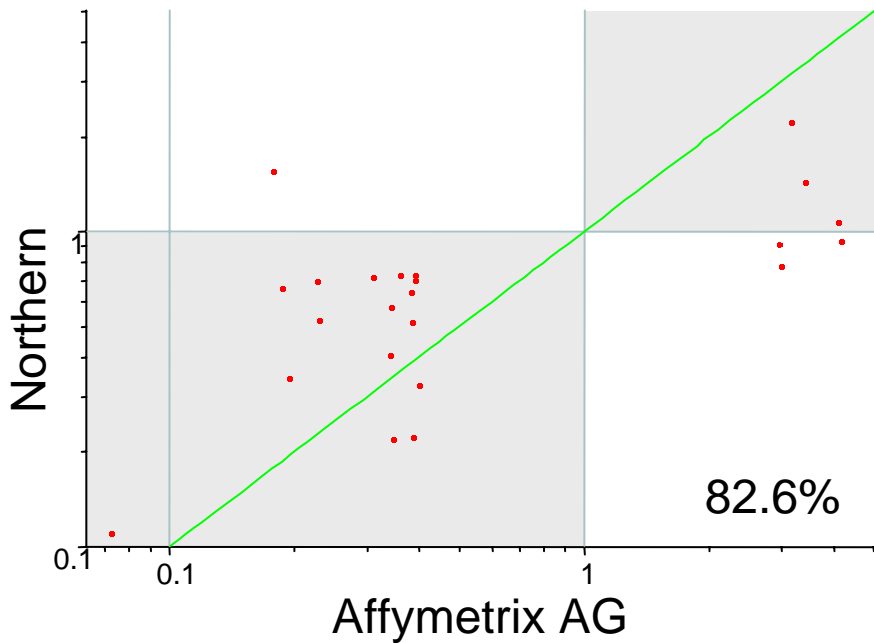
8 hour experiment



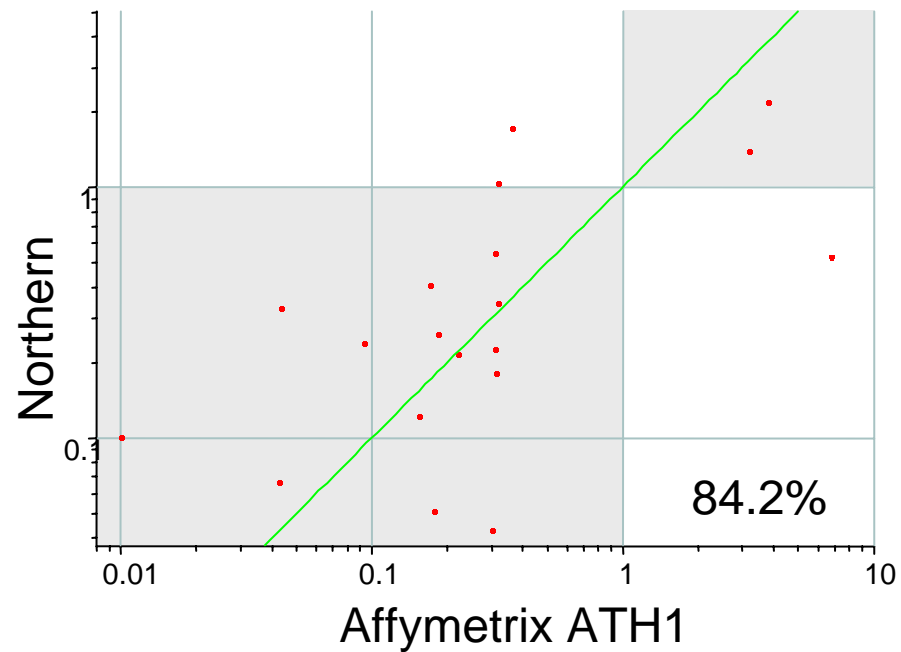
Results

Affymetrix

2 hour experiment



8 hour experiment





Conclusions

- Similarity of results between platforms is dependent on relative channel intensity.
- Very low level of significantly regulated genes conserved across all 3 platforms.
- High level of success validating significant genes by Northern analysis, regardless of platform.
- Amplicon platform performs as well as other platforms.

Acknowledgements

- Pierre Fobert – Group Leader
 - Maureen Anderson
- Bioinformatics
 - Jacek Nowak
 - Mark Wilkinson (UBC)
 - Kevin Koh
- Funding
 - GHI 1 and 2
- Montreal Genome Center
 - André Ponton
- Amplicon Arrays
 - Bill Crosby (U of S)
 - Donald Pokatong (Sam Ratulangi University, Indonesia)
 - David Carter (Robarts, London, ON)