

# *Genetics and diversity*

**John Williams**

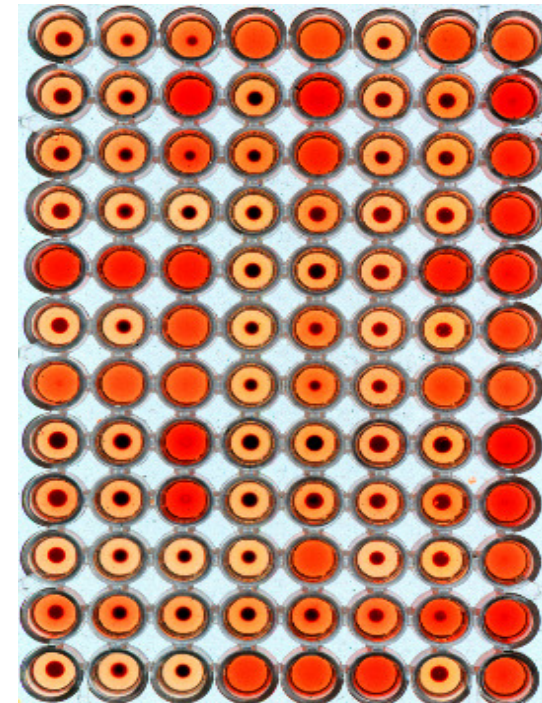
**Parco Tecnologico Padano  
Lodi**

# Basic tools

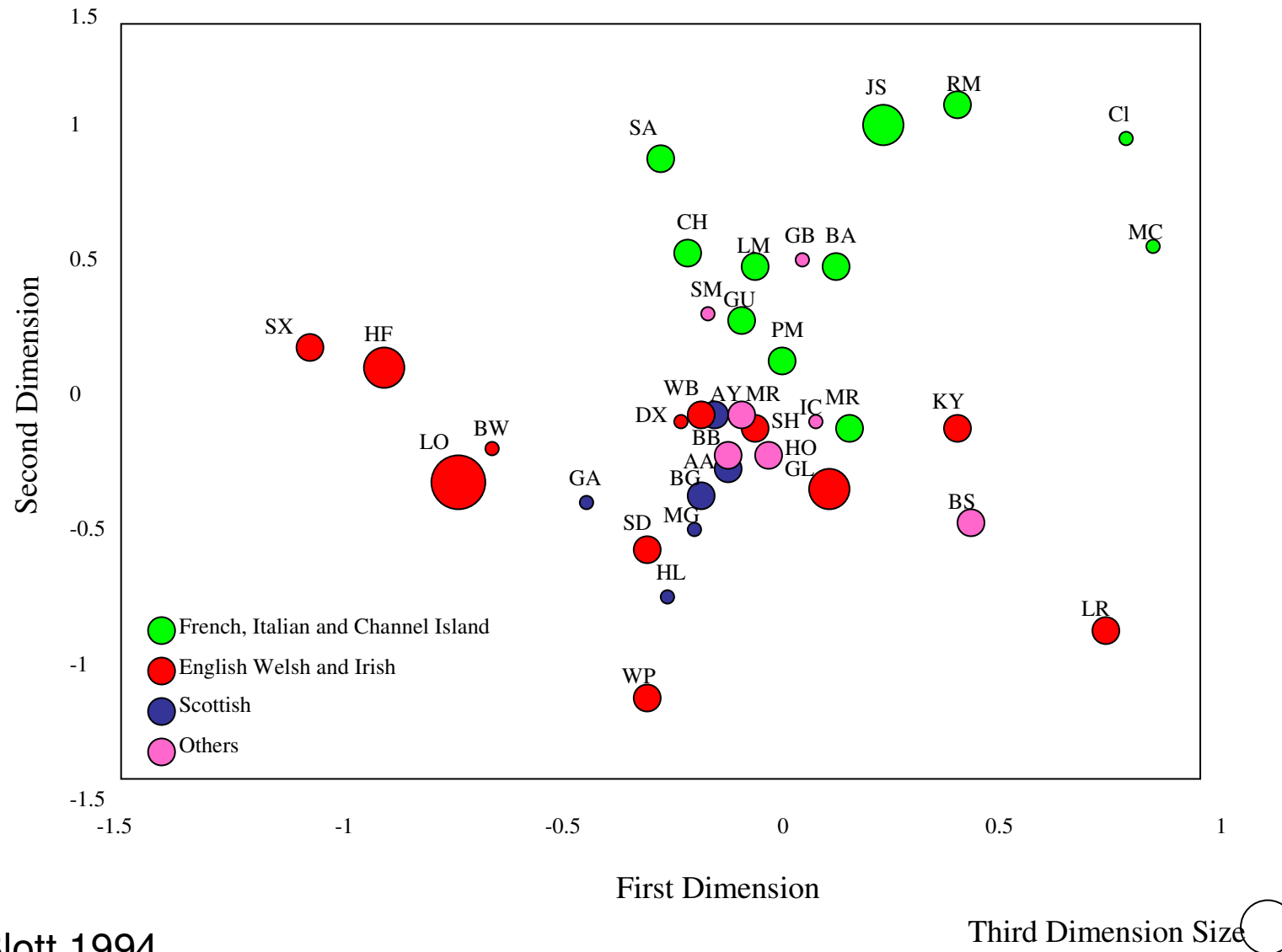
## ***Blood group systems in cattle***

System A	A D H Z'
System B	B G I K O P Q T Y A' B' D' E' F' I' J' K' O' P' Q' Y' A'' B'' D'' F'' G'' I'' J'' O''
System C	C E R W X X' C' L'' C''
System F	F V N'
System J	J
System L	L
System M	M M'
System S	S U H' U' H'' S'' U''
System Z	Z
System R'	R' S'
System T'	T'

11 'systems'



# *Genetic distance vs location*



# *Hereford cattle*

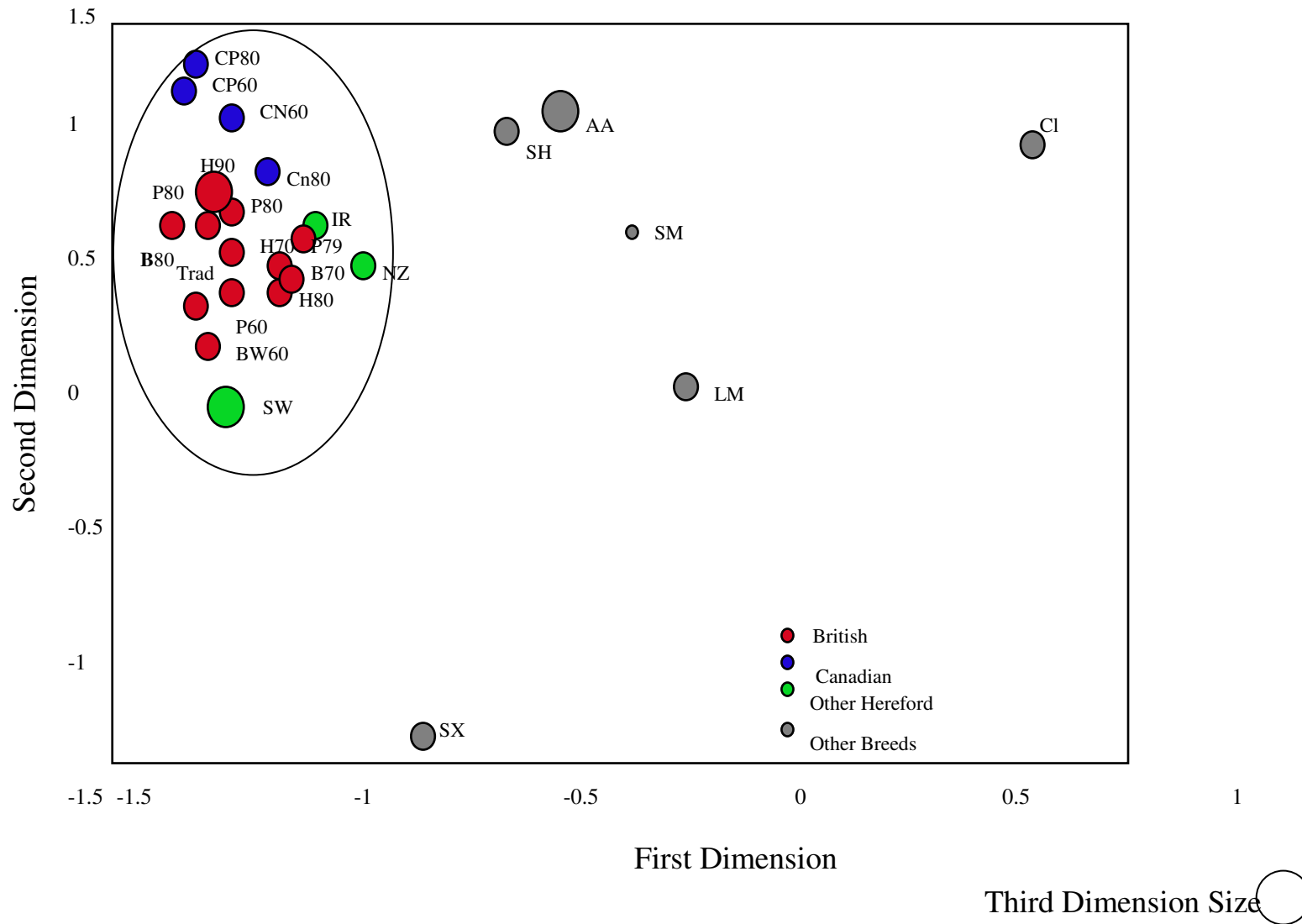


UK traditional

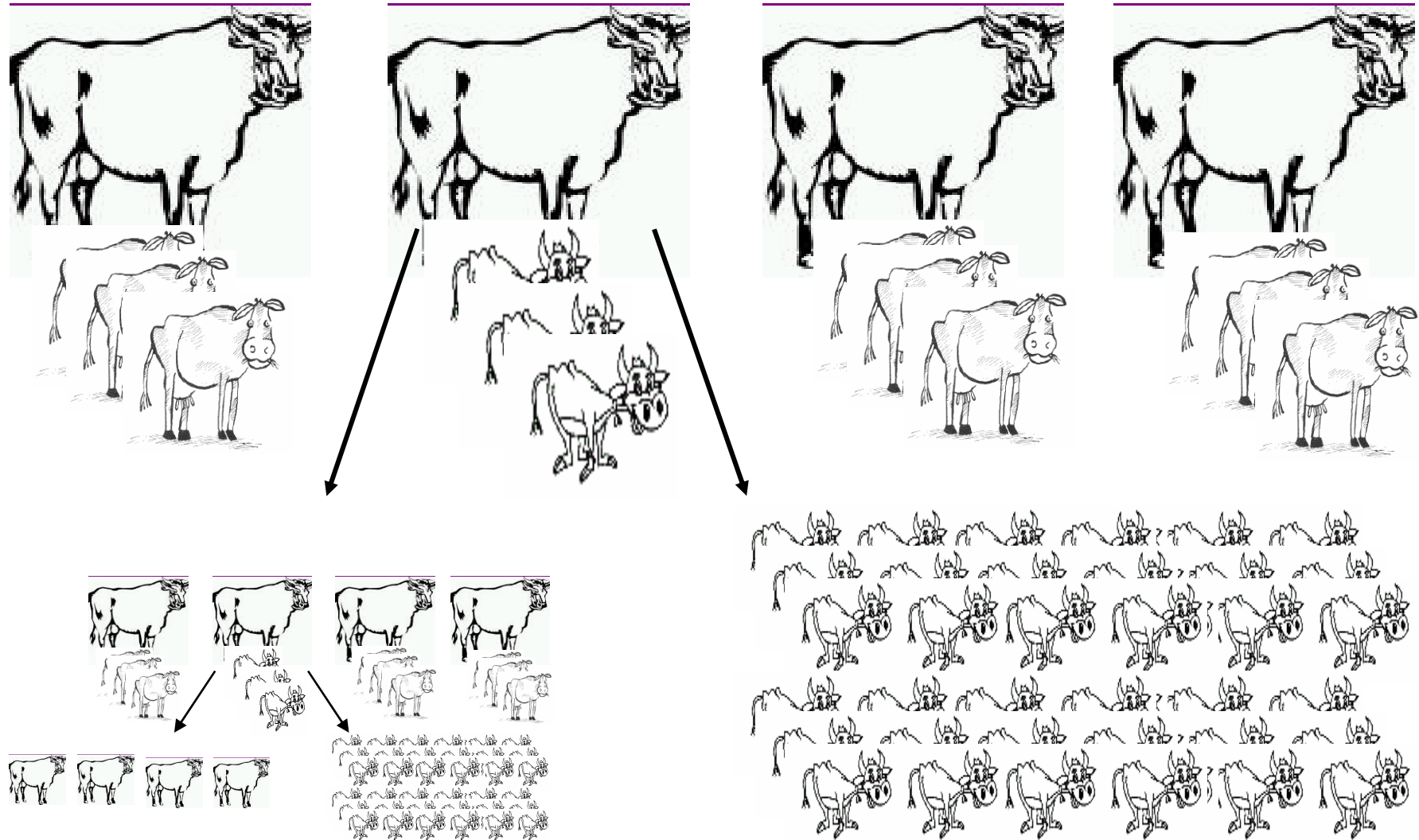


Canadian

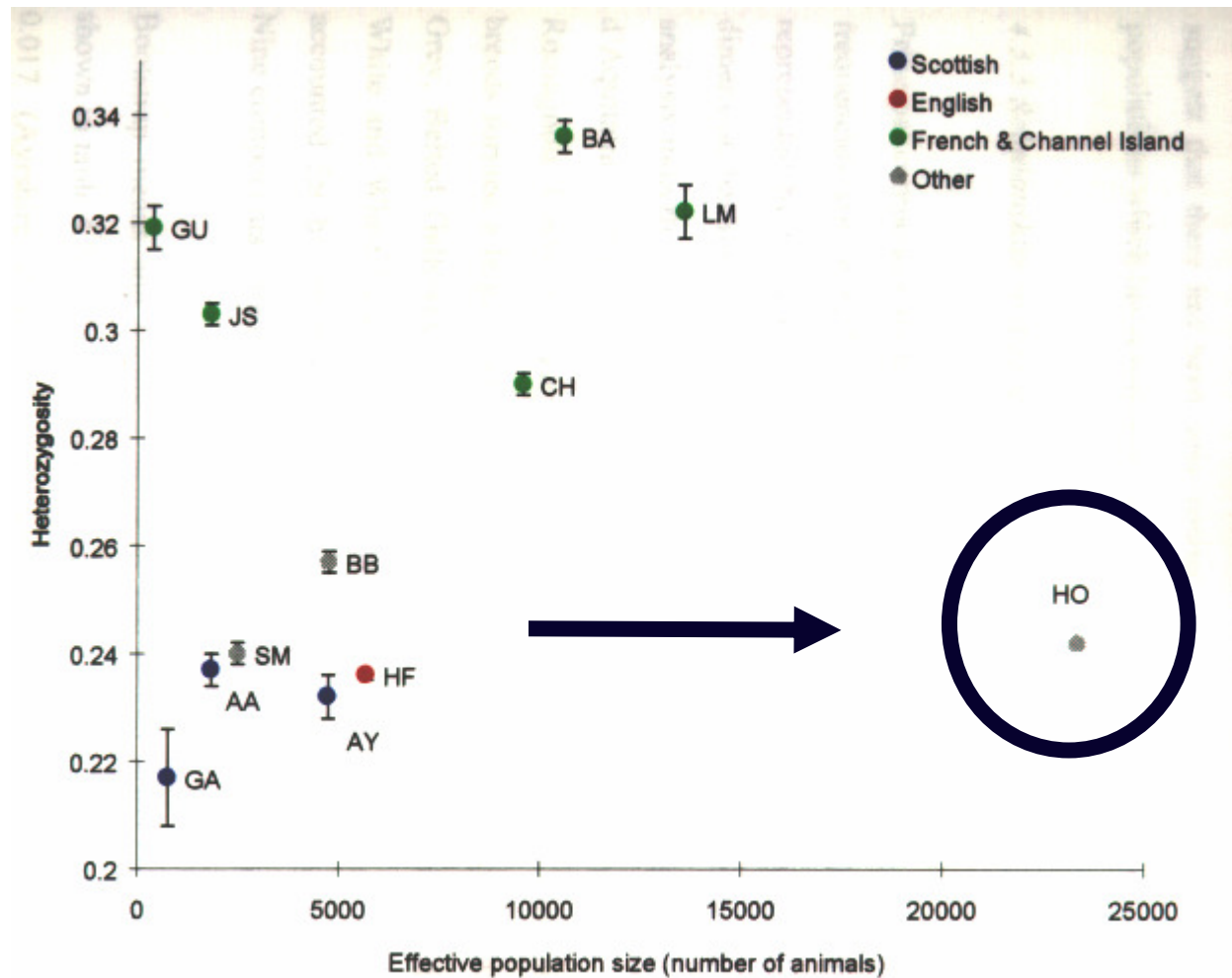
# *Hereford across the world*



# *Effects of selection*



# Effects of selection on diversity

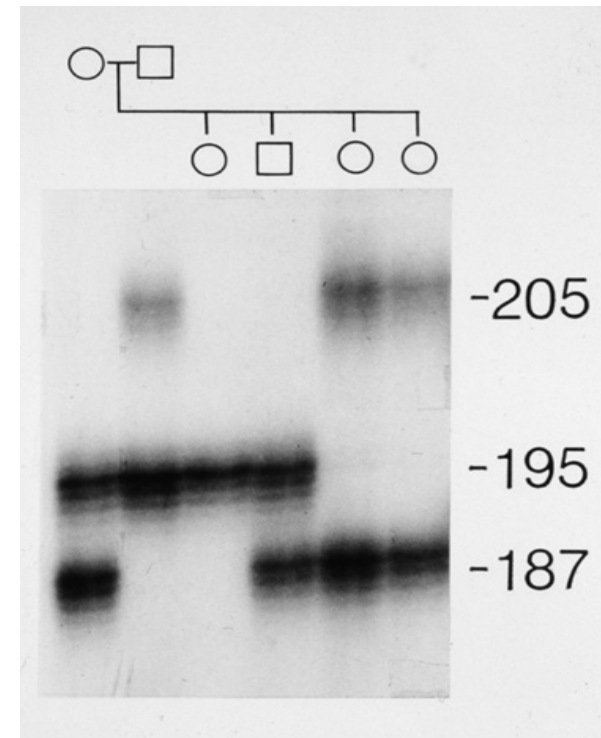
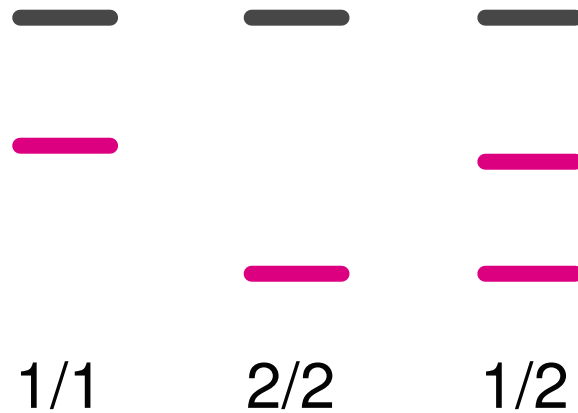




# Then came microsatellite loci

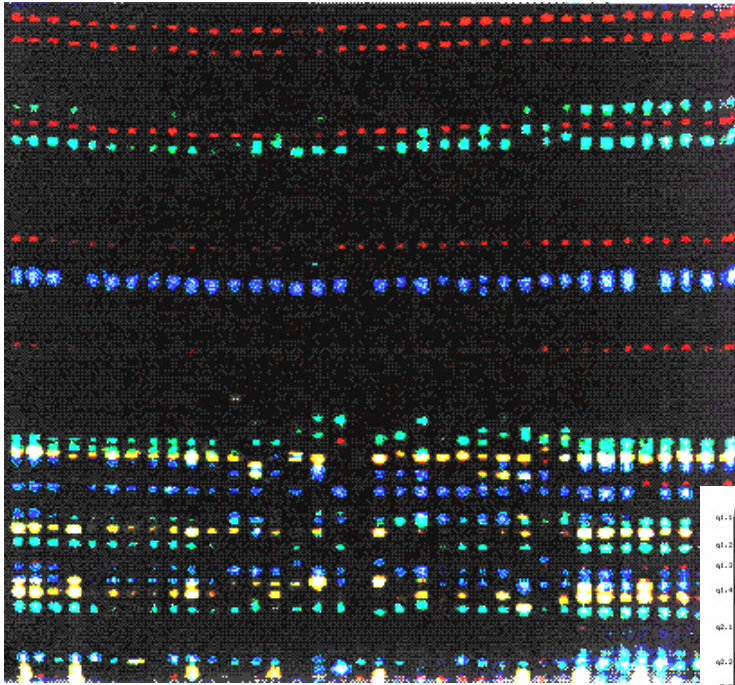
1 gatccagcaCACACACACACACACACAcggttacgatcatg

2 gatccagcaCACACACACAacggttacgatcatg

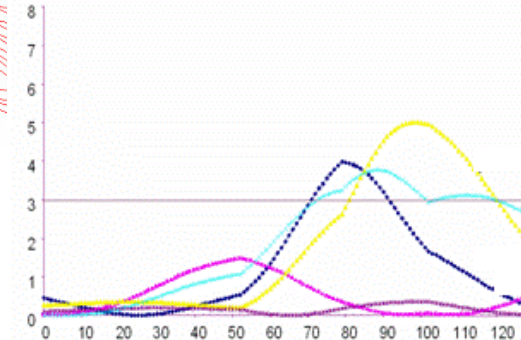
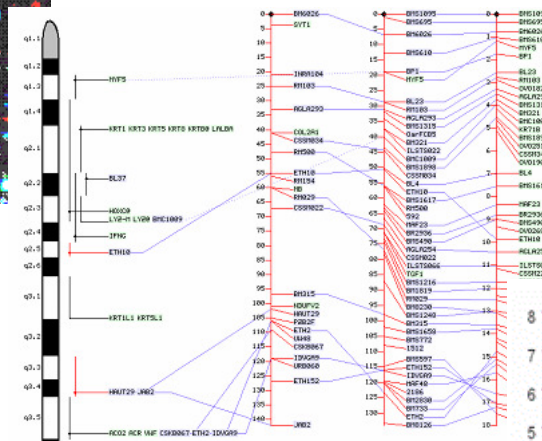




# The QTL decade (1995 – 2005)

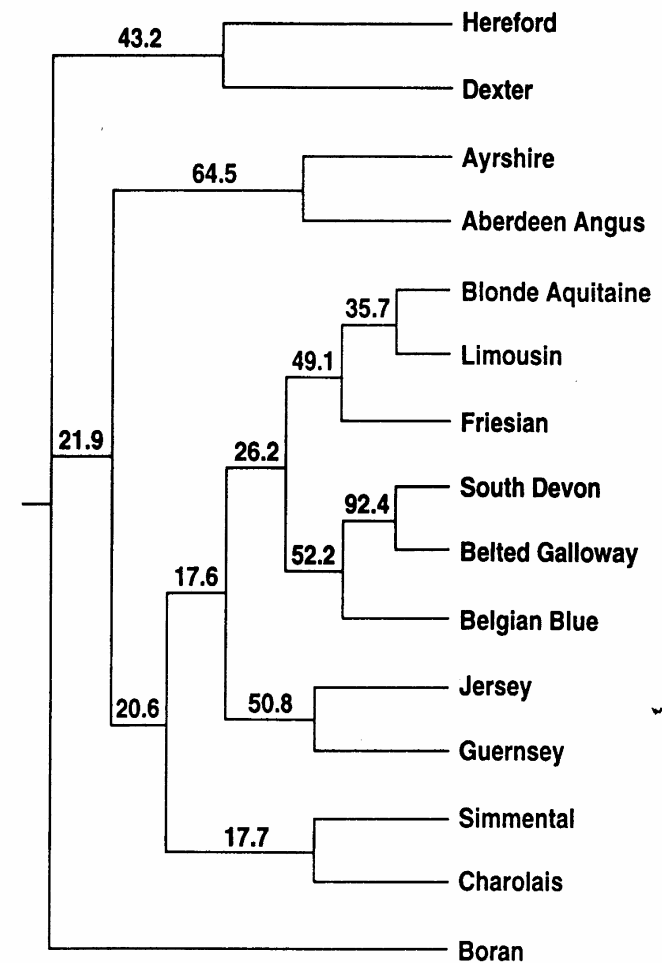


- Markers built into genetic maps
- Maps used to localize “Quantitative trait loci”
- A few major genes identified

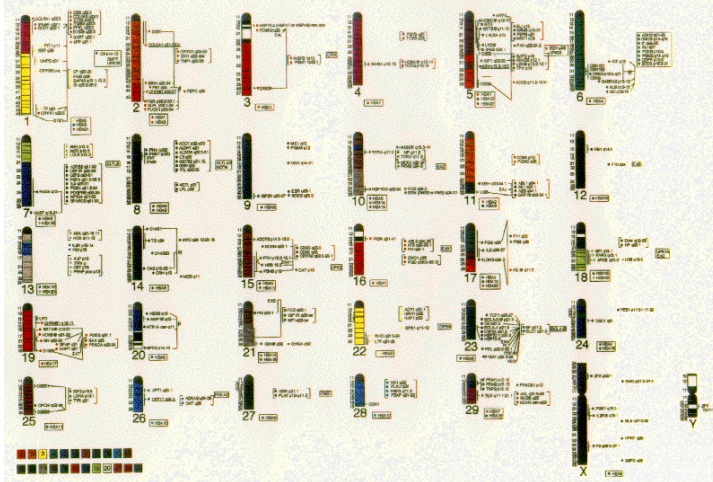


# Setting the standard

- *European ResGen Project*
  - (AIR3 CT94-20665, 1994)
- Definition of breed by DNA markers
- 60 Breeds
- 30 Microsatellite markers
  - Adopted by ISAG &FAO



# No information...



- In cattle 1 marker per chromosome
- No data on genome wide diversity

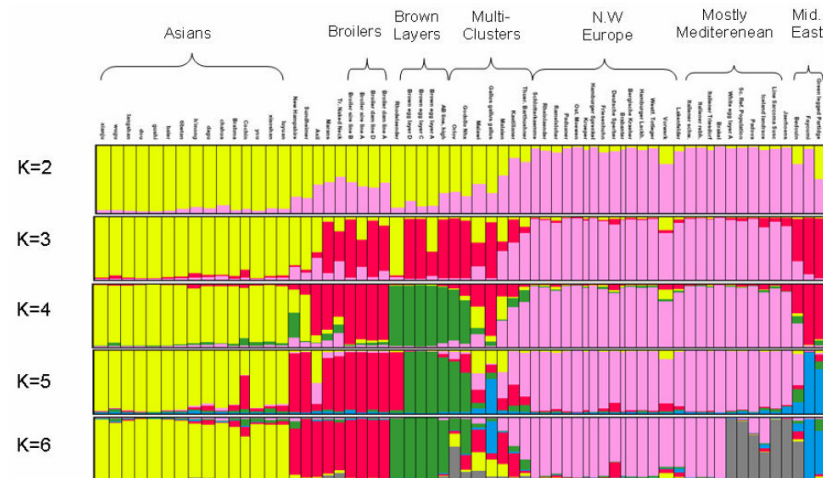
- No information on phenotypes



# No progress for 10 years!

- Lenstra, J.A., et al (2006) Marker-Assisted Conservation of European Cattle Breeds: An Evaluation. *Animal Genetics* **37** 475-481

- Hillel J. et al (2007)  
Molecular markers for  
the assessment of  
chicken biodiversity.  
*World Poultry Science  
Journal* **63** 34-35



29 microsatellite markers  
65 populations



# Genome sequencing of “agricultural” species

*Chicken*      2005

*Bovine*      2006

*Pig*      2008

*Sheep*      2009

*+ plants and micro-organisms.....*



NIH NEWS ADVISORY

National Institutes of Health

National Human Genome Research Institute

## **Bovine Genome Assembled**

*International Effort Makes Data Freely Available to Scientists Worldwide*



## **NEWS RELEASE**

United States Department of Agriculture • Office of Communications • 1400 Independence Avenue, SW  
Washington, DC 20250-1300 • Voice: (202) 720-4623 • Email: [oc.news@usda.gov](mailto:oc.news@usda.gov) • Web: <http://www.usda.gov>

Release No. 0295.06

Contact:

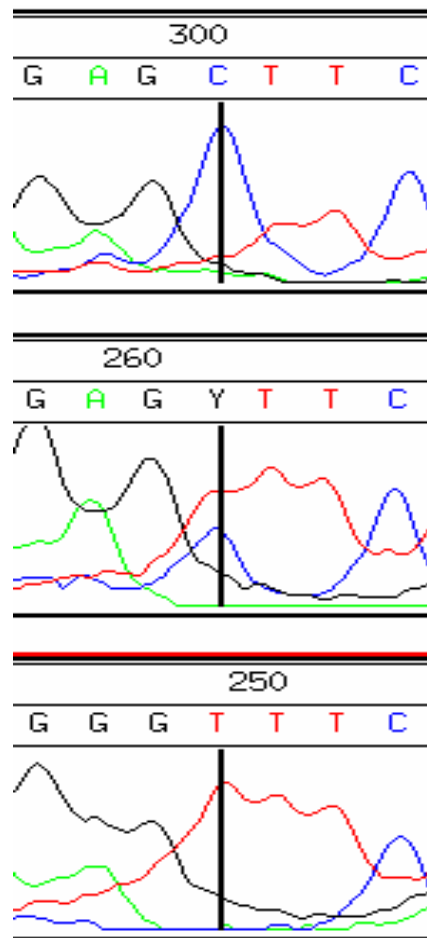
Ed Loyd (202) 720-4623

Jennifer Martin (202) 720-8188

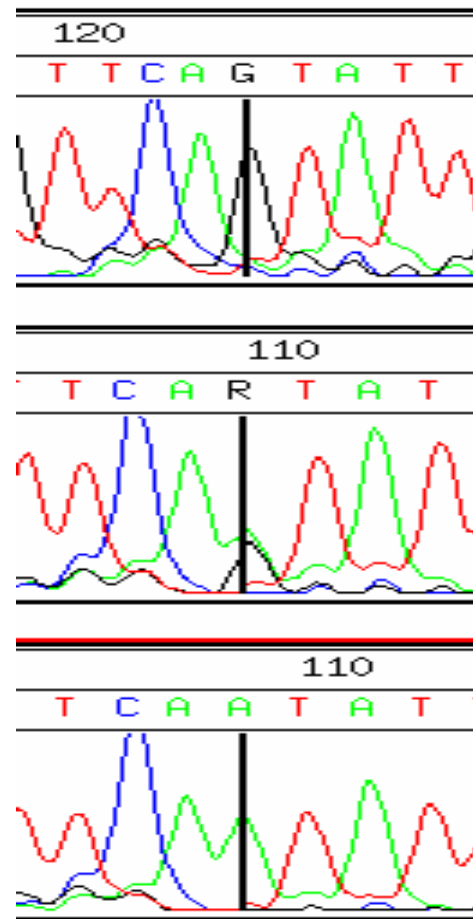
**INTERNATIONAL BOVINE GENOME SEQUENCING PROJECT RELEASES NEW  
GENETIC RESOURCES**

WASHINGTON, April 14, 2006 — The International Bovine Genome Sequencing Project (IBGSP) has released the first draft of the bovine genome sequence, a major milestone in the effort to understand the genetics of the cow and its role in human health and the environment.

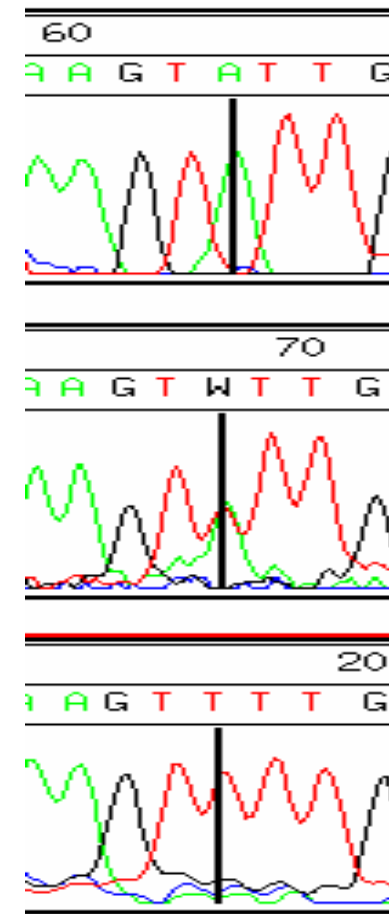
# *And so came SNPs*



**170/171**



**220/221**



**230/231**

# Single Nucleotide Polymorphism (SNP) discovery



Brahman



Limousin



Angus



Hereford



Holstein



Jersey



Norwegian Red

	Sequence Reads
Holstein	143,498
Norwegian Red	154,347
Angus	26,170
Brahman	30,000
Jersey	10,400
Limousin	3,548



# *SNP validation and use*

- **SNP discovery**

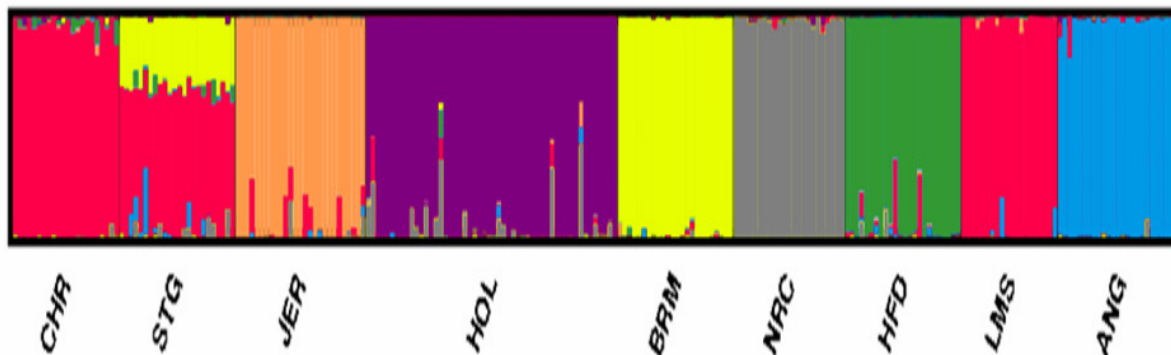
- Angus
- Brahman
- Limousin
- Holstein
- Jersey
- Norwegian Red



150K Putative SNPs

- **30K SNPs tested**

- Charolais
- Beefmaster
- Brown Swiss
- Gir
- Guernsey
- Holstein
- N'Dama
- Nellore
- Piedmontese
- Red Angus
- Romagnola
- Santa Gertrudis
- Sheko



# 2.3M bovine SNPs in dbSNP

NCBI dbSNP BUILD 127

ENTREZ **SNP**  
Single Nucleotide Polymorphism

My NCBI  
[Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure Popset Taxonomy SNP

Search SNP for bos Go Clear Save Search

Limits Preview/Index History Clipboard Details

Display Graphic Summary Show 20 Sort by Send to

All: 2323150 Human: 0 Mouse: 0 NEW: 17341 Other Organisms: 2225118 UPDATE: 31

Un no Map No Gene Not on mRNA No 3D No OMIM

20: [rs17873847](#) [*Bos indicus* x *Bos taurus*]  
AAGATGTGGACTAATTGGAGGGTGGG[A/G] GTTGGGGGGTGGAAATATGAGGTTTA

Un no Map No Gene Not on mRNA No 3D No OMIM

Items 1 - 20 of 2323150 Page 1 of 1161

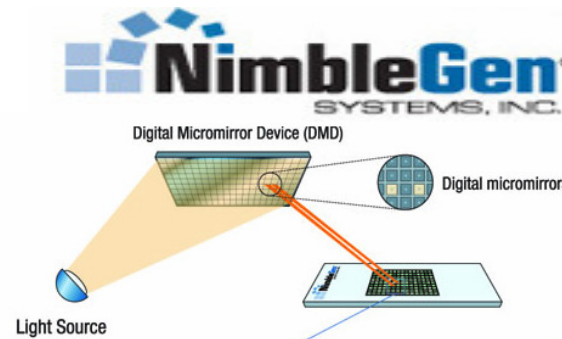
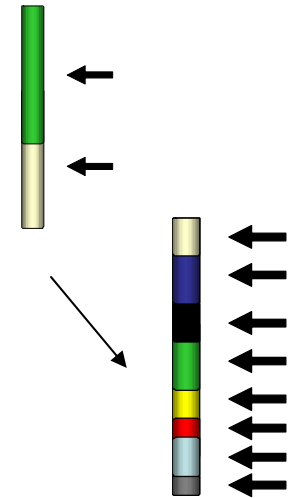
Display Graphic Summary Show 20 Sort by Send to

[Restrictions on Use](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)

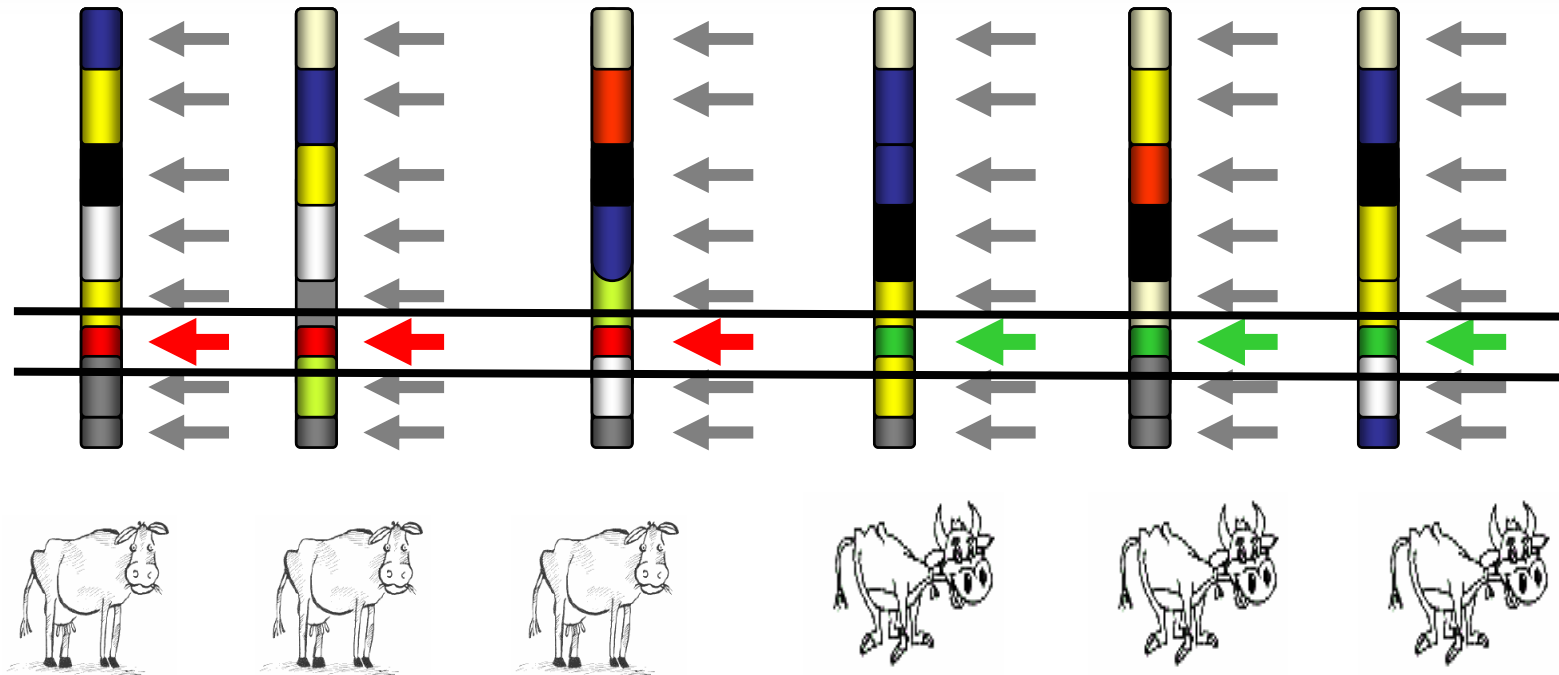
- 2323150 SNPs
- 2.1M SNPs from Hereford sequence
- 123K interbreed SNP
- 100K other sources

# *New opportunities*

- High density, high throughput, genotyping.
  - Several “platforms” available to genotype many thousands of markers simultaneously



# *Associations at the population level*

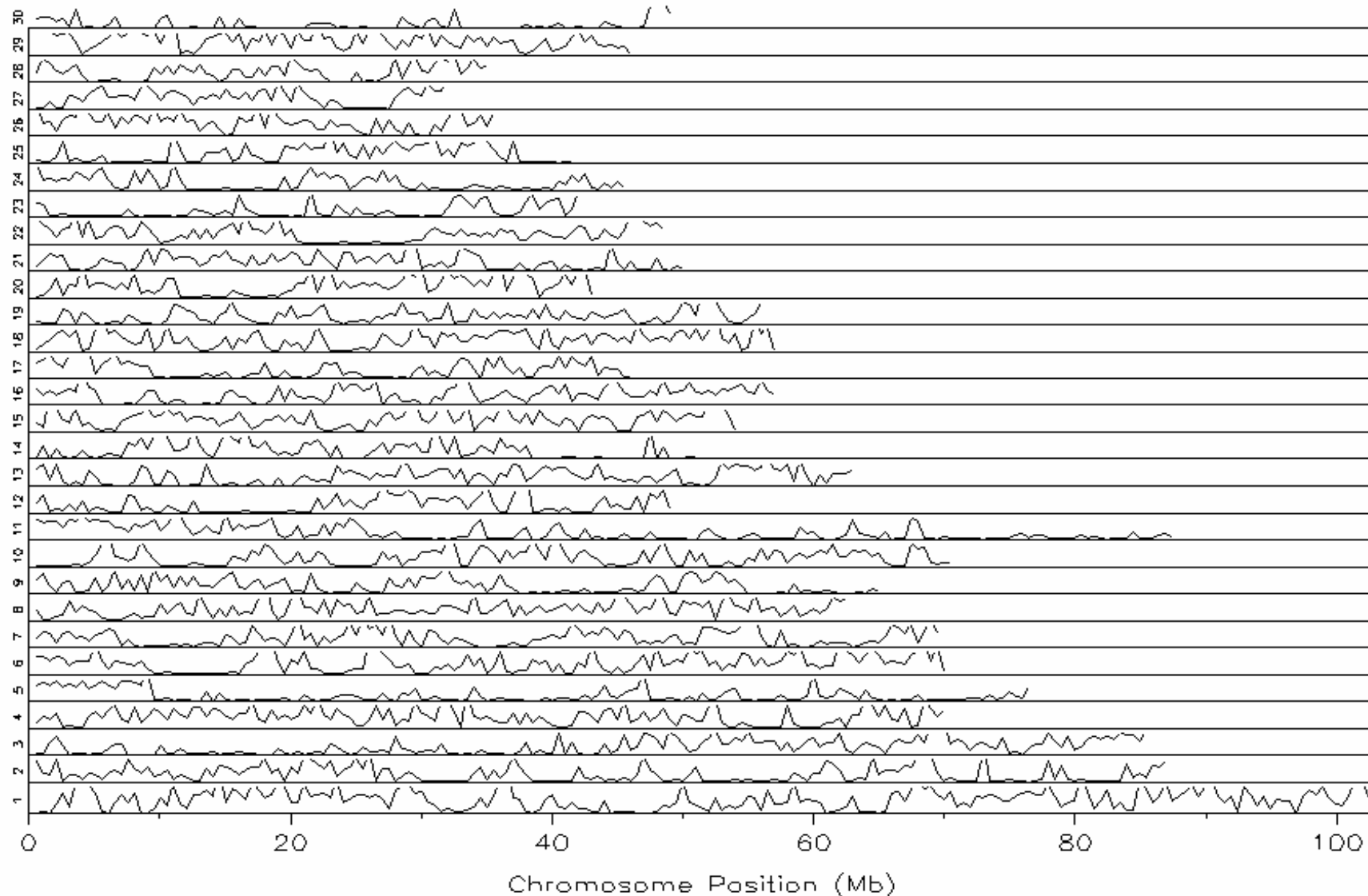


***Illumina 50K SNP panel :***

**DGAT1**                      **6 years to discover**  
(acyl-CoA:diacylglycerol acyltransferase)

**2007 mapped to within 3000bp in 2 months**

# Signatures of selection



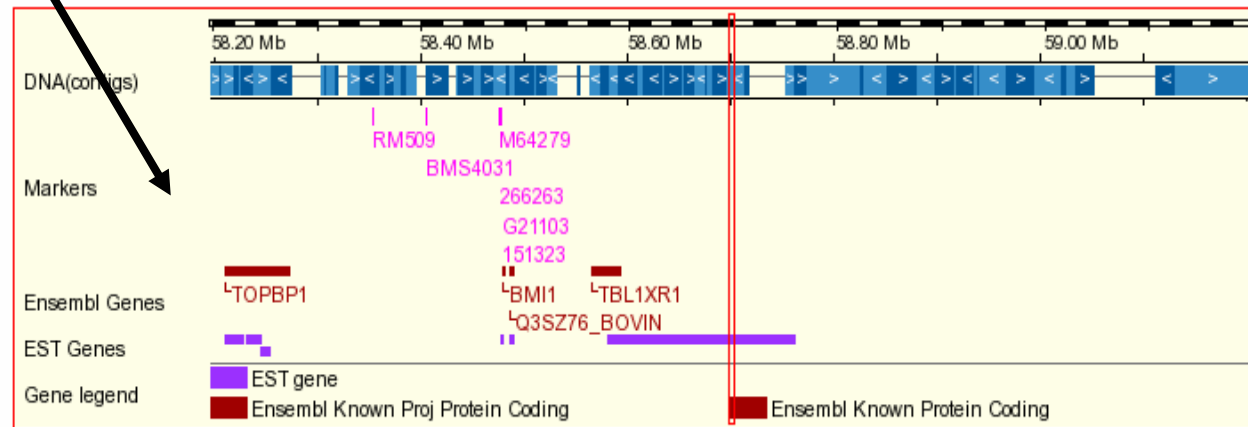
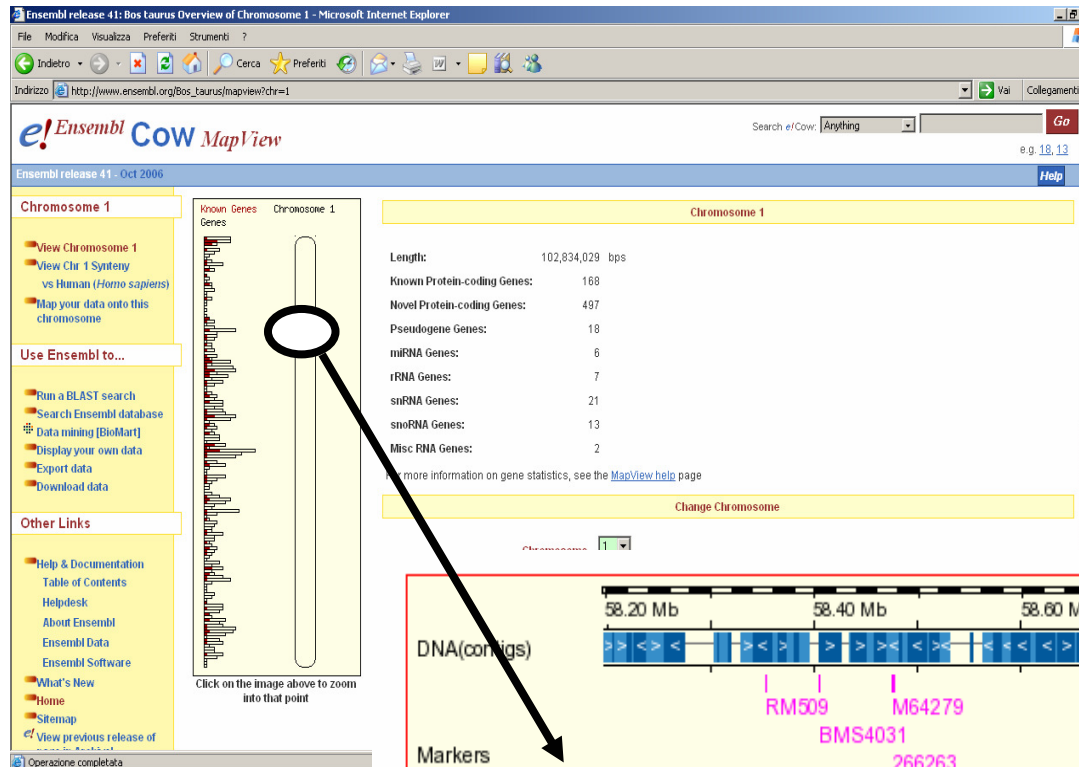
# Knowledge of the genes

## Genome annotation

Ensembl *e!*

- Semi-automated gene identification

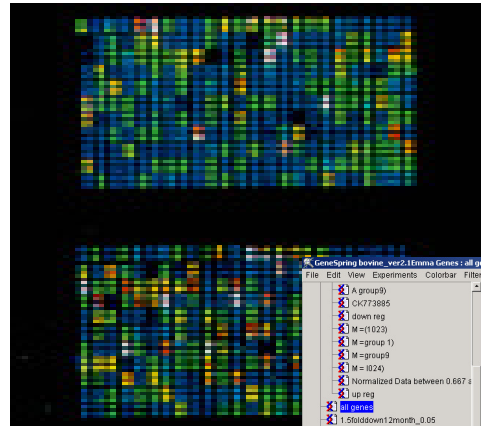
[www.ensembl.org](http://www.ensembl.org)



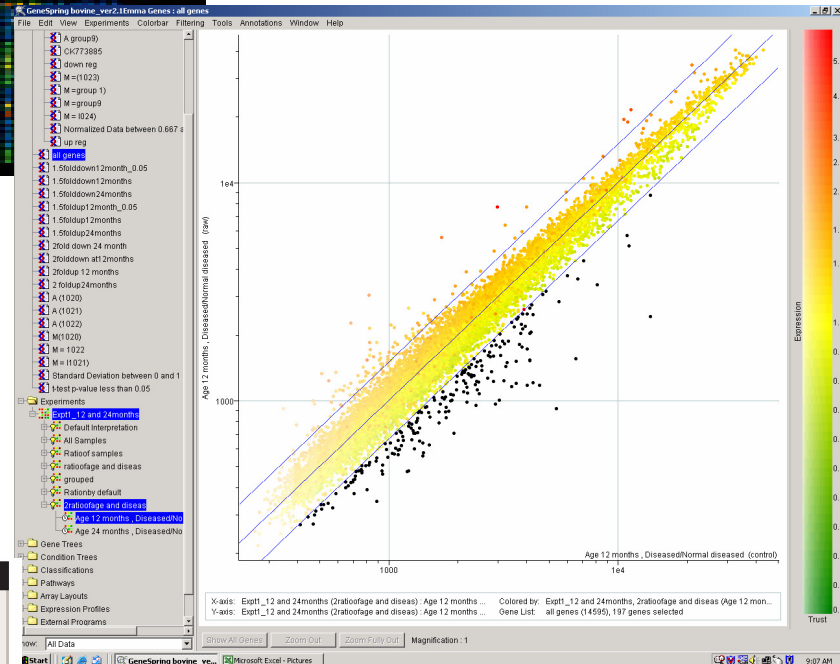
# Responses to environment

Comparison:  
of gene expression

environment 1  
vs  
environment 2



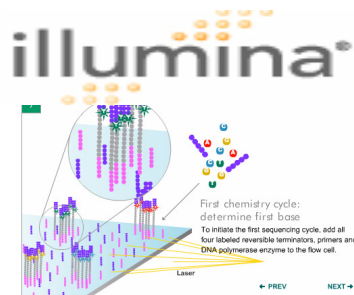
AFFYMETRIX® PRODUCT FAMILY >	
ARRAYS >	
AFFYMETRIX®	
Data Sheet	
GeneChip® Bovine Genome Array	
<b>Critical Specifications</b>	
<i>Bos taurus</i> (Bovine) probe sets	24,072
<i>Bos taurus</i> (Bovine) transcripts	approximately 23,000
UniGene clusters	approximately 19,000
<b>Unique probe sets to single species:</b>	
Number of arrays in set	one
Array format	100
Feature size	11 µm
Oligonucleotide probe length	25-mer
Probe pairs/sequence	11
Hybridization controls:	<i>bioB</i> , <i>bioC</i> , <i>bioD</i> , from <i>E. coli</i> and <i>cre</i> from <i>P1 B. subtilis</i>
Poly-A controls:	<i>dap</i> , <i>lys</i> , <i>phe</i> , <i>thr</i> , <i>trp</i> from <i>B. subtilis</i>
Housekeeping/Control genes:	actin, GAPDH, <i>eflc</i> , 5.8S rRNA, 12S rRNA, 18S rRNA, cyclophilin B, glutathione S-transferase, lactophorin, translation initiation factor eIF-4E
Detection sensitivity	1:100,000 <sup>1</sup>





# *State of the art*

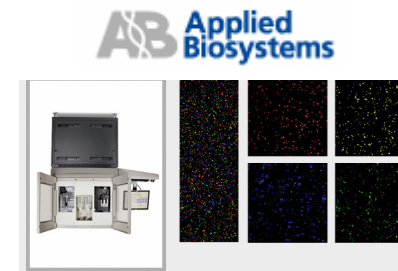
- Re-sequencing .... in days
- Genetic variations
- Gene expression
- Also new species without genome sequence



Solexa



Roche 454



SOLiD

# Opportunities

- Understanding functional diversity
  - Genotype by phenotype
  - Genotype by environment by phenotype
- Selection based on genomes  
(High densities of SNPs)
  - More rapid progress (quicker loss of diversity)
  - Genome wide conservation of diversity
  - Transfer of genes between populations
- Understanding micro-environments  
(Identification by sequence)
  - eg micro-organisms