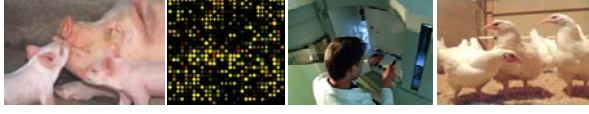


EADGENE **SABRE**
 CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

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



EADGENE mastitis data meta analysis

Elisabetta Giuffra

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.

Parco Tecnologico Padano
 La ricerca è la risposta
 L'innovazione è la certezza

EADGENE **SABRE**
 CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Starting points of meta-analysis project

- A “genomic” review on important livestock diseases
 - Davies, Genini, Bishop, Giuffra 2008 Animal, in press
- Choice of project from EADGENE work: **gene expression of mastitis in ruminants**
- Availability of data sets: an ongoing process (e.g. from SABRE)
- Knowledge transfer agreement (KTA), or simply people willing to collaborate

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

Parco Tecnologico Padano
 La ricerca è la risposta
 L'innovazione è la certezza

EADGENE **SABRE**
 CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Structure of the project

1. Collection of data with contrasting

- designs and sizes
- host species (cattle, sheep, goat) and tissues
- pathogens (*E. coli*, *S. aureus*, *S. uberis*)
- arrays (Ark-Genomics, NBFGC)
- time points PI
- ... **but re-analyzed in the same way!!!**

2. Evaluation of data sets and methods

- tools: meta-analysis with pointillist

3. Inter- & intraspecies specific comparisons

4. Evaluation of results

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

Parco Tecnologico Padano
 La ricerca è la risposta
 L'innovazione è la certezza

EADGENE **SABRE**
 CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Meta-analysis

A combination/integration of data across studies addressing the same biological question (*R. A. Fisher (1944) Source of the idea of cumulating probability values; W. G. Cochran (1953) Discusses a method of averaging means across independent studies*)

- It can be applied on raw data or on the results of single analyses

Benefits:

- HIGHER power to detect affected genes, reliability of results - LOWER number of false positives

Problems:

- Accuracy = f [technology]
- Bias = f [data set] (abundance, scope, mode of analysis)

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

Meta-analysis and microarrays

- Pennings et al. 2008 Identification of a common gene expression response in different lung inflammatory diseases in rodents and macaques. PLoS ONE. Jul 9;3(7):e2596.
- Edwards et al. 2008 Meta-analysis of microarray gene expression in mouse stem cells: redefining stemness. PLoS ONE. 16;3(7):e2712.
- Greco et al. 2008 Physiology, pathology and relatedness of human tissues from gene expression meta-analysis. PLoS ONE. 3(4):e1880.
- Sohal et al 2008 Meta-analysis of microarray studies reveals a novel hematopoietic progenitor cell signature and demonstrates feasibility of inter-platform data integration. PLoS ONE. 13;3(8):e2965.
- Pihur et al. 2008 Finding common genes in multiple cancer types through meta-analysis of microarray experiments: A rank aggregation approach. Genomics.

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

1. Datasets for mastitis

Different cDNA microarrays (ARK 17k bovine cDNA array, ARK 20k cDNA bovine cDNA array, NBFGC bovine cDNA array).

Group(s)	Pathogen(s)	Challenge system	Sample type
RI/RIBFA	<i>S. aureus</i> / <i>E. coli</i>	Udder (cattle)	Udder tissues
NSVS	<i>S. aureus</i>	Macrophages (cattle)	Primary cells
ID L./RIBFA	<i>S. uberis</i>	Udder (cattle)	Udder tissues
UNIMI/PTP	<i>S. aureus</i>	Udder (goat)	Leukocytes
INRA	<i>S. aureus</i>	DC (sheep)	Primary cells

Evaluation of the experimental designs, critical interpretation of previous analyses, identification of comparisons to be included in meta-analysis...

Analysis: Bluefuse, Bioconductor

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

2. Re-analysis of the microarray raw data

The harmonization of the gene names across these different arrays was a crucial step to enable meta-analysis.

- **nomenclature**
ARK 17k cDNA array ≠ ARK 20k cDNA array
- **identity**
ARK cDNA arrays ≠ NBFGC bovine cDNA array

By blast comparison between the sequences (threshold for similarity >100 bp without mismatches): clones in common = 5,772/29,413

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

Parco Tecnologico Padovano
 La ricerca è la impresa

3. Choice of tool for meta-analysis: pointillist

A data integration methodology for systems biology

Daehee Hwang*, Alistair G. Rust*, Stephen Ramsey*, Jennifer J. Smith*, Deena M. Leslie*, Andrea D. Weston*, Pedro de Atauri*, John D. Aitchison*, Leroy Hood*, Andrew F. Siegel†, and Hamid Bolouri††

A data integration methodology for systems biology: Experimental verification

Daehee Hwang*, Jennifer J. Smith*, Deena M. Leslie*, Andrea D. Weston*, Alistair G. Rust*, Stephen Ramsey*, Pedro de Atauri*, Andrew F. Siegel†, Hamid Bolouri†*, John D. Aitchison*, and Leroy Hood*†

PNAS | November 29, 2005 | vol. 102 | no. 48

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

Parco Tecnologico Padovano
 La ricerca è la impresa

Pointillist: evidence-weighted inferer

- is a data integration methodology that can handle multiple and different data sets to determine which elements are affected by a perturbation
- false positives / negatives are minimized
- assigns to each type of data a weight (reliability parameter), based on consistency with the other types of evidences ("power" option)
- is efficient for detecting similarities or commonalities between different datasets
- very good for situations where different genomic platforms have been used
- weak datasets can be beneficially combined into the analyses without hindering the overall interpretation of the results

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

Parco Tecnologico Padovano
 La ricerca è la impresa

Current efforts and final workplan

- **Milestone** - Complete annotation!!!
- Complete the meta-analysis of the first set of data: *core group* (NSVS, INRA, RI, PTP) – *IPA* (Ingenuity Package Analysis)
- New datasets arriving soon: sheep and goats diverging for SSC (+ bovine ref?).
- Improving meta-analysis tools (optimization of the meta-analysis criteria and tools, advanced data mining (Univ. Milan - Dept. Information Technologies)
- A multi-disciplinary workshop together with EADGENE DAYS 2009 (contributors to the study of the mastitis traits from medicine, bioinformatics, microbiology, veterinary sciences...)

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

Parco Tecnologico Padovano
 La ricerca è la impresa

Contributors so far, and thanks to

<p>PTP</p> <ul style="list-style-type: none"> • S. Genini, G. Sclep, E. Giuffra 	<p>UNIMI</p> <ul style="list-style-type: none"> • G. Pisoni, B. Castiglioni, P. Moroni
<p>RI</p> <ul style="list-style-type: none"> • S.C. Bishop, D. Waddington, D.J. de Koning, C. Haley, S. Sechi 	<p>RIBFA/RI</p> <ul style="list-style-type: none"> • H.-M. Seyfert, K. Jensen
<p>Ark-Genomics</p> <ul style="list-style-type: none"> • M. Fell, D. Mouzaki 	<p>NSVS</p> <ul style="list-style-type: none"> • I. Olsaker, G. Boman
<p>EADGENE-SABRE</p> <ul style="list-style-type: none"> • S. Ayuso, C. Channing 	<p>ID-L</p> <ul style="list-style-type: none"> • A. de Greeff, H. Smith

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

Parco Tecnologico Padovano
 La ricerca è la impresa