



European Animal Disease Genomics Network of Excellence
 for Animal Health and Food Safety



Animal Disease Genomics: Opportunities and Applications 10th - 11th June 2008, Edinburgh, UK



MicroArray annotation within EADGENE Christophe Klopp



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.



EADGENE Oligo-sets

→ 3 oligo-sets are made available through the EADGENE network :

Species	Origin	Nb oligos	Design year	Number of slides
Chicken	ARK-Genomics	20,460	2005	RI 445 / INRA 200
Cattle	BOMC / ARK Genomics	24,000	2003	RI 85
Pig	DIAS	25 210	2003	0

More information : <http://www.eadgene.info/>
 Integration / Biological resources and facilities

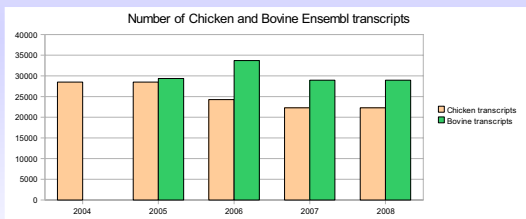


Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 2



Why should we re-annotated the oligo-sets?

- Genomic information evolves (new genome assembly versions, new gene build versions,..)
- Annotation information changes (GO, orthologous genes, pathways, ...)



Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 3



Ideal conditions to design an oligo-set

- Complete genome sequence, with highly specific sequences for the transcripts.
- Complete transcript annotation.
- Other type of RNA annotations (like ncRNA).



Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 4



Quality criteria of an oligo-set

- Representation of the transcriptome : each oligo represents an exon, a transcript or a gene
- Kane specificity criteria :
 - >75% similarity over 50 bp for 50 mers
 - No stretch of more than 15 contiguous bp

4552-4557 *Nucleic Acids Research*, 2000, Vol. 28, No. 22
 © 2000 Oxford University Press
Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays
 Michael D. Kane, Timothy A. Jatkoe, Craig R. Stumpff, Jia Lu¹, Jeffrey D. Thomas and Steven J. Madore²
 Department of Molecular Biology and Genomics and ¹Department of Infectious Diseases, Pfizer Global Research and Development, Ann Arbor, MI 48105, USA
 Received July 5, 2000; Revised and Accepted September 21, 2000

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 5



Annotation pipeline

- Inputs :
 - Oligo fasta file
 - Ensembl transcript fasta file
 - Genome chromosome fasta file
- Analysis steps
 - Specificity analysis
 - Annotation retrieval for oligos uniquely linked to a gene
- Outputs :
 - Oligo set description (pdf file)
 - Annotation files (text file)



Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 6



EADGENE

Oligo set description file (pdf)

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 7

EADGENE

Specificity analysis : general overview

- Aims :
 - Link each oligo to a gene
 - Give information about possible noise (cross hybridization)
- Process :
 - Alignment of the oligos versus the transcripts and versus the genome
 - Oligo classification
- Outputs :
 - Oligo classification

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 8

EADGENE

Specificity analysis : Classification

- Ideal :
 - One full-length match
 - No noise (cross hybridization)
- Processing steps
 - Distinguish information from noise
 - Classify the oligos

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 9

EADGENE

Specificity analysis : Class system

- Table of classes : based on transcript alignments

Noise / Number of Hits	No Hit	One hit	More than one hit
No noise	5	1	6
Noise	4	2	
	3		

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 10

EADGENE

Specificity analysis : Chicken oligo set

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 11

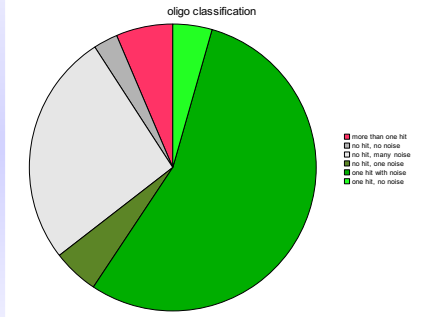
EADGENE

Specificity analysis : Bovine oligo set

Animal Disease Genomics: Opportunities and Applications
 10th - 11th June 2008, Edinburgh, UK / page 12



Specificity analysis : Pig oligo set



Animal Disease Genomics: Opportunities and Applications
10th - 11th June 2008, Edinburgh, UK / page 13



Annotation retrieval

- Oligo location (alignment result)
- Ensembl API (perl)
 - Retrieve orthologous genes for human, mouse and rat
 - Retrieve GO annotation
 - Retrieve Identifier in other nomenclatures (HGNC,...)
- Kegg database
 - Enzyme
 - Pathway ID



Animal Disease Genomics: Opportunities and Applications
10th - 11th June 2008, Edinburgh, UK / page 14



Future work

- Annotation workshop in November 2008 (NL)
- Add description and annotation file access to the EADGENE web site.
- Add Kegg information retrieval module to the pipeline.
- Generate description and annotation files for new versions of Ensembl.
- Create an annotation database and tools to retrieve the annotations in a chosen format.

Animal Disease Genomics: Opportunities and Applications
10th - 11th June 2008, Edinburgh, UK / page 15



Acknowledgments

- Pierrot Casel / INRA
- Pieter Neerincx / WU
- Haisheng Nie / WU
- Jack Leunissen / WU
- Martien Groenen / WU
- Dennis Prickett : IAH

Animal Disease Genomics: Opportunities and Applications
10th - 11th June 2008, Edinburgh, UK / page 16

