Module optionnel X "Génomique/Protéomique
Université Pierre et Marie Curie
Mercredi 17 novembre 2004

Application des approches d'analyse du transcriptome à l'étude des étapes précoces d'autoimmunité

….de la Génétique à l'Immunologie …

Evie Melanitou

LISTE DES MALADIES AUTOIMMUNES

Cette liste est établie en allant des maladies les plus spécifiques d'organes à celles qui on sont le moins spécifiques

MALADIES	ANTIGENES CIBLES
Thyroïdite de Hashimoto	Thyroglobuline, microsomes
Maladie de Basedow	Récepteur de la TSH *(5)*
Maladie d'Addison (1)	Cortico surrénale
Insuffisance hypophysaire	Hypophyse
Anémie de Biermer	Muqueuse gastrique
Spondylarthrite ankylosante	Enthèses
Arthrites réactionnelles	Enthèses
Rhumatisme psoriasique	Enthèses, cartilage, synoviale
Uvéite antérieure aigus	Chambre antérieure de l'oeil
Rétinochoroïdopathie Birdshot	Choroïde, rétine
Polychondrite atrophiante	Cartilage
Certaines stérilités	Spermatozoïdes, ovaires
Diabète juvénile	**ilôts de Langherans, insuline**
Syndrome sucré de Goodpasture	Membrane basale glomérulaiœ
Myasthénie	Muscle strié, récepteur de l'Ach (6)
Maladie de La Peyronie	Aponévroses du pénis
Rhumatisme articulaire aigu	Myocarde, streptocoques
Pemphigus	Ponts intercellulaires de l'épiderme
Dematomyosite	Noyaux, Jo1, muscles
Sclérodermie	Tissu conjonctif, noyaux, Sc170
Connectivite mixte	Noyaux, RNP
Lupus érythémateux discoïde	Noyaux

MALADIES	ANTIGENES CIBLES
Pemphigoïde bulleuse	Membrane basale de l'épiderme
Epidermolyse bulleuse acquise	Fibroblastes, kératinocytes
Dermatite heq~étifomie	Gliadine, réticuline
Maladie coeliaque	Gliadine, réticuline
Pelade	Follicule pileux
Néphropathie membraneuse idiopathique	Glomérules rénaux
Néphrose lipoïdique de l'enfant	Glomérules rénaux
Néphropathie à IgA	Glomérules rénaux
Sclérose en plaques	Oligodendrocytes, myéline
Narcolepsie	Certaines cellules cérébrales
Certaines anémies hémolytiques	Hématies
Certaines granulopénies	Granulocytes
Purpura thrombocytopénique idiopathique	Plaquettes
Cirrhose biliaire primitive	Mitochondries
Hépatite chronique active (2)	Muscles lisses, noyaux, mitochond microsomes
Syndrome de Gougerot-Sjögren	Glandes lacrymales, salivaires, no SSA, SSB
Maladie de Horton (3) et PPR	Artère temporale, muscles des ceint
FaScilte de Shulman	Peau, aponévroses
Arthrite chronique juvénile (4)	Cartilage, synoviale, oeil, noyaux
Polyarthrite tiiumatoïde	Cartilage, synoviale, IgG, noyaux
Lupus érythémateux disséminé	Noyaux, ADN, Sm, cardiolipine

Outlines

Part I

- **The Genetics of T1D in the NOD mouse**

Part II

- **Gene Expression Networks in Early Autoimmunity**

a. Subphenotype: -characterization
-significance relative to diabetes

b. Gene signatures: -Description of the approach
-Biological significance

c. Correlation: Genetics ⟷ Genomics

Genetics of Type 1 Diabetes

Prevalence in the human population	0.3%
Identical twin concordance	36-75%
MHC-identical sibling concordance	12%
Average to siblings of affected probands	6%

Complex inheritance:
T1D is a multifactorial disease with a strong genetic component and environmental factors contributing to its aetiology

INSTITUT PASTEUR

Physiopathology

Type 1 Diabetes: a chronic autoimmune destruction of the insulin producing β cells of the pancreas

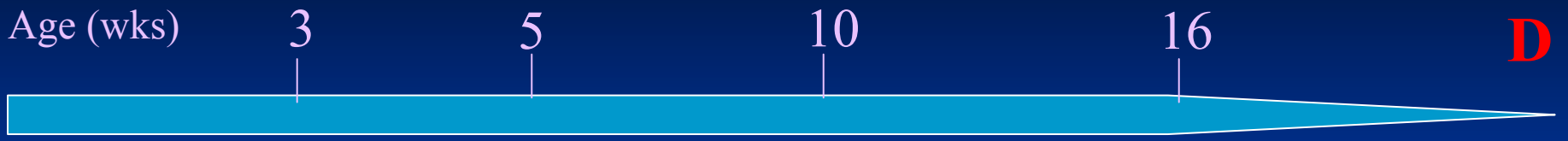
- Appearance of autoantibodies
- Presence of insulitis
- Defects in T lymphocyte activity
- Sensitivity to immunosuppression
- Autoimmune disease

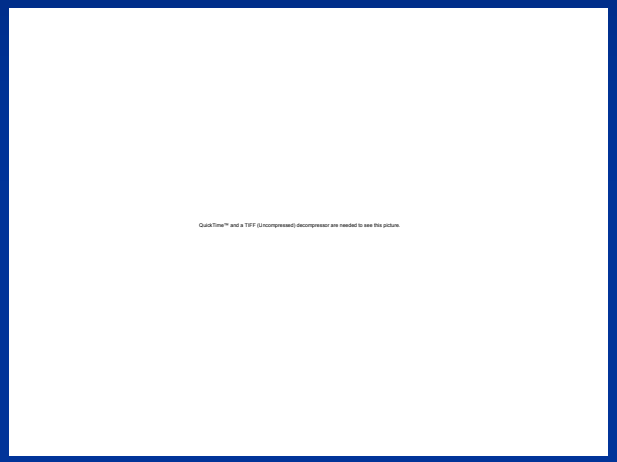
Human Diabetes versus mouse model

Characteristics	Human	NOD
•Insulitis	present	present
•Autoantibodies	present	present
•Defects in T-lymphocyte activity	present	present
•Sensitivity to immunosuppression	present	present
•MHC susceptibility genes	present	present
•Disease frequency in both sexes	same	females>males

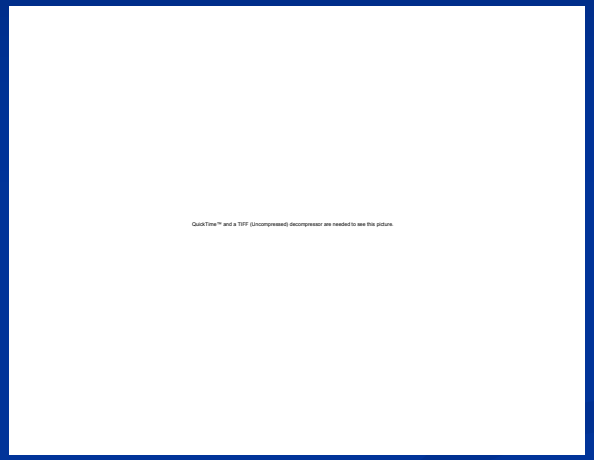
INSTITUT PASTEUR

Disease progression

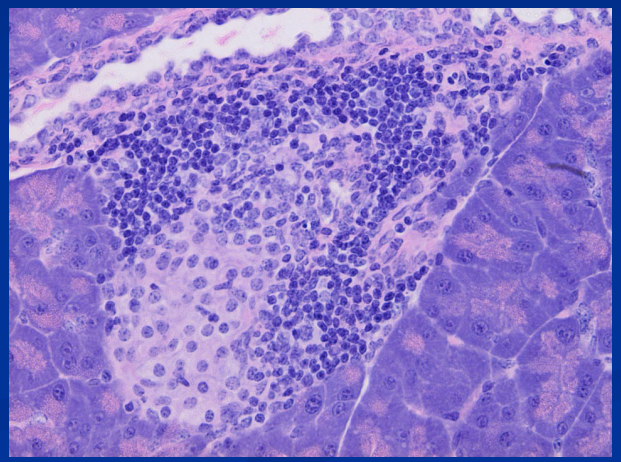




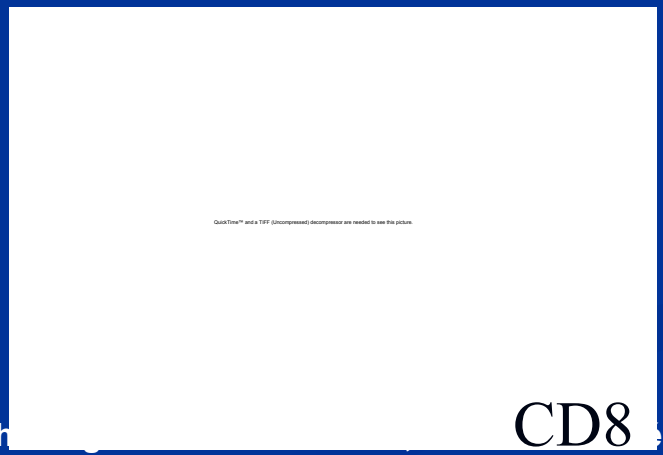
Intact islet



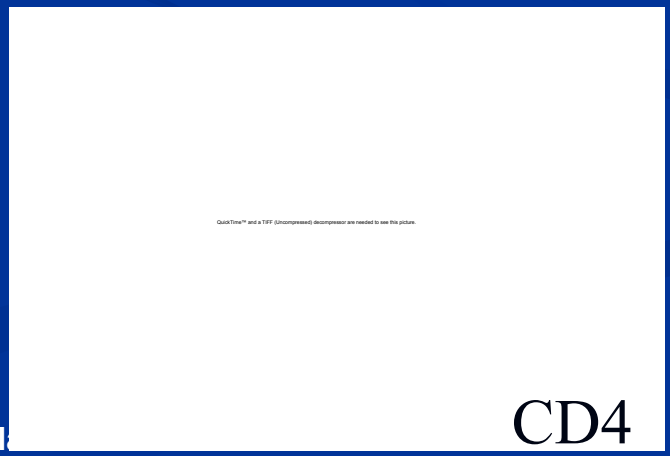
Periinsulitis



Insulitis



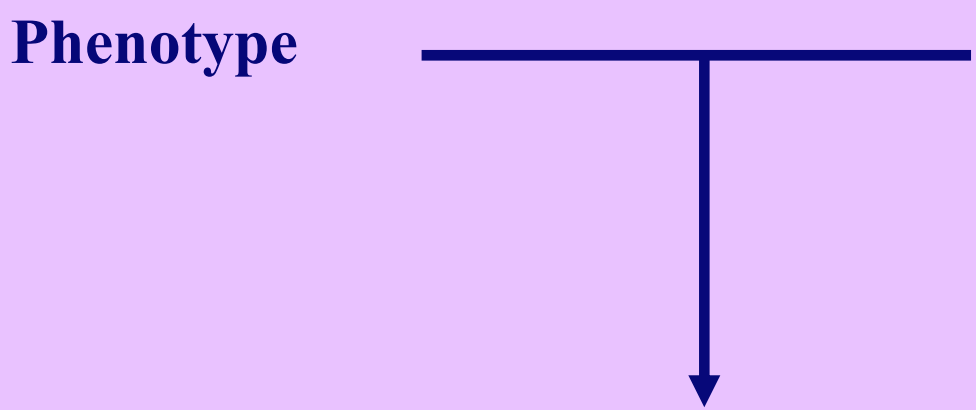
CD8



CD4

INSTITUT PASTEUR

GENETIC ANALYSIS



Localisation of genes of unknown/known function responsible for the phenotype

If there are obstacles, the shortest line between two points maybe the crooked line.
Bertold Brecht (1898-1956)

INSTITUT PASTEUR

Positional cloning

Genetic disease

↓ *Linkage analysis*

Map to a chromosome site

↓

Identify functional genes and ESTs that map to the same site as the disease gene

↙ ↘

Implement mutation detection assays

Select and characterize Genomic clones

↓

Implement mutation detection assays

INSTITUT PASTEUR

STRATEGY FOR THE RESISTANT PARTNER CHOICE

a. Intra-specific crosses

NOD	x	*Mus musculus domesticus*
Susceptible		Resistant

* Less polymorphic
* Less genetic heterogeneity
* Higher phenotype incidence
* Stonger statistics (significance)?

C57Bl/6
C57Bl/10
NON

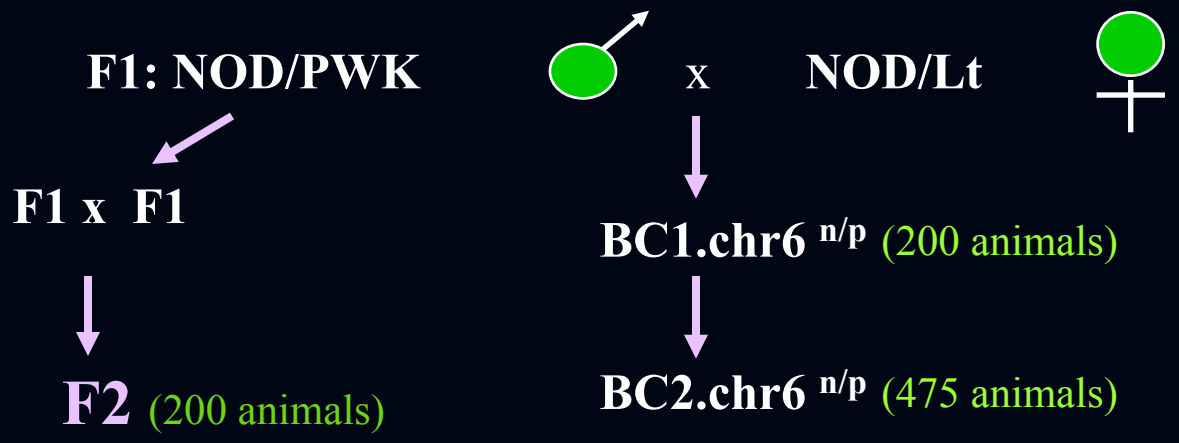
b. Inter-(sub) specific crosses

NOD	x	*Mus musculus musculus*
Susceptible		Resistant

* Higher polymorphism
* Higher genetic heterogeneity
* Lower phenotype incidence
* Segregation of resistance loci
* Closer to human families analysis?

SPE
PWK

Construction of Crosses for Genetic Analysis (NOD. PWK^{Chr6})



GENETIC ANALYSIS OF BC1 PROGENIES

CHROMOSOME	LOCUS	M χ^2	F χ^2	TOTAL χ^2
3	D3Mit7	0.18	13.41	4.5
14*	D14Mit5	8.38	1.74	9.27
17	D17Mit9	0.58	9.01	6.33
6*	D6Mit11	11.32	2.49	12.15
6*	Ly4	13.29	1.76	12.29
6*	D6Mit25	13.08	1.48	12.11
6*	D6Mit14	14.05	0.5	10.31

(Melanitou *et al*, 1998)

***Resistance loci**

INSTITUT PASTEUR

Two loci with resistance alleles

Chrom	Marker	GRR
6	D6*Mit*53	1.6
6	D6*Mit*25	1.6
6	D6*Mit*14	1.3
14	D14*Mit*5	2.8

GENOTYPIC RISK RATIO

GRR = 0.5 : susceptibility locus

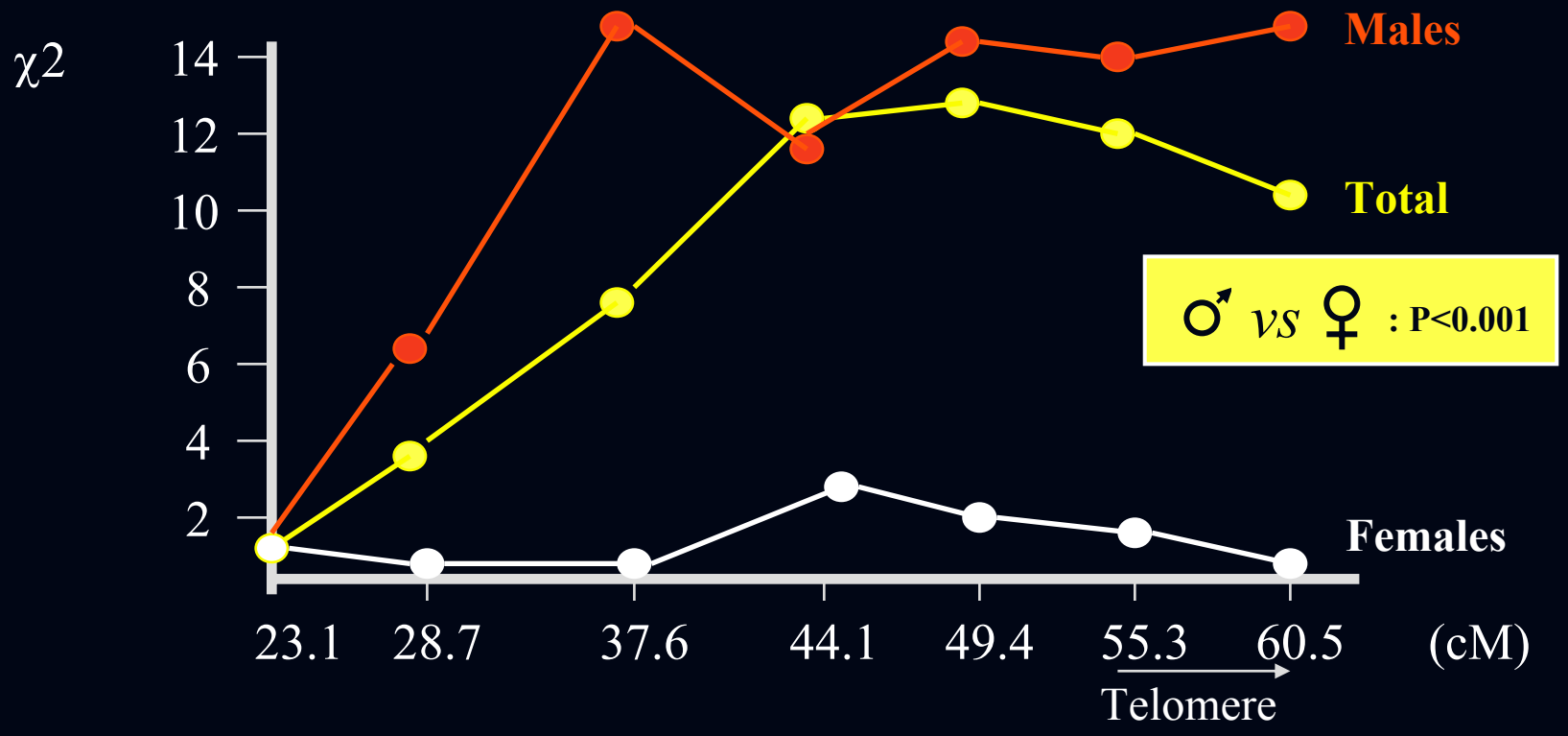
GRR > 1 : resistance locus

$$GRR = \frac{\textit{Total Number of BC1 diabetic}}{\textit{2x number of BC1 diabetic homozygous animals}}$$

(Melanitou *et al*, 1998)

INSTITUT PASTEUR

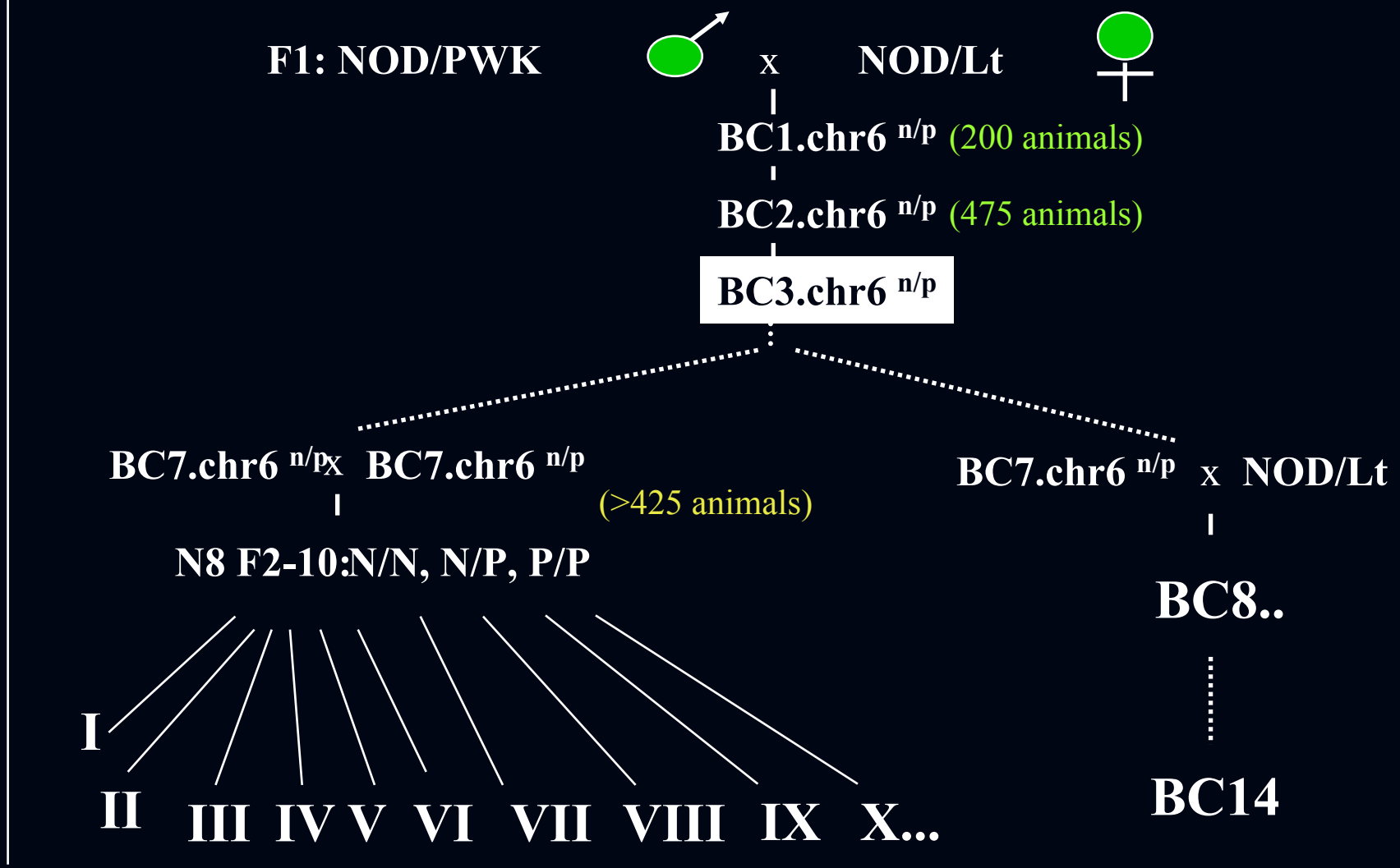
Genetic distance of the PWK Idd locus on distal chromosome 6 in the BC1 cross



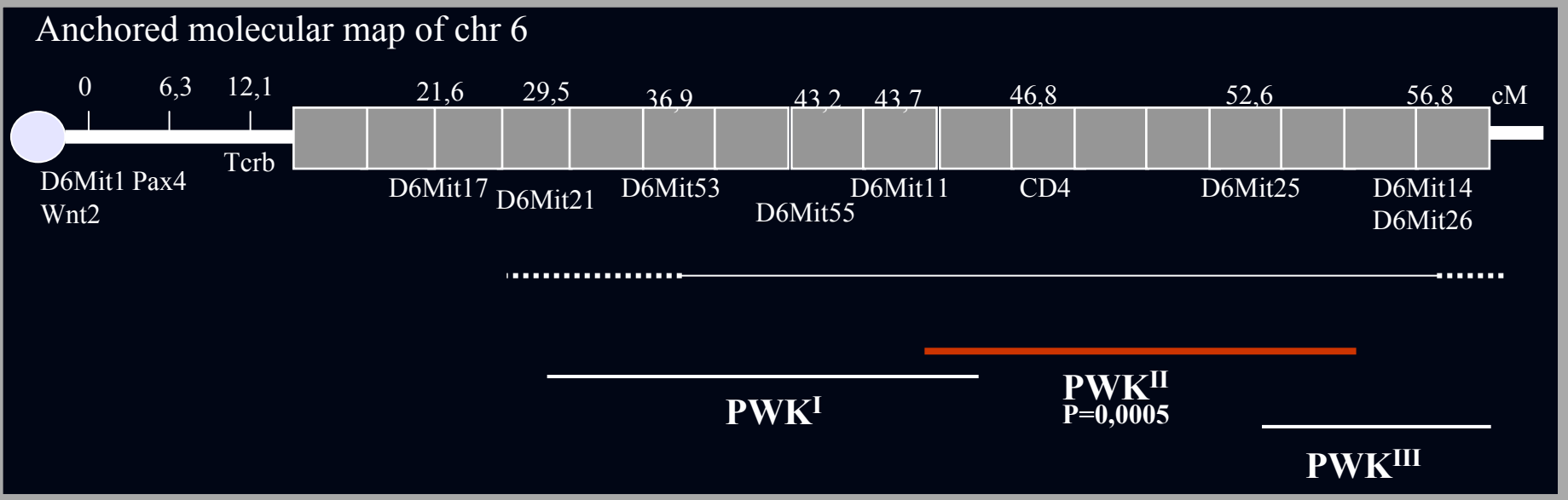
NOD/PWK cross	6.5 **11rec**	6.5 **11rec**	0.6 **1rec**	10 **17rec**	3.5 **6rec**	**27.1 cM**
MIT map	11.2	1.89	1.2	6.72	10.8	**31.89 cM**

(Melanitou *et al*, 1998)

Construction of Congenic lines (NOD. PWK^{Chr6})



NOD.PWK chr6 (6/21-26) BC7 F1 Congenics and genetic interval analysis



Interval	Locus	Normal N/N:N/P:P/P	Insulitis N/N:N/P:P/P	Severe phenotype N/N:N/P:P/P	$\chi 2$ (pvalue)
I	D6Mit21-11	75:26:11	11:9:1	63:28:17	1.55 (ns)
II	**D6Mit11-25**	**113:24:11**	**16:11:0**	**77:36:18**	**11.39 (0.0005)**
III	D6Mit25-26	128:25:18	20:8:2	99:40:15	2.82 (ns)

(Melanitou *et al*, 1998)

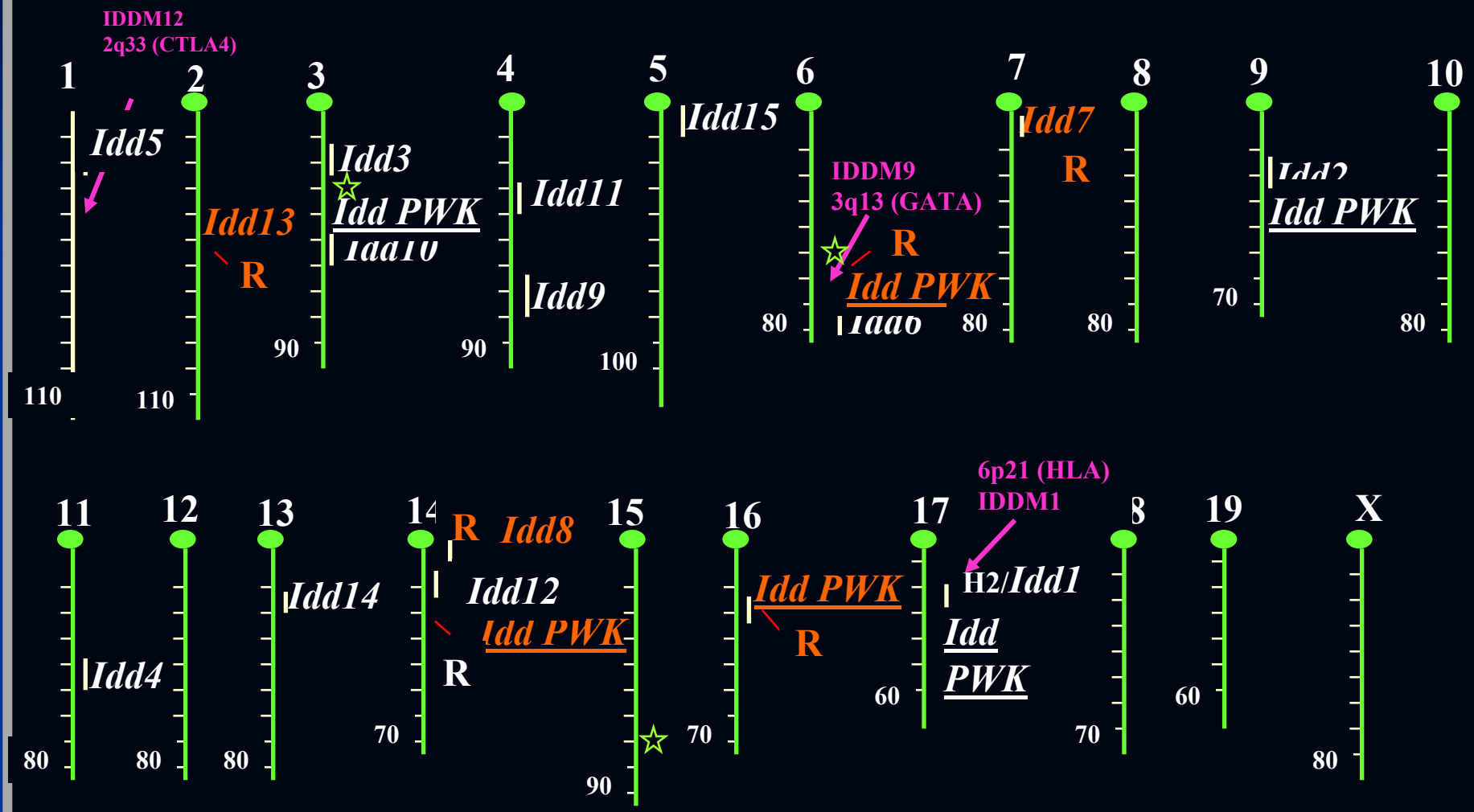
INSTITUT PASTEUR

Phenotypic analysis of congenic lines with fixed segments from the Resistant strain, is compatible with the hypothesis of the presence of more than one gene implicated in type 1 diabetes within this region.

Genetic interaction (or epistasis) between genes may explain diabetes phenotypes within the mouse chromosome 6 locus

Melanitou, 1998
Rogner and Melanitou, 2001

Chromosomal locations of *Idd* loci identified by genetic linkage analysis

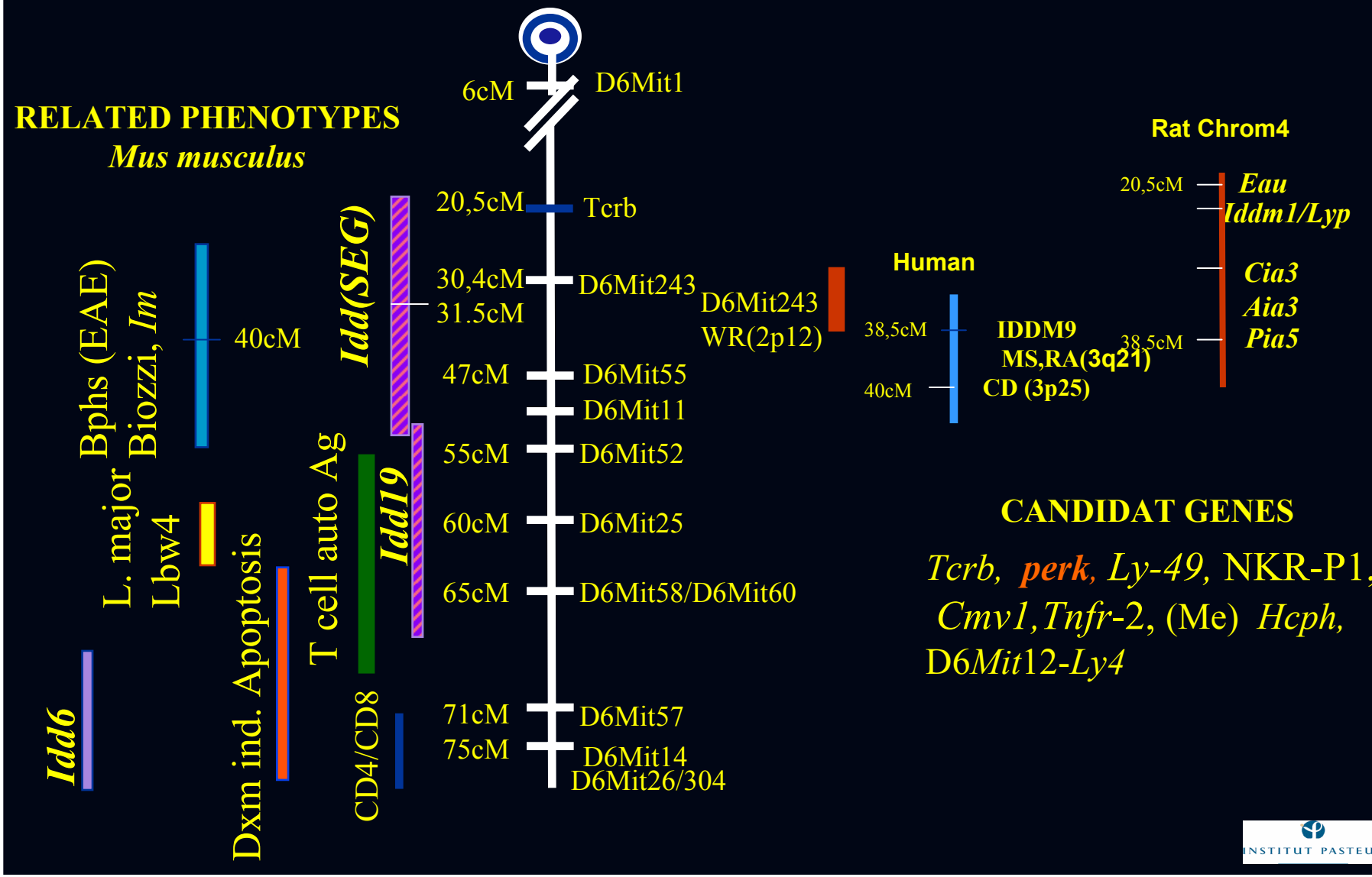


IddPWK : NOD/PWK cross

☆ *Idd*SEG : NOD/SEG cross

INSTITUT PASTEUR

•CHROMOSOME 6
Candidate genes and associated phenotypes



SUMMARY I

* *Genetics:*

1. Susceptibility (S) gene is localised in a genetic interval of 3 to 7 cM, at the most distal part of the locus.

2. The resistance (R) gene is localised within a 0.1 to 2 cM genetic interval proximal to the S locus.

Epistatic interactions take place between the two loci.

* *Phenotypes:*

1. The presence of R inhibits diabetes onset at T1; a second stimulation is necessary for final phenotype.

2. Epistatic interactions are stronger in a male environment.

S/R respond in a differential manner according to the sex of the animals (hormonal status?)

Interests and limits of the genetic analysis of a complex trait

What did we learn from the Genetics?

1. There is an *Idd* locus on chromosome 6
2. There are possibly several diabetogenic genes within this locus
3. There is interaction between these genes (epistasis)
4. Congenic lines are useful tools for establishing a gene-phenotype correlation

- Construction of additional crosses is necessary for identifying recombinants segregating the genes of interest.
- Positional cloning might be laborious and with uncertain results due to the complexity of the phenotype.

Part II : **Gene Expression Networks in Early Autoimmunity**

AIM

Towards the identification of molecular mechanisms underlying early steps of autoimmune diabetes

METHOD

Analysis of "gene profiles" specific for the Insulin AutoAntibodies (IAA) diabetes sub-phenotype

TARGET

Lymphoid tissue : Pancreatic Lymph nodes

Outlines

Part II

- **Gene Expression Networks in Early Autoimmunity**

a. Subphenotype: -characterization
-significance relative to diabetes

b. Gene signatures: -Description of the approach
-Biological significance

c. Correlation: Genetics ⟷ Genomics

1. The phenotype:

Insulin Auto Antibodies

DISEASE TIME CHART

Birth

3 weeks

4 weeks

5 weeks

16 weeks

Sub phenotype

Ante partum

Post partum

Maternal IAA

IAA

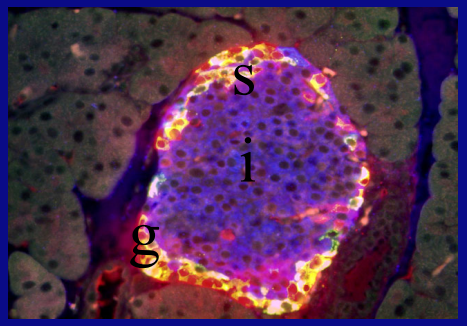
IAA

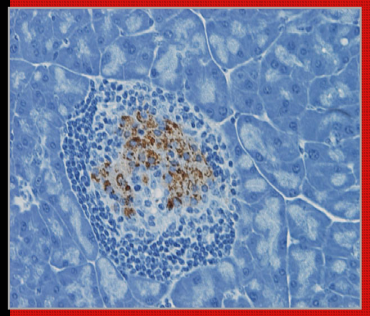
IAA

Initiation of apparent autoimmune destruction

Diabetes onset

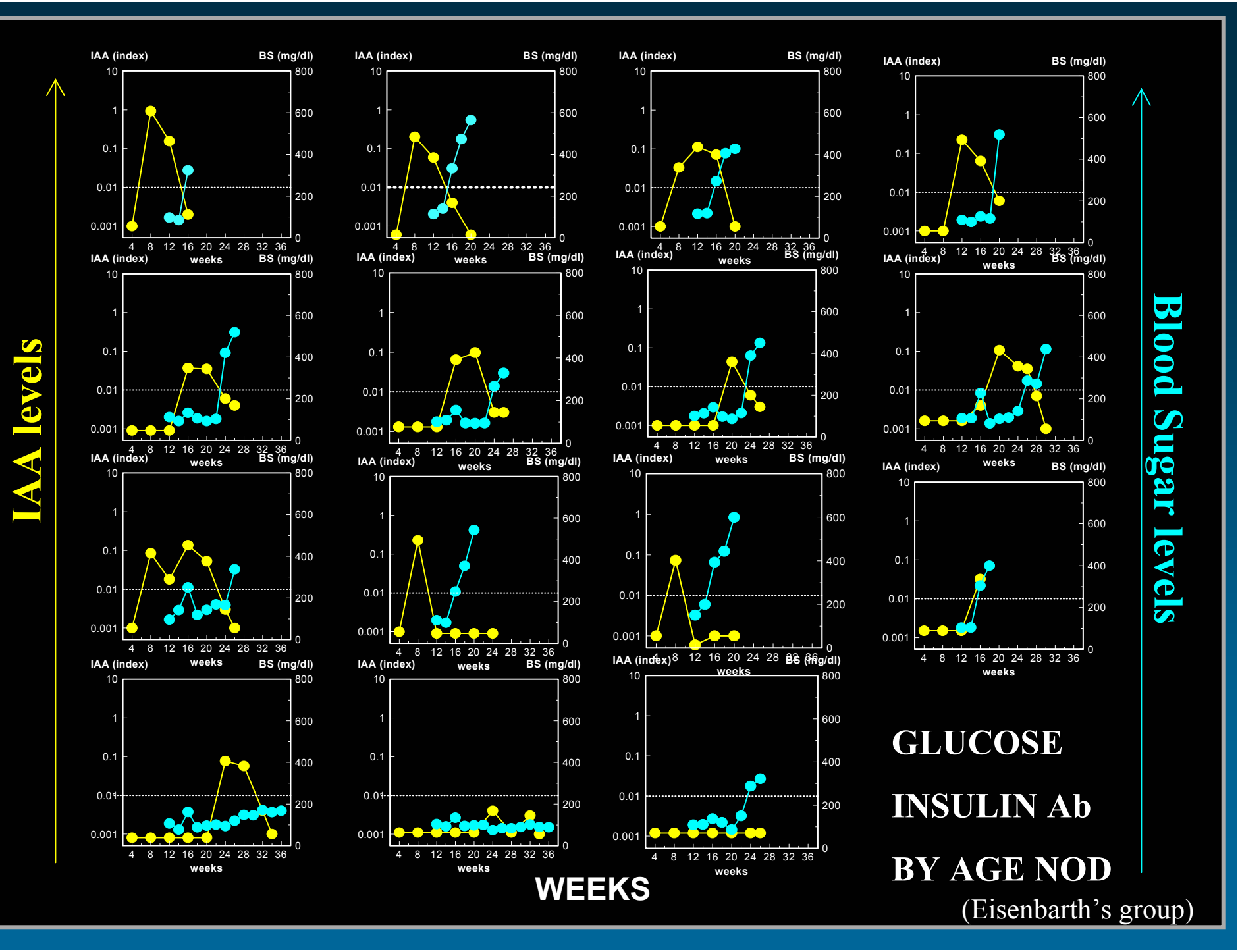
Phenotype





Intact islet

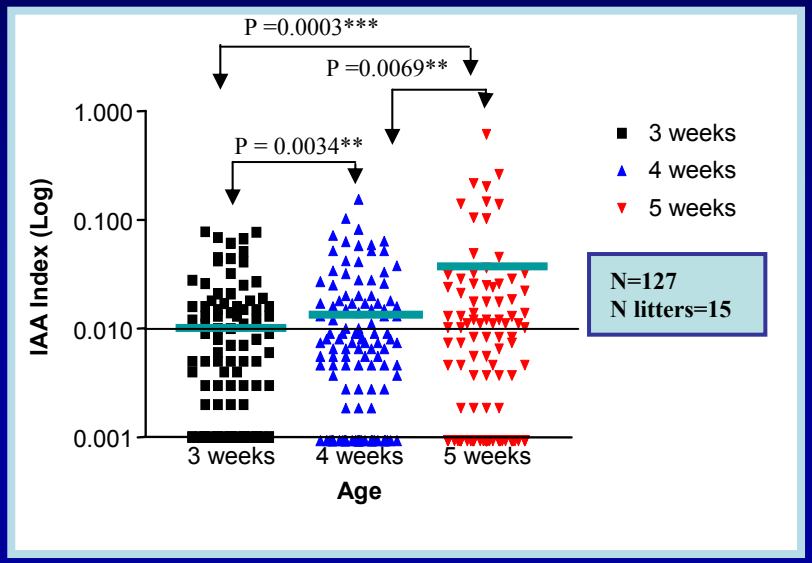
Peri-insulitis/Insulitis

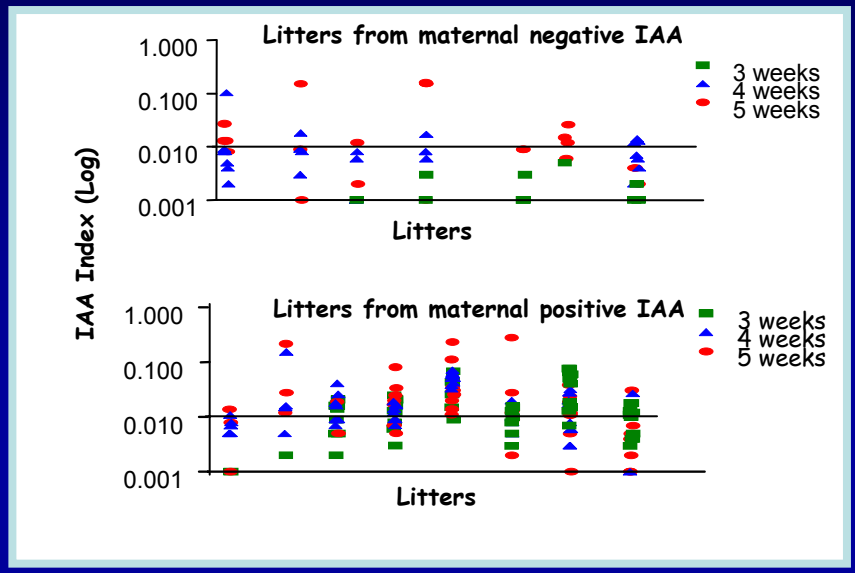

IAA levels
IAA (index)
BS (mg/dl)
weeks
WEEKS
Blood Sugar levels
GLUCOSE
INSULIN Ab
BY AGE NOD
(Eisenbarth's group)

Insulin Auto Antibodies

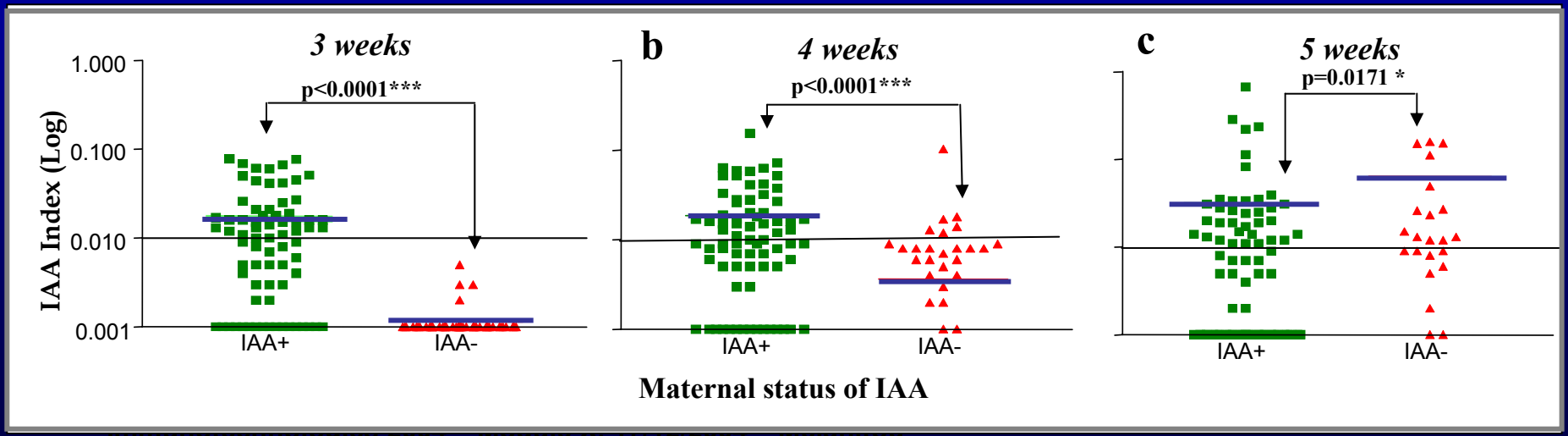
➡ can be detected at 3 weeks (E-IAA)

➡ E-IAA in offspring correlate with matIAA *ante partum* status





➡ Higher IAA levels are observed in matIAA+ than in matIAA- offspring



Melanitou *et al*, 2004 J.Immunol *in press*

Evolution of the E-IAA phenotype

IAA + offspring at 3wks remain IAA + at 8 weeks

	IAA+ 8wk	IAA- 8wk
IAA+ 3wk	5/8 (62.5%)	3/8
IAA– 3wk	5/14	9/14 (64%)

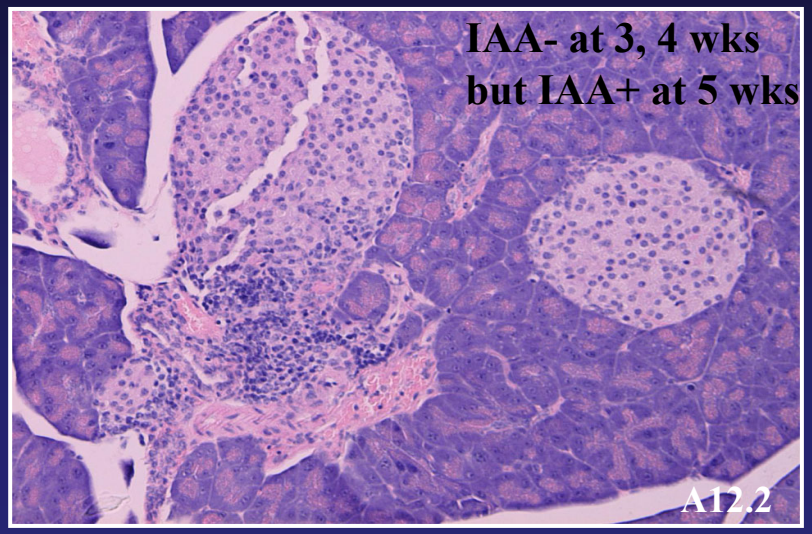
This data strongly argue for a biological role etiologic or correlative of this E-IAA sub phenotype in later disease phenotype

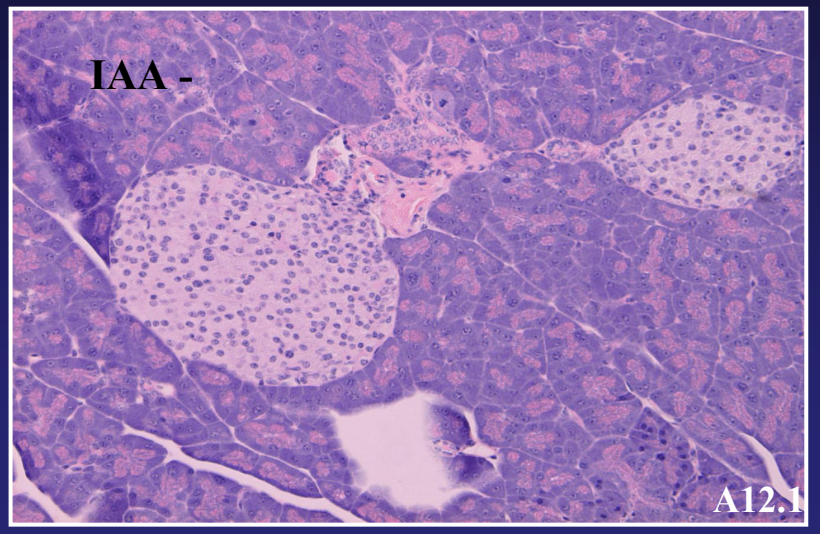
Melanitou *et al*, 2004 J.Immunol *in press*

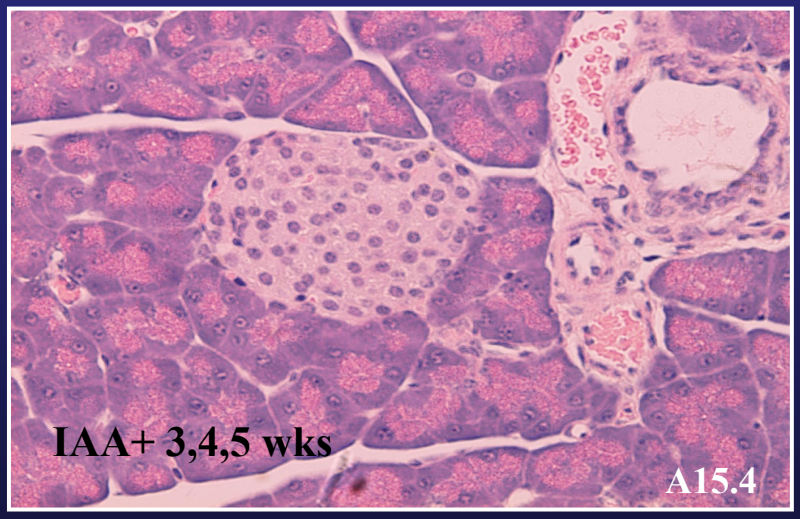
Significance relative to diabetes:

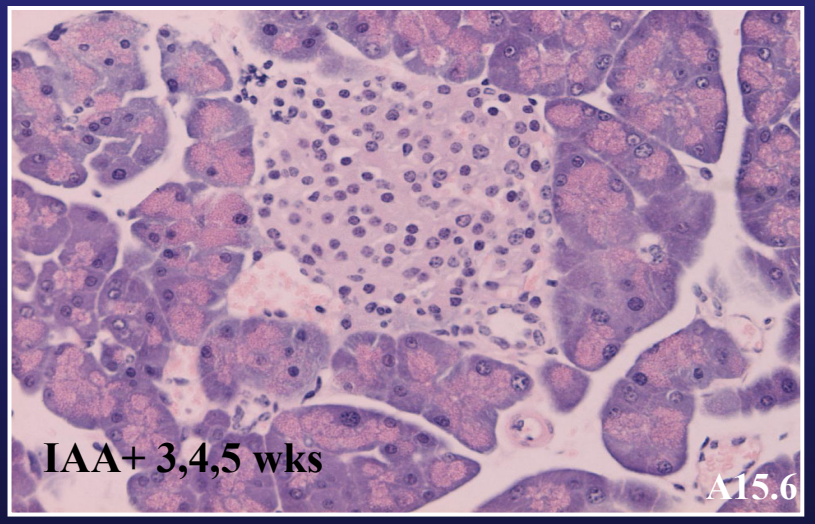
Does early IAA reflects also early appearance of diabetes in the NOD mouse ?

IAA can be present without any apparent insulitis at 5 weeks



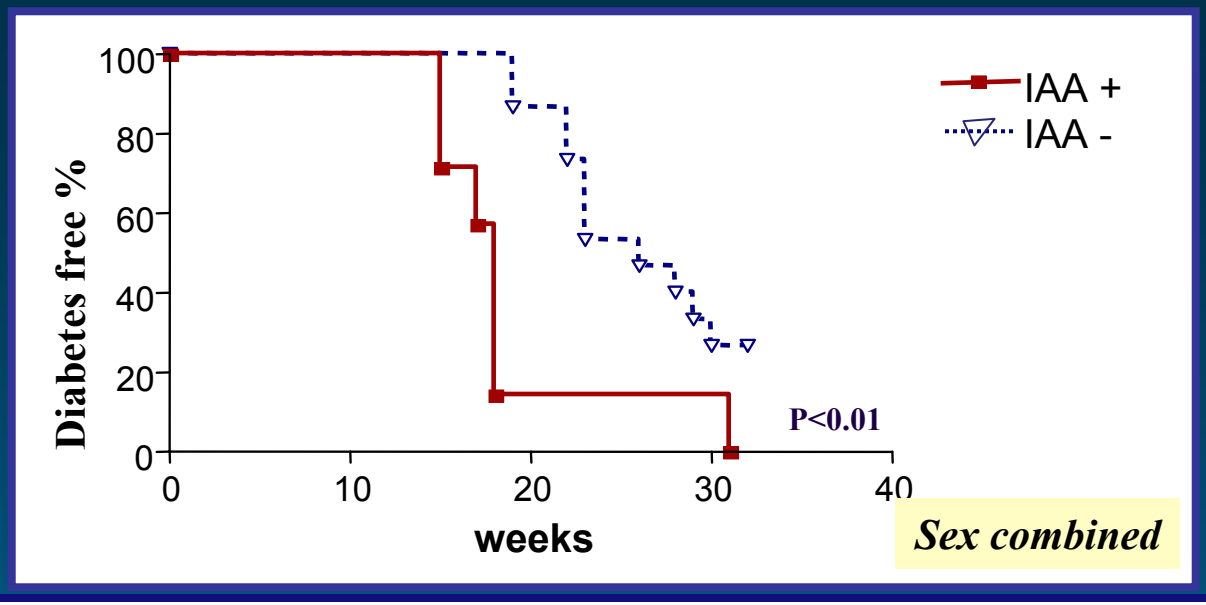


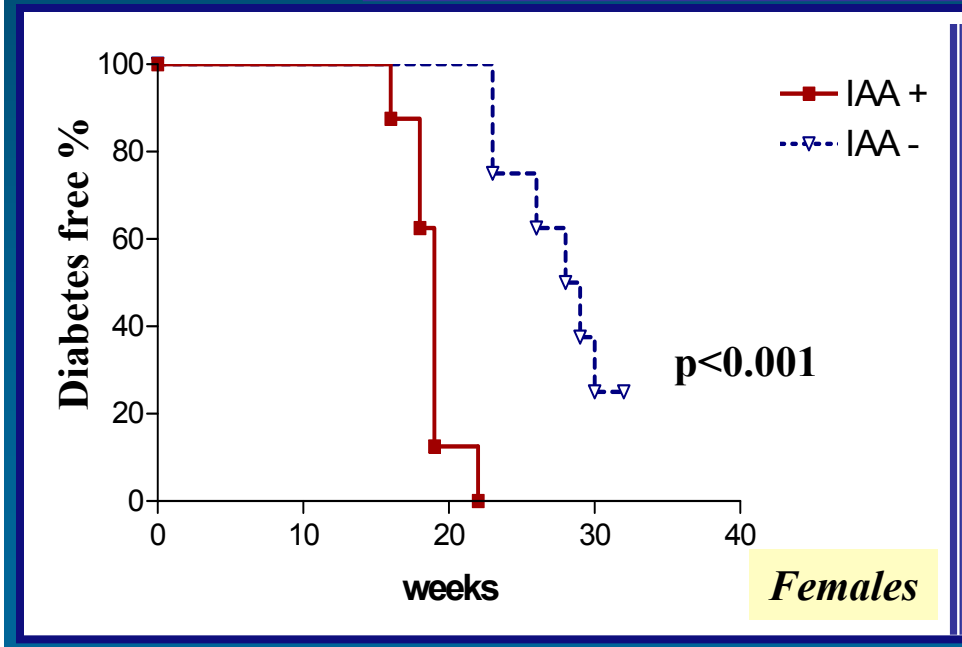


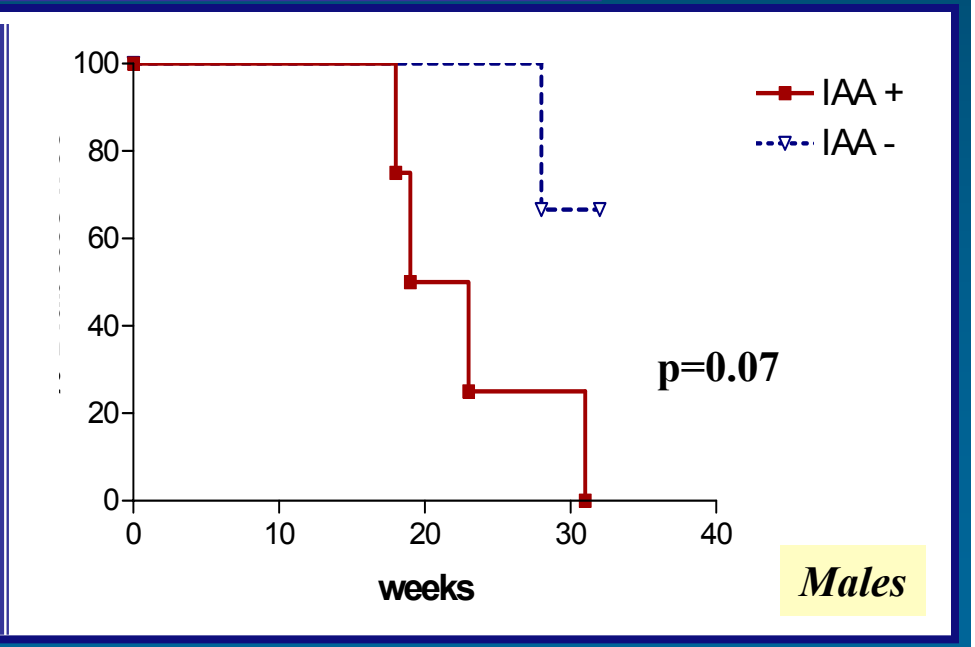


Melanitou *et al*, 2004 J.Immunol *in press*

Early diabetes incidence correlates with E-IAA (3-5 wks of age)



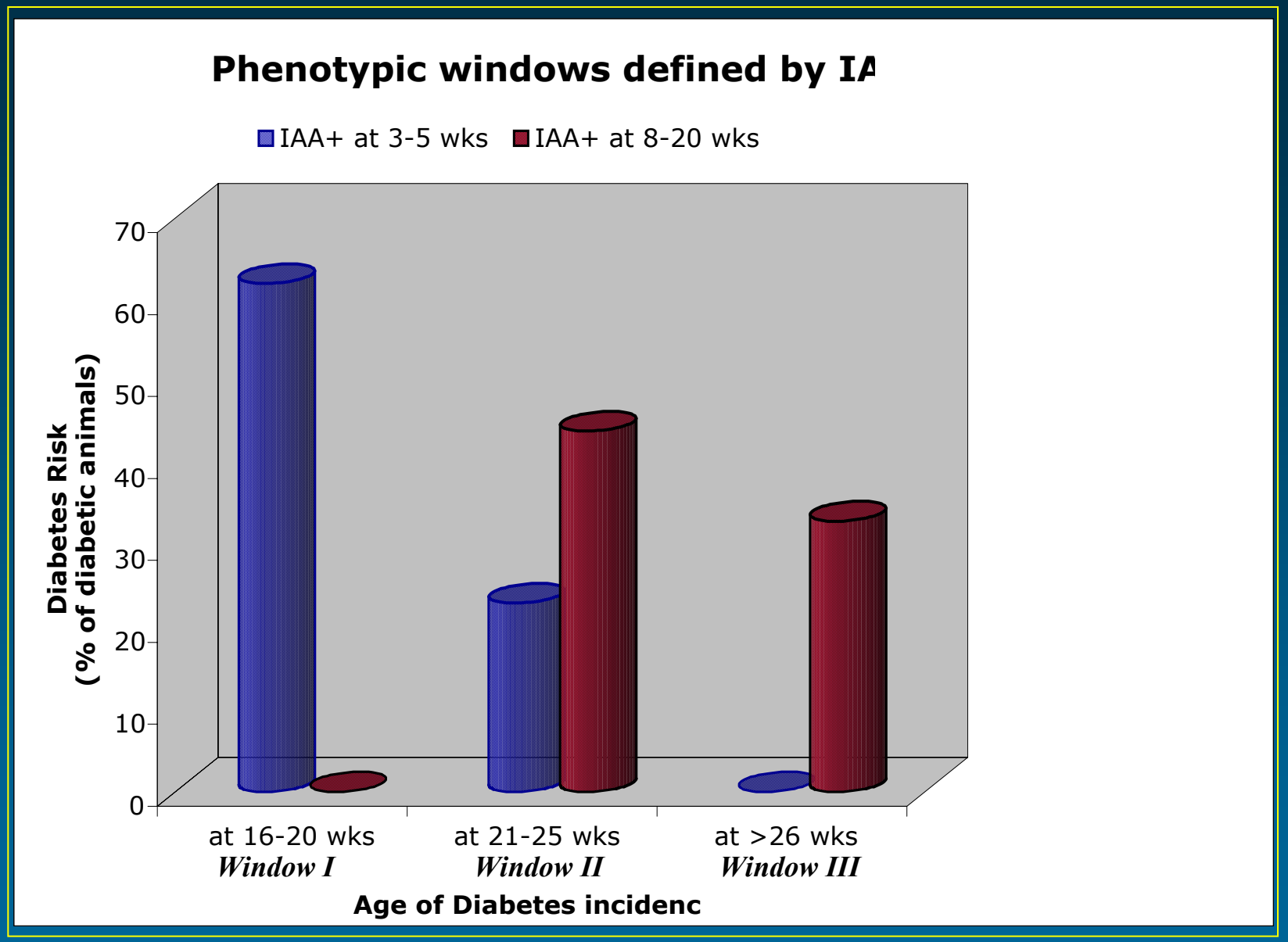




Melanitou *et al*, 2004 J.Immunol *in press*

Transplacental Auto Antibodies?

- Maternal IAA showed lower levels than the IAA in the pups
- Rarely were transient as pups usually remained IAA+
- The presence of IAA at 3 weeks correlated with IAA at 8 weeks
- E-IAA correlates with Early Type 1 diabetes (16 - 20 weeks)



Melanitou *et al*, 2004 J.Immunol *in press*

SUMMARY II

- Maternal IAA influences the IAA frequency in the offspring in a significantly different manner
- The presence of IAA at 3 weeks correlated with IAA at 8 weeks
- E-IAA is quantal and correlates with Early Type 1 diabetes (16 - 20 weeks)
- Maternal immune imprinting takes place predisposing the litters to early autoimmunity
- Autoimmune windows can be defined by the presence of E-IAA

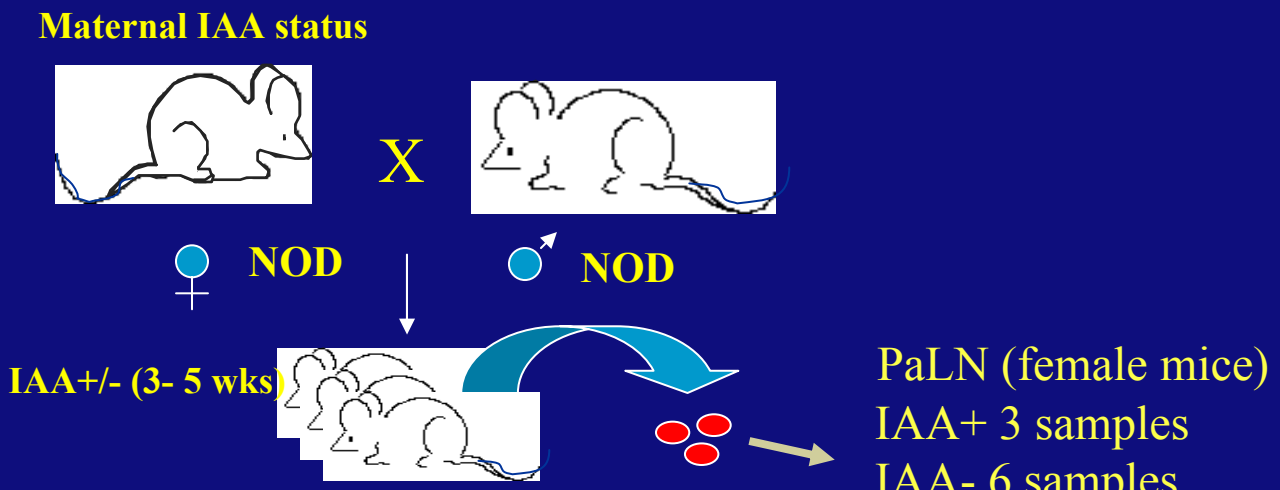
Finally

in an analogy to the transferable maternal immunologic memory the period before and after birth might be the key in understanding also autoimmunity

Second part :

Is there a molecular signature that distinguishes at an early stage individuals which will develop diabetes ?

Experimental design



Pancreatic Lymph Node gene profiles specific to early IAA, sub-phenotype of type diabetes

Gene signatures : Description of the approach

Affymetrix Microarray Analysis

- The Murine Genome U74v2 Set is comprised of 3 arrays
- This represents ~36,000 full-length genes and EST clusters
- The sequences represented are derived from sequence clusters in Build 74 of the UniGene Database. UniGene clusters are represented by one or more consensus sequences derived directly from cluster members

<u>Define Relationships Between Known Genes</u>

- The first array in the set (MG-U74Av2) represents all sequences (~6,000) in the Mouse UniGene database that have been functionally characterized
- In addition, ~6,000 EST clusters are also represented on this single array

Melanitou E

Microarray flowchart

- Isolate total RNA by Quiagen kit (~ 5 μg)
- Quantitate RNA by spectrophotometry

- Check RNA quality by Agilent RNA chip

(high quality of RNA is necessary)

- Synthesize double stranded cDNA

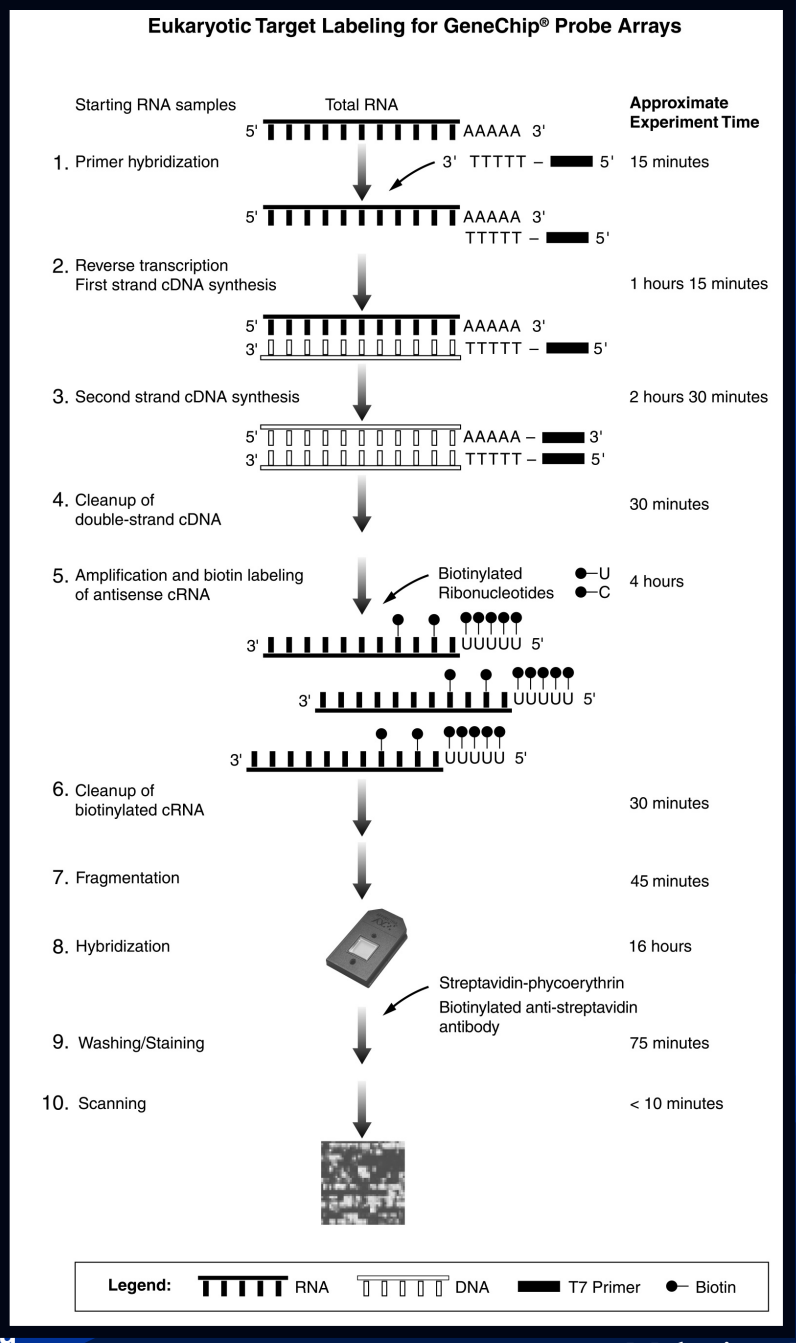
T7-(dT)24 primer + Reverse Transcriptase: first strand DNA polymerase-I /T4 DNA polymerase

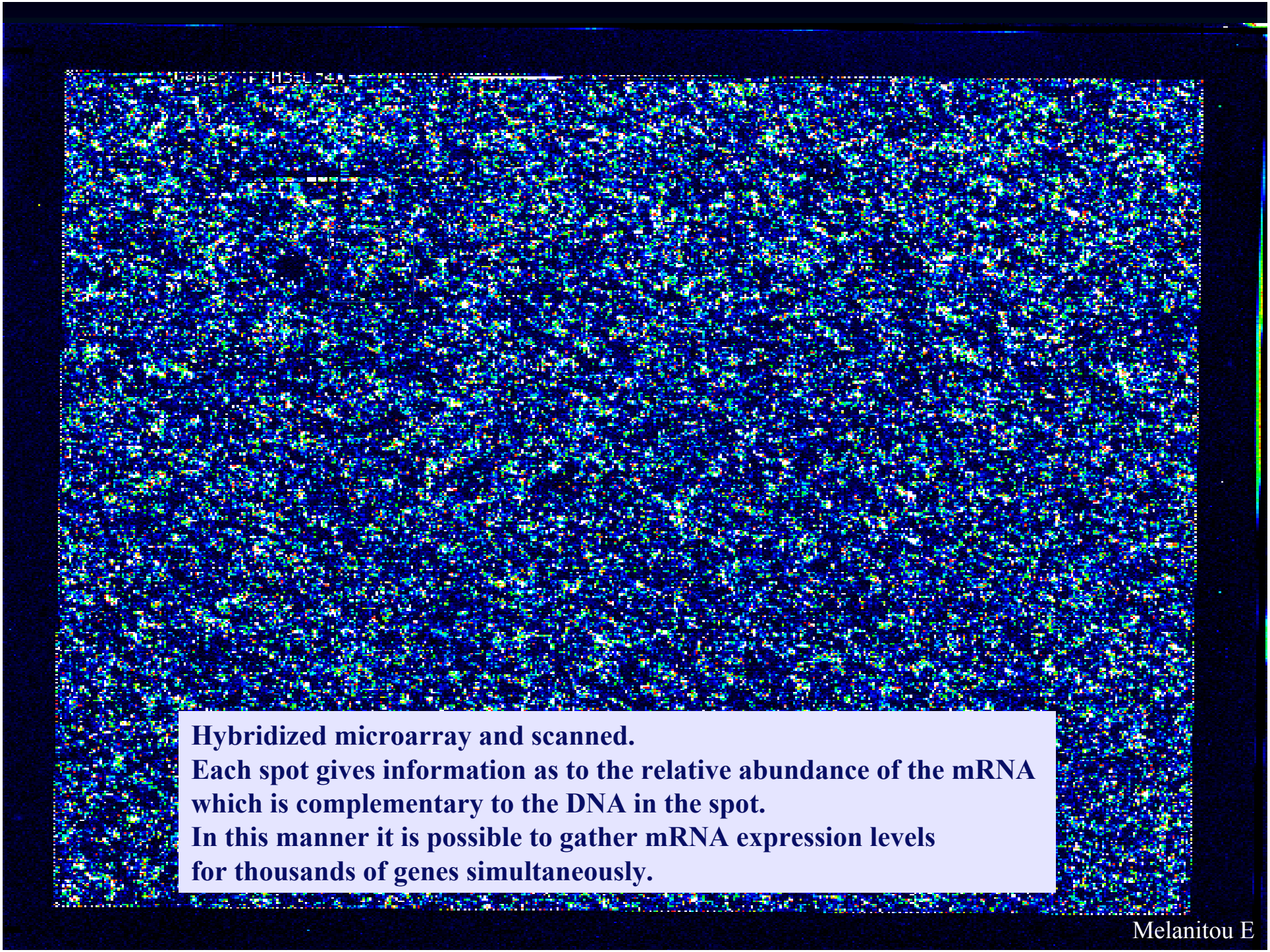
- Generate biotinylated RNA by in vitro transcription

Biotin-labeled Ribo nucleotides by T7 RNA polymerase

- Fragment RNA **(RNA hydrolysis 200-300bp)**
- Hybridize to array
- Stain with streptavidin-phycoerythrin

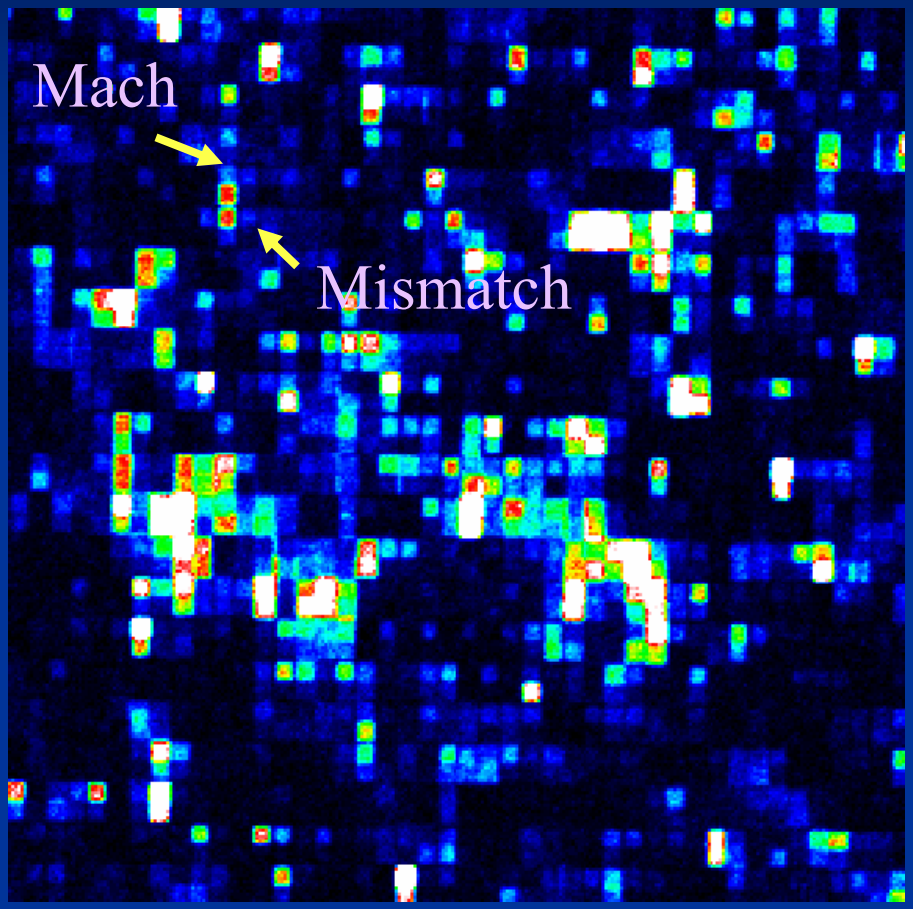
- Scan chip
- Normalize fluorescence
- Analyze results

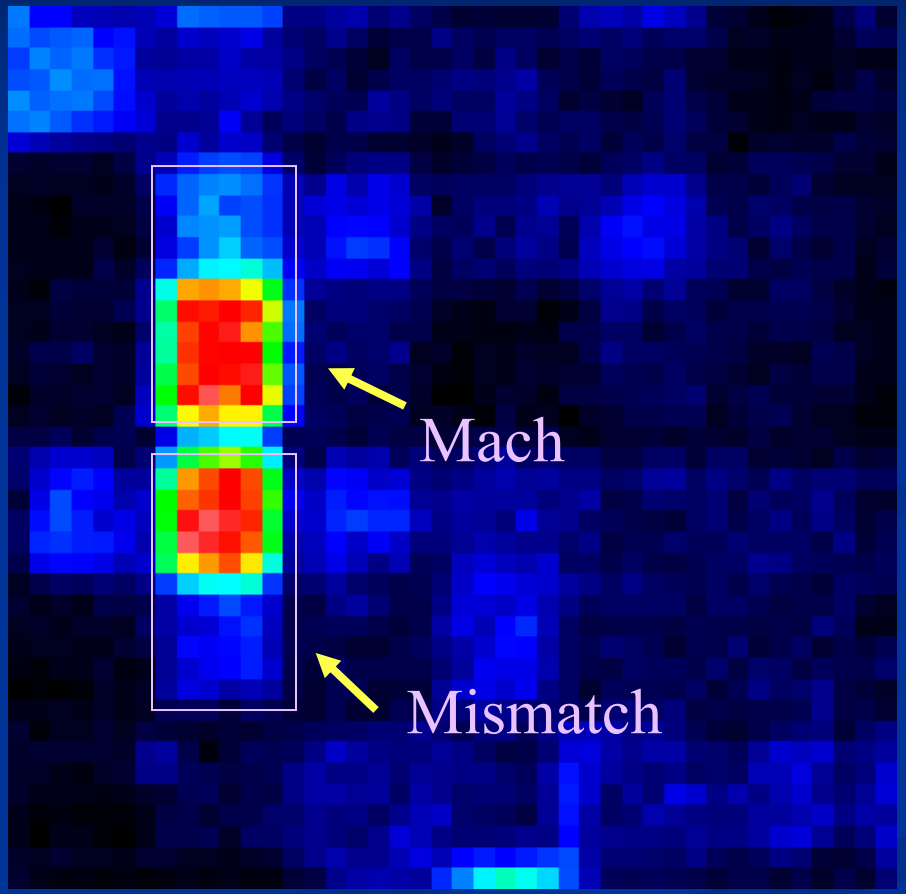




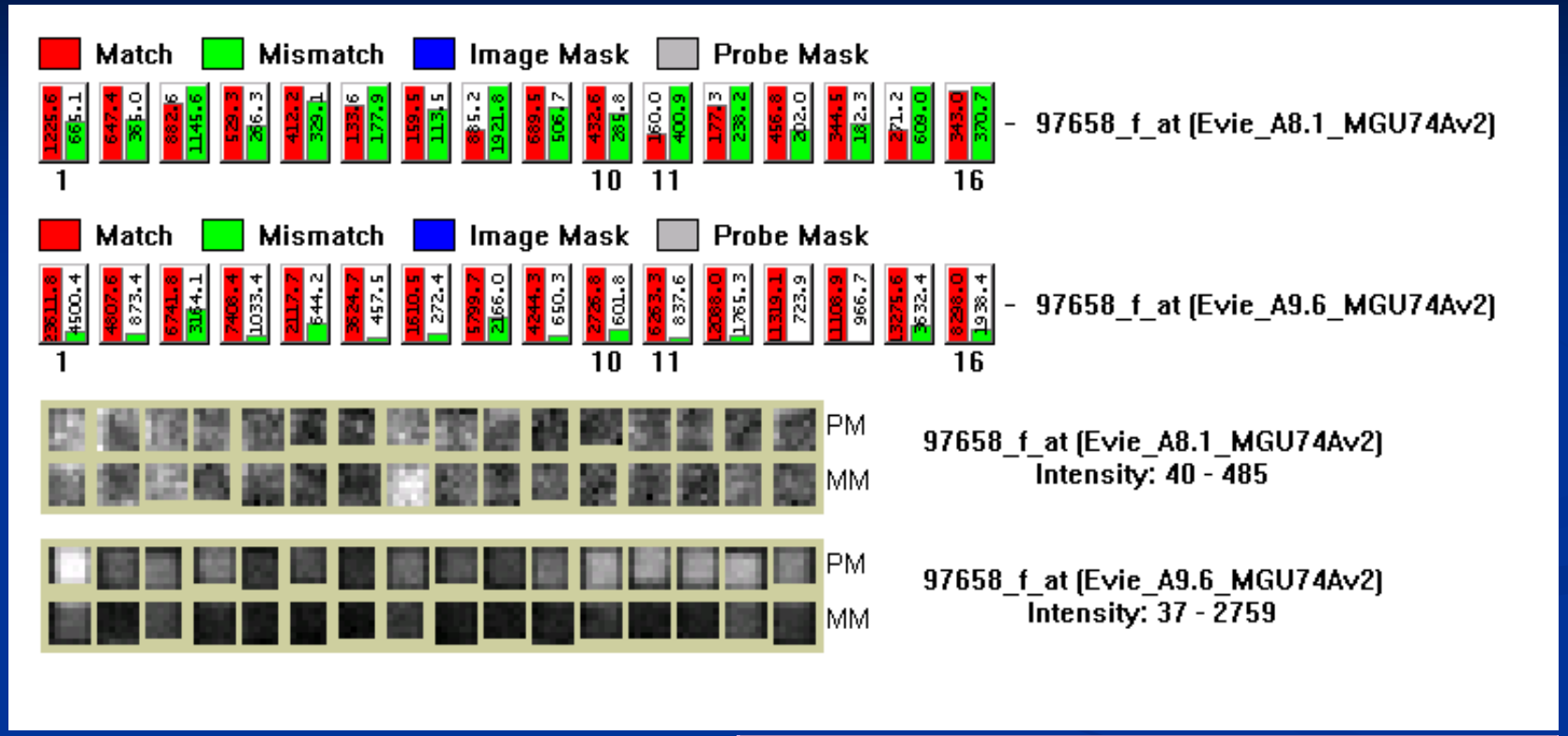
Hybridized microarray and scanned.
Each spot gives information as to the relative abundance of the mRNA which is complementary to the DNA in the spot.
In this manner it is possible to gather mRNA expression levels for thousands of genes simultaneously.

Melanitou E


Mach
Mismatch


Mach
Mismatch

Data analysis in Microarray Suite



Sample	Signal	Detection	Detection p-val
A8.1	115.6	A	0.418069
A9.6	4862.3	P	0.000219

$$D = (PM - MM) / (PM + MM)$$

This Discrimination Score (D) for each probe set is used to calculate a Detection p-value and to assign genes a Present, Marginal or Absent call

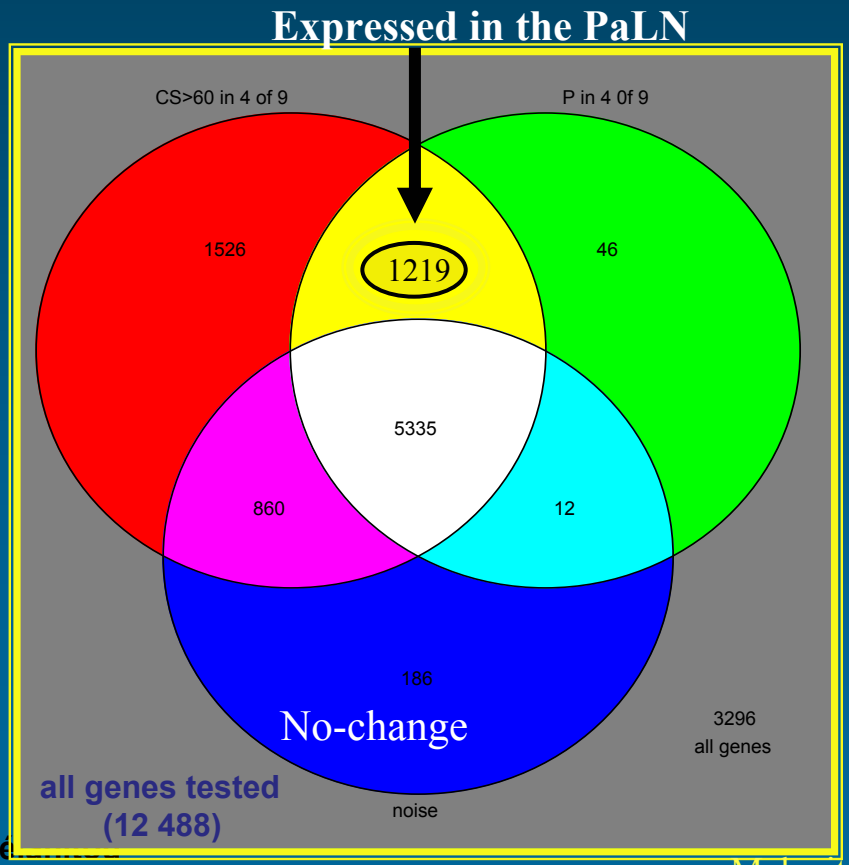
Data analysis GS6.0

NORMALIZATION

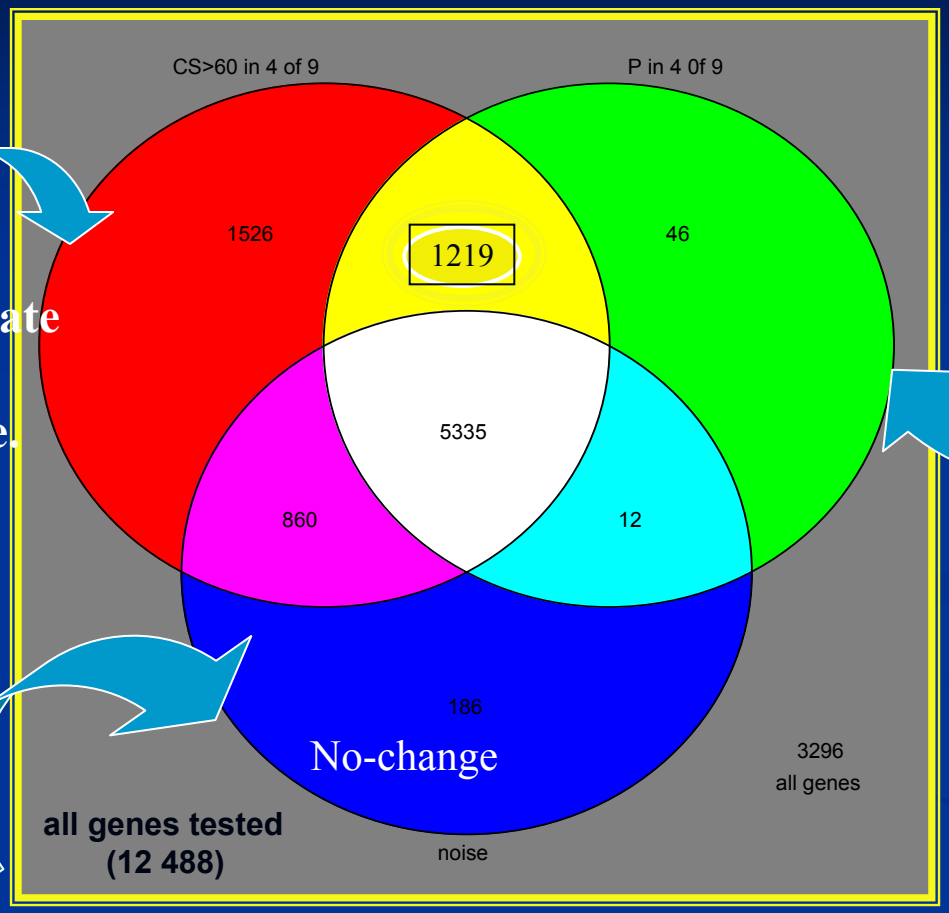
- Per chip normalization: Normalize to the distribution of all genes using 50th percentile
- Per gene normalization: Normalize to the median for each gene (a cut-off value of 0.01)
- Make minimum value 0

FILTERING

- Control Strength CS>60
- Genes not changing (generated by using the drawn gene function in which this artificial gene was a horizontal line)
- Genes present in at least 4 out of 9 samples
- Plot data by Venn diagram



Data Analysis - Genespring



1. Calculate an estimate of noise, and find genes above the noise.

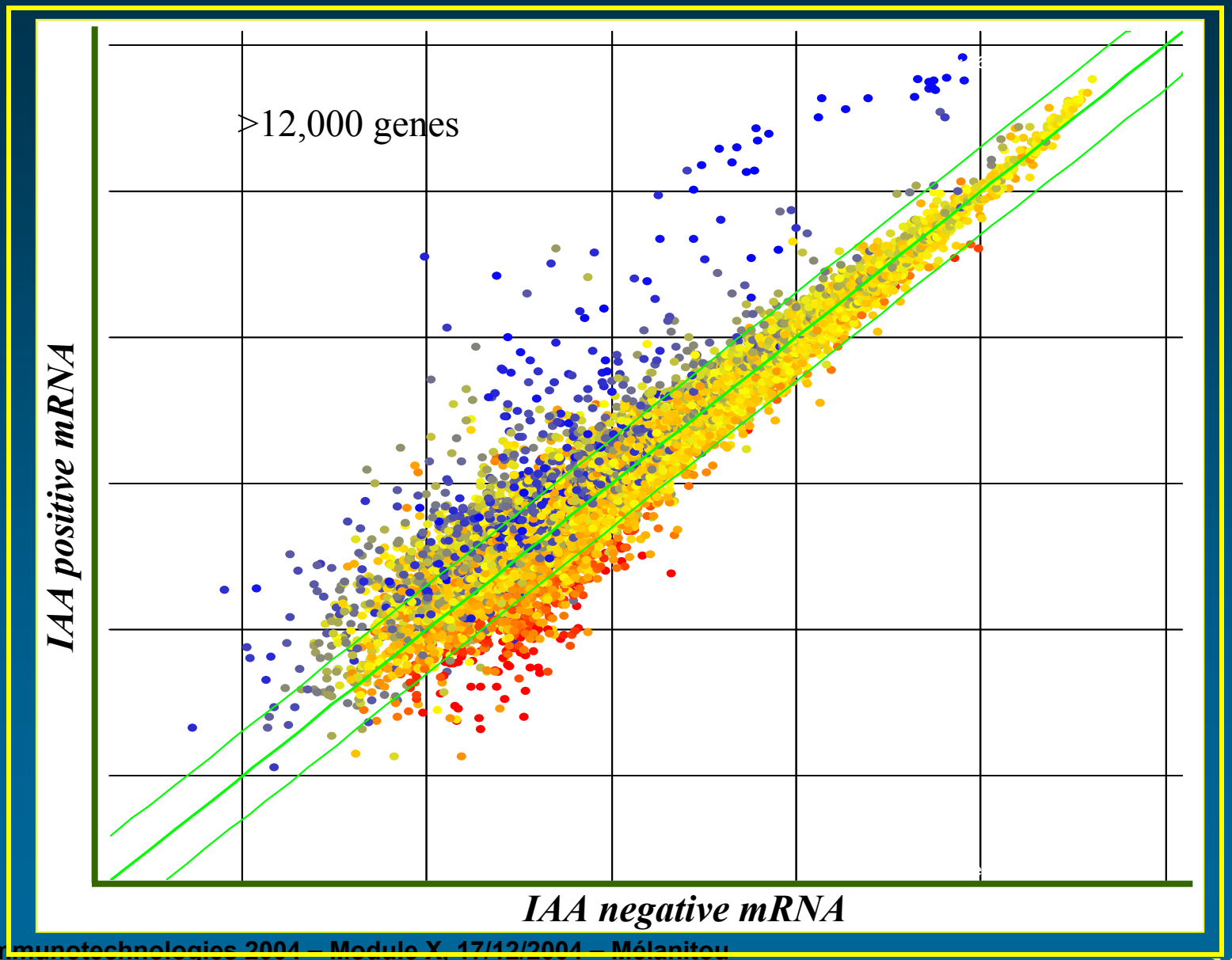
2. Find genes that are "Present"in at least 4 of the 9 samples.

3. Find genes that do not change either:

- **a) Considering each sample separately (no grouping by phenotype) <u>OR</u>**
- **b) Grouping by phenotype**

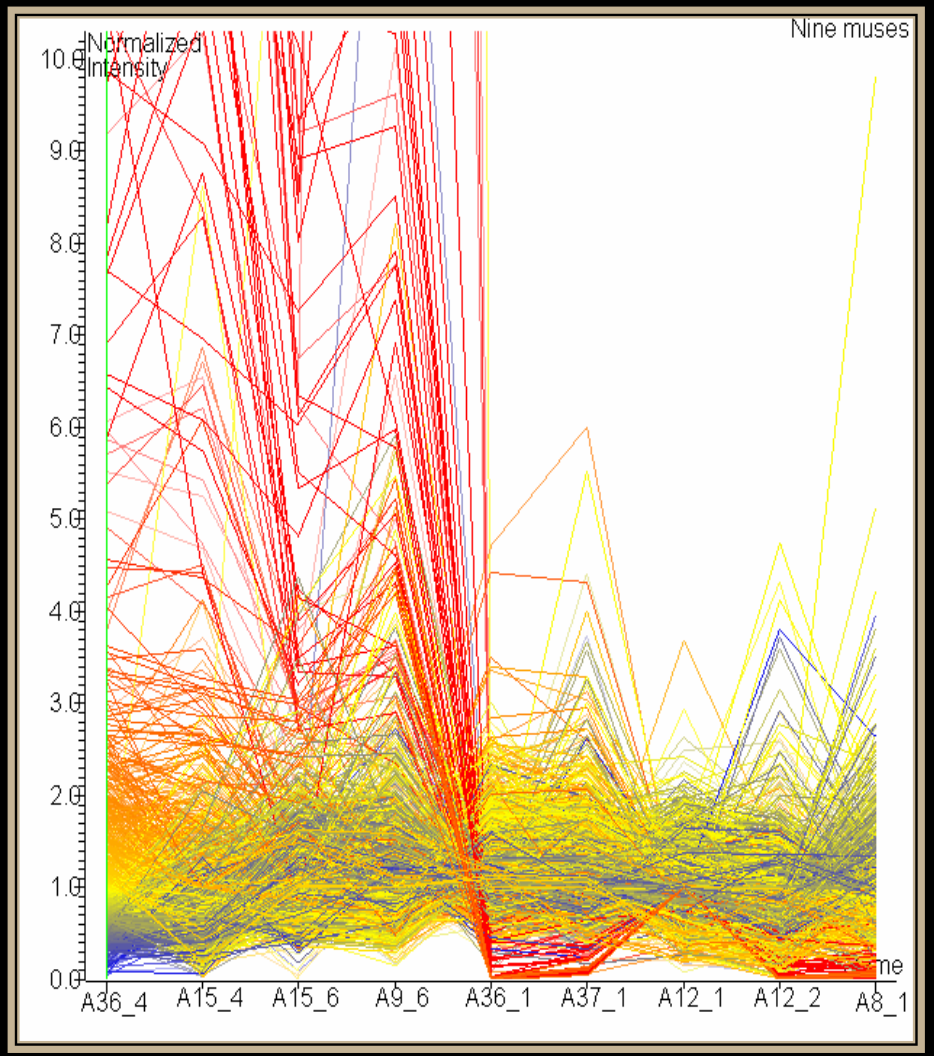
Then, using a Venn diagram, find the genes in 1 and 2, but not 3.

Correlation of all genes expression for IAA positive vs IAA negative in PaLN (Scatter plot Arrays MU74AVersion 2, GS6.0)



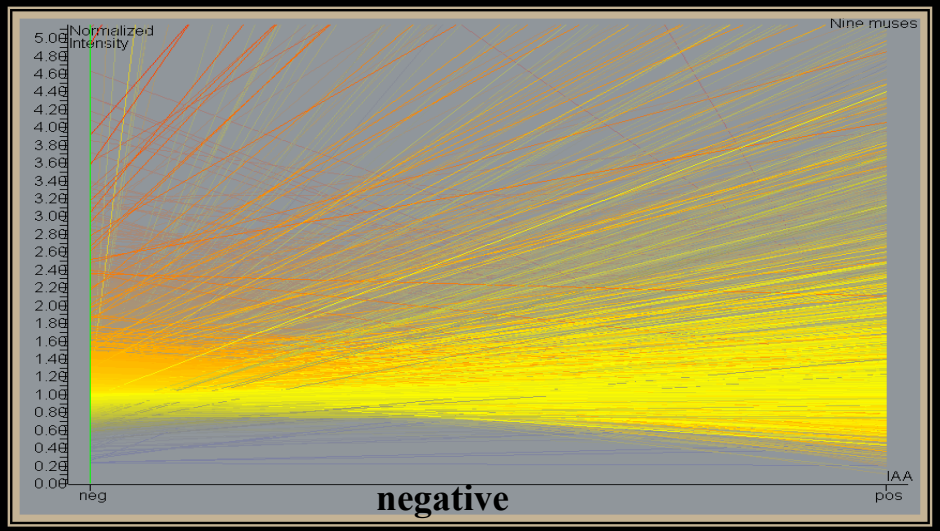
Distribution of gene expression patterns in PaLN of IAA+ and IAA- NOD mice

Line diagram of genes changing by expression

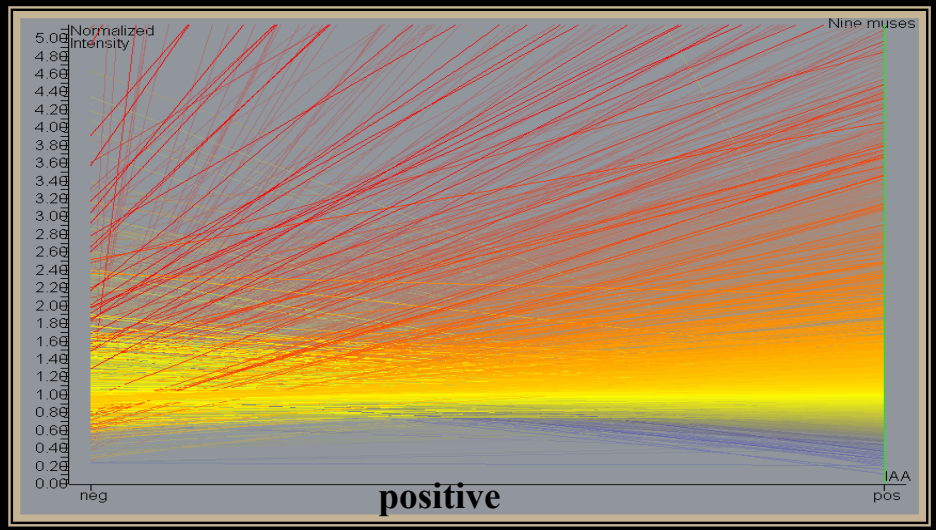


IAA+ gene patterns **IAA- gene patterns**

Line diagrams of differences Of 12,488 genes expression patterns



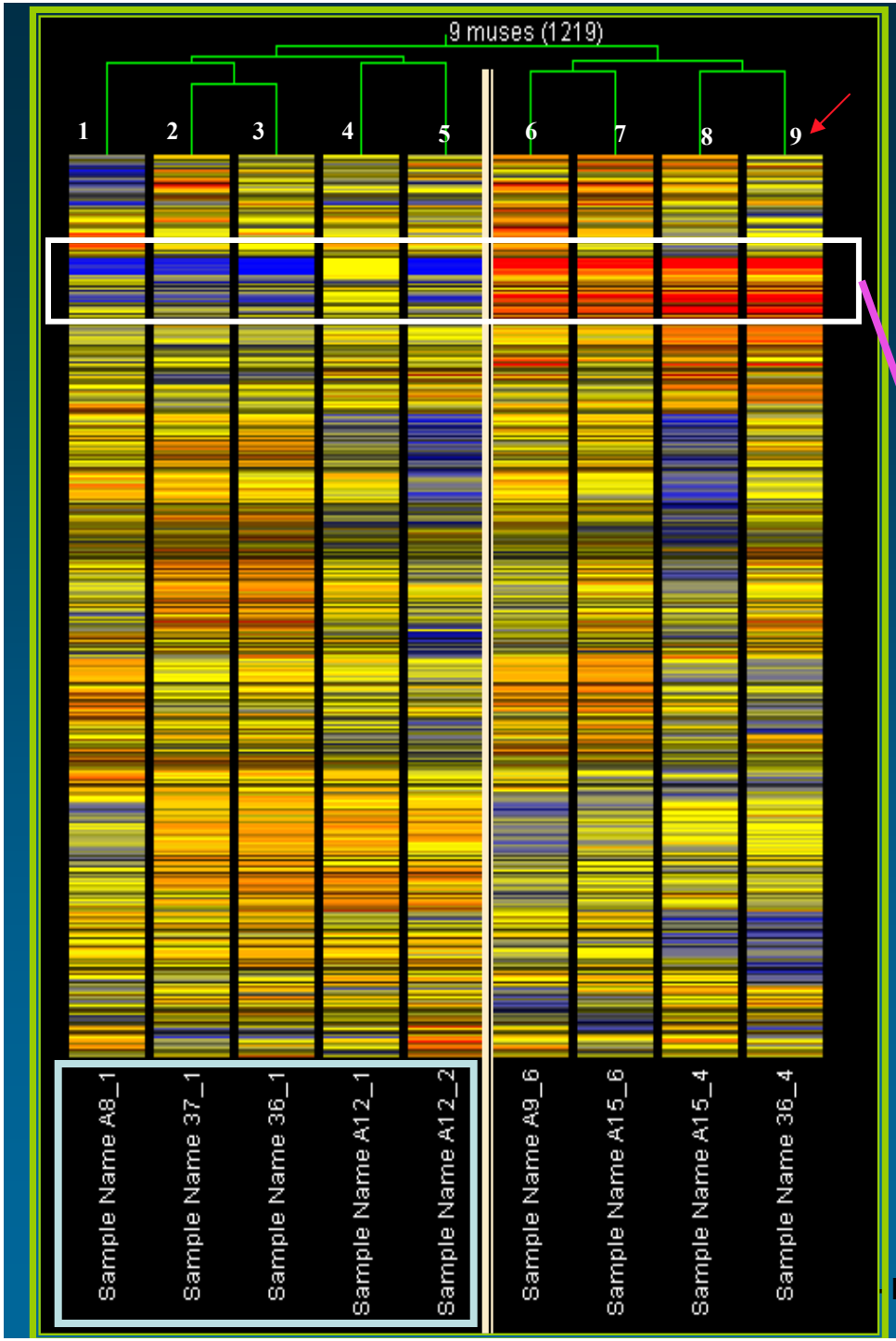
negative

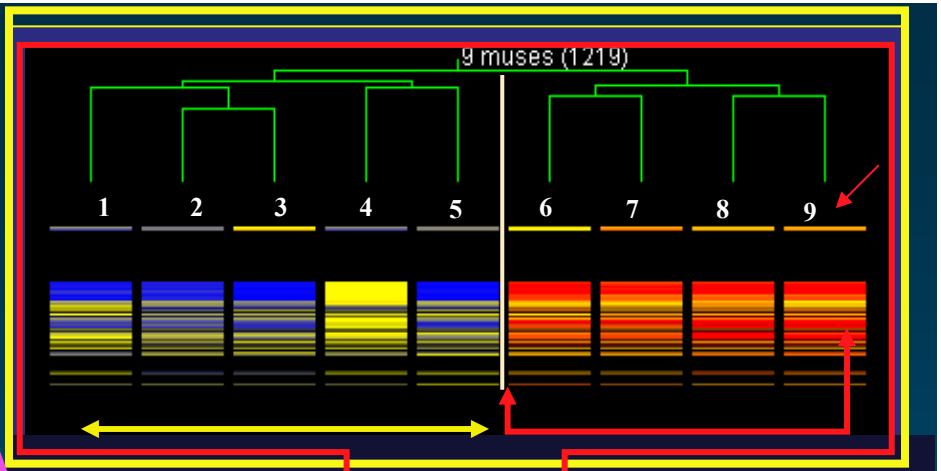


positive

All genes 12,488

Melanitou E





IAA- gene patterns

IAA+ gene patterns

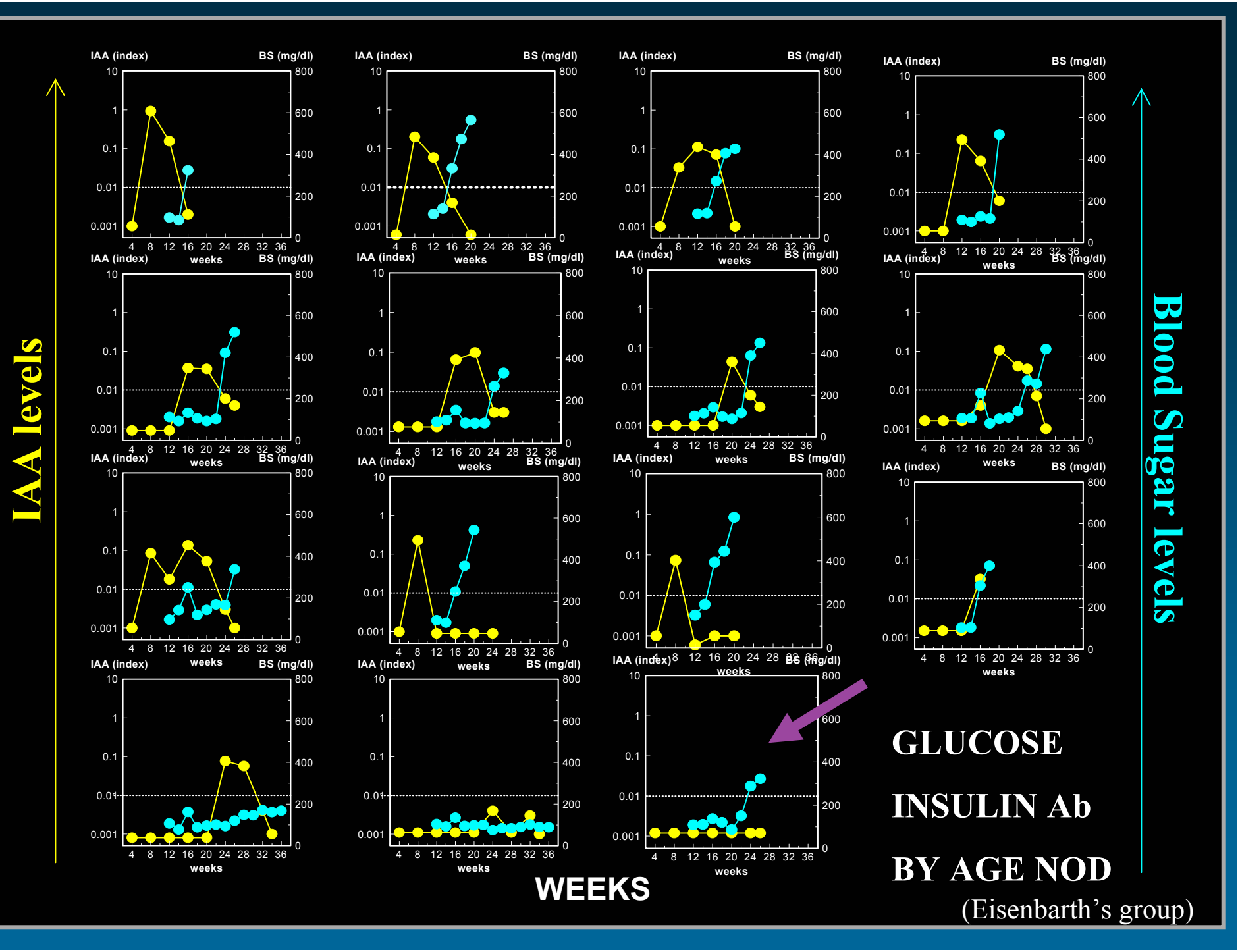
90 genes that are over expressed in IAA+ samples

125 genes

Phenotype non-considered

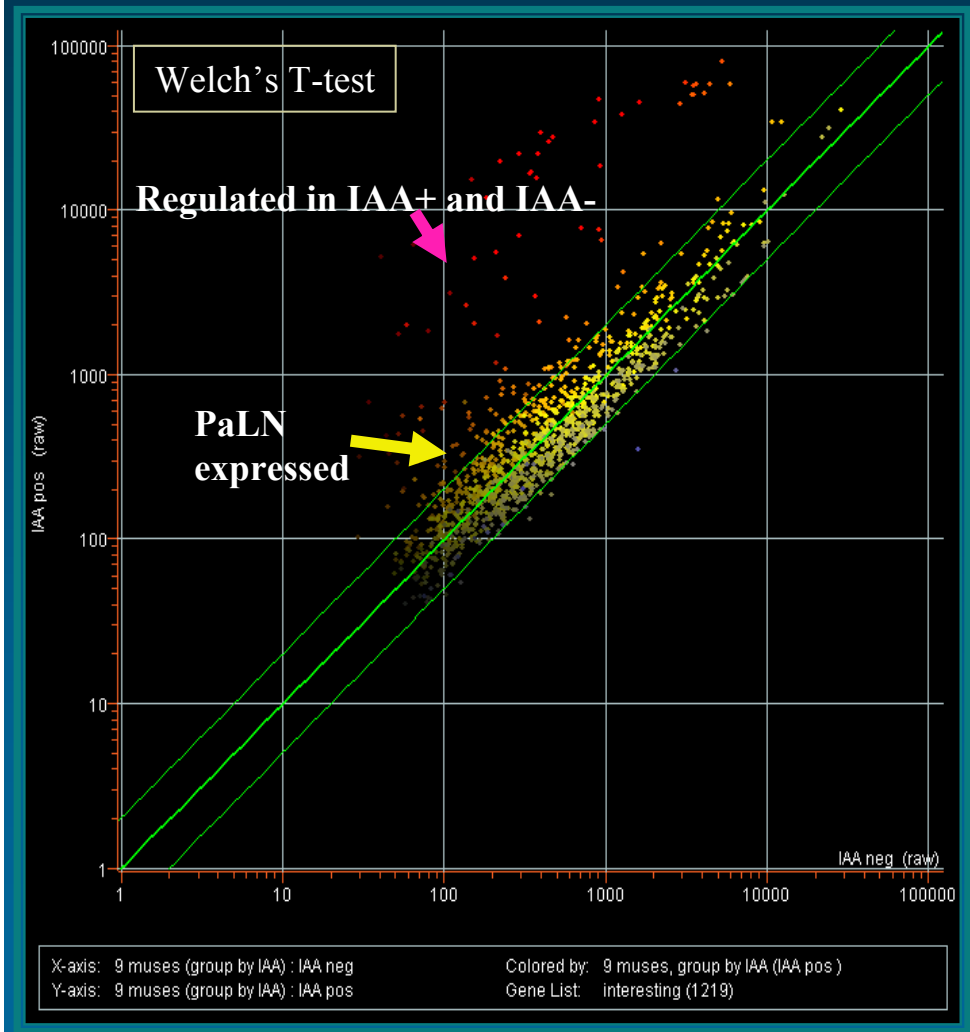
Gene expression clustering
Dendrogram based on the 1219 genes

Melanitou E

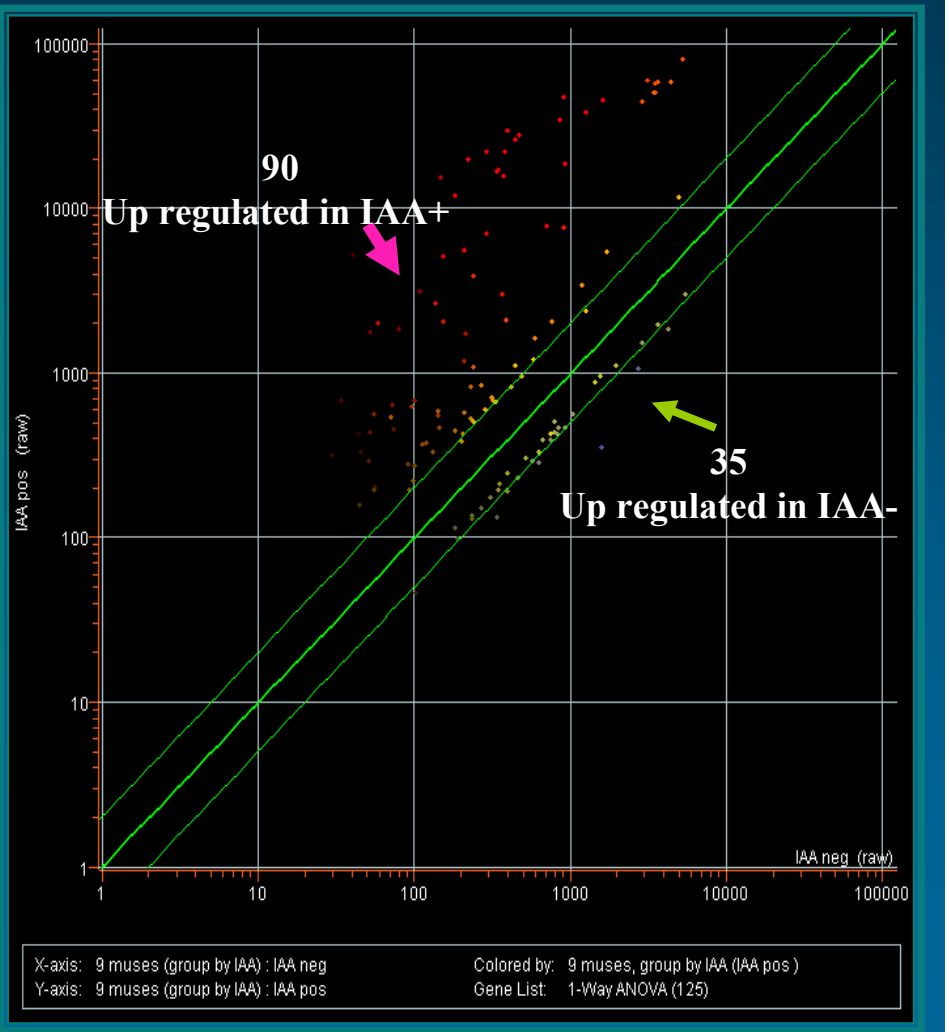

IAA levels
IAA (index)
BS (mg/dl)
weeks
WEEKS
Blood Sugar levels
GLUCOSE
INSULIN Ab
BY AGE NOD
(Eisenbarth's group)

The 1219 genes have been analyzed for differences with the
Welch's T-test (does not assume that variances are equal, p value cut-off =0.10)
Multiple correction method by the Benjamini and Hochberg false discovery rate

125 genes

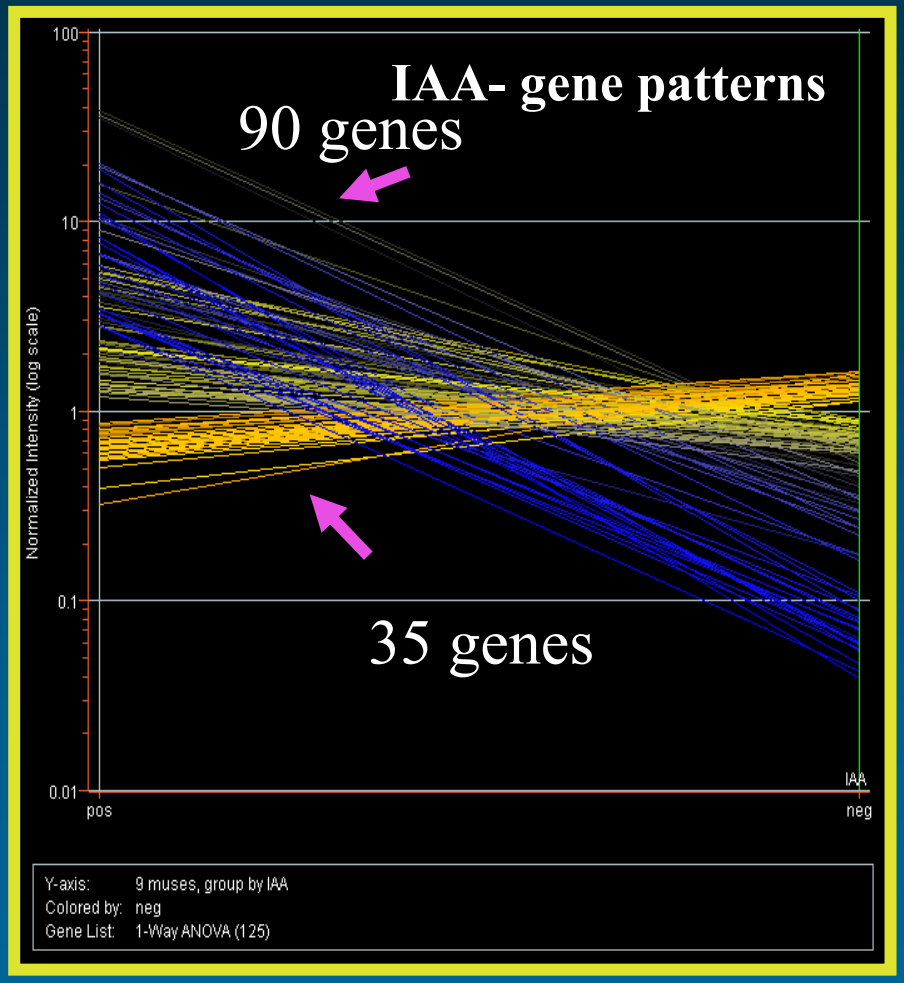


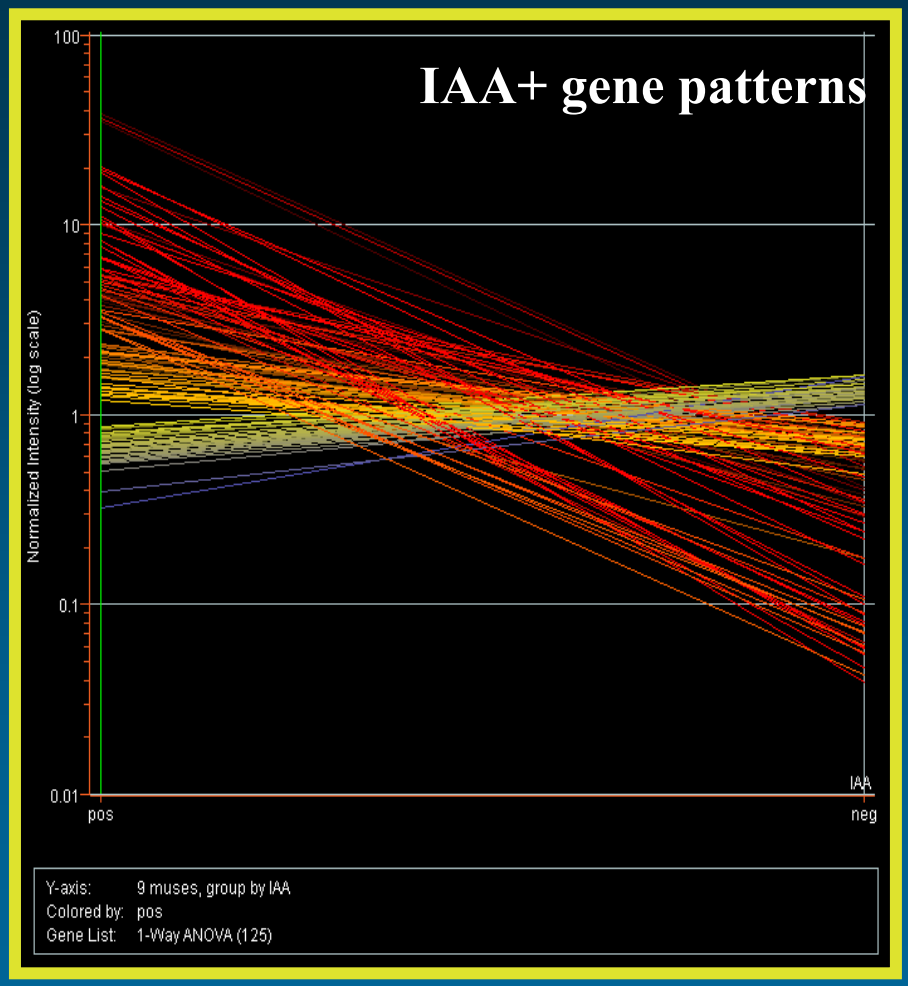
Scatter plot of 1219 genes



Scatter plot of 125 genes

Line diagrams of differences in gene expression patterns
125 significantly different genes



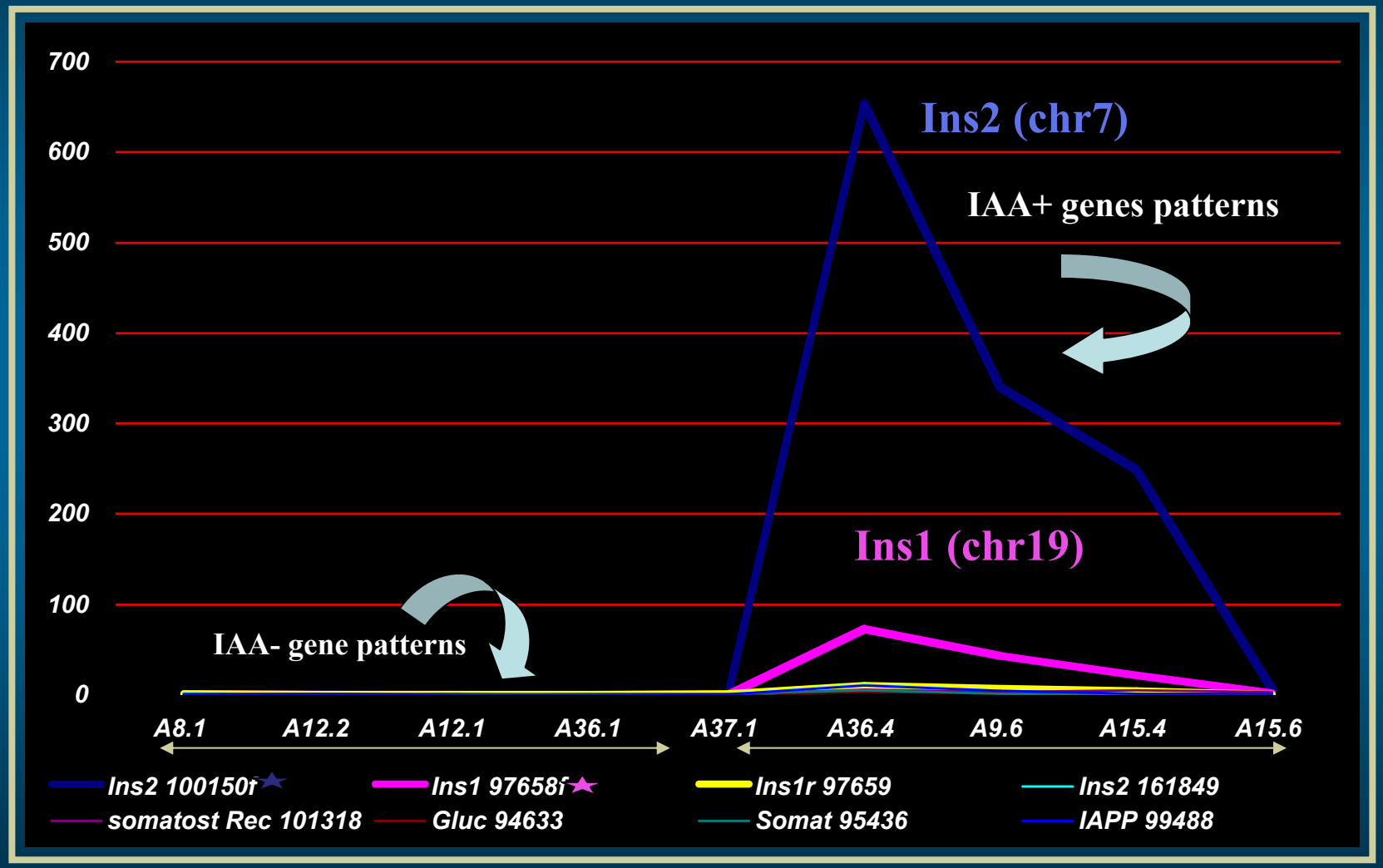


Biological Significance of gene profiling

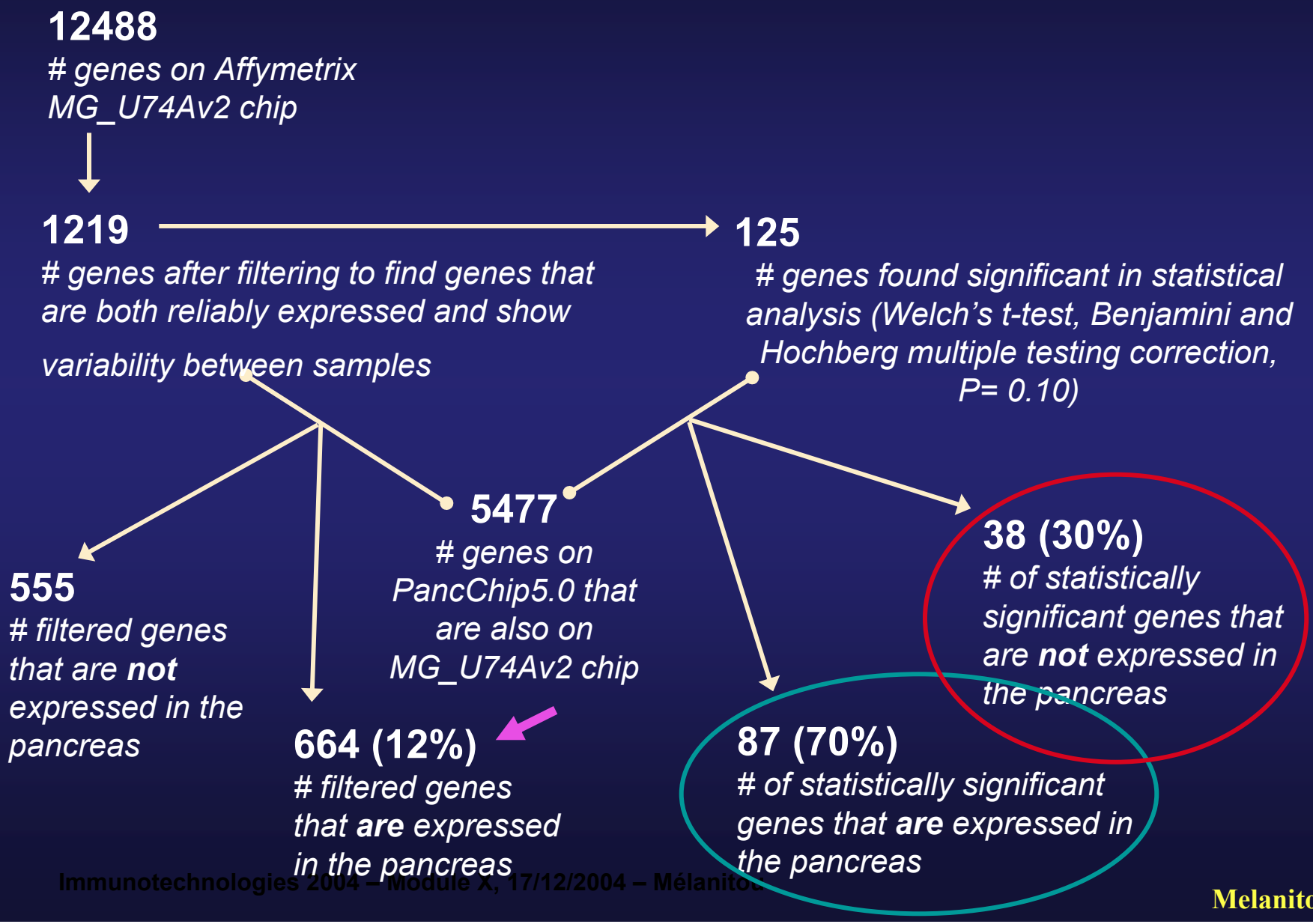
Functional analysis of the microarray data

1. How many islet-specific genes are found also expressed in the PaLN?
2. How many genes expressed in the PaLN are also expressed in the pancreas?
3. Are there gene networks specific to the presence or absence of IAA?
4. Hierarchy of genes expressed, by differences intensity.

1. How many islet-specific genes are found also expressed in the PaLN?



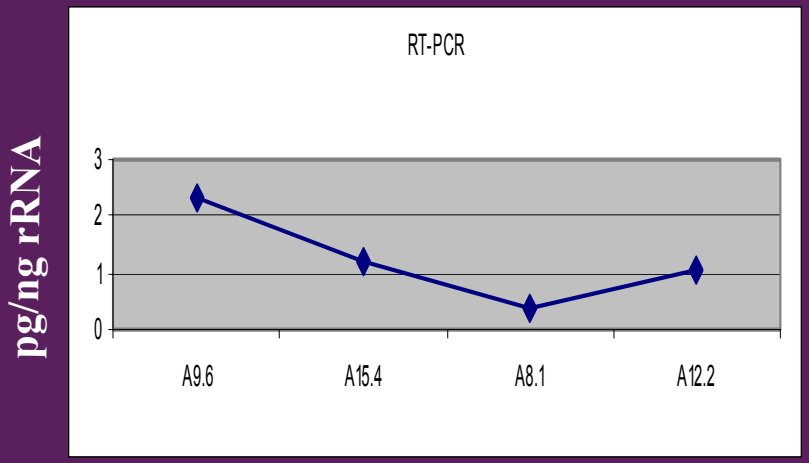
2. How many genes expressed in the PaLN are also expressed in the pancreas?



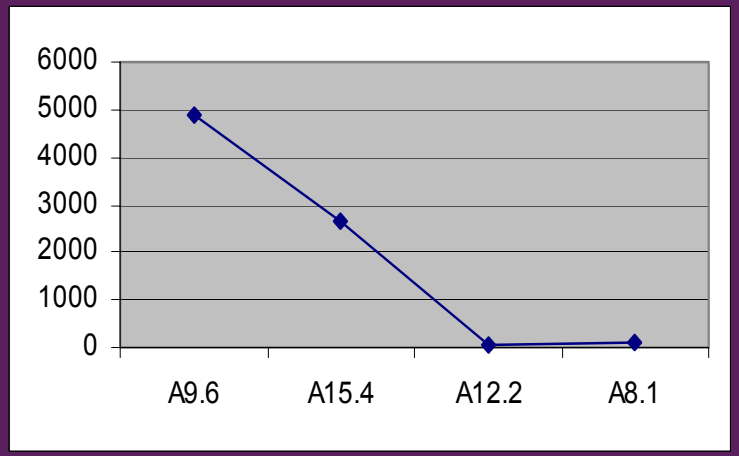
Q-RT-PCR vs microarrays: insulin genes expression

Insulin 1

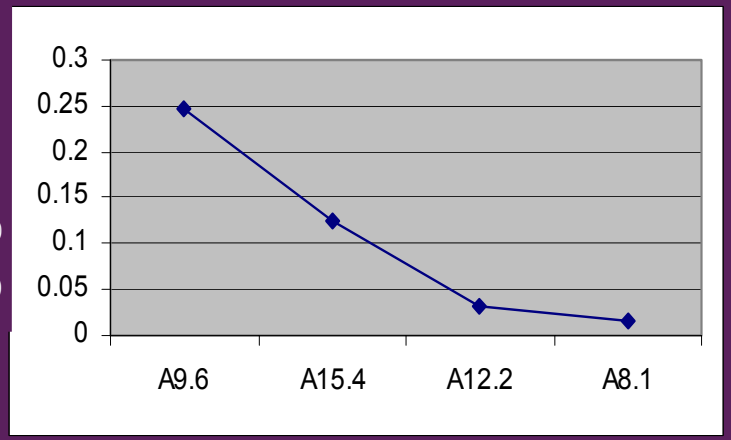
Real-Time PCR

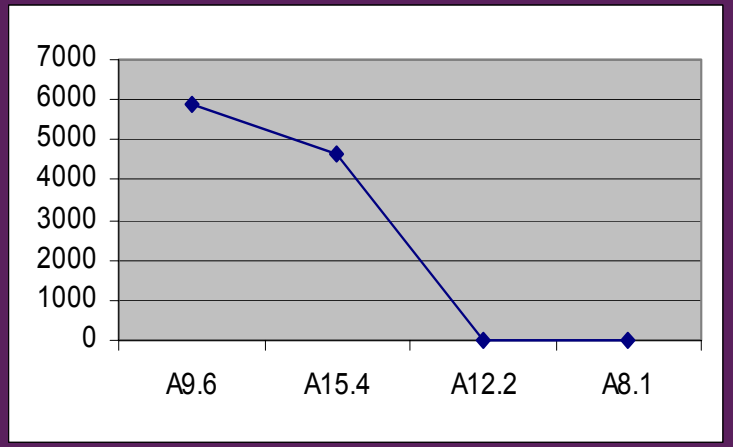


Arrays Row data

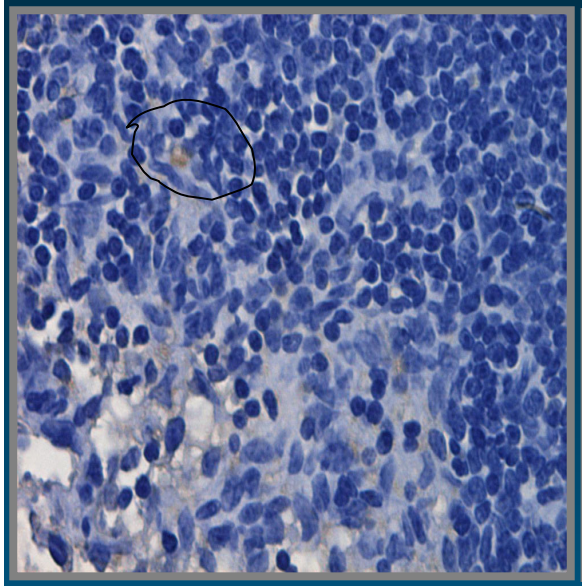


Insulin 2

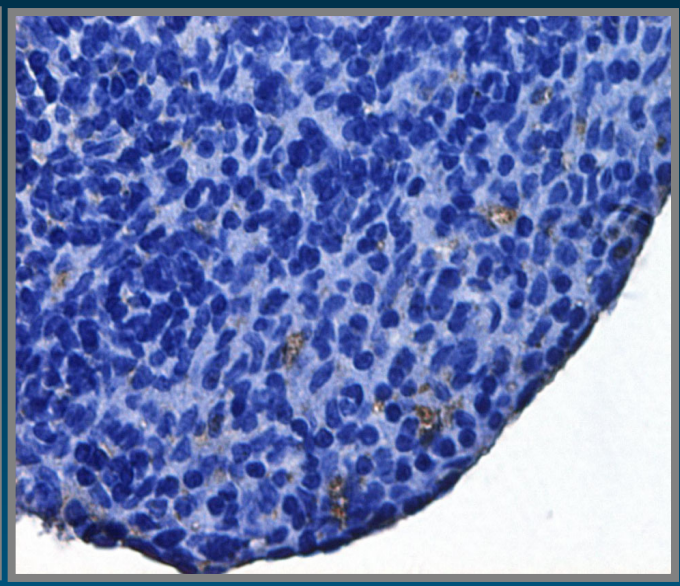




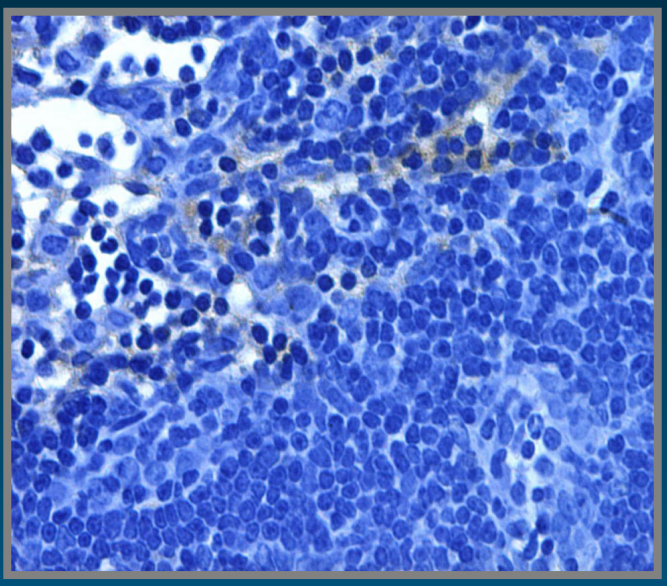
Pancreatic lymph Nodes : Insulin 1:200 (5 weeks, females)



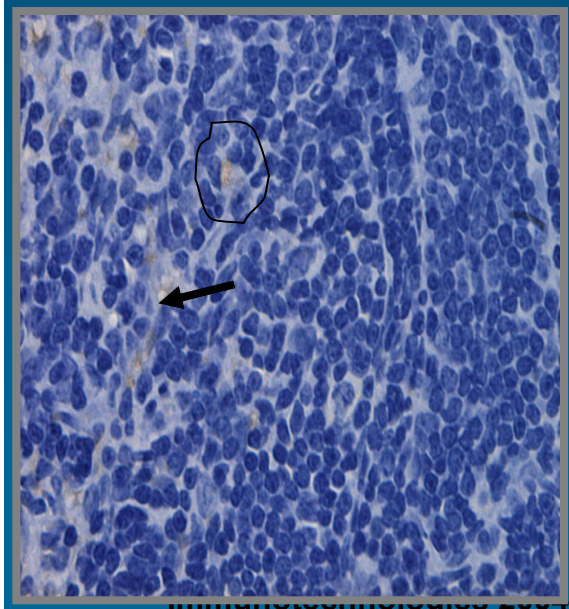
Balb/c

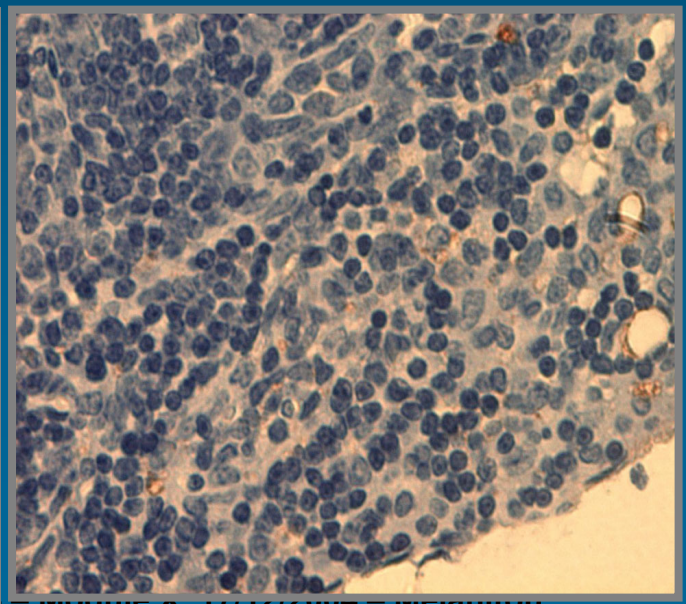


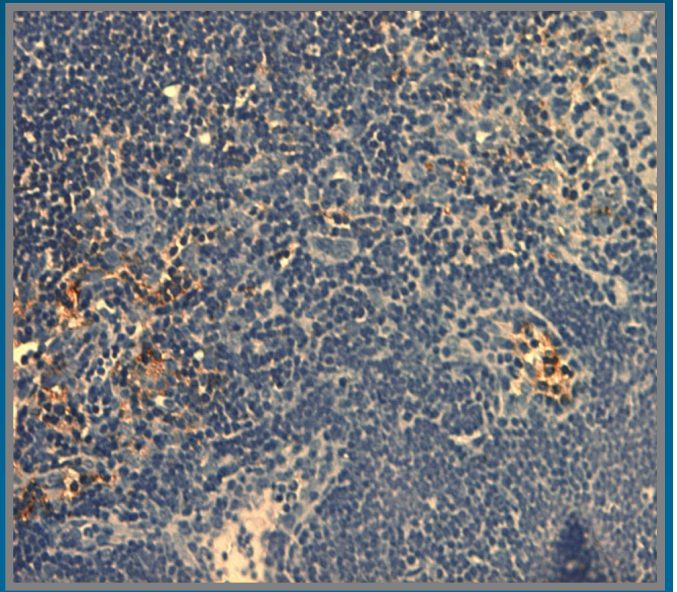
C57Bl6



NOD/Tac







3. Are there gene networks specific to the presence or absence of IAA?

90 IAA+ up regulated genes

Biological process	%genes from GO	% genes from list
•Physiological process	0.7%	49%
•**Digestion**	**66.6%**	**9%**
•Metabolism	0.7%	34%
•**Catabolism**	**0.3%**	**21%**
•**Macromolecular catabolism**	**2.9%**	**16%**
•**Protein metabolism**	**1%**	**18%**
•**Protein catabolism**	**3%**	**16%**
•**Proteolysis and peptidolysis**	**3%**	**16%**
•Cellular process	0.4%	20%
•Cell growth and/or maintenance	0.6%	14%

35 IAA+ down regulated genes

Biological process	%genes from GO	% genes from list
•Physiological process	0.3%	54%
•Metabolism	0.3%	40%
•Catabolism	0.4%	8.5%
•Macromolecular catabolism	0.6%	8.5%
•Protein metabolism	0.2%	8.5%
•Protein catabolism	0.6%	8.5%
•Proteolysis and peptidolysis	0.6%	8.5%
•Cellular process	0.2%	23%
•Cell growth and/or maintenance	0.2%	14.2%
•Cell communication	0.1%	8.5%
•**Nucleobase, nucleoside and nucleic acid metabolism**	**0.4%**	**20%**
•**RNA metabolism**	**3%**	**11%**
•**RNA processing**	**3.1%**	**11%**
•**mRNA processing**	**3.5%**	**8.5%**

Melanitou E

SUMMARY III

- There are gene network differences in the PaLN at 5 weeks of age between IAA+ and - NOD mice
- A stochastic expression of pancreas related genes is associated with early autoimmunity
- Selective genes involved in protein metabolism/catabolism are up-regulated in the IAA+ samples
- In contrast, genes involved in RNA processing are up-regulated in the IAA- samples
- Several genes playing a role in T cell activation/regulation are differentially expressed
- IAA is a good marker for early autoimmune changes in T1D reflecting gene profiles changes