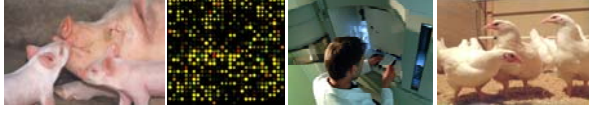


GOEAST and Globaltest:
Two different methods for Gene Ontology analysis
Ina Hulsegge, Arun Kommadath
Animal Sciences Group, Wageningen UR, Lelystad, NL

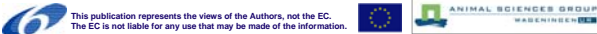



EADGENE and SABRE Post-Analyses Workshop
12-14th November 2008, Animal Sciences Group, Wageningen UR, Lelystad



GOEAST and Globaltest:
Two different methods for Gene Ontology analysis
Ina Hulsegge, Arun Kommadath



This publication represents the views of the Authors, not the EC. The EC is not liable for any use that may be made of the information.

Overview

- **Globaltest**
 - General idea
 - Chicken infection data
- **GOEAST**
 - Hypergeometric test for GO term over representation
- **Discussion**

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad





Globaltest



Testing whether a given group of genes is significantly associated with a response

Goeman JJ. Et al. (2003)

- A global test for groups of genes: Testing association with a clinical outcome, *Bioinformatics*, 20:93-99
- Bioconductor package: *globaltest* (version 4.12)
- <http://www.bioconductor.org>



EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad






Globaltest: general idea

- **Direct gene set testing**
No need for first testing single genes
- **Basic idea**
Based on a regression model: predict group label from expressions
- **Models**
Logistic, linear, Cox model
- **Null hypothesis**
Group label cannot be predicted from the gene expressions
- **Alternative**
Group label can (partly) be predicted from the gene expressions

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

What gene sets?



Any externally defined set that has something in common

Some examples:

- **Pathways**
KEGG, Biocarta
- **Gene Ontology terms**
Biological process, Molecular function, Cellular component
- **Chromosomal regions**
Chromosome arms, cytobands, linkage peaks
- **Published gene sets**
Predictive signatures, gene lists

Goeman, 2008


EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

Informative plots:

- **Sample plot:** how good fits a sample to its phenotype
- **Checkerboard:** Correlation between samples
- **Gene plot:** Influence of single genes to test statistics


EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad



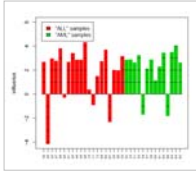
GOEAST and Globaltest:
Two different methods for Gene Ontology analysis
Ina Hulsegge, Arun Kommadath
Animal Sciences Group, Wageningen UR, Lelystad, NL

EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Informative plots:



Checkerboard plot: looking at similarity
Black box: Samples have dissimilar expression profile
White box: Samples have similar expression profile



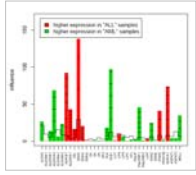
Sampleplot: looking at influential samples
Positive bar: More similar to samples with same condition
Negative bar: More similar to samples with different condition

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Informative plots:



Geneplot: look at influential genes
Line gives expected influence

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad


EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Immune reactions after a homologous or heterologous challenge of broilers primed with *Elmerla maxIma*.

Association of gene expression with MM8_PM8 infection

Global Test result:

Genes Tested	Statistic Q	Expected Q	sd of Q	P-value
all 20458	20458	35.166	11.111	4.5702
				0.0079365



Conclusions

- Infection associated with gene expression profile
- Potential in predicting infection from gene expression
- Chickens with similar gene expression values tend to have similar infection

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Association with MM8_PM8 infection

Top 5	Genes	z.score >2	P-value	FDR.adjusted
Biological Process				
GO:0032831 positive regulation of CD4-positive..	1	1	0.0020	0.0579
GO:0051017 actin filament bundle formation	6	2	0.0023	0.0579
GO:0006996 organelle organization and biogenesis	3	2	0.0024	0.0579
GO:0048739 cardiac muscle fiber development	1	1	0.0025	0.0579
GO:0015808 L-alanine transport	1	1	0.0025	0.0579
Molecular function				
GO:0019976 interleukin-2 binding	2	2	0.0020	0.0532
GO:0005280 hydrogen:amino acid symporter activity	1	1	0.0025	0.0532
GO:0015180 L-alanine transmembrane transporter...	1	1	0.0025	0.0532
GO:0015187 glycine transmembrane transporter activity	1	1	0.0025	0.0532
GO:0015193 L-proline transmembrane transporter...	1	1	0.0025	0.0532
Cellular component				
GO:001891 phagocytic cup	1	1	0.0019	0.0399
GO:0001726 ruffle	18	9	0.0030	0.0399
GO:0005719 nuclear euchromatin	2	1	0.0031	0.0399
GO:0005652 nuclear lamina	1	1	0.0032	0.0399
GO:0043626 L-alanine transport	1	1	0.0032	0.0399

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Association with MM8_PM8 infection

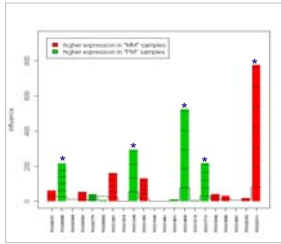
Top 5	Genes	z.score >2	P-value	FDR.adjusted
Biological Process				
GO:0051017 actin filament bundle formation	6	2	0.0023	0.0469
GO:0006996 organelle organization and biogenesis	3	2	0.0024	0.0469
GO:0015816 glycine transport	2	1	0.0025	0.0469
GO:0016042 lipid catabolic process	7	6	0.0027	0.0469
GO:0009113 purine base biosynthetic proces	4	2	0.0029	0.0469
Molecular function				
GO:0019976 interleukin-2 binding	2	2	0.0020	0.0404
GO:0015187 glycine transmembrane transporter activity	2	1	0.0025	0.0404
GO:0031013 troponin 1 binding	2	1	0.0029	0.0404
GO:0003847 1-alkyl-2-acetylglycerophosphocholine....	2	2	0.0030	0.0404
GO:0004438 phosphatidylinositol-3-phosphatase activity	2	2	0.0031	0.0404
Cellular component				
GO:0001726 ruffle	18	9	0.0030	0.0339
GO:0005719 nuclear euchromatin	2	1	0.0031	0.0339
GO:0005884 actin filament	12	6	0.0032	0.0339
GO:0000307 cyclin-dependent protein kinase...	5	3	0.0037	0.0339
GO:0016529 sarcoplasmic reticulum	4	3	0.0037	0.0339

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

EADGENE SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Gene plot: GO:0001726 (example)



* Significant in LIMMA

Goeman, 2008

EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad

GOEAST and Globaltest:
Two different methods for Gene Ontology analysis
Ina Hulsegge, Arun Kommadath
Animal Sciences Group, Wageningen UR, Lelystad, NL

I. MM8_PM8 – upregulated genes

Cat.	Terms	q	m	t	k	p
BP	cell structure disassembly during apoptosis	4	14	7099	303	0.029
	cellular component disassembly	4	15	7099	303	0.041
	phospholipid biosynthetic process	4	14	7099	303	0.029
	membrane lipid biosynthetic process	4	15	7099	303	0.041
	glutathione metabolic process	3	8	7099	303	0.029
	regulation of intracellular pH	2	5	7099	303	0.095
MF	RNA binding	14	138	7099	303	0.099
	nucleotidase activity	2	5	7099	303	0.095
	endonuclease activity	3	5	7099	303	0.003
	myosin binding	3	10	7099	303	0.076
CC	cyclin-dependent protein kinase holoenzyme complex	2	5	7099	303	0.095

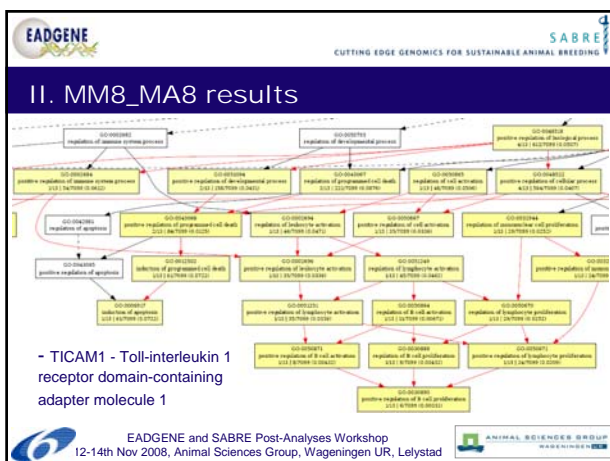
I. MM8_PM8 – downregulated genes

Cat.	Terms	q	m	t	k	p
BP	cardioblast differentiation	2	5	7099	216	0.054
	regulation of epidermis development	2	7	7099	216	0.096
	polysaccharide biosynthetic process	2	6	7099	216	0.096
	cellular polysaccharide biosynthetic process	2	6	7099	216	0.096
	positive regulation of bone mineralization	2	5	7099	216	0.054
	regulation of lipoprotein lipase activity	2	6	7099	216	0.096
	cation:amino acid symporter activity	2	6	7099	216	0.096
MF	LIM domain binding	2	5	7099	216	0.054
	tumor necrosis factor receptor binding	2	7	7099	216	0.096
	low-density lipoprotein receptor binding	4	12	7099	216	0.004
	endodeoxyribonuclease activity	2	7	7099	216	0.096
CC	ruffle	4	21	7099	216	0.068

- I. MM8_PM8 – downregulated genes**
- GO:0010002 - cardioblast differentiation
 - RIGG02337 - Integrin beta-1 precursor (**ITGB1**)
 - Signaling in Immune System
 - RIGG04390 - Transforming growth factor beta-2 precursor (**TGFβ2**)
 - suppressive effects on IL-2 dependent T-cell growth

II. MM8_MA8 – downregulated genes

GO Terms (BP)	q	m	t	k	probes	p
defense response to virus	1	8	7099	13	TICAM-1	0.004
immune response-activating signal transduction	1	12	7099	13	TICAM-1	0.007
positive regulation of B cell proliferation	1	6	7099	13	TICAM-1	0.003
positive regulation of chemokine biosynthetic process	1	5	7099	13	TICAM-1	0.002
positive regulation of NF-kappaB transcription factor activity	1	12	7099	13	TICAM-1	0.007
positive regulation of nitric oxide biosynthetic process	1	7	7099	13	TICAM-1	0.004
response to lipopolysaccharide	1	5	7099	13	TICAM-1	0.002
positive regulation of I-kappaB kinase/NF-kappaB cascade	2	54	7099	13	TICAM-1 RHOA	0.005

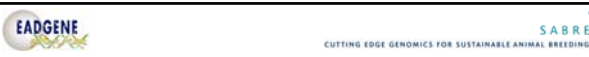


II. MM8_MA8 results (contd.)

GO Terms (BP)	q	m	t	k	probes	p
integrin-mediated signaling pathway	1	23	7099	13	ITGA1	0.019
Neutrophil [hetero-] chemotaxis	1	10	7099	13	ITGA1	0.005
positive regulation of protein ubiquitination	1	10	7099	13	LRRK2	0.005

- Low count in study set
- Processes suggestive of innate immune system and inflammatory response


GOEAST and Globaltest:
Two different methods for Gene Ontology analysis
Ina Hulsegge, Arun Kommadath
Animal Sciences Group, Wageningen UR, Lelystad, NL



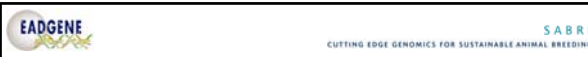
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

III. MM8_MM24 results

GO Terms (BP)	q	m	t	k	p
regulation of protein binding	3	11	7099	146	0.09
regulation of cytokine biosynthetic process	5	37	7099	146	0.07
positive regulation of interleukin-2 biosynthetic process	2	7	7099	146	0.04
microtubule bundle formation	2	6	7099	146	0.03




EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad



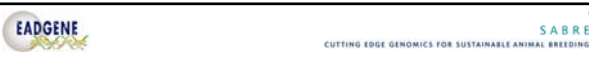
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Conclusion

- Results of Globaltest and GOEAST hypergeometric testing could not be compared
- GOEAST results – indicated immune processes – low count in the study set




EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad



CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

- Thank you
- QUESTIONS



EADGENE and SABRE Post-Analyses Workshop
12-14th Nov 2008, Animal Sciences Group, Wageningen UR, Lelystad