

Analysis of honey bee microarray gene expression data

Sandra Rodriguez Zas
and Heather Adams

Department of Animal Sciences
Institute for Genomic Biology

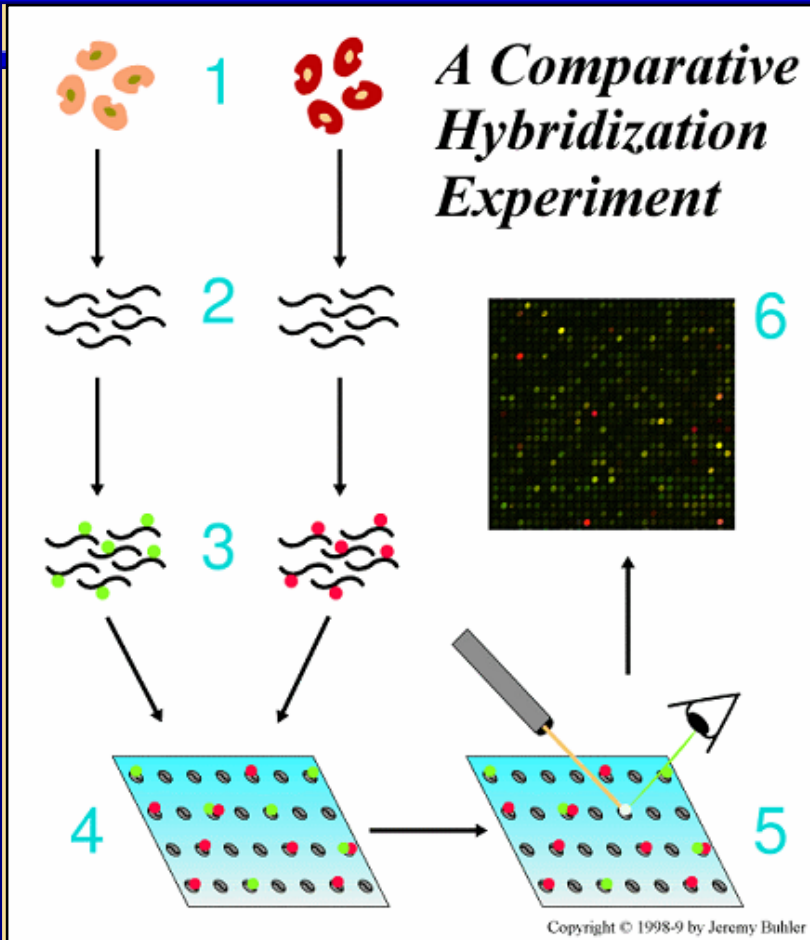


Outline

- **Microarray technology**
- **Experimental design**
- **Normalization and analysis**
- **Robinson's lab experiments**
 - **Ismail et al. (4 treatments)**
 - **Sen Sarma et al. (2 behaviors)**
 - **Alaux et al. (2-3 treatments/age)**
- **Beehive: integrated microarray data workflow**



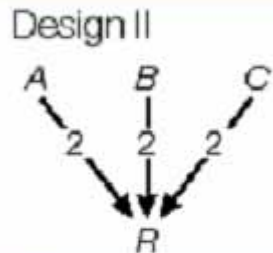
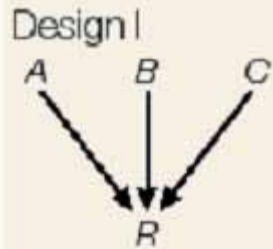
Spotted microarray technology



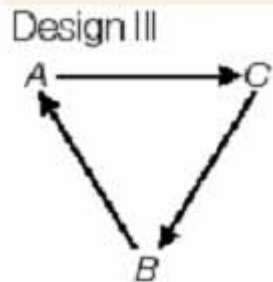
- Two samples hybridized to the same array
- Each sample labeled with a different dye



Microarray designs



Direct design



Design I – T1 as common reference



Design II – direct: sequential



Design III – common reference



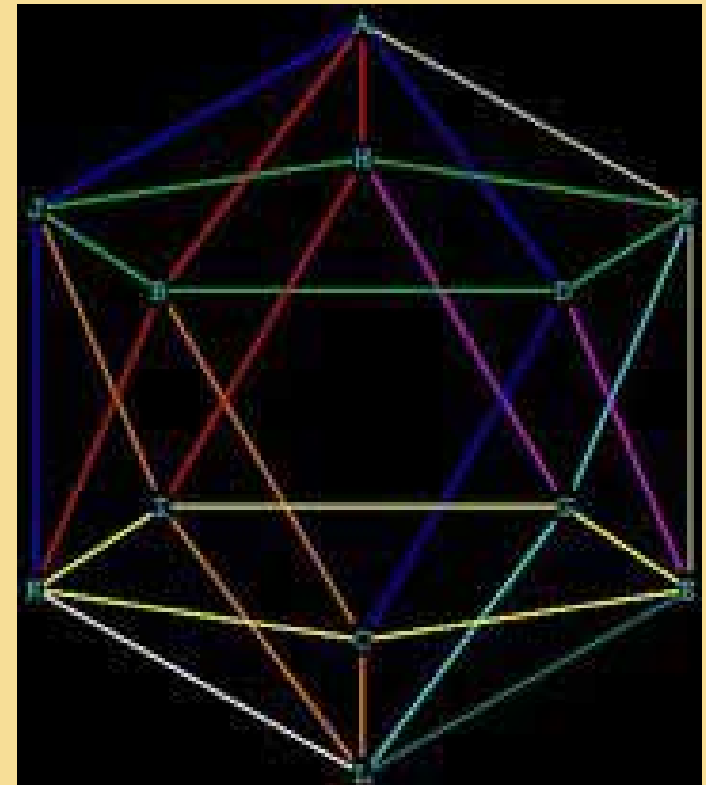
Design IV – T1 as common reference



Design V – direct: loop



Design VI – direct: mixed



(Yang and Speed, 2002)



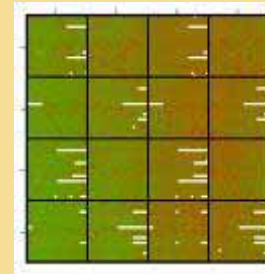
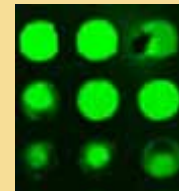
Considerations

- **Microarray gene expression experiments**
 - Goal is to detect interesting genes
 - Can have potentially complex designs
 - Produce vast amounts of data
 - Exhibit technical sources of variation
 - Require thousands of analyses



Filtering and Normalization

- **Removal of unreliable spots and microarray elements**
 - Weak or variable
- **Normalization: removal of systematic technical noise**
 - Dye effects
 - Microarray effects
- **Combination of replicate spots**



Analysis

- **Linear mixed effects models**
 - dye + factor(s) + covariate(s) + interactions + microarray + other blocking factors
- **Multiple-test adjustment of P-values**
- **Consideration of specific contrasts among factor levels**
 - P-value, fold change, observation num.
- **Clustering, PCA, DA**



Sen Sarma & Ismail experiments

- 9K cDNA microarray platform
- Sen Sarma:
 - One-day old nurse and forager bees
 - 4 species (AC, AD, AF, AM)
 - direct comparison, dye swap, 44 microarrays
 - 3 colonies, 2 pools/behavior/species
- Ismail:
 - 4 treatments (1wk, 2wk, CC, treated)
 - 3 colonies, 60 microarrays
 - direct comparison, partial loop and dye swap



Results ($P < 0.0001$)

- Sen Sarma's experiment: number of elements with differential expression
- Ismail's experiment: 13 elements

	AC	AD	AF	AM
AC	47	5	18	19
AD	-	43	7	13
AF	-	-	60	22
AM	-	-	-	132



Alaux et al. experiments

- 13K oligo microarray platform
- Experiment 1: QMP vs NT (2 ages)
 - direct comparison
 - 32 arrays, 2 colonies, 2 ages, 4 bees/treatment
- Experiment 2: BP, Brood, control (2 ages)
 - 4 loop designs/age/colony
 - 48 arrays, 2 colonies, 2 ages, 4 bees/treatment



Alaux's results

- Number of elements with differential expression ($P < 0.0001$ & 1.25 fold change):
 - 6 elements QMP vs NT @ Age 3
 - 22 elements QMP vs NT @ Age 4
 - 18 [5, 1, 4: BP-Br, BP-C, Br-C] @ Age 5
 - 20 [3, 3, 3: BP-Br, BP-C, Br-C @ Age 15



Functional annotation of results

- **Sen Sarma:**
 - lipid transport and metabolism; ion transport / homeostasis/defense response
 - nervous system development/mesoderm and ectoderm development/cell proliferation/cell-cell signaling/regulation of protein transport /nucleic acid binding and metabolism
- **Ismail:**
 - neuromuscular junction development/synaptic growth/dendrite development/ axonogenesis /eye development
 - transmission of nerve impulse/acetyltransferase activity/drug transporter activity/amino acid metabolism
- **Alaux:**
 - 2 muscle proteins; response to virus; mitochondrial ribosomal protein; mating behavior, pupal development, pigmentation



NIH-NIGMS Beehive system

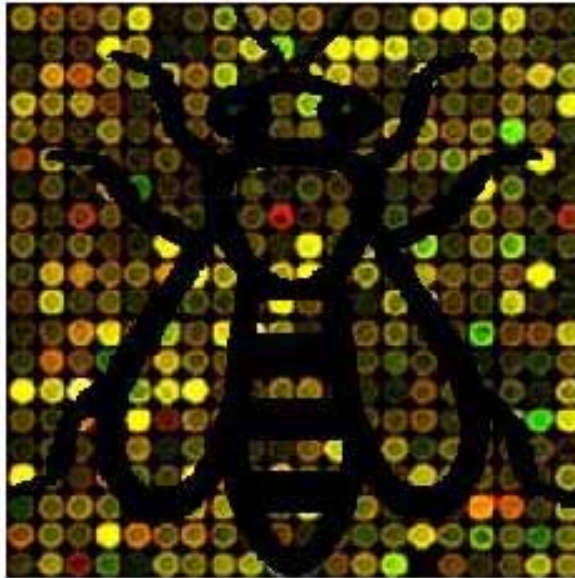
- <http://stagbeetle.animal.uiuc.edu/Beehive3.0>
- Web-based public resource
- Facilitates and supports storage, analysis and interpretation of microarray gene expression data
- Integrated, one-stop, workflow including:
 - Simplified uploading of microarray experiments
 - Storage/download MIAME compliant experiments
 - Normalization and analyze data
 - Visualization of results
 - Query of results
 - Integrated annotations



NIH Beehive data repository

<http://stagbeetle.animal.uiuc.edu/Beehive3.0>

Beehive



User Id

Password

Login

[New User? \(Free Account\)](#)



Uploading Instructions

First, create ZIP archive containing the gpr files with the gene expression data and an experiment design .tdf file (see [Manual](#))
Note: Pop-us must be enabled.

[Browse...](#)

Login, data uploading and, MIAME specification

Platform Information

Title GPL2

GPL2

Organism apis mellifera

[Details](#)

Series Information

Title Eight_Array

Eight_Array

Type

[Details](#)

Sample Information

1132

1132

1132

1133

1133

1133

1134

1134

1134

1135

1135

1135

1136

1136

1136

1137

1137

1137

1138

1138

1138

1139

1139

1139

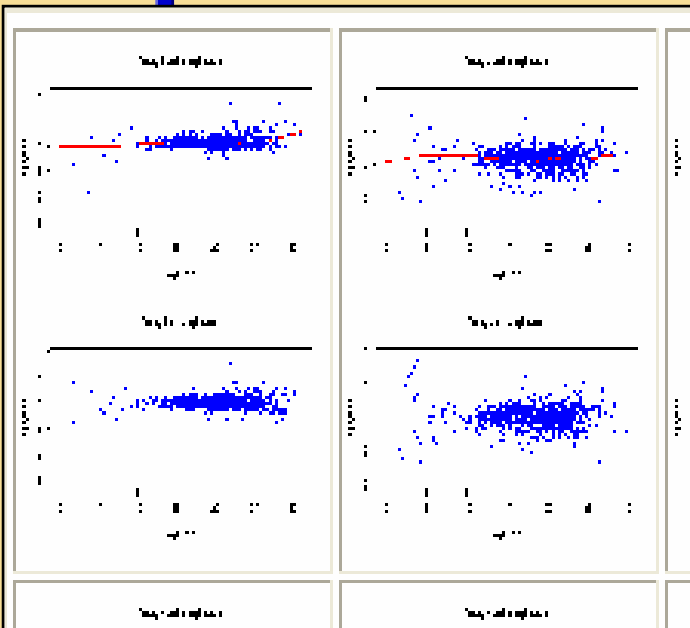
[Details](#)

Export

[Validate Fields and Export](#)

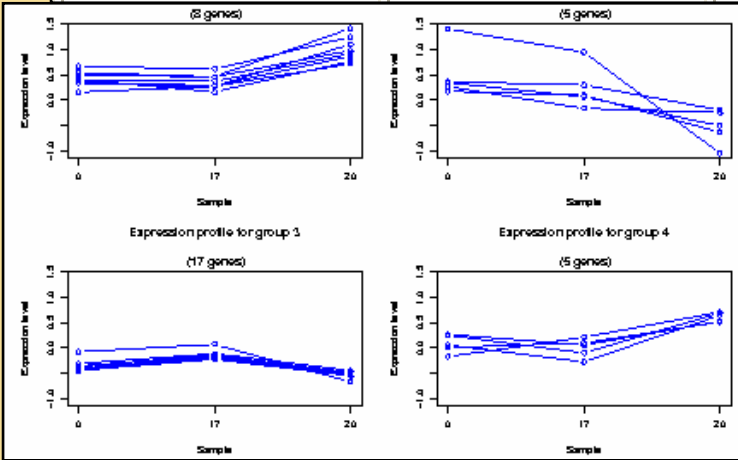
NIH Beehive analysis/visualization

<http://stagbeetle.animal.uiuc.edu/Beehive3.0>



Normalization, clustering and analysis specification

Analysis Name	mod2				
Normalization	<input type="radio"/> shift <input checked="" type="radio"/> glowess <input type="radio"/> rlowess <input type="radio"/> linglog <input type="radio"/> linglogshift				
Number of Spots per Element	2				
Collapsing Method	1				
	<i>* 0 for no average; 1 for mean of the replicates; 2 for median of the replicates.</i>				
List of Elements to Exclude Before Normalization	C:\Documents and Settings\... Browse...				
List of Elements to Exclude After Normalization	Browse...				
Count of Flags not to Exceed	<input checked="" type="checkbox"/> 5				
Minimum Safe Flag Level	-75				
Minimum Level of Gene Expression	200				
	Use in Model	Test term?	Test with contrast?	Assign Reference?	Test an interaction?
Array	Random	<input type="checkbox"/>	<input type="checkbox"/>	Not Supported	<input checked="" type="checkbox"/>
Dye	Fixed	<input type="checkbox"/>	<input type="checkbox"/>	Not Supported	
Age	Not Used	<input type="checkbox"/>	<input type="checkbox"/>	None	
Treatment	Fixed	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	None	
	Return	Reset	Generate R Script		



NIH Beehive query/annotation

<http://stagbeetle.animal.uiuc.edu/Beehive3.0>

GetGeneInfo

<input type="checkbox"/>	Data Name	Analysis Name	cDNA ID	Overall F Value	Overall p-Value	Overall Adj p-Value	Contrast F Value	Contrast p-Value	Contrast Adj p-Value	Contrast	Estimate	Flag Count	Low Expression Count
<input type="checkbox"/>	test	testanal	Empty	7.24593	0.0359651	0.9997833	7.2459331	0.03597	0.99978	ab	1.03831	0	32
	test	testanal	Empty	4.15813	0.0875474	0.9997833	4.1581324	0.08755	0.99978	ab	0.97493	0	32
	test	testanal	Empty	3.85531	0.0972231	0.9997833	3.8553101	0.09722	0.99978	ab	1.37849	0	32
	test	testanal	Empty	4.24831	0.0849221	0.9997833	4.2483054	0.08492	0.99978	ab	1.27412	0	32
	test	testanal	Empty	16.7164	0.0064396	0.9997833	16.7163998	0.00644	0.99978	ab	1.41557	0	32
<input type="checkbox"/>	test	testanal	BB170032A10G07	6.61011	0.0422773	0.9997833	6.6101147	0.04228	0.99978	ab	-0.49508	0	0
<input type="checkbox"/>	test	testanal	BB170030B10B09	7.49689	0.0338232	0.9997833	7.4968915	0.03382	0.99978	ab	-0.32639	0	0
<input type="checkbox"/>	test	testanal	BB170029A20H06	7.88739	0.030819	0.9997833	7.8873885	0.03082	0.99978	ab	-0.29411	0	31

Query of results and annotations

Gene_ID	Official_ID	Fly_ID	GO_ID	GO_Descip	GENBANK_ID
BB170030B10B09	GB18501-PA	CG5809	GO:0006464	protein modification	BI505914
BB170030B10B09	GB18501-PA	CG5809	GO:0003756	protein disulfide isomera	BI505914
BB170030B10B09	GB18501-PA	CG5809	GO:0005489	electron transporter acti	BI505914
BB170030B10B09	GB18501-PA	CG5809	GO:0006118	electron transport	BI505914



Acknowledgements

- Heather Adams
- Bruce Southey
- Younhee Ko
- Brandon Smith
- Gene Robinson and lab
- Charlie Whitfield
- Chris Elsik
- NSF FIBR
- NIH NIGMS

