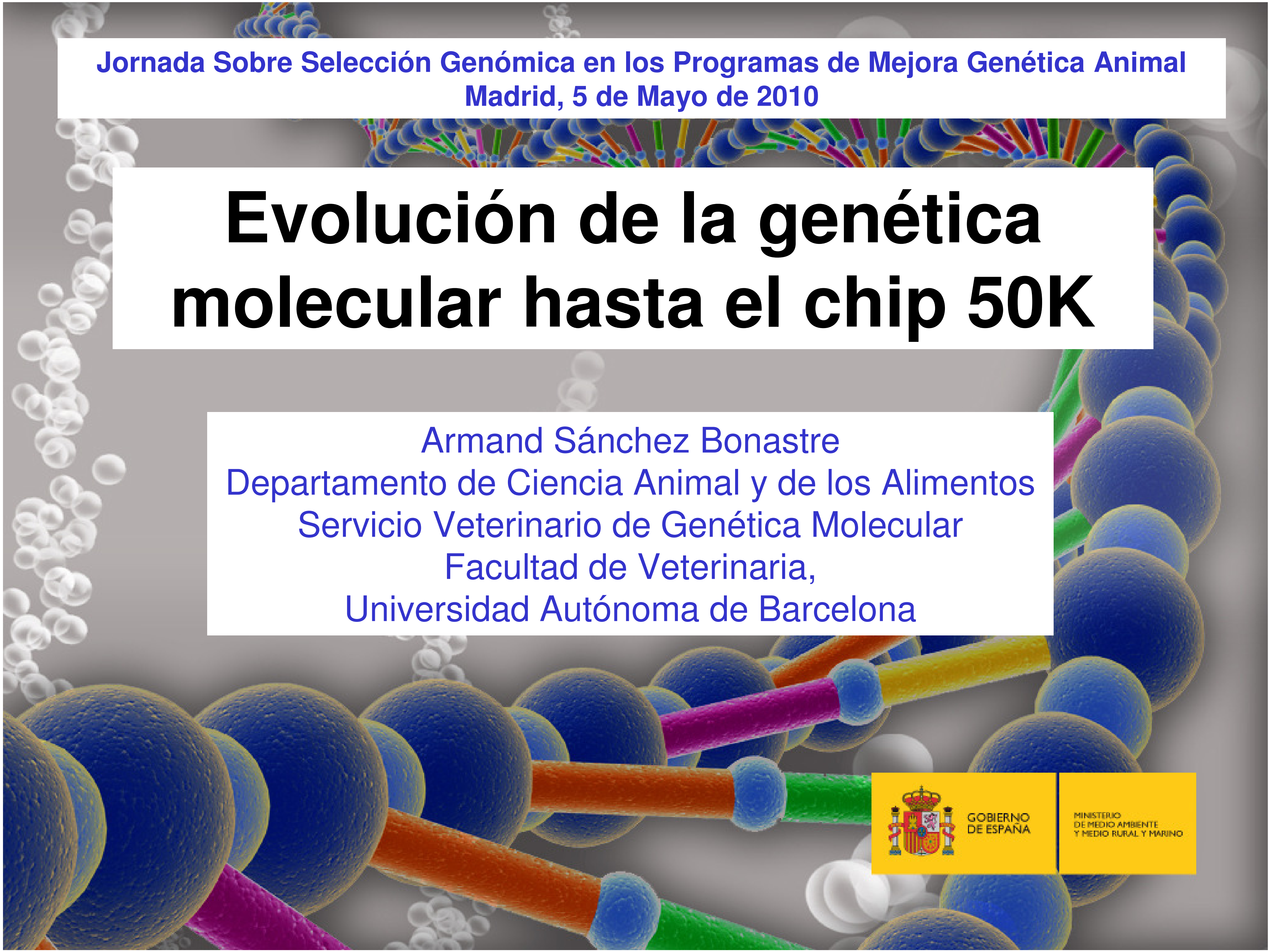


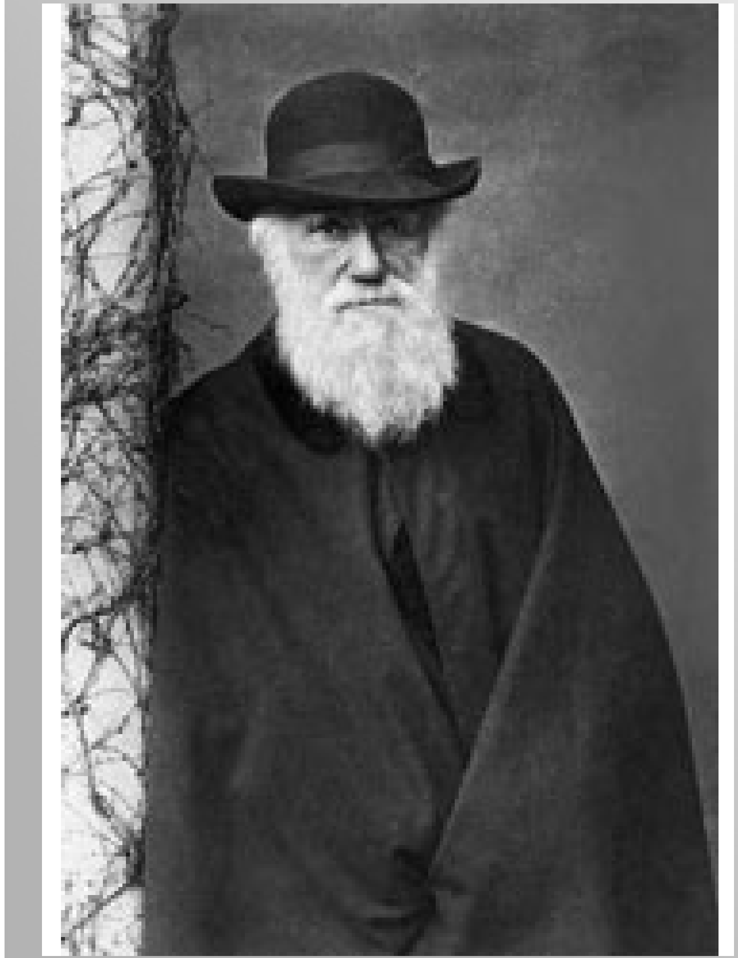
Jornada Sobre Selección Genómica en los Programas de Mejora Genética Animal
Madrid, 5 de Mayo de 2010

Evolución de la genética molecular hasta el chip 50K

Armand Sánchez Bonastre
Departamento de Ciencia Animal y de los Alimentos
Servicio Veterinario de Genética Molecular
Facultad de Veterinaria,
Universidad Autónoma de Barcelona

GOBIERNO DE ESPAÑA

MINISTERIO DE MEDIO AMBIENTE Y MEDIO RURAL Y MARINO

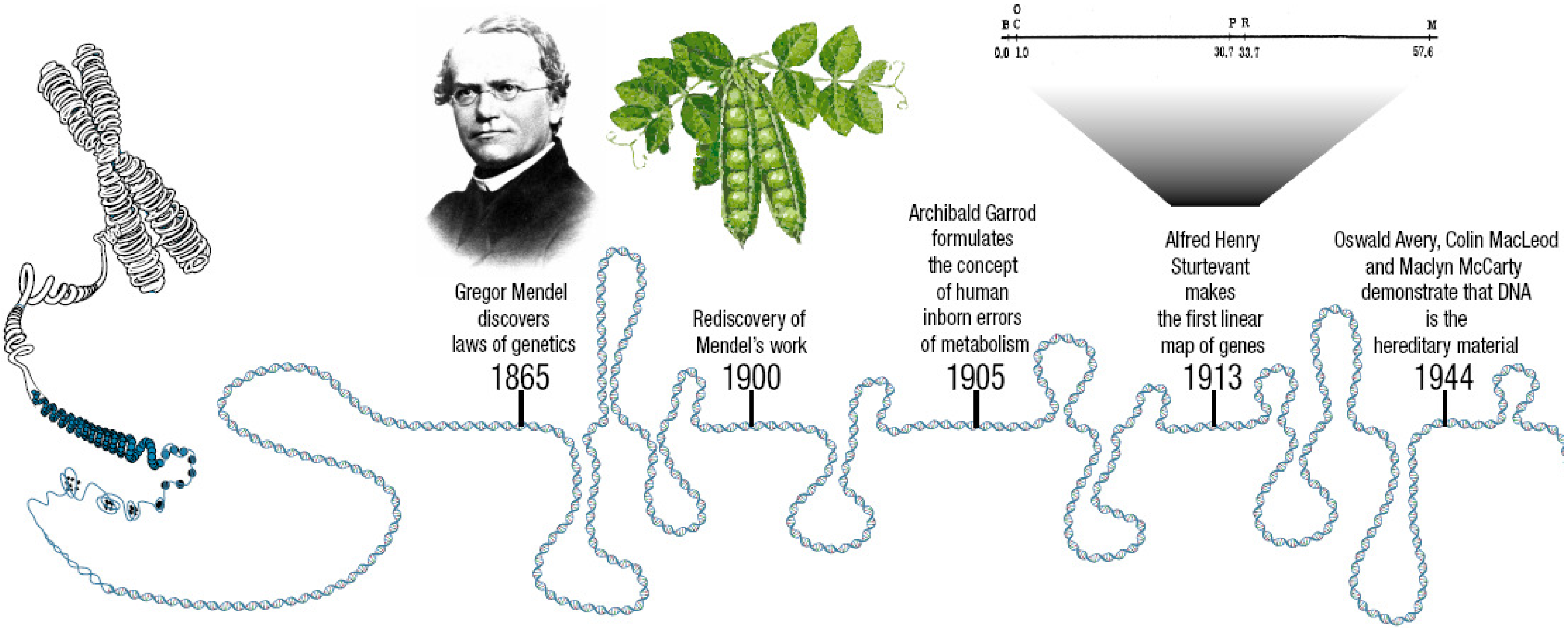


“The key to domestic breeding is man’s power of accumulative selection: nature gives successive variations; man adds them up in certain directions useful to him”

Charles Darwin
The Origin of Species
(1859)

De los genes a los genomas…

- Genes cómo concepto abstracto (Mendel, 1865)



Gregor Mendel discovers laws of genetics 1865

Rediscovery of Mendel's work 1900

Archibald Garrod formulates the concept of human inborn errors of metabolism 1905

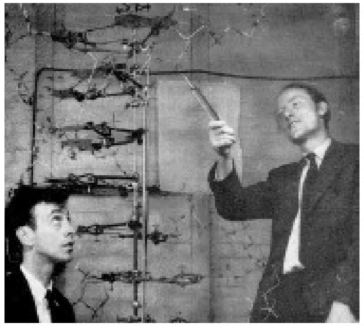
Alfred Henry Sturtevant makes the first linear map of genes 1913

Oswald Avery, Colin MacLeod and Maclyn McCarty demonstrate that DNA is the hereditary material 1944

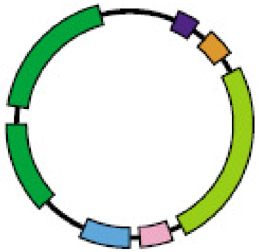
- Genes alineados en los cromosomas (Morgan, 1910 ; Sturtevant 1913)
- Genes en el ADN (Avery 1940s)

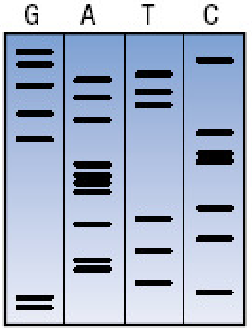
De los genes a los genomas…

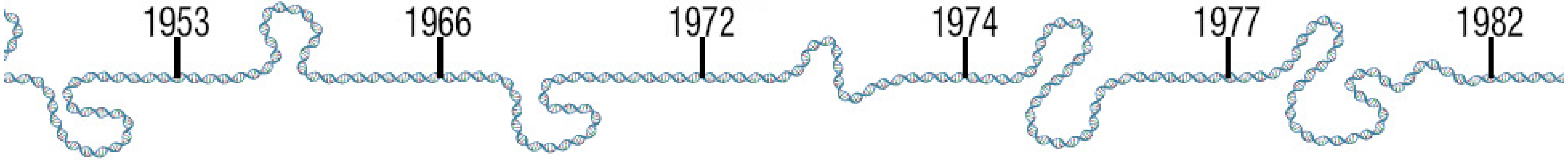
- **Estructura del ADN, Código Genético … (1950-60s)**



	U	C	A	G	
U	Phe Leu	Ser	Tyr stop stop	Cys stop Trp	U C A G
C	Leu	Pro	His Gln	Arg	U C A G
A	Ile Met	Thr	Asn Lys	Ser Arg	U C A G
G	Val	Ala	Asp Glu	Gly	U C A G







James Watson and Francis Crick describe the double-helical structure of DNA
1953

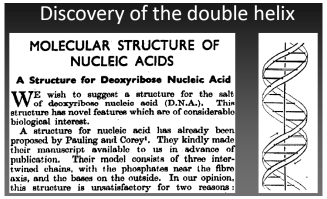
Marshall Nirenberg, Har Gobind Khorana and Robert Holley determine the genetic code
1966

Stanley Cohen and Herbert Boyer develop recombinant DNA technology
1972

The Belmont Report on the use of human subjects in research is issued
1974

Frederick Sanger, Allan Maxam and Walter Gilbert develop DNA-sequencing methods
1977

GenBank database established
1982



Discovery of the double helix

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

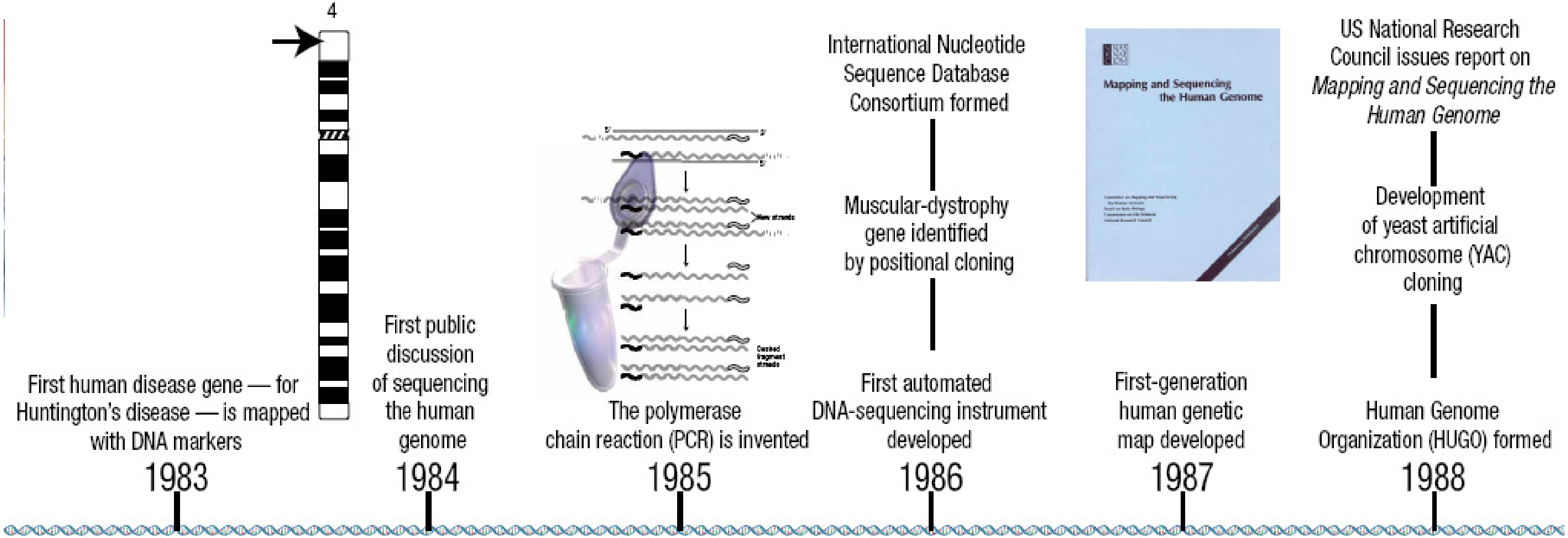
WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons:

- **Genes (secuencia, promotores, intrones… 1970-80s)**

De los genes a los genomas…

- **PCR (Kary Mullis, 1985)**



First human disease gene — for Huntington's disease — is mapped with DNA markers
1983

First public discussion of sequencing the human genome
1984

The polymerase chain reaction (PCR) is invented
1985

International Nucleotide Sequence Database Consortium formed

Muscular-dystrophy gene identified by positional cloning

First automated DNA-sequencing instrument developed
1986

First-generation human genetic map developed
1987

US National Research Council issues report on *Mapping and Sequencing the Human Genome*

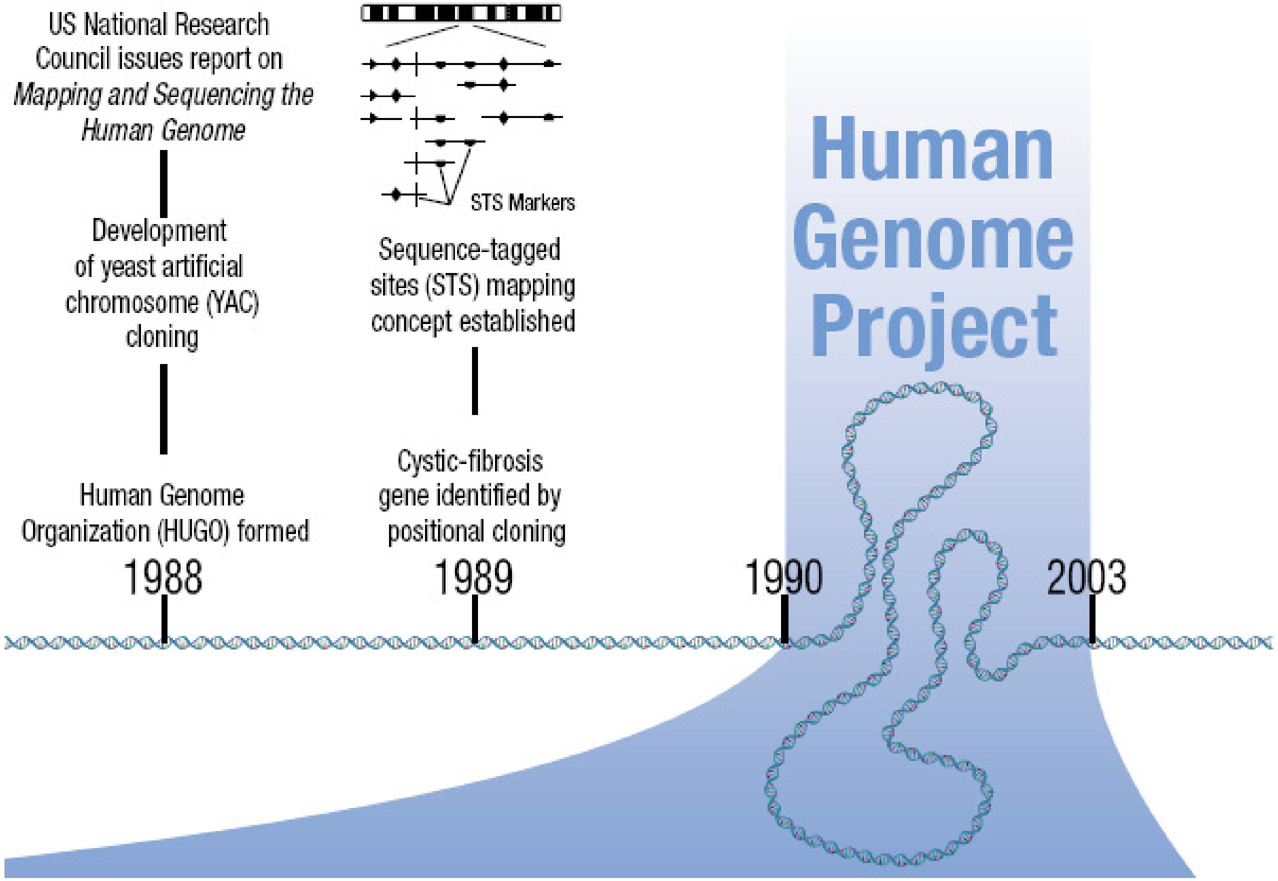
Development of yeast artificial chromosome (YAC) cloning

Human Genome Organization (HUGO) formed
1988

- **Secuenciación automática del ADN (Applied Biosystems, 1986)**

De los genes a los genomas…

- Genomas secuenciados por completo (1990s - …)



2005

…

Análisis masivo de genomas

- Análisis funcional multigénico (2000s)
- Secuenciación y genotipado masivos (2005 -…)

Universitat Autònoma de Barcelona

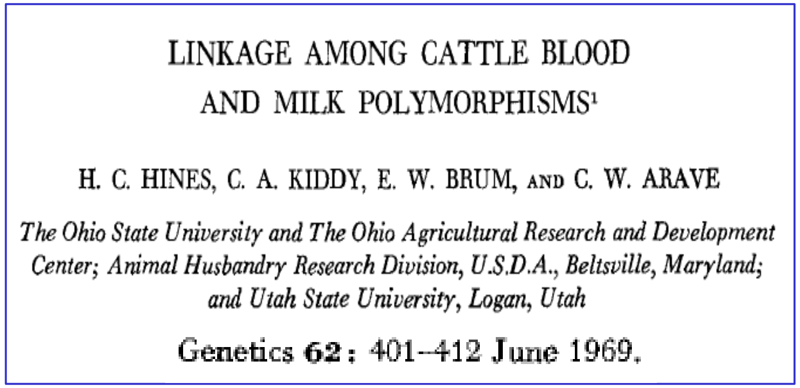
El estudio de la variabilidad genética

1900 - 1970	Cruzamientos y cromosomas
años 70	Electroforesis de proteínas y grupos sanguíneos
años 80	ADN mitocondrial
años 90	ADN nuclear
2000 - …	Genómica

Marcadores Genéticos

- **Proteínas e isoenzimas**
- **ADN mitocondrial**
- **Polimorfismos de fragmentos de restricción (RFLP)**
- **Minisatélites**
- **Secuenciación del ADN (Sanger)**
- **Microsatélites**
- **SNPs**
- **2ª generación de secuenciadores (HTS)**
- **3ª generación de secuenciadores (próximamente)**

Grupos sanguíneos y polimorfismos bioquímicos cómo marcadores genéticos (1970 – 1990)

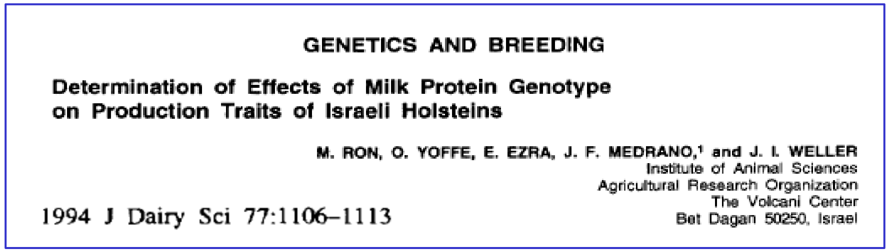


LINKAGE AMONG CATTLE BLOOD AND MILK POLYMORPHISMS[1]

H. C. HINES, C. A. KIDDY, E. W. BRUM, AND C. W. ARAVE

The Ohio State University and The Ohio Agricultural Research and Development Center; Animal Husbandry Research Division, U.S.D.A., Beltsville, Maryland; and Utah State University, Logan, Utah

Genetics **62**: 401–412 June 1969.

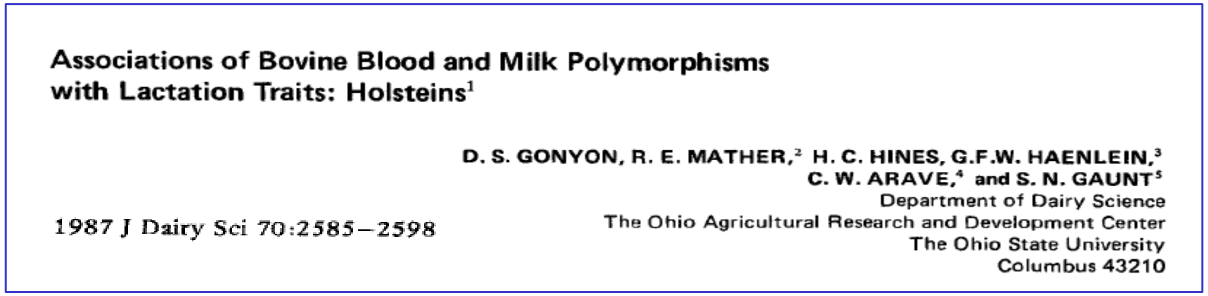


GENETICS AND BREEDING

Determination of Effects of Milk Protein Genotype on Production Traits of Israeli Holsteins

M. RON, O. YOFFE, E. EZRA, J. F. MEDRANO,[1] and J. I. WELLER
Institute of Animal Sciences
Agricultural Research Organization
The Volcani Center
Bet Dagan 50250, Israel

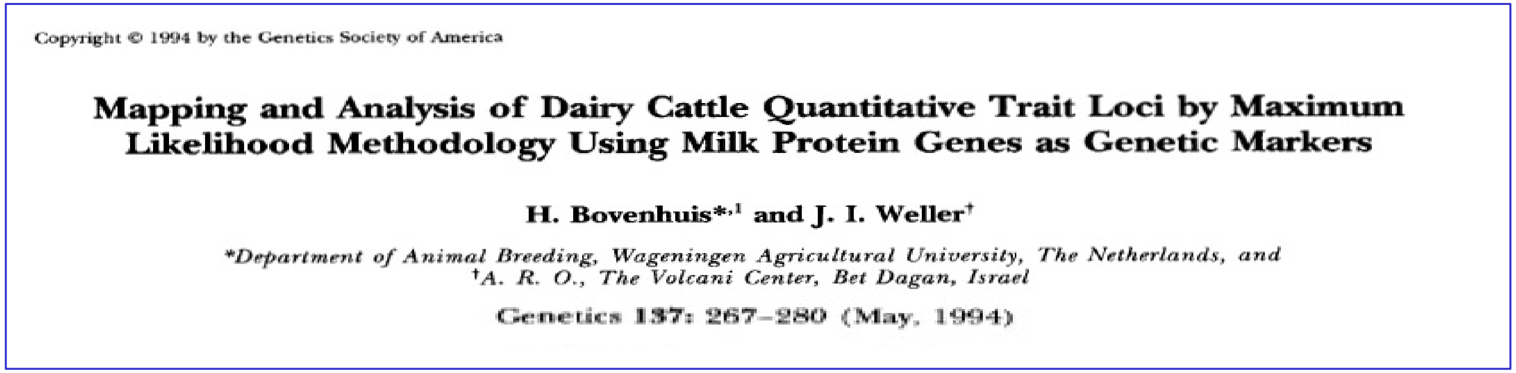
1994 J Dairy Sci 77:1106–1113



Associations of Bovine Blood and Milk Polymorphisms with Lactation Traits: Holsteins[1]

D. S. GONYON, R. E. MATHER,[2] H. C. HINES, G.F.W. HAENLEIN,[3] C. W. ARAVE,[4] and S. N. GAUNT[5]
Department of Dairy Science
The Ohio Agricultural Research and Development Center
The Ohio State University
Columbus 43210

1987 J Dairy Sci 70:2585–2598



Copyright © 1994 by the Genetics Society of America

Mapping and Analysis of Dairy Cattle Quantitative Trait Loci by Maximum Likelihood Methodology Using Milk Protein Genes as Genetic Markers

H. Bovenhuis*,[1] and J. I. Weller†

**Department of Animal Breeding, Wageningen Agricultural University, The Netherlands, and †A. R. O., The Volcani Center, Bet Dagan, Israel*

Genetics 137: 267–280 (May, 1994)

TIPOS DE POLIMORFISMOS EN EL DNA

DE REPETICIÓN→ **para un fragmento equivalente de ADN, el número de veces que se repite una secuencia núcleo no es idéntico.**

MINISATÉLITES

MICROSATÉLITES
Repeticiones en tándem de 2 a 5 bases.

CNVs **(variaciones en el número de copias): incluye inserciones y deleciones de diversa magnitud.**

DE SECUENCIA→ **para un fragmento equivalente de ADN, la secuencia de nucleótidos no es idéntica.**

SNP (***single nucleotide polymorphism***)

Variación que afecta a un nucleótido en la secuencia del ADN.

Los SNPs son los actuales marcadores de elección en genómica animal

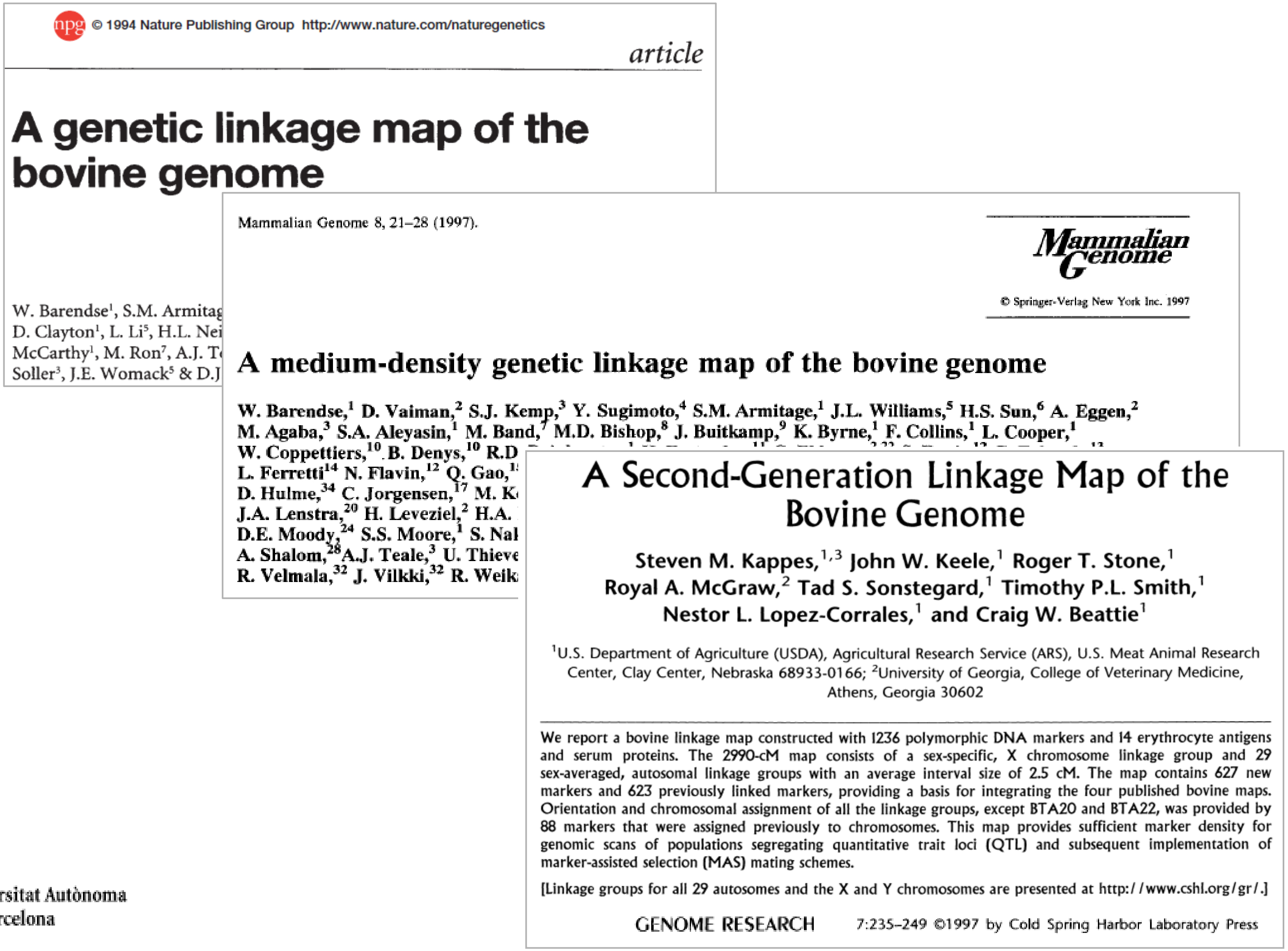
SNP: *Single Nucleotide Polymorphism*, un cambio de una base nucleotídica en la secuencia del ADN

- **Constituyen la fuente más abundante de variabilidad genética.**
- **Aproximadamente un SNP cada 300 – 500 bases.**
- **Millones de SNPs caracterizados en las principales especies domésticas y disponibilidad de chips para su análisis masivo.**
- **Fáciles de analizar**

..GAATCTTATGCTATACATAATTATATACTAATCGGGTATTGTTCTTAT..
..CTTAGAATACGATATGTATTAATATATGATTAGCCCATAACAAGAATA..

..GAATCTTATGCTATACATAATTATATACTAATAGGGTATTGTTCTTAT..
..CTTAGAATACGATATGTATTAATATATGATTATCCCATAACAAGAATA..

Microsatélites cómo marcadores para la construcción de mapas genéticos (1990 – 1997)



© 1994 Nature Publishing Group http://www.nature.com/naturegenetics

article

A genetic linkage map of the bovine genome

W. Barendse[1], S.M. Armitag
D. Clayton[1], L. Li[5], H.L. Nei
McCarthy[1], M. Ron[7], A.J. T
Soller[3], J.E. Womack[5] & D.J

Mammalian Genome 8, 21–28 (1997).

Mammalian Genome

© Springer-Verlag New York Inc. 1997

A medium-density genetic linkage map of the bovine genome

**W. Barendse,[1] D. Vaiman,[2] S.J. Kemp,[3] Y. Sugimoto,[4] S.M. Armitage,[1] J.L. Williams,[5] H.S. Sun,[6] A. Eggen,[2]
M. Agaba,[3] S.A. Aleyasin,[1] M. Band,[7] M.D. Bishop,[8] J. Buitkamp,[9] K. Byrne,[1] F. Collins,[1] L. Cooper,[1]
W. Coppettiers,[10] B. Denys,[10] R.D
L. Ferretti[14] N. Flavin,[12] Q. Gao,
D. Hulme,[34] C. Jorgensen,[17] M. K
J.A. Lenstra,[20] H. Leveziel,[2] H.A.
D.E. Moody,[24] S.S. Moore,[1] S. Nal
A. Shalom,[28] A.J. Teale,[3] U. Thieve
R. Velmala,[32] J. Vilkki,[32] R. Weik**

A Second-Generation Linkage Map of the Bovine Genome

**Steven M. Kappes,[1,3] John W. Keele,[1] Roger T. Stone,[1]
Royal A. McGraw,[2] Tad S. Sonstegard,[1] Timothy P.L. Smith,[1]
Nestor L. Lopez-Corrales,[1] and Craig W. Beattie[1]**

[1]U.S. Department of Agriculture (USDA), Agricultural Research Service (ARS), U.S. Meat Animal Research Center, Clay Center, Nebraska 68933-0166; [2]University of Georgia, College of Veterinary Medicine, Athens, Georgia 30602

We report a bovine linkage map constructed with 1236 polymorphic DNA markers and 14 erythrocyte antigens and serum proteins. The 2990-cM map consists of a sex-specific, X chromosome linkage group and 29 sex-averaged, autosomal linkage groups with an average interval size of 2.5 cM. The map contains 627 new markers and 623 previously linked markers, providing a basis for integrating the four published bovine maps. Orientation and chromosomal assignment of all the linkage groups, except BTA20 and BTA22, was provided by 88 markers that were assigned previously to chromosomes. This map provides sufficient marker density for genomic scans of populations segregating quantitative trait loci (QTL) and subsequent implementation of marker-assisted selection (MAS) mating schemes.

[Linkage groups for all 29 autosomes and the X and Y chromosomes are presented at http://www.cshl.org/gr/.]

GENOME RESEARCH 7:235–249 ©1997 by Cold Spring Harbor Laboratory Press

Universitat Autònoma de Barcelona

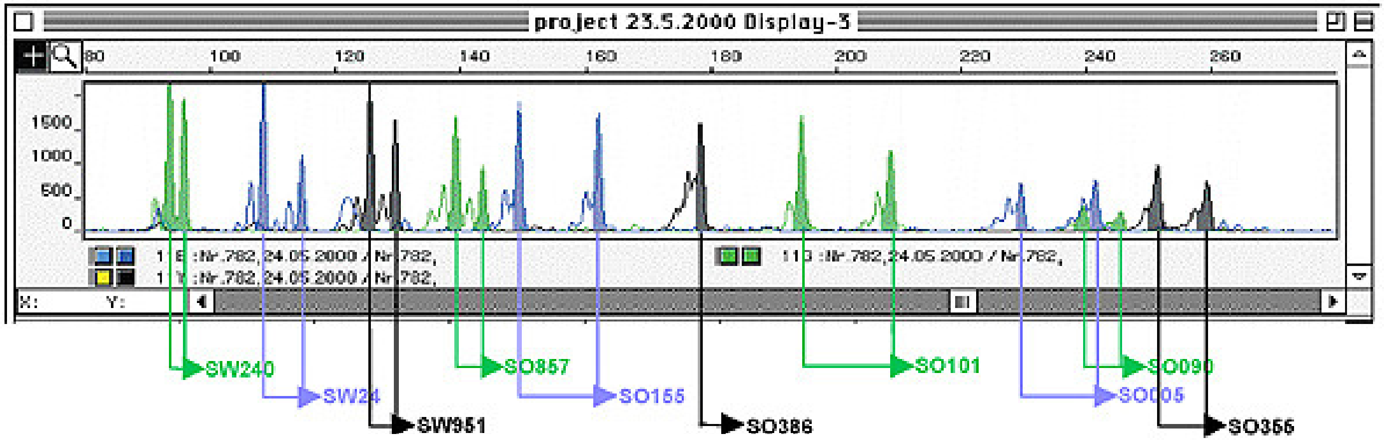
Estrategia molecular básica

DETECCIÓN DE QTLS PARA CARACTERES PRODUCTIVOS

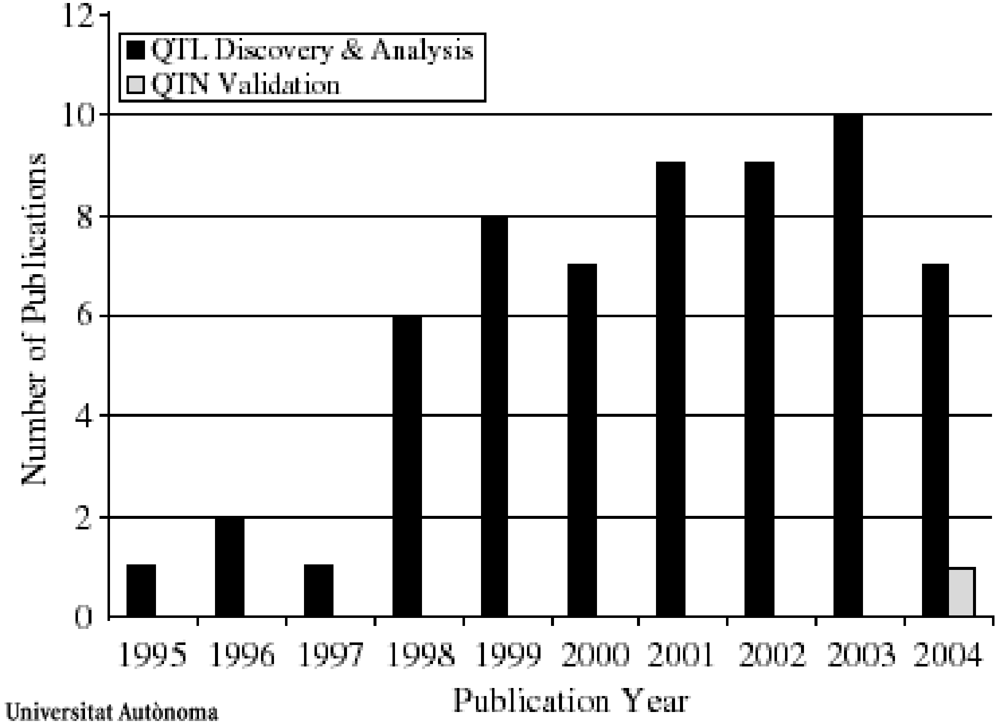
- **Usar marcadores para monitorizar la herencia y obtener probabilidades de los posibles genotipos y su localización en el genoma.**
- **Asociar genotipos con probabilidades para determinados fenotipos en cada familia.**

En su inicio (años 90) mediante marcadores del tipo microsatélite. Actualmente mediante SNPs

Detección de QTLs en especies domésticas



Microsatélites utilizados cómo marcadores para la detección de QTLs

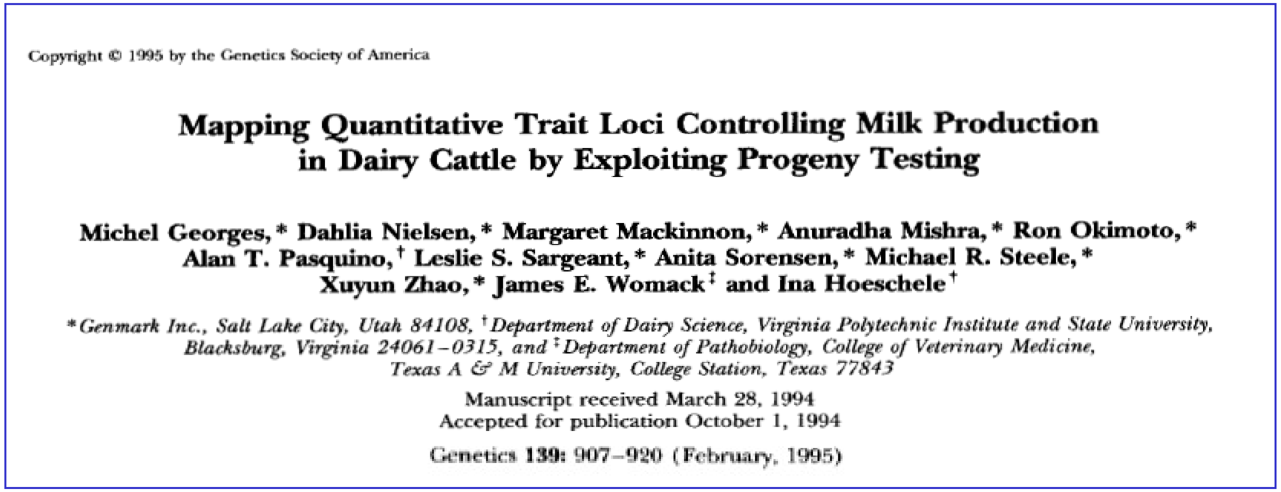


QTLs en el ganado bovino

(Sonstegard T.S. y C. P. Van Tassell, 2004)

Universitat Autònoma de Barcelona

Del QTL al QTN en el ganado bovino (1995 – 2002)



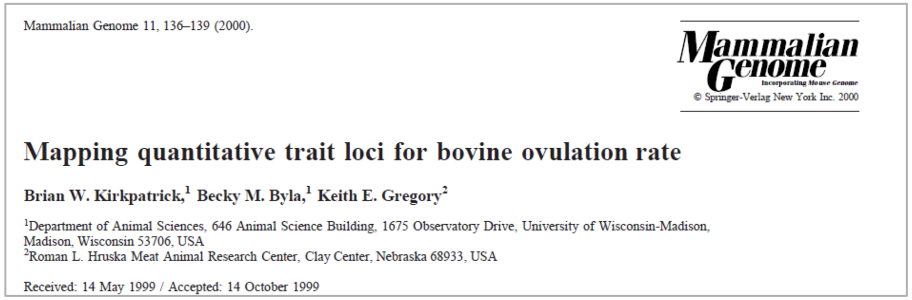
Mapping Quantitative Trait Loci Controlling Milk Production in Dairy Cattle by Exploiting Progeny Testing

Michel Georges,* Dahlia Nielsen,* Margaret Mackinnon,* Anuradha Mishra,* Ron Okimoto,*
Alan T. Pasquino,† Leslie S. Sargeant,* Anita Sorensen,* Michael R. Steele,*
Xuyun Zhao,* James E. Womack‡ and Ina Hoeschele†

**Genmark Inc., Salt Lake City, Utah 84108, †Department of Dairy Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061–0315, and ‡Department of Pathobiology, College of Veterinary Medicine, Texas A & M University, College Station, Texas 77843*

Manuscript received March 28, 1994
Accepted for publication October 1, 1994

Genetics **139:** 907–920 (February, 1995)



Mammalian Genome 11, 136–139 (2000).

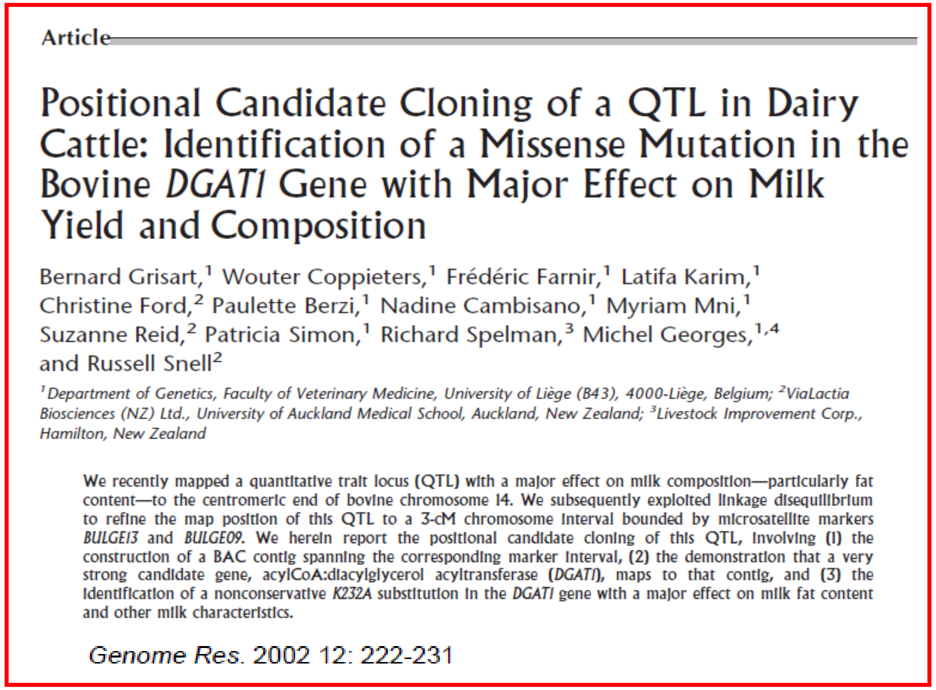
Mammalian Genome
Incorporating Mouse Genome
© Springer-Verlag New York Inc. 2000

Mapping quantitative trait loci for bovine ovulation rate

Brian W. Kirkpatrick,[1] Becky M. Byla,[1] Keith E. Gregory[2]

[1]Department of Animal Sciences, 646 Animal Science Building, 1675 Observatory Drive, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA
[2]Roman L. Hruska Meat Animal Research Center, Clay Center, Nebraska 68933, USA

Received: 14 May 1999 / Accepted: 14 October 1999



Article

Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine *DGAT1* Gene with Major Effect on Milk Yield and Composition

Bernard Grisart,[1] Wouter Coppieters,[1] Frédéric Farnir,[1] Latifa Karim,[1]
Christine Ford,[2] Paulette Berzi,[1] Nadine Cambisano,[1] Myriam Mni,[1]
Suzanne Reid,[2] Patricia Simon,[1] Richard Spelman,[3] Michel Georges,[1,4]
and Russell Snell[2]

[1]Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 4000-Liège, Belgium; [2]ViaLactia Biosciences (NZ) Ltd., University of Auckland Medical School, Auckland, New Zealand; [3]Livestock Improvement Corp., Hamilton, New Zealand

We recently mapped a quantitative trait locus (QTL) with a major effect on milk composition—particularly fat content—to the centromeric end of bovine chromosome 14. We subsequently exploited linkage disequilibrium to refine the map position of this QTL to a 3-cM chromosome interval bounded by microsatellite markers *BULGE13* and *BULGE09*. We herein report the positional candidate cloning of this QTL, involving (1) the construction of a BAC contig spanning the corresponding marker interval, (2) the demonstration that a very strong candidate gene, acylCoA:diacylglycerol acyltransferase (*DGAT1*), maps to that contig, and (3) the identification of a nonconservative *K232A* substitution in the *DGAT1* gene with a major effect on milk fat content and other milk characteristics.

Genome Res. 2002 12: 222-231

Importancia de las nuevas técnicas?

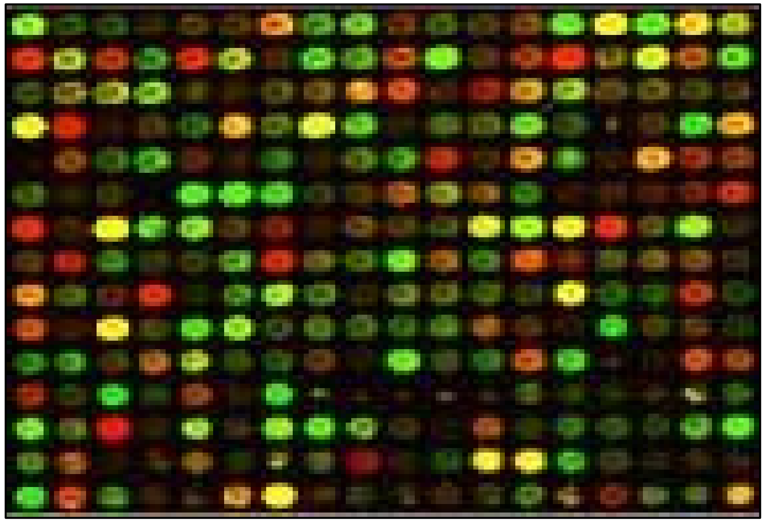
“Progress in science depends on new techniques, new discoveries, and new ideas, probably in that order.”

Sydney Brenner, Nature, June 5, 1980

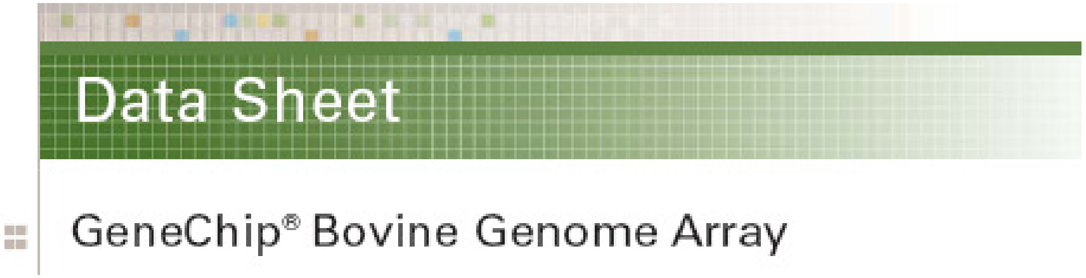
Genómica funcional mediante microarrays

Genes ↔ Funciones

Plataforma de Affymetrix



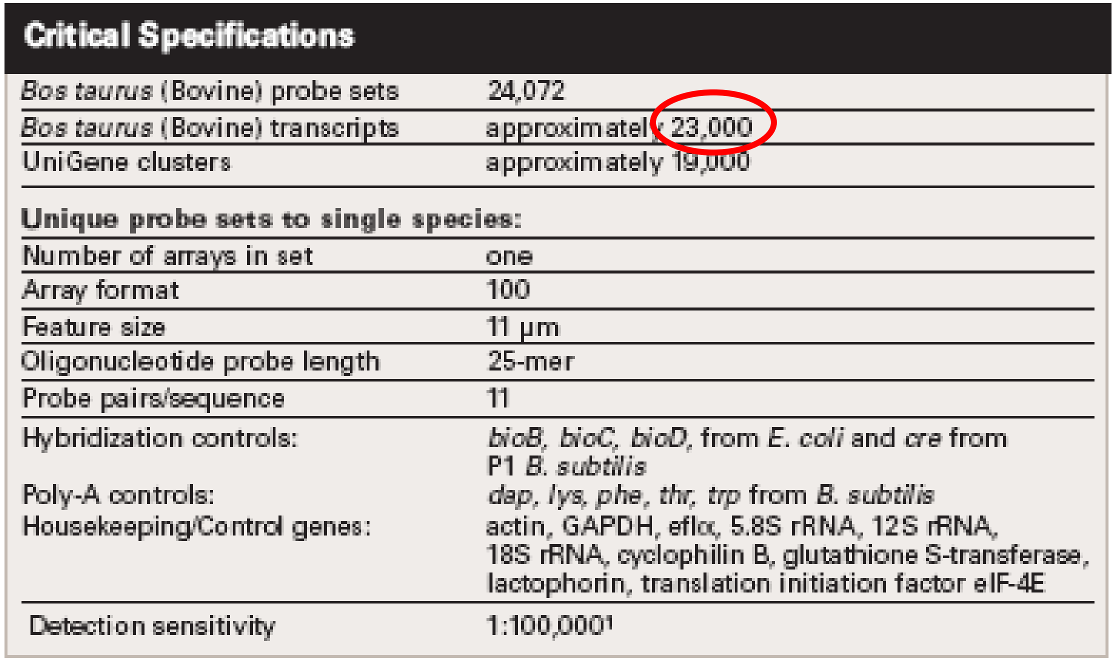
AFFYMETRIX®



Data Sheet

GeneChip® Bovine Genome Array

Permite analizar la expresión de unos 23.000 tránscritos



Critical Specifications

Bos taurus (Bovine) probe sets	24,072
Bos taurus (Bovine) transcripts	approximately 23,000
UniGene clusters	approximately 19,000
Unique probe sets to single species:	
Number of arrays in set	one
Array format	100
Feature size	11 µm
Oligonucleotide probe length	25-mer
Probe pairs/sequence	11
Hybridization controls:	*bioB, bioC, bioD*, from *E. coli* and *cre* from P1 *B. subtilis*
Poly-A controls:	*dap, lys, phe, thr, trp* from *B. subtilis*
Housekeeping/Control genes:	actin, GAPDH, ef1α, 5.8S rRNA, 12S rRNA, 18S rRNA, cyclophilin B, glutathione S-transferase, lactophorin, translation initiation factor eIF-4E
Detection sensitivity	1:100,000[1]

[1] As measured by detection in comparative analysis between a complex target containing spiked control transcriptions and a complex target with no spikes

Secuenciación del ADN

ABI

1986 / 2004

Roche 454

2005 /2007

Illumina

2006 / 2010

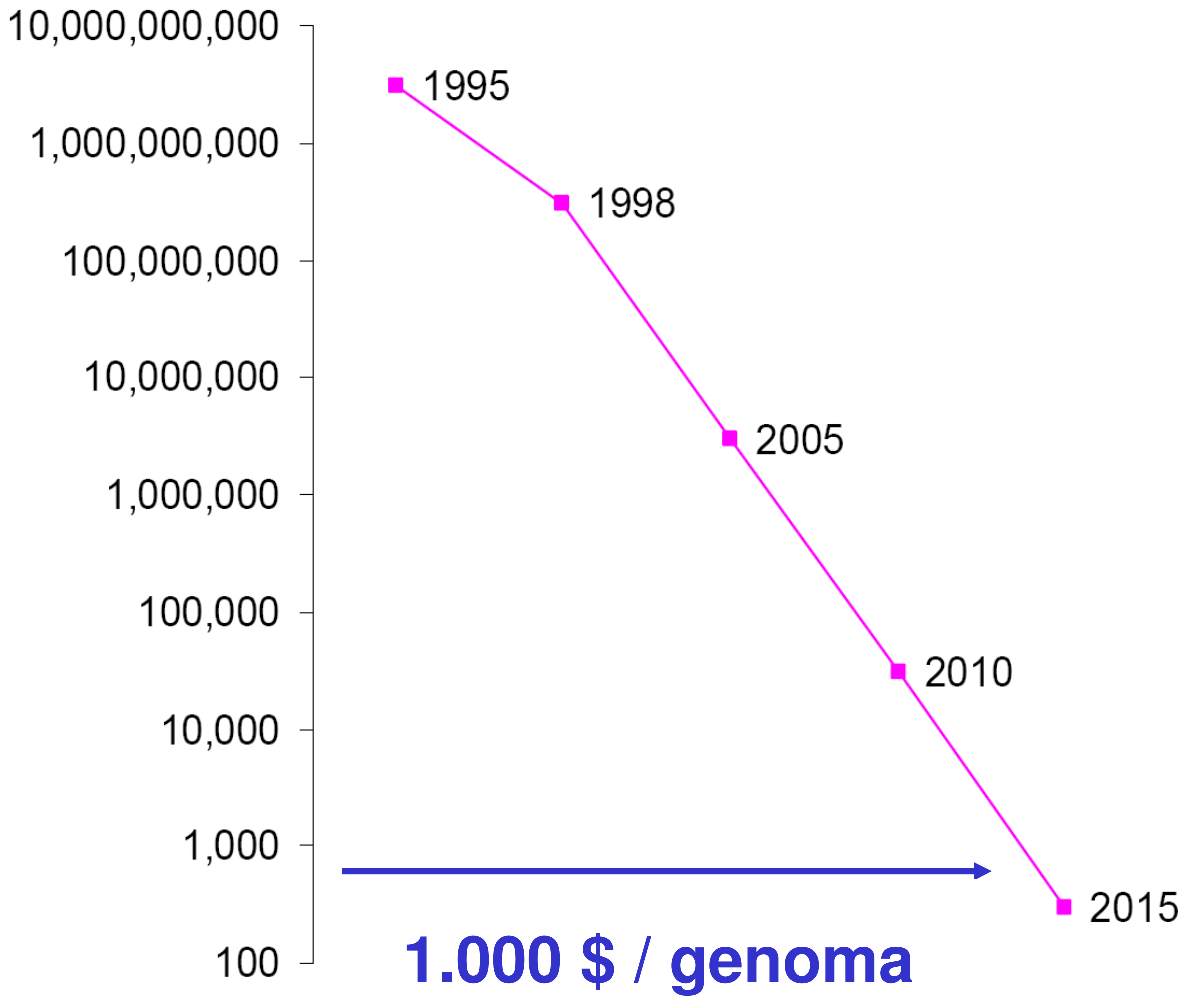
ABI SOLiD

2007

	Año
Humano (~ 3 x 10^9 bp) ~ $ 3.000 millones	2000
Bovino (~ 2,9 x 10^9 bp) ~ $ 53 million	2004/2009
Pollo (~ 1,2 x 10^9 bp) - $?	2004
Porcino (~ 2,9 x 10^9 bp) < $ 20 millones	2010 ?

Costes comparativos de los proyectos de secuenciación de algunos genomas

Coste de (re)secuenciar un genoma de mamífero



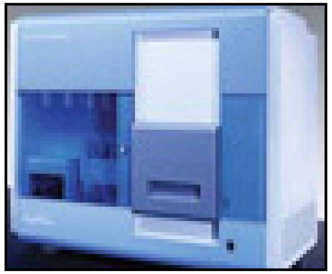
Costes de secuenciación

ABI 3730XL

Roche 454

2005

Illumina



2006

ABI SOLiD

2007

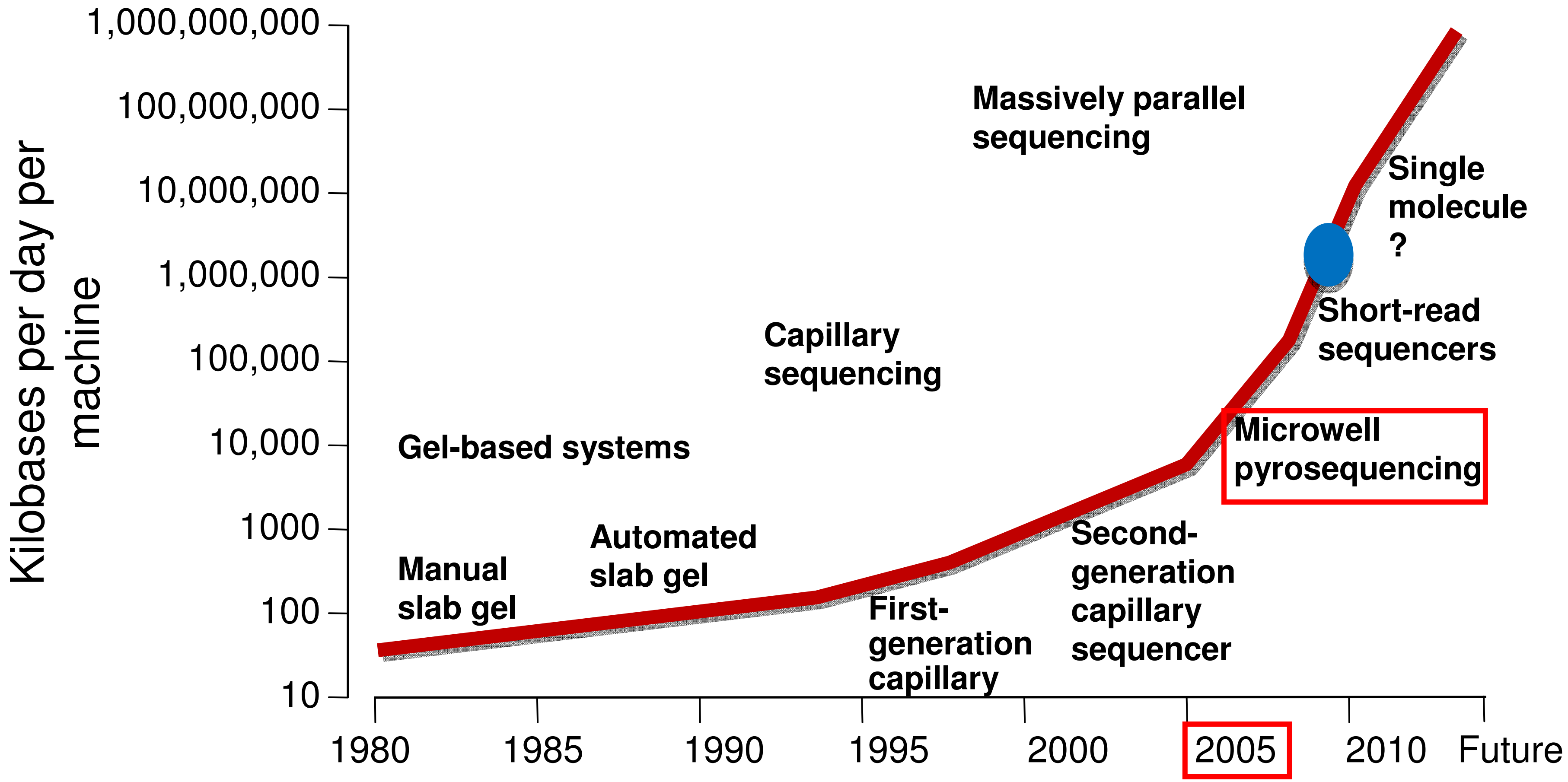
?

Coste de un genoma de mamífero: ~ 3 x 10^9 pb (cobertura 30x)

100.000 $ a finales del 2009

30.000 $ a finales del 2010

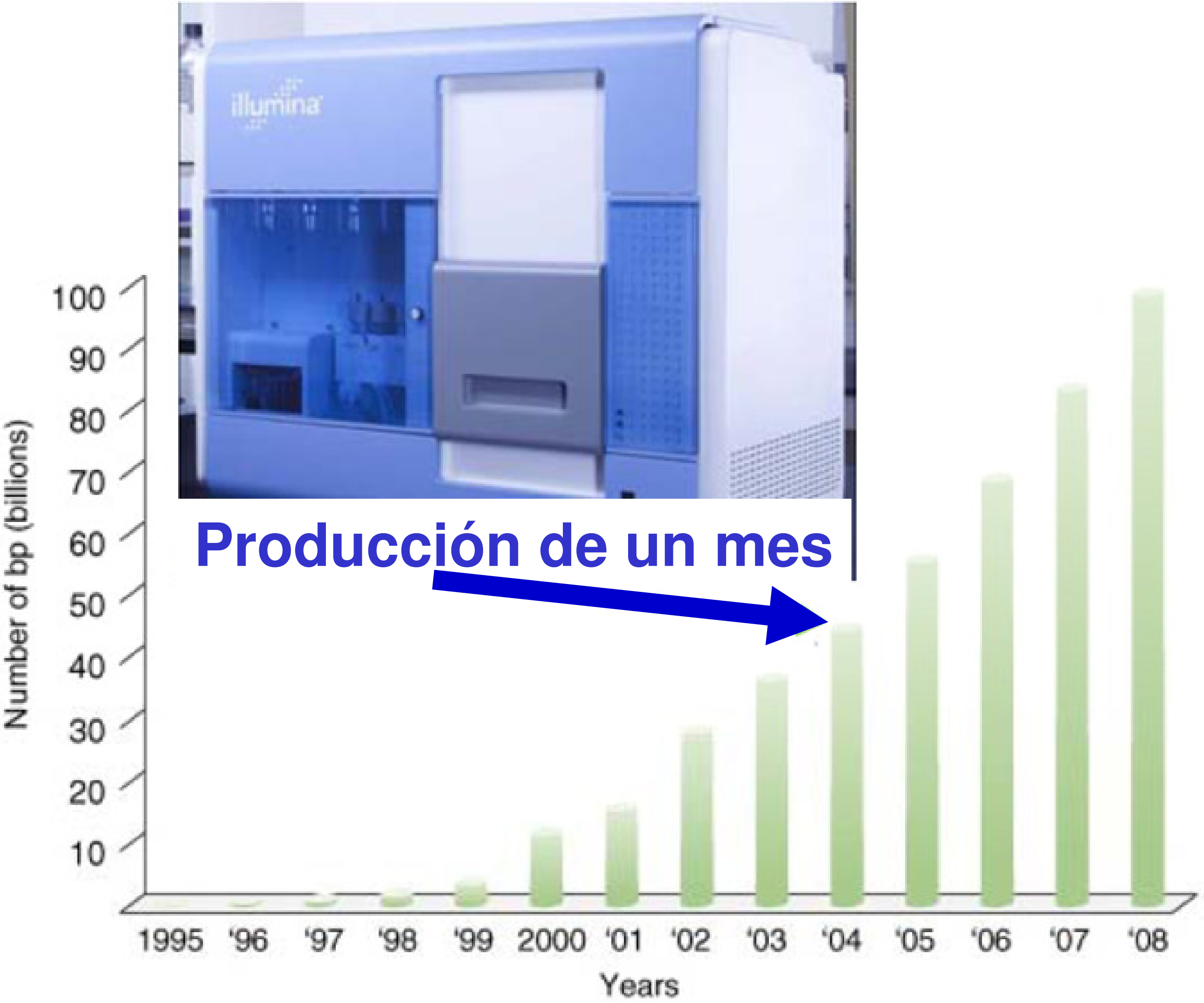
Secuenciación del ADN
(actualmente 1-2 Gb por maquina por día)



MR Stratton *et al. Nature* 458, 719-724 (2009).25

Universitat Autònoma de Barcelona

Rendimiento de los secuenciadores de segunda generación…



Producción de un mes

Secuencias de ADN (en pb) depositadas en GenBank desde 1995

Gilad et al.(2009) Trends Genet.

Secuenciación del genoma de especies domésticas

2004

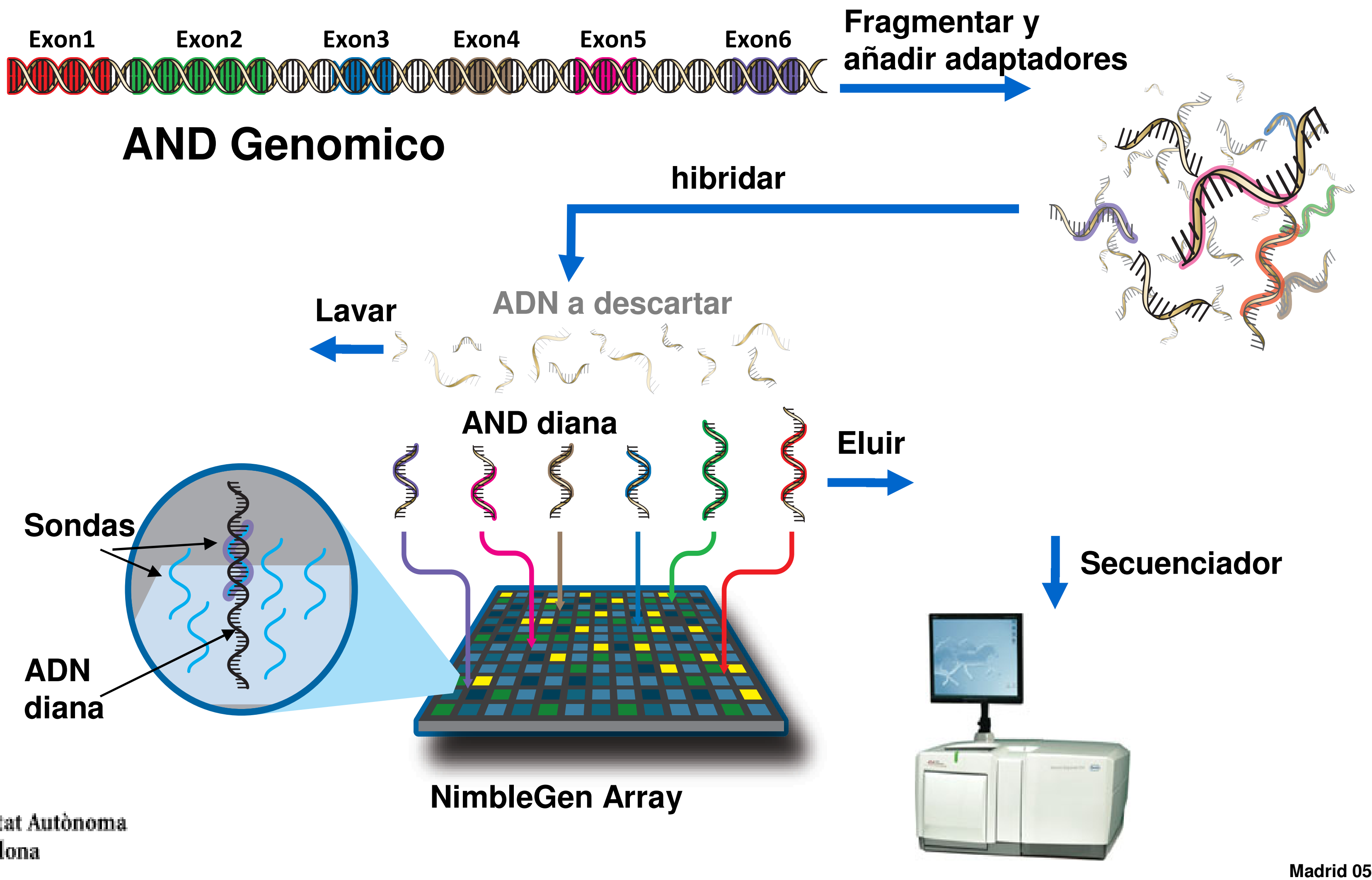
2005

2009

Cow	*Bos taurus*	Mammal	Draft Assembly (7x)	BCM-HGSC	Completed
Cat	*Felis catus*	Mammal	Draft Assembly	WUGSC	Completed
Chicken	*Gallus gallus*	Non-Mammalian Vertebrate	Draft Assembly (6,6x)	WUGSC	Completed
Dog	*Canis familiaris*	Mammal	Draft Assembly	BI/MIT	Completed
Horse	*Equus caballas*	Mammal	Draft Assembly (7x)	BI/MIT	Completed
Pig	Sus scrofa	Mammal	Draft Assembly (BAC to BAC)	Sanger	In Process
Rabbit	*Oryctolagus cuniculus*	Mammal	Low Coverage WGS (~2X)	BI/MIT	Completed
Sheep	*Ovis aries*	Mammal	Draft Assembly	BCM-HGSC	In progress

2010 ?

“Arrays de captura” para secuenciación



Exon1
Exon2
Exon3
Exon4
Exon5
Exon6

AND Genomico

Fragmentar y
añadir adaptadores

hibridar

ADN a descartar

Lavar

AND diana

Eluir

Secuenciador

Sondas

ADN
diana

NimbleGen Array

Caracterización de SNPs bovinos por secuenciación masiva

March 2008 | volume 5 | number 3

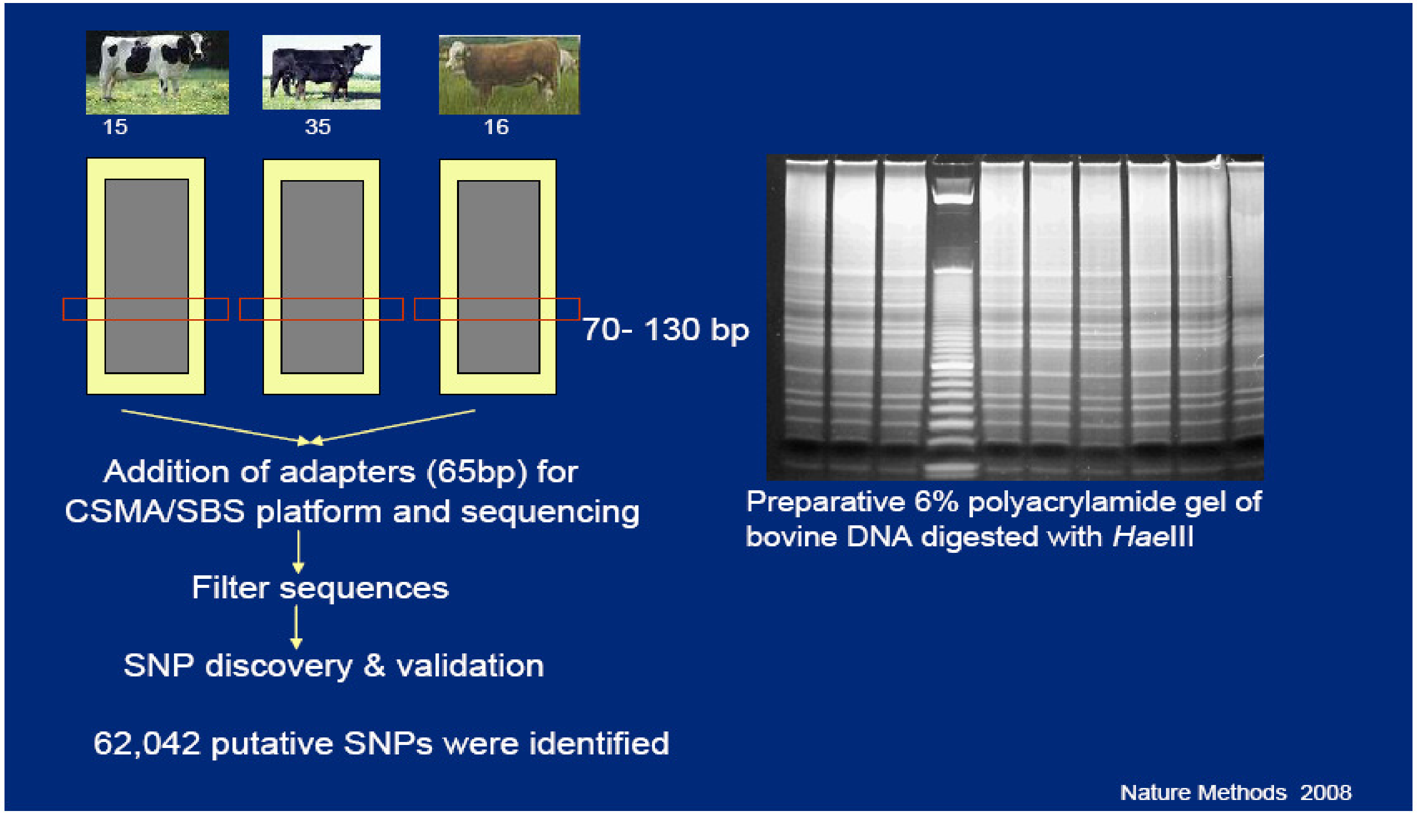
nature methods

www.nature.com/naturemethods

Techniques for life scientists and chemists

- Bovine SNP map
- Following stochastic tumor initiation
- High-content TF analysis
- A yeast genome library
- Resolving RNA dynamics with NMR

SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries.
Van Tasell et al. Nature Methods, 5:247-252 (2008)



Addition of adapters (65bp) for CSMA/SBS platform and sequencing

Filter sequences

SNP discovery & validation

62,042 putative SNPs were identified

Preparative 6% polyacrylamide gel of bovine DNA digested with *Hae*III

Nature Methods 2008

Estudios de asociación a nivel genómico (GWAS)

Utilización de chips de genotipado de SNPS

Actualmente capaces de analizar hasta un millon de SNPs en una muestra

illumina

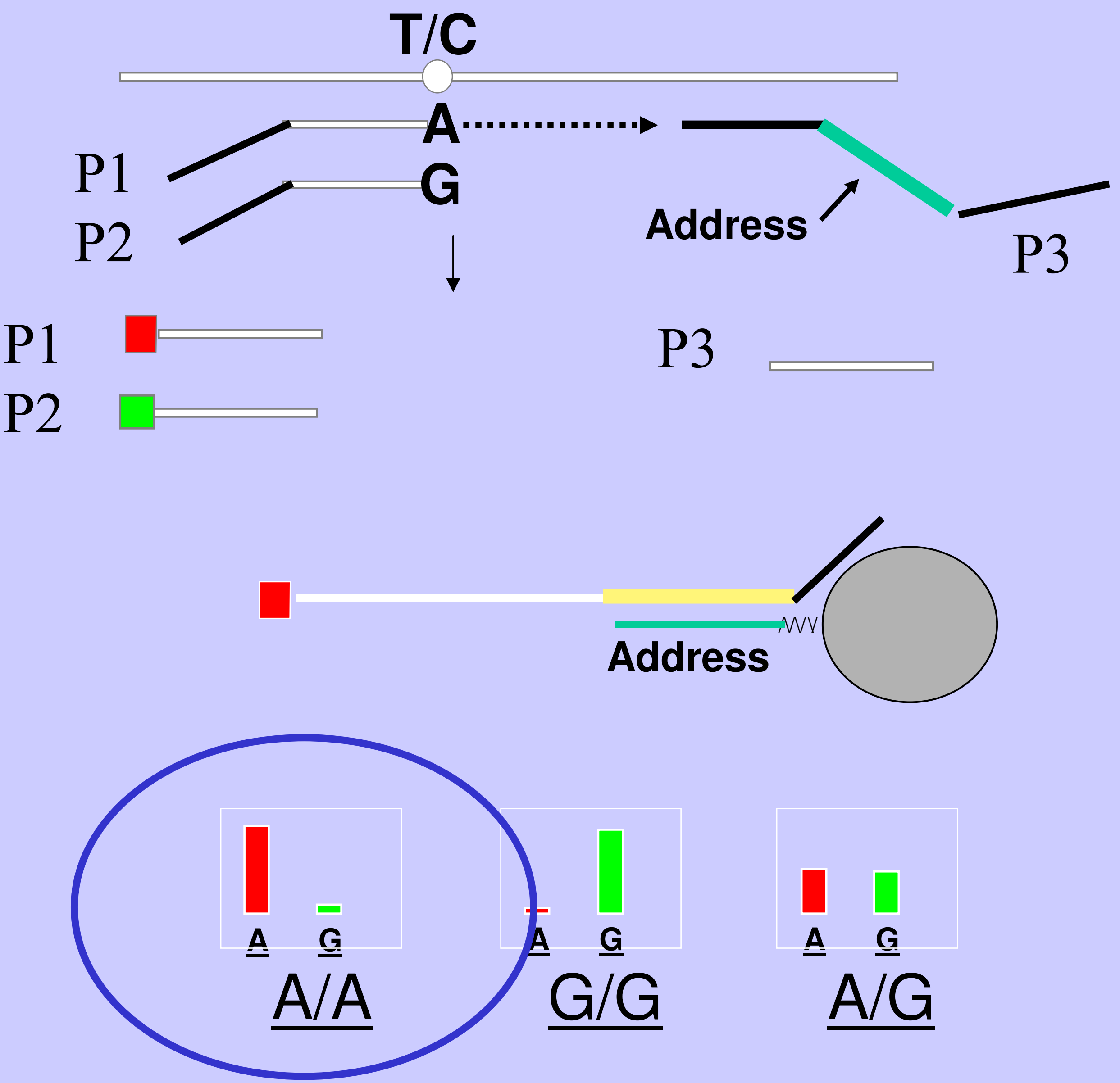
Infinium Assay

"SNP chips"

AFFYMETRIX

GeneChip Array

Illumina BeadArray (Golden Gate)



Allele Specific Extension, Ligation

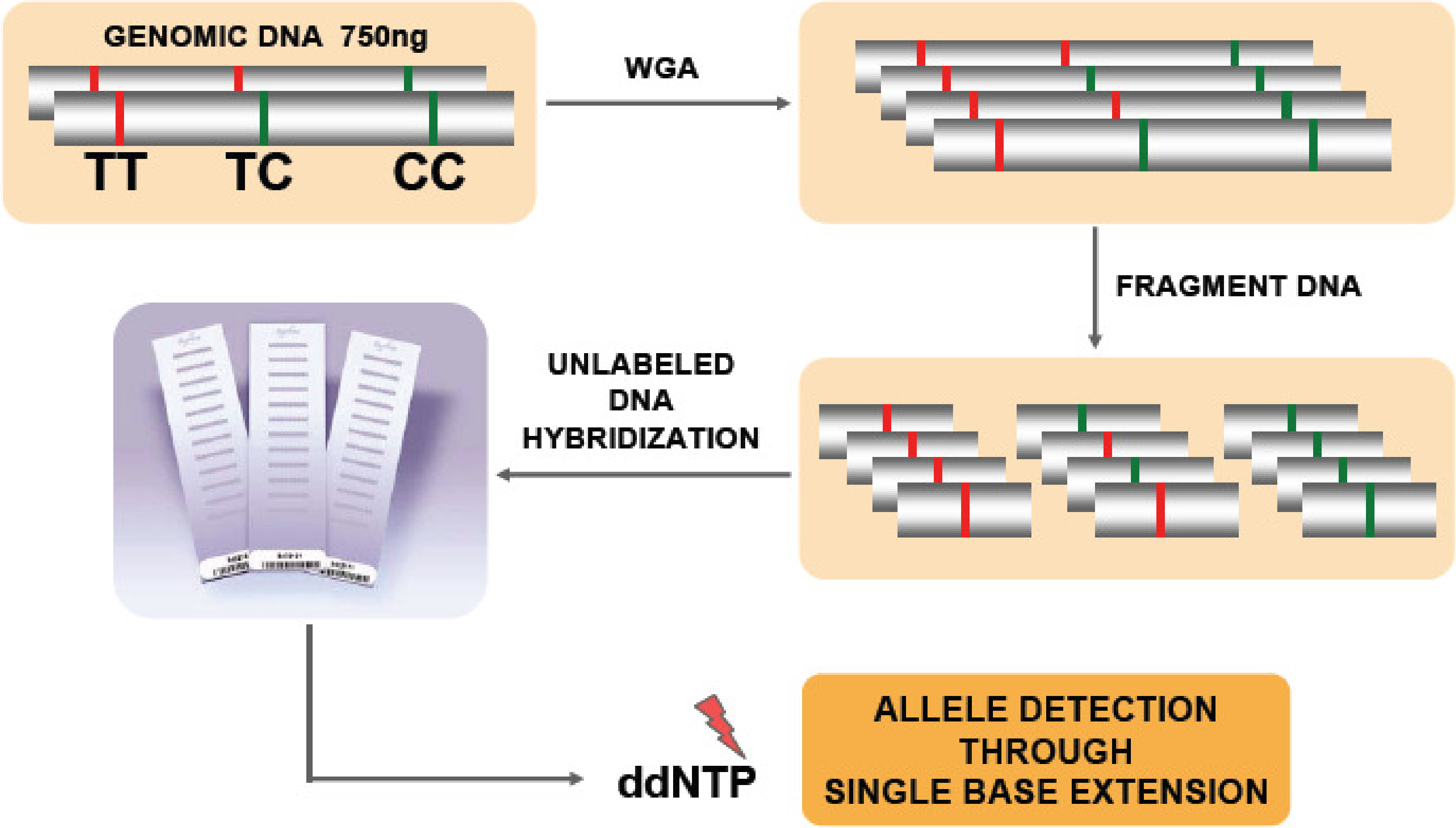
PCR with common primers

Product capture by hybridization to array

Readout

Illumina Infinium II SNP genotyping

From 7K to 1M SNPs – One Assay - Infinium



GENOMIC DNA 750ng

TT TC CC

WGA

FRAGMENT DNA

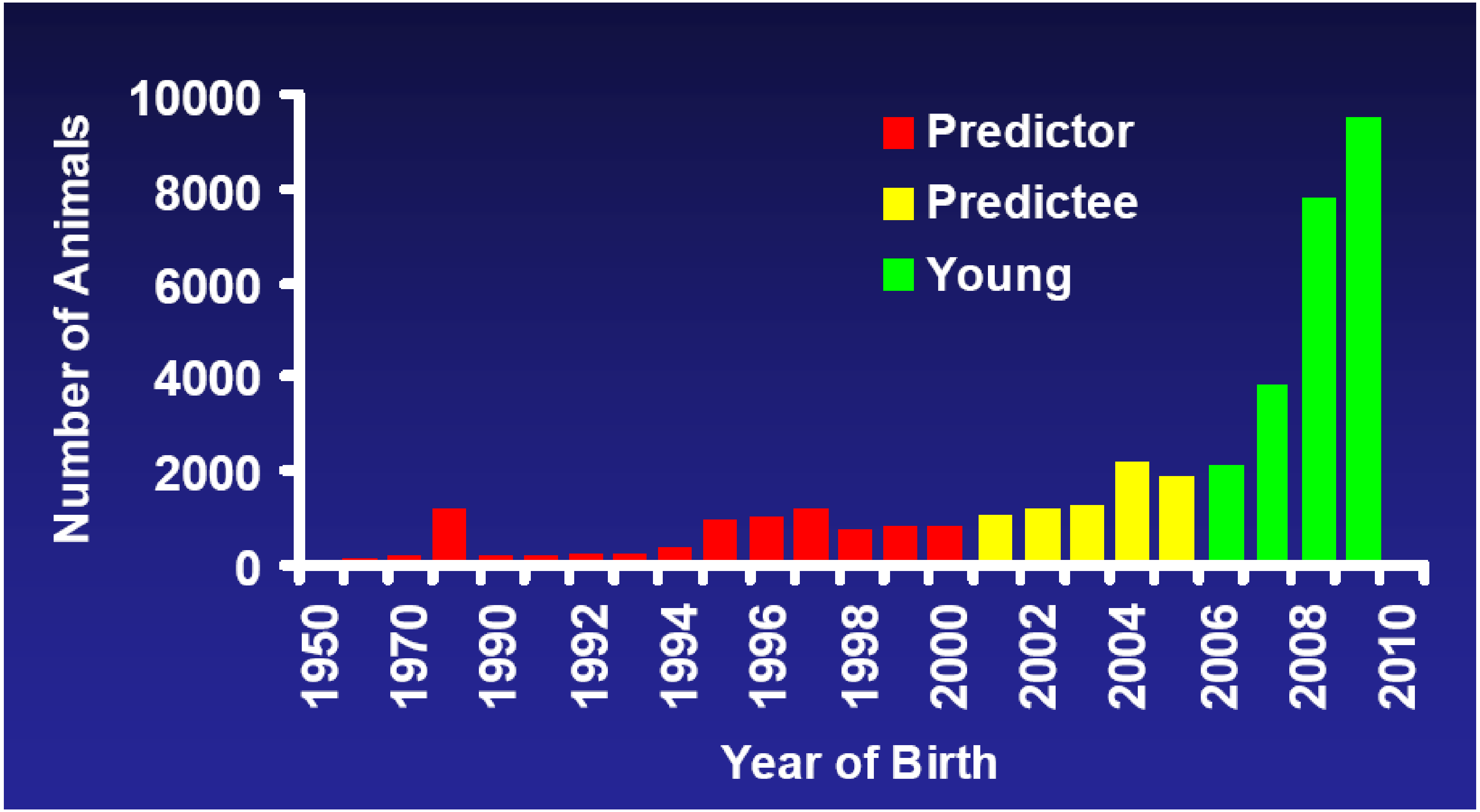
UNLABELED DNA HYBRIDIZATION

ddNTP

ALLELE DETECTION THROUGH SINGLE BASE EXTENSION

Animales Genotipados con el chip 50K (Illumina) en Norteamérica (Febrero 2010)

n=42.653



Council on Dairy Cattle Breeding, April 2010 (4)

Paul VanRaden

Madrid 05/05/10

Características del chip de genotipado 50K

Illumina SNP50 BeadChip	58,336
Insufficient number of beads	1,389
Unscorable SNP	4,360
Monomorphic in Holsteins	5,734
Minor allele frequency <5%	6,145
Not in H-W equilibrium	282
Highly correlated	2,010
Used for genomic prediction	38,416

Van Raden P. *Gordon Conference on Quantitative Genetics*. (2009)

Nuevo chip de genotipado bovino de alta densidad (700K ?)
Bovine HD (Illumina)

- Whole genome research array to complement the current BovineSNP50 for the identification of agriculturally important genes and for a deeper implementation of genomic selection in cattle
  - Developed in collaboration
    - USDA-ARS
    - Pfizer Animal Genetics
    - University of Missouri
    - UNCEIA, INRA
    - Other high-profile agricultural organizations

BovineHD : Groups, institutions and companies involved in the development

USA
- USDA
- George Mason University
- University of Missouri
- University of Illinois
- Pfizer

Brasil
- UNESP, Aracatuba
- Embrapa, Brasil

France
- UNCEIA-INRA

Netherlands
- CRV

Republic of Korea
- Chungbuk National University
- Yeungnam University, Gyeongsan

Danemark
- University of Aarhus, Danemark

Germany
- Technical University of Munich

Italy
- University of Piacenza
- University of Milano

UK
- Roslin Institute

Interbull Technical Workshop
Use of Genomic Information in Genetic evaluations
Paris, March 4-5, 2010

A. Eggen (2010)

Nuevo chip de genotipado bovino de alta densidad (700K ?)
Bovine HD (Illumina)

- Representing more than **20 diverse breeds**
  - *Bos taurus taurus, Bos taurus indicus,* Tropically adapted *Bos taurus taurus*
- Average sequence read depth >100x from over 120 animals
  - >50 million putative SNPs discovered
  - SNPs will be submitted to dbSNP
- Built from an algorithm developed for the BovineSNP50
  - Additions to algorithm include: breed weighting, breed-specific gaps and MAF, differential spacing based on region of genome (i.e. exon, repetitive, segmental duplication/CNV, other)
- Commercial launch is anticipated before the end of Q2 2010
- The BovineHD BeadChip will be on an 8-sample format

Interbull Technical Workshop
Use of Genomic Information in Genetic evaluations
Paris, March 4-5, 2010

A. Eggen (2010)

Nuevo chip de genotipado bovino de alta densidad (700K ?)
Bovine HD (Illumina)

- **All of BovineSNP50 SNPs** will be attempted
- ~800,000 bead types after 60,800 beads of BovineSNP50 positioned
- 795,000 SNP positioned on BTA1-29,X
- 5,000 beads represent unknown contigs, BTA Y, and mitochondrial SNP
- Breed groups used:
  - Holstein, Angus, Nelore, Taurine dairy, Taurine beef, Indicine, tropically adapted Taurine
- 852,645 total gaps
  - 850,816 <20kb
  - 1795 >20kb, < 100kb
  - 34 > 100 kb

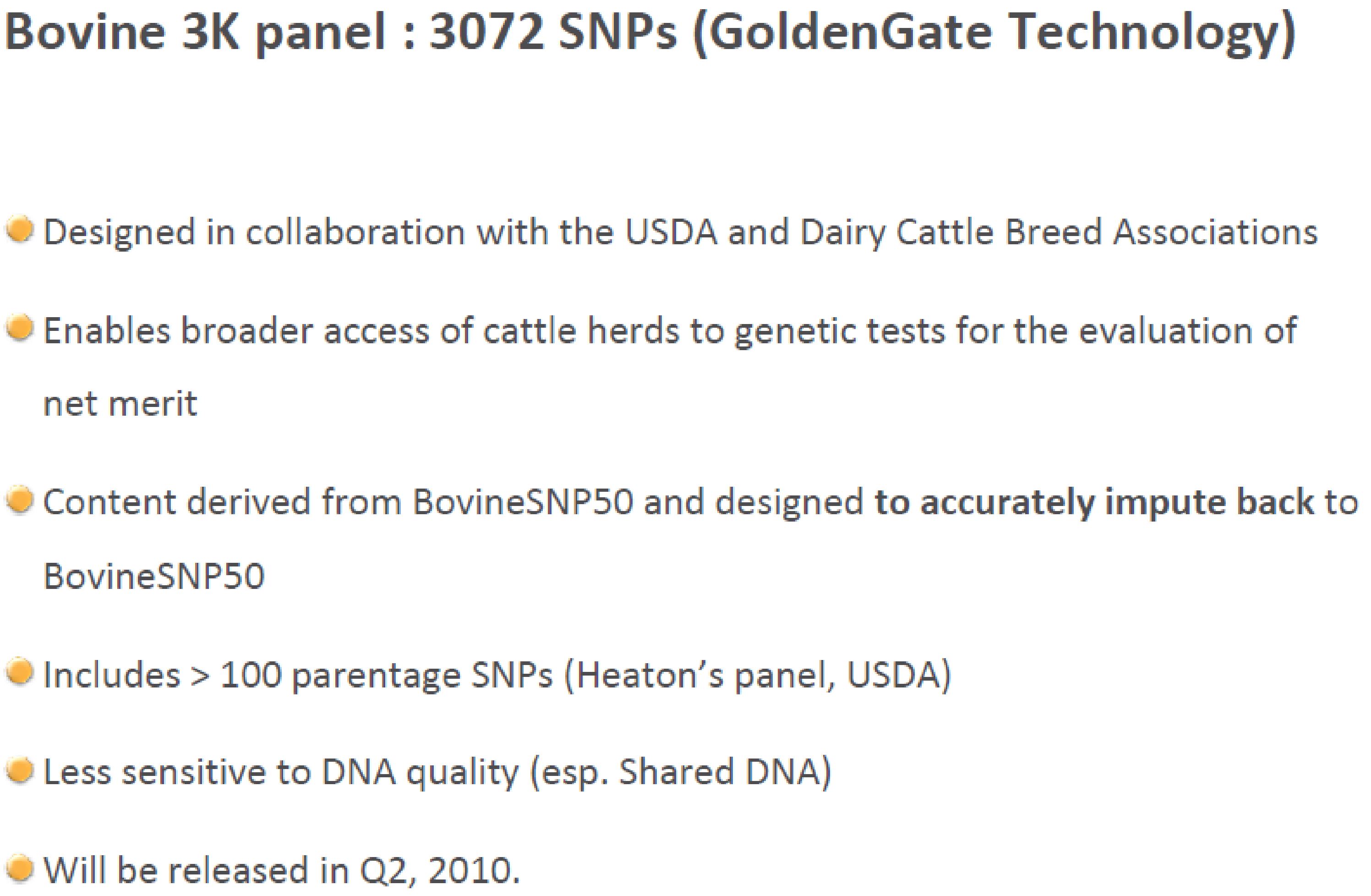
Interbull Technical Workshop
Use of Genomic Information in Genetic evaluations
Paris, March 4-5, 2010

A. Eggen (2010)

Hitos recientes en Genómica Bovina

	Año
Microsatélites (inicio de tests con ADN)	1992
Primera detección de QTLs	1995
Primer QTN (DGAT1) caracterizado	2002
Borrador de secuencia del genoma bovino	2004
58,000 SNPs seleccionados	2007
Chips de genotipado (10k, 25K Affymetrix; 50K Illumina) en el mercado	2007
Predicciones genómicas preliminares	Abril 2008
Predicciones genómicas oficiales	Enero 2009
Bovine Genome sequence completed	Abril 2009
Nuevos chips (700K i 3K) anunciados	2010

Nuevo panel reducido de SNPs para genotipado en el ganado bovino



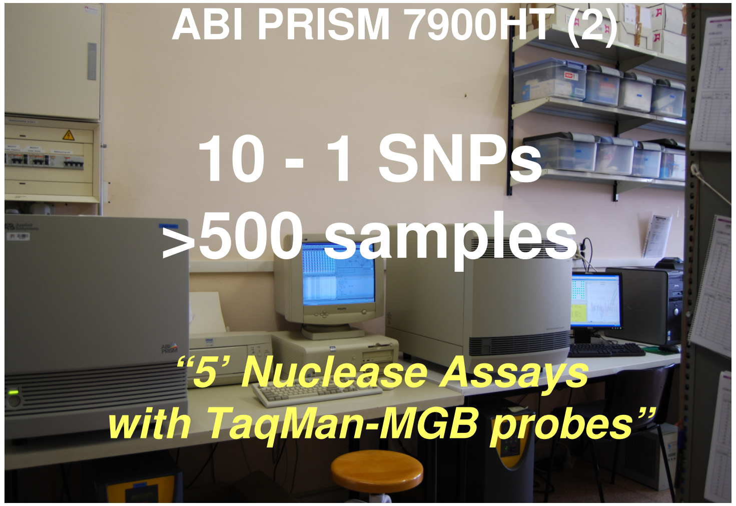
Bovine 3K panel : 3072 SNPs (GoldenGate Technology)

- Designed in collaboration with the USDA and Dairy Cattle Breed Associations
- Enables broader access of cattle herds to genetic tests for the evaluation of net merit
- Content derived from BovineSNP50 and designed **to accurately impute back** to BovineSNP50
- Includes > 100 parentage SNPs (Heaton's panel, USDA)
- Less sensitive to DNA quality (esp. Shared DNA)
- Will be released in Q2, 2010.

Interbull Technical Workshop
Use of Genomic Information in Genetic evaluations
Paris, March 4-5, 2010

A. Eggen (2010)

Plataformas de genotipado en el SVGM (UAB)


ABI PRISM 7900HT (2)

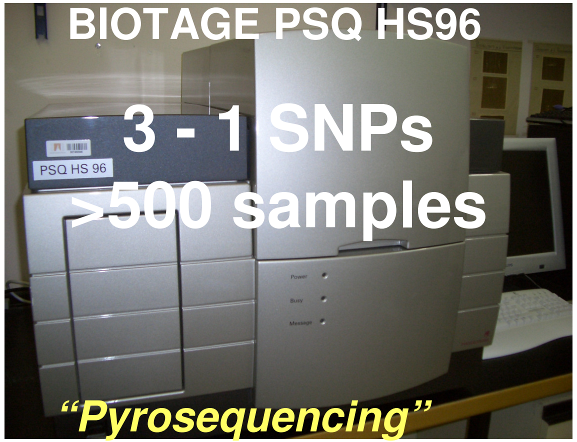
10 - 1 SNPs
>500 samples

"5' Nuclease Assays with TaqMan-MGB probes"


ABI 3100 / ABI 3730

10 - 1 SNPs
< 500 samples

"Primer Extension (SnapShot)"
"SNPlex"



BIOTAGE PSQ HS96

3 - 1 SNPs
>500 samples

"Pyrosequencing"


Illumina BeadXpress

48 - 1360 SNPs

"VeraCode Golden Gate"


Illumina BeadStation

10^6 - 1360 SNPs

"Golden Gate Assay"
"Infinium Assay"

Pruebas analíticas

Determinación de patógenos

Patologías hereditarias

Genes de interés productivo

Sexado de aves

Identificación individual y control de filiación. Trazabilidad.

Identificación de especies animales

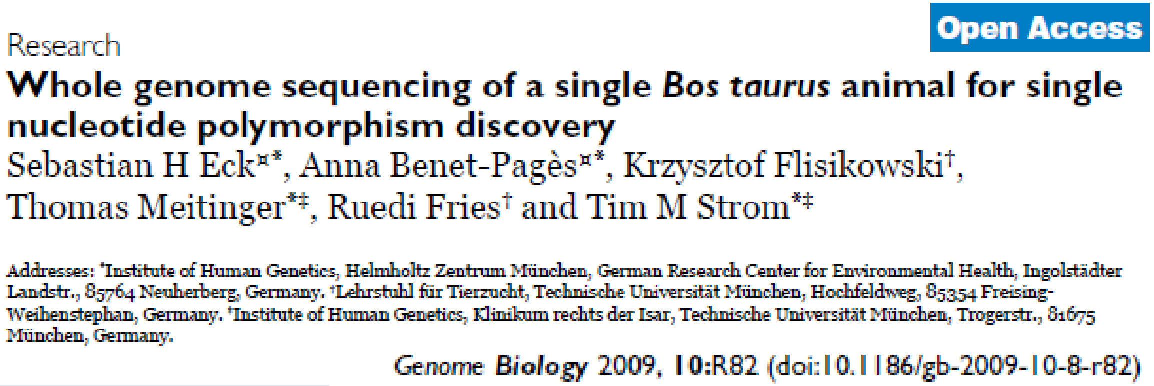
Origen y contaminación en productos derivados

Identificación de microorganismos

http://svgm.uab.es

Consorcio CSIC – IRTA - UAB

Secuenciar el genoma de un animal ya empieza a ser técnicamente factible… aunque aún demasiado costoso !



Research

Open Access

Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery

Sebastian H Eck[¤,*], Anna Benet-Pagès[¤,*], Krzysztof Flisikowski[†],
Thomas Meitinger[*,‡], Ruedi Fries[†] and Tim M Strom[*,‡]

Addresses: [*]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr., 85764 Neuherberg, Germany. [†]Lehrstuhl für Tierzucht, Technische Universität München, Hochfeldweg, 85354 Freising-Weihenstephan, Germany. [‡]Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Trogerstr., 81675 München, Germany.

Genome ***Biology*** 2009, **10:**R82 (doi:10.1186/gb-2009-10-8-r82)

- **24 Gb of sequence**
- **7.4-fold coverage**
- **2.44 millions SNPs**
- **115’000 indels**

HiSeq2000 launched

in January 2010

- 200 Gb per run or over 2 bovine genomes at a quality allowing the detection of single base pair differences (to the SNP level)
- 25 Gb/day; 7-8 days for 2 x 100 bp

A. Eggen (2010)

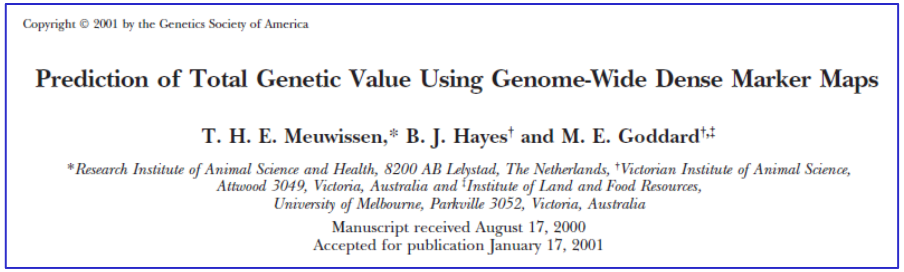
La (r)evolución genómica en mejora animal

- Secuenciación del genoma en especies domésticas

- Paneles masivos para el genotipado de SNPs

HACIA UN CAMBIO DE PARADIGMA ?

De la búsqueda del mejor animal a la búsqueda del mejor Gen(es) / Genoma

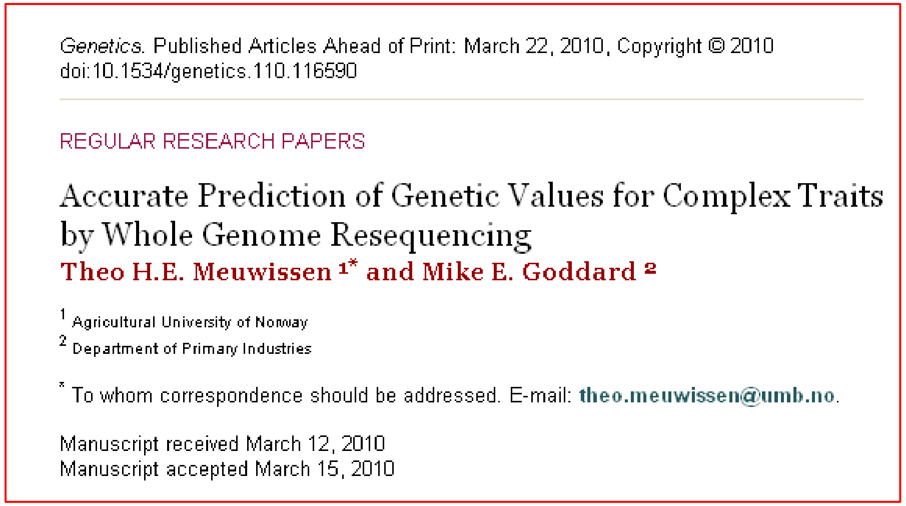


Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,* B. J. Hayes[†] and M. E. Goddard[†,‡]

*Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science,
Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources,
University of Melbourne, Parkville 3052, Victoria, Australia

Manuscript received August 17, 2000
Accepted for publication January 17, 2001



Genetics. Published Articles Ahead of Print: March 22, 2010, Copyright © 2010
doi:10.1534/genetics.110.116590

REGULAR RESEARCH PAPERS

Accurate Prediction of Genetic Values for Complex Traits by Whole Genome Resequencing

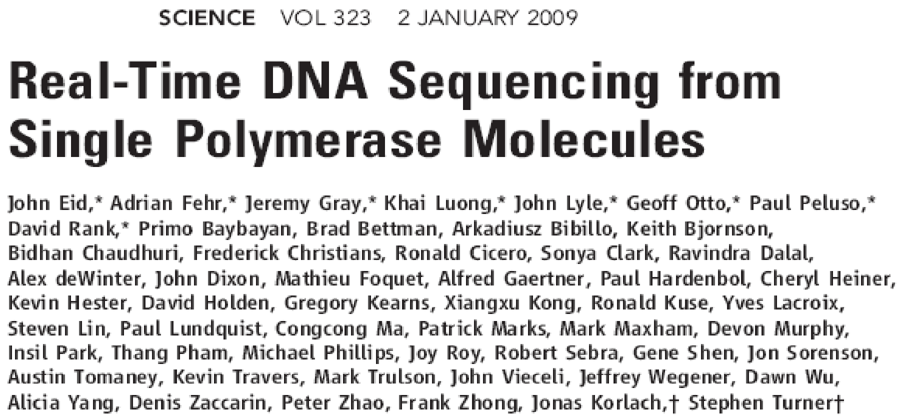
Theo H.E. Meuwissen [1*] and Mike E. Goddard [2]

[1] Agricultural University of Norway
[2] Department of Primary Industries

[*] To whom correspondence should be addressed. E-mail: theo.meuwissen@umb.no.

Manuscript received March 12, 2010
Manuscript accepted March 15, 2010

Futura tercera generación de secuenciadores

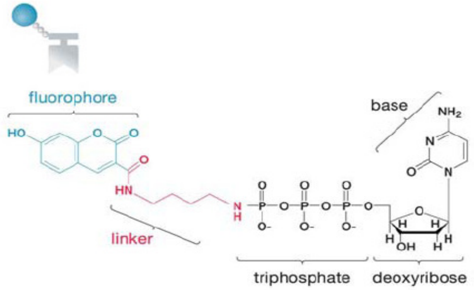


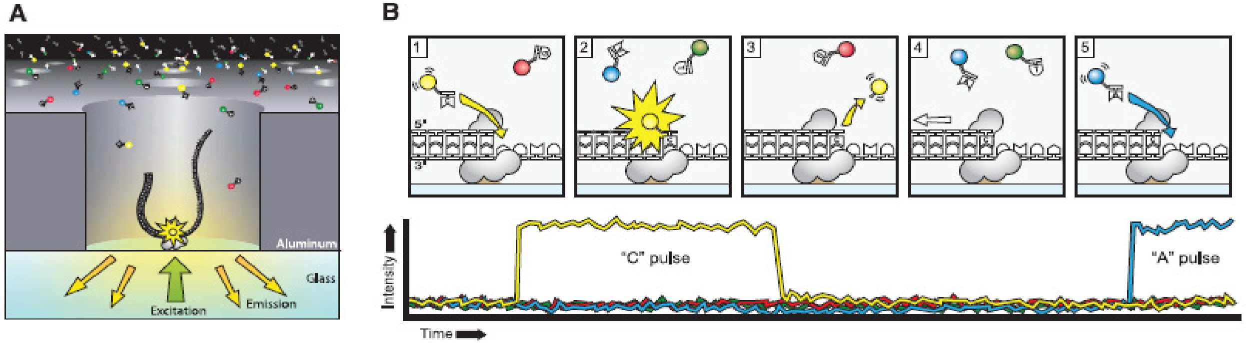
SCIENCE VOL 323 2 JANUARY 2009

Real-Time DNA Sequencing from Single Polymerase Molecules

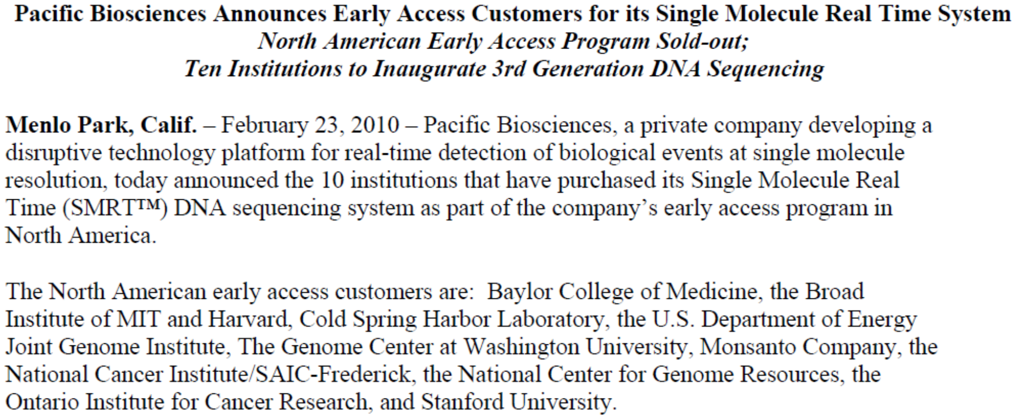
John Eid,* Adrian Fehr,* Jeremy Gray,* Khai Luong,* John Lyle,* Geoff Otto,* Paul Peluso,* David Rank,* Primo Baybayan, Brad Bettman, Arkadiusz Bibillo, Keith Bjornson, Bidhan Chaudhuri, Frederick Christians, Ronald Cicero, Sonya Clark, Ravindra Dalal, Alex deWinter, John Dixon, Mathieu Foquet, Alfred Gaertner, Paul Hardenbol, Cheryl Heiner, Kevin Hester, David Holden, Gregory Kearns, Xiangxu Kong, Ronald Kuse, Yves Lacroix, Steven Lin, Paul Lundquist, Congcong Ma, Patrick Marks, Mark Maxham, Devon Murphy, Insil Park, Thang Pham, Michael Phillips, Joy Roy, Robert Sebra, Gene Shen, Jon Sorenson, Austin Tomaney, Kevin Travers, Mark Trulson, John Vieceli, Jeffrey Wegener, Dawn Wu, Alicia Yang, Denis Zaccarin, Peter Zhao, Frank Zhong, Jonas Korlach,† Stephen Turner†

PACIFIC BIOSCIENCES™






http://www.pacificbiosciences.com



Pacific Biosciences Announces Early Access Customers for its Single Molecule Real Time System

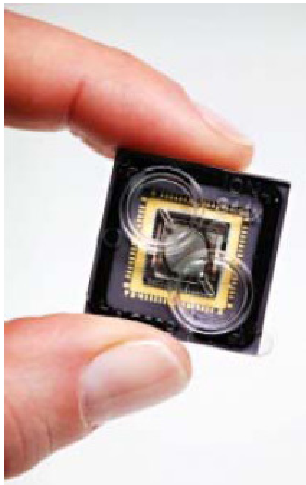
North American Early Access Program Sold-out;
Ten Institutions to Inaugurate 3rd Generation DNA Sequencing

Menlo Park, Calif. – February 23, 2010 – Pacific Biosciences, a private company developing a disruptive technology platform for real-time detection of biological events at single molecule resolution, today announced the 10 institutions that have purchased its Single Molecule Real Time (SMRT™) DNA sequencing system as part of the company's early access program in North America.

The North American early access customers are: Baylor College of Medicine, the Broad Institute of MIT and Harvard, Cold Spring Harbor Laboratory, the U.S. Department of Energy Joint Genome Institute, The Genome Center at Washington University, Monsanto Company, the National Cancer Institute/SAIC-Frederick, the National Center for Genome Resources, the Ontario Institute for Cancer Research, and Stanford University.

Futura tercera generación de secuenciadores

ion torrent Semiconductor Sequencing for Life™



- **Low cost.** The instrument is tentatively priced at $45K, and disposables (chips) are $500 per run.

- **Throughput.** The standard chip, called the 314, contains 1.5 million wells. It's one read per well, they expect ~50% of wells to be filled. That's 750,000 reads; at the current read length (100 bp), we're talking 75 Mbp per 1h run.

http://www.iontorrent.com/

Entrando en la era *Petabyte*… de la Genética

- Un Petabyte equivale a un millón de Gigabytes

"The Petabyte Age: Sensors everywhere. Infinite storage. Clouds of processors. Our ability to capture, warehouse, and understand massive amounts of data is changing science, medicine, business, and technology. As our collection of facts and figures grows, so will the opportunity to find answers to fundamental questions.
Because in the era of big data, more isn't just more. More is different. "

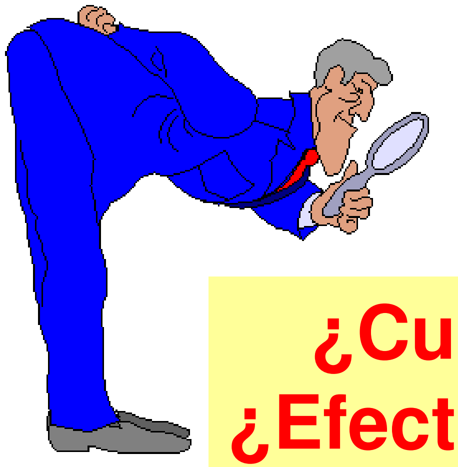
Chris Anderson 23.06.08

•In 2020 it will be more convenient to switch to EXABYTES!
Exabyte is a billion of Gigabytes

Kilo	10^3	a typical gene
Mega	10^6	microbial genome
Giga	10^9	human genome
Tera	10^{12}	bp in Genbank
Peta	10^{15}	humanity's hapmap

Universitat Autònoma de Barcelona

Abriendo la “caja negra”



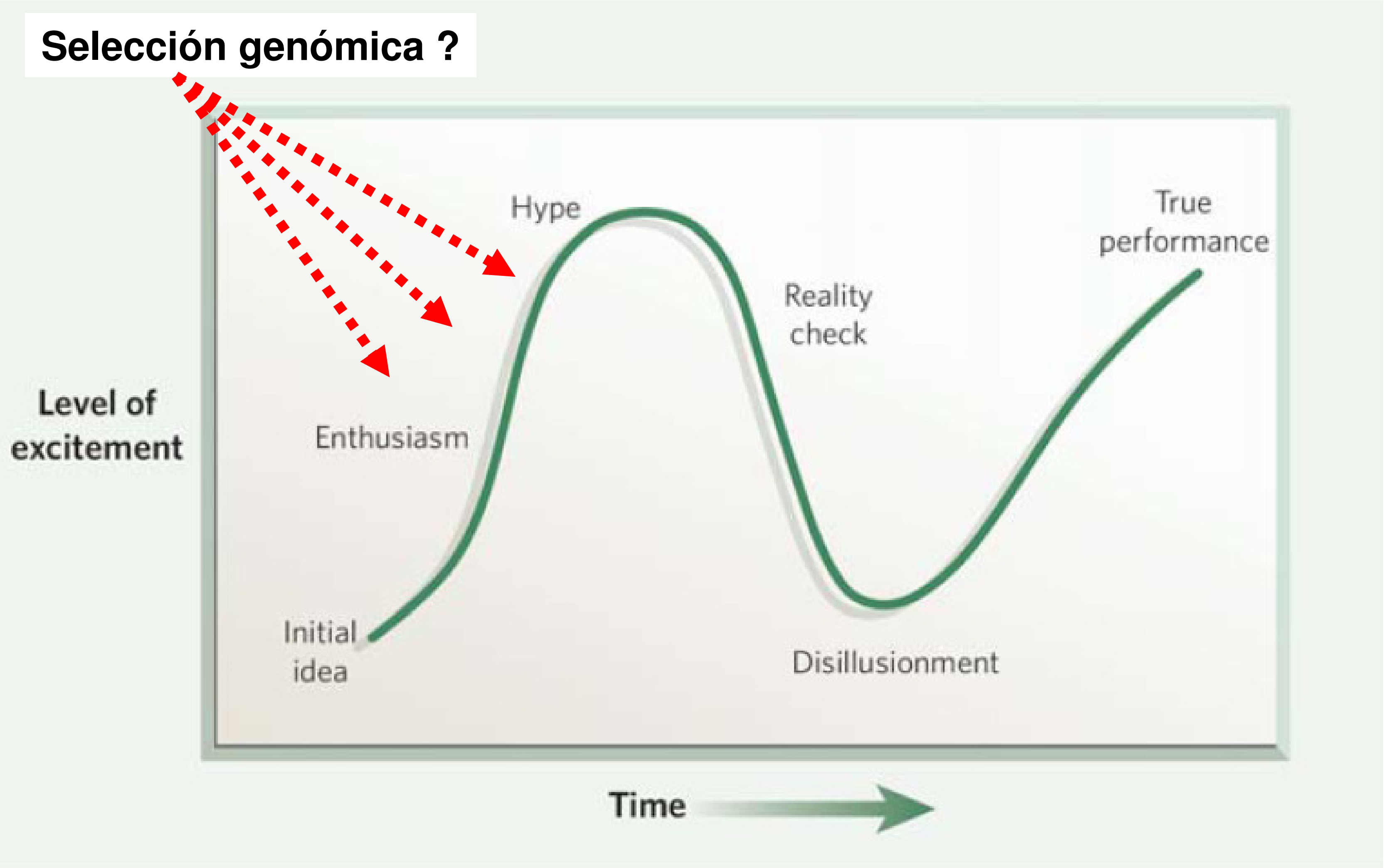
Antes: modelo infinitesimal

Ahora: además información genómica

¿Cuantos QTLs?
¿Efecto de los QTLs?
¿QTNs?

Universitat Autònoma de Barcelona

"Niveles de entusiasmo ante las nuevas tecnologías"



Selección genómica ?

Collins (2006) Nature

Muchas gracias por su atención

http://svgm.uab.es