

# *“Applied Genomics in Cattle – Identification of the SLICK locus in tropically adapted cattle”*



Tad Sonstegard

USDA, ARS, Bovine Functional Genomics Lab,  
Beltsville, MD

[Tad.Sonstegard@ars.usda.gov](mailto:Tad.Sonstegard@ars.usda.gov)

# Overview



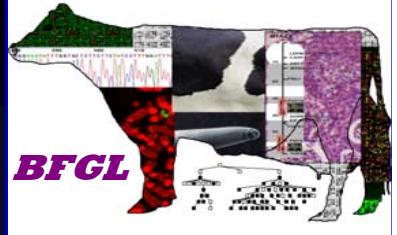
- Research Mission
- Background on Cattle genomics
- Tool development in cattle
- Unraveling the SLICK locus

# *Beltsville Agricultural Research Center*

- ~6,600 acres adjacent to the D.C. Beltway
  - \$121 million, ~300 permanent PhD scientists
  - 100+ postdocs & visiting scientists, 200+ undergraduate, graduate students
- Animals & natural resources
  - 12 Laboratories w/ ~110 Scientists
  - Dairy, Beef, Swine, Transgenic Mouse and Poultry facilities



# *Bovine Functional Genomics Laboratory*



- To improve the efficiency of ruminants through fundamental research on gene expression and genome selection.
  - Develop molecular biological and bioinformatic tools to facilitate genomic studies
  - Enhance selection through genome selection
  - Identify genes - for enhanced productivity and disease resistance
  - Develop fundamental knowledge - physiology & development of the mammary gland & gastrointestinal tract



# Applying genomics to the hunger



World Food  
Programme

wfp.org

Category	1	2	3	4	5	Incomplete data
Undernourished	<5%	5-9%	10-19%	20-34%	≥35%	
Description	Extremely low	Very low	Moderately low	Moderately high	Very high	

Source: *The State of Food Insecurity in the World 2010*, Food and Agriculture Organization of the United Nations. Please note that the SOFI 2010 data in some cases dates back to 2005 so may not always reflect the present-day situation in individual countries.

© 2011 World Food Programme

The designations employed and the presentation of material on this map do not imply the expression of any opinion whatsoever of WFP concerning the legal or constitutional status of any country, territory or sea area, or concerning the delimitation of frontiers.

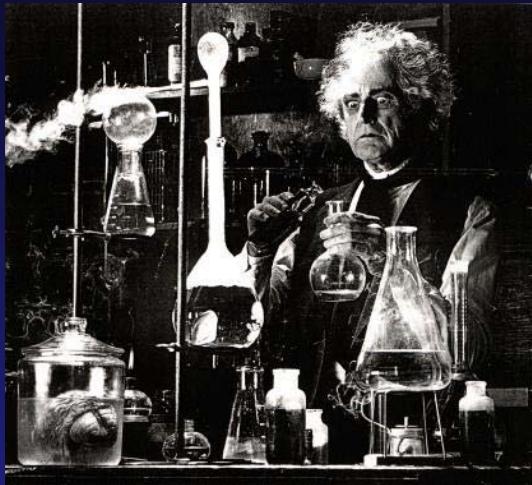
\* This Line of Control in Jammu and Kashmir agreed on by India and Pakistan is represented approximately by a dotted line. The Line of Control in Jammu and Kashmir has not yet been agreed upon by the parties.

\*\* A dispute exists between the governments of Argentina and the United Kingdom of Great Britain and Northern Ireland concerning sovereignty over the Falkland Islands (Malvinas).

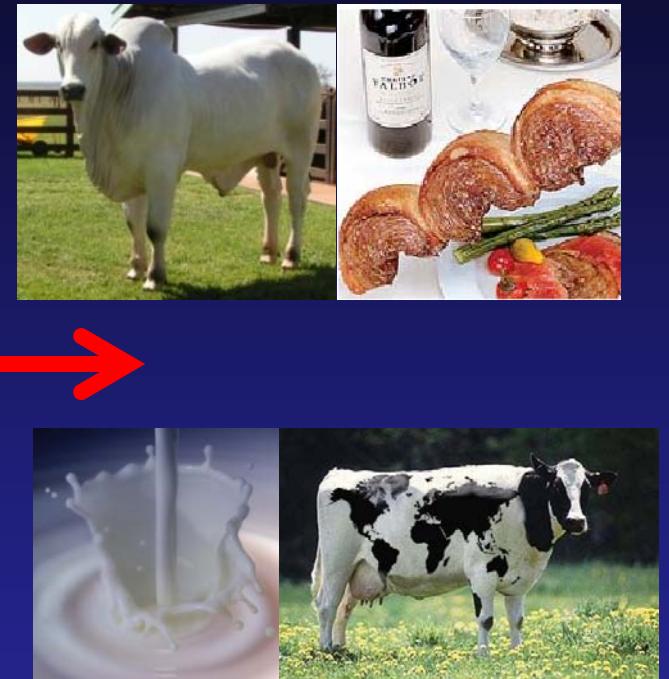
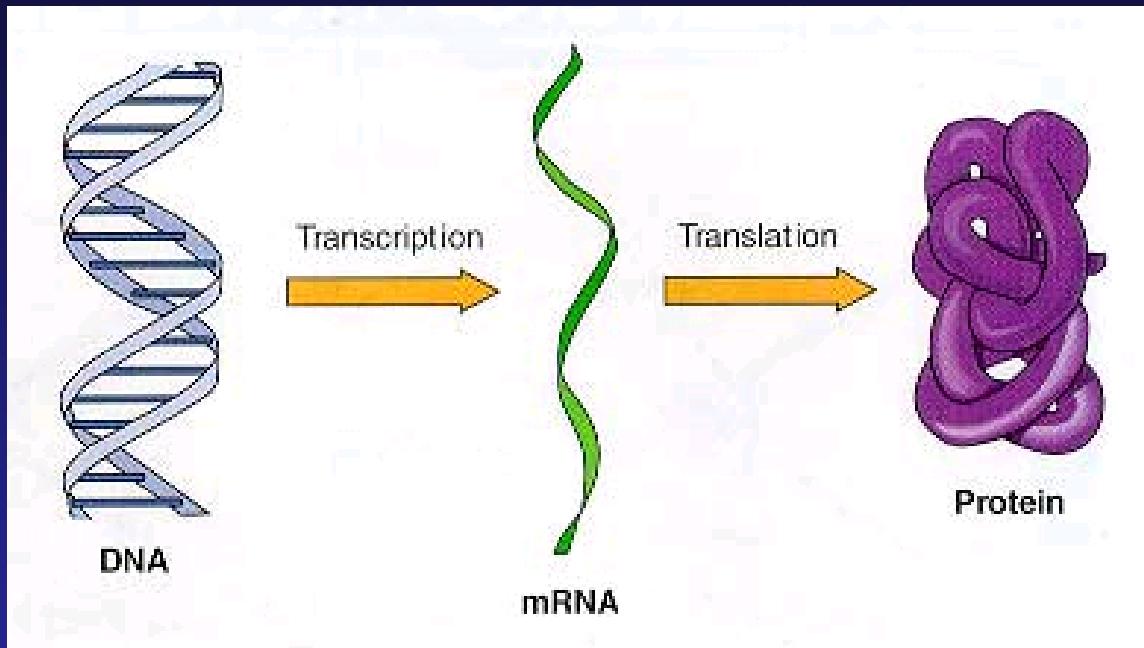
# **BFGL Applied Genomics Research Group**

- **4 Scientists specializing in genetics, genomics, & bioinformatics**
  - 3 technicians & 4 post-docs, students, and visiting scientists
  - ~\$75-150K in all other spending/year
  - 7 active “official” sub-projects to support genomics
  - Work closely with Animal Improvement Programs Lab & Industry & many, many international partners
- **CPU & Instrumentation (~\$2 million)**
  - 3 next generation sequencers & supporting equipment
  - 1 SNP chip scanner
  - >\$250K in bio-computing infrastructure

# What is genomics?



# Central Dogma



**GENOTYPE**

**PHENOTYPE**

# Genomics

- a branch of biotechnology concerned with applying the techniques of genetics and molecular biology to the **genetic mapping** and **DNA sequencing** of sets of genes or the complete genomes of selected organisms, with organizing the results in databases, and with applications of the data (as in medicine or biology)

# Bioinformatics

- is an interdisciplinary field that develops and improves upon methods for storing, retrieving, organizing and analyzing biological data. A major activity in bioinformatics is to develop software tools to generate useful biological knowledge

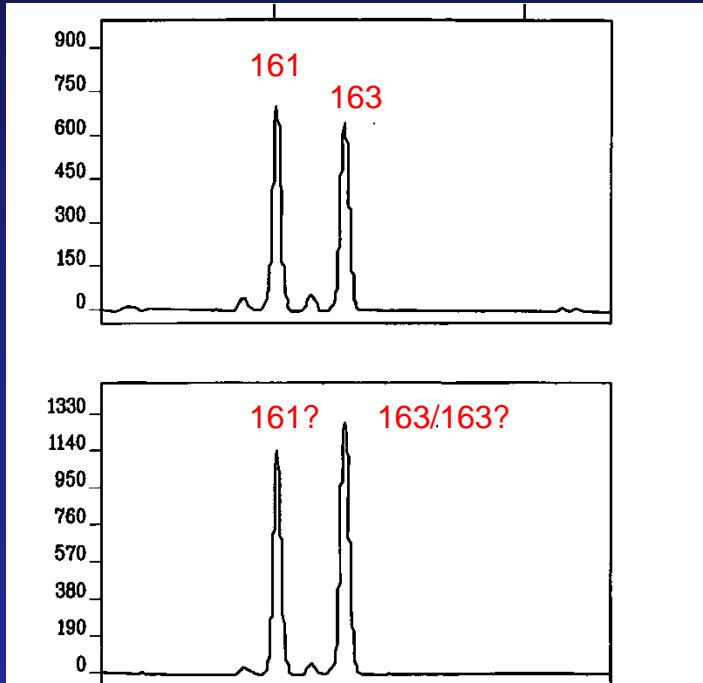
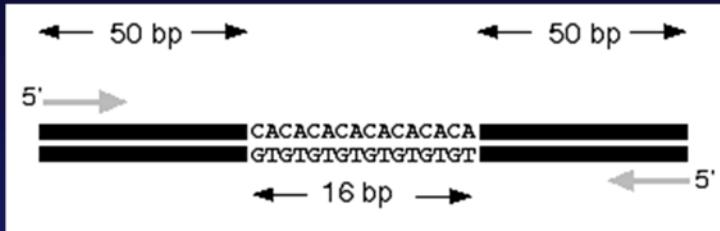
# History of Cattle Genomics

- Mapping traits since the early 1990s
- Most mapping was done with microsatellite markers – a slow and tedious process
- First cattle linkage maps in late 90s
- Information was not good enough for marker assisted selection...

# What is a microsatellite (MS)?

**Microsatellite locus  
sequence – typically not  
found in genes**

- **The problem...**
  - Sometimes cannot tell the true genotype
  - Example would be 163 allele of SLICK marker test
  - The error rate in marker scoring for MS parentage markers is about 5%

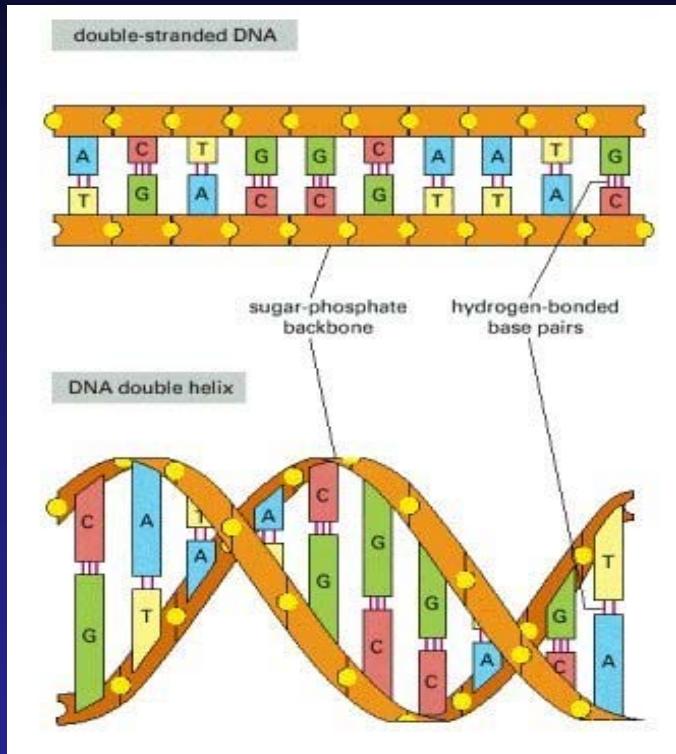


## *The big breakthrough in cattle genomics...*

- The bovine reference genome(s).
  - >\$50 million invested with Baylor
  - Source of 2.3 million **SNP** to enable high resolution mapping and **Genome Selection**
- Result – 30 Chr assemblies/2.85 billion bp

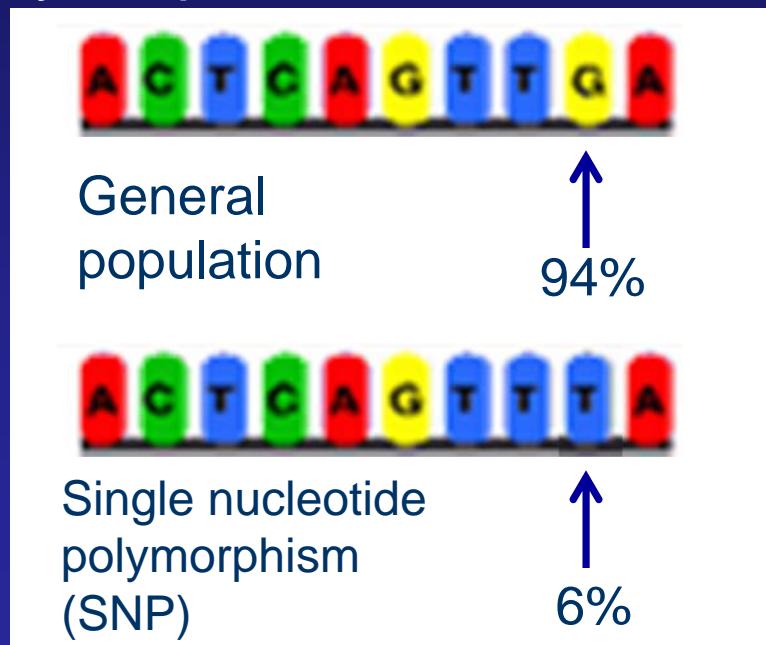


# DNA & Genetic Markers



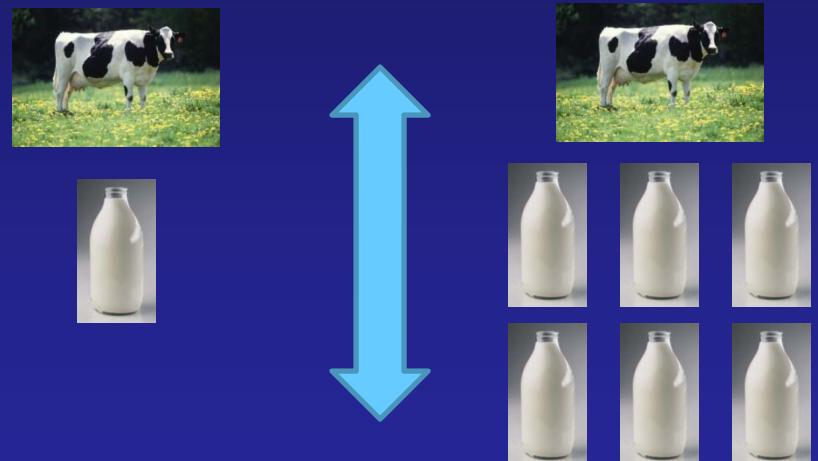
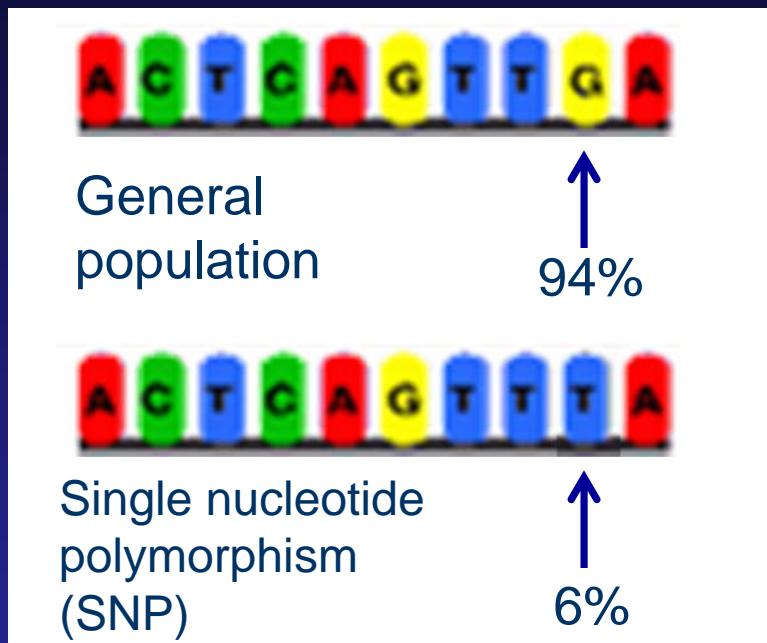
- Allow inheritance to be followed in a region across generations

- **Polymorphism**= “many + forms”
- Millions of genetic markers
- **SNP**= Single nucleotide polymorphism



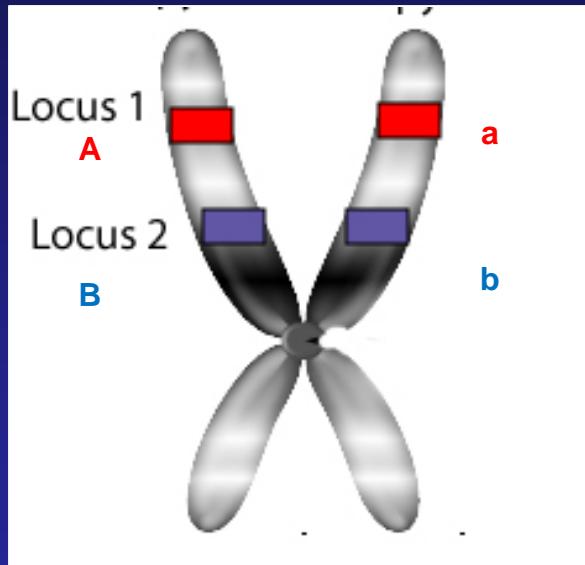
# SNP associations

## SNP Frequencies



# Linkage Disequilibrium

**Linkage Disequilibrium (LD):** the non-random association of alleles at two or more loci, that may or may not be on the same chromosome



49% **AB**

49% **ab**

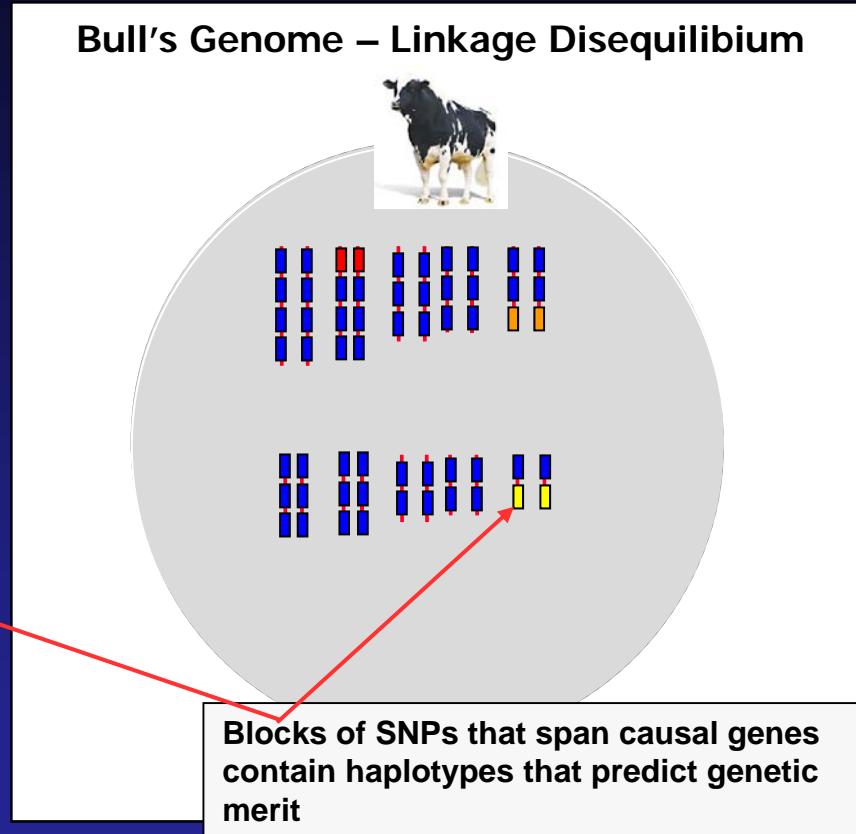
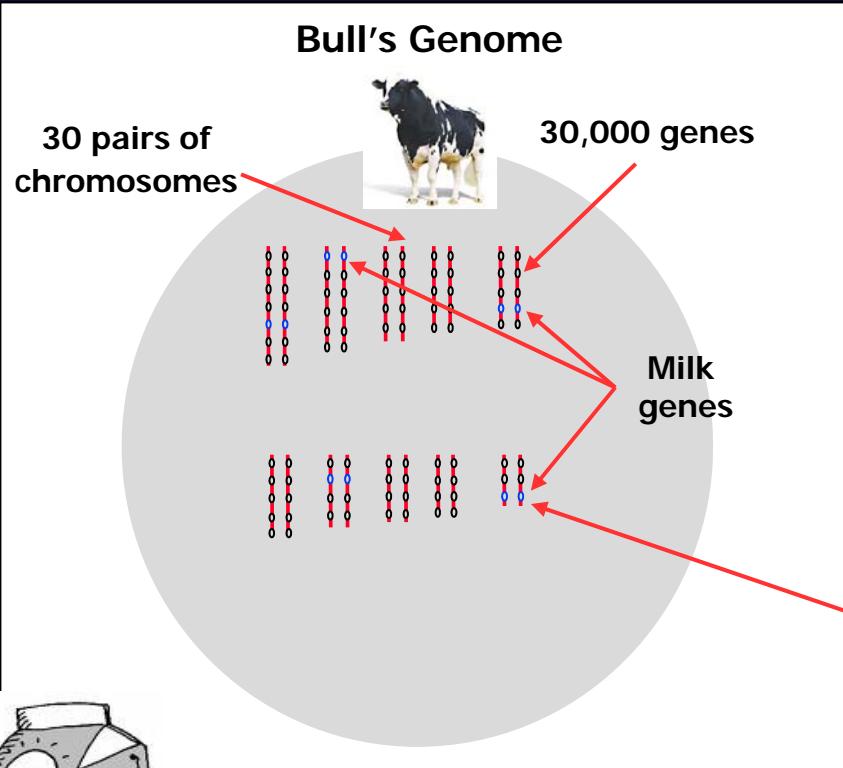
1%

**Ab**

1%

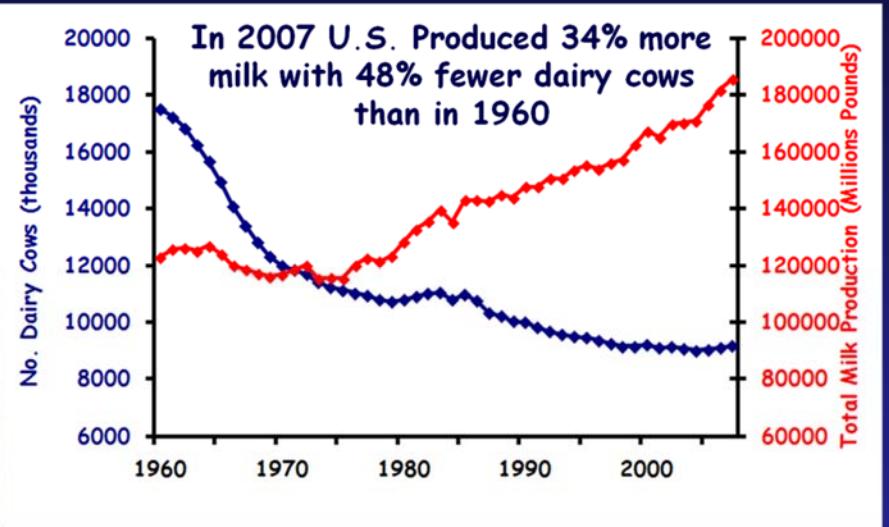
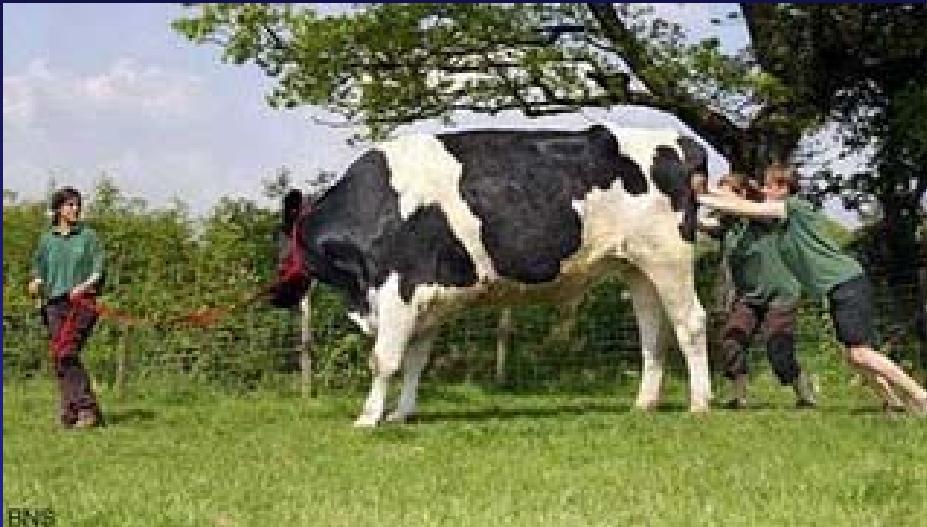
**aB**

# Whole Genome Selection



The **concept of WGS** is to use high resolution map (50K SNPs) to divide the genome into small segments and estimate the contribution of each chromosomal segment and haplotype to genetic merit – form of WGA analysis

# *A previous challenge...*



*How do we move the NA dairy industry forward with genomics?*



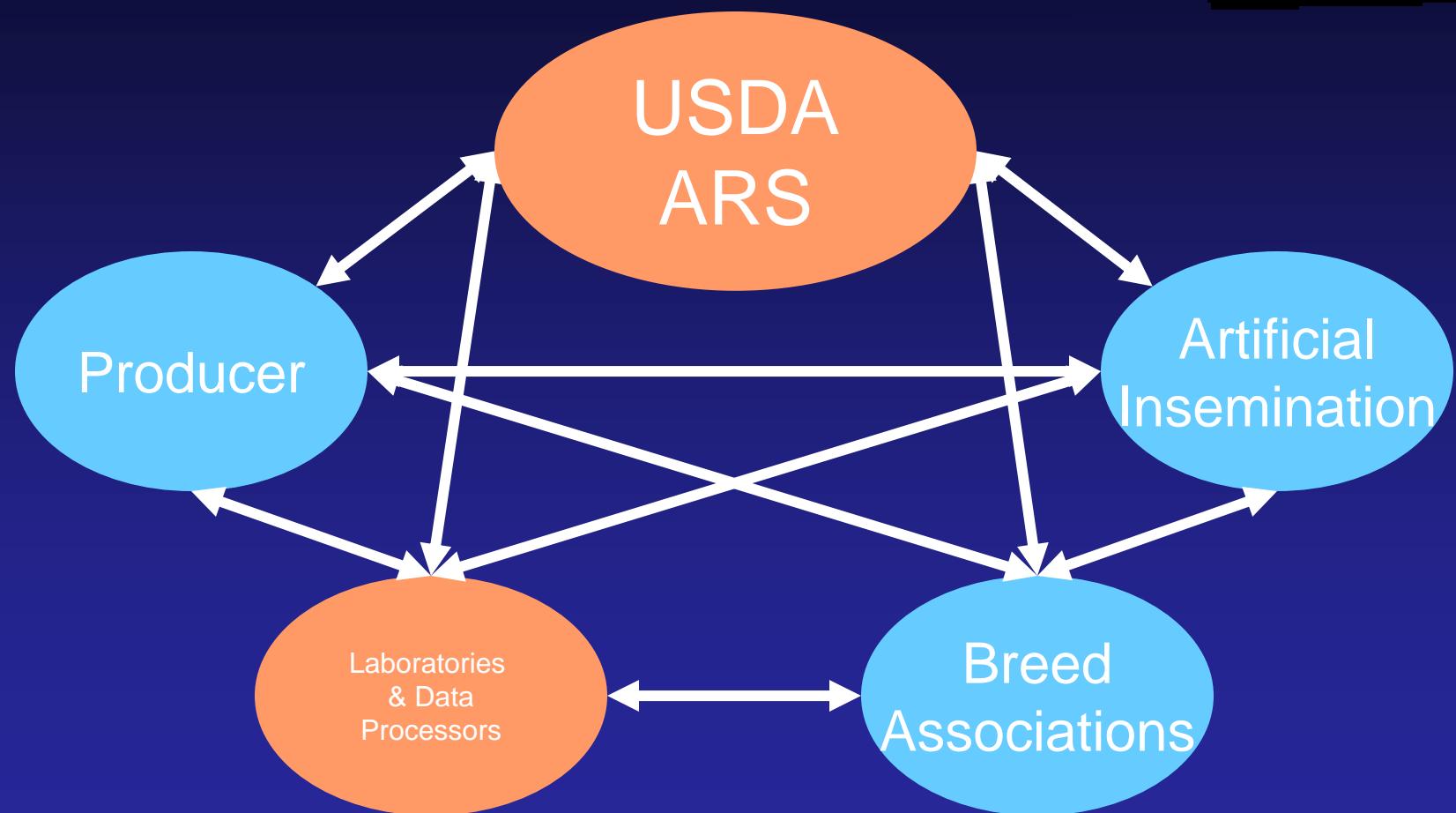
# *Traditional Selection*



- USDA has been collecting dairy data for 100 years
- Estimate genetic merit for animals
- Genetic improvement approached theoretical optimum (200+ lb milk per yr) – but expensive



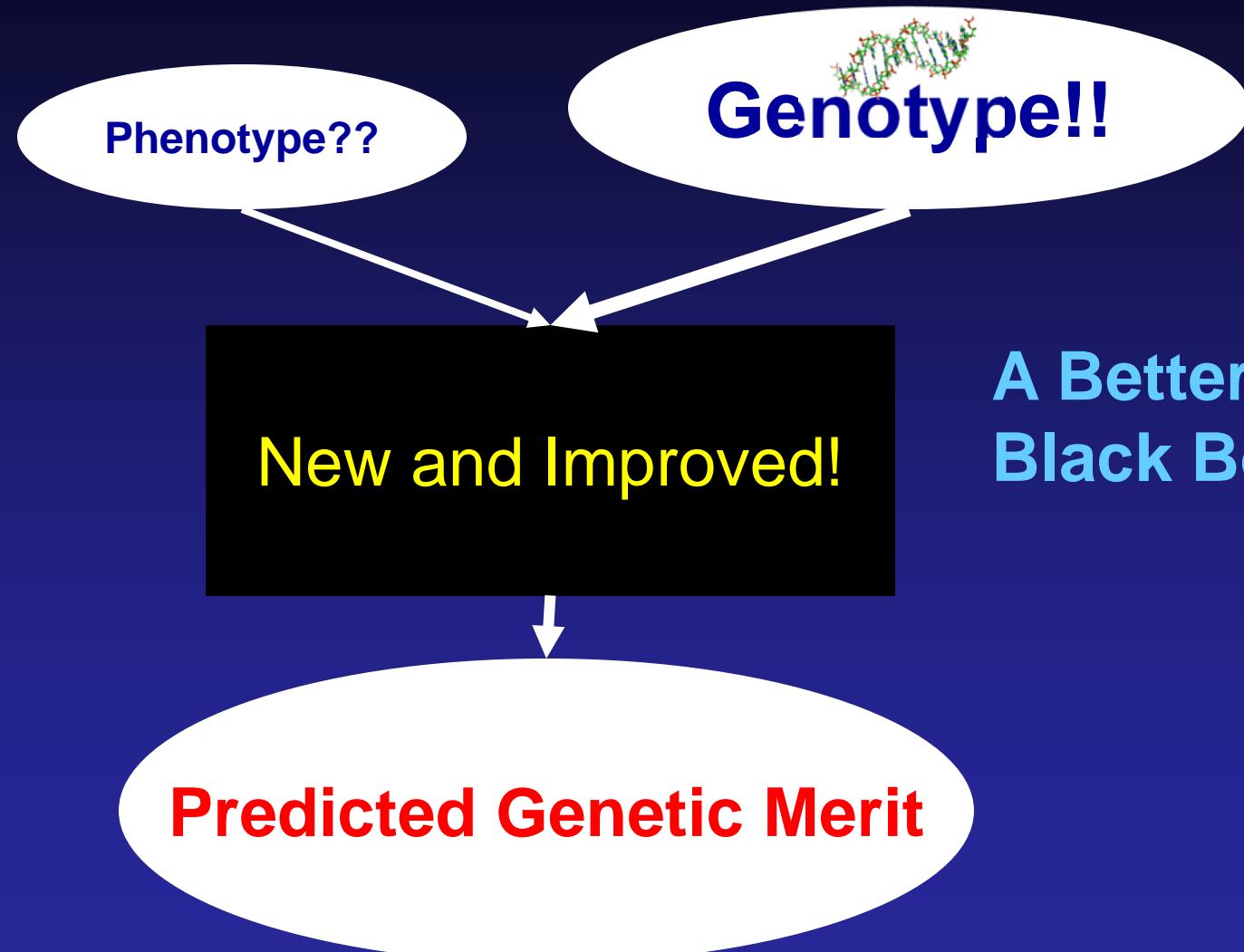
# Dairy Cattle Genetics Industry Cooperation



# Historic Data!

- **68,270,792** identification records
- **334,402** animal genotypes
- **142,157,859** lactation records (1960+)
- **558,425,959** daily yield records (1990+)
- **139,043,355** reproduction event records
- **25,223,471** calving difficulty scores
- **21,971,890** stillbirth scores

# *More efficient genetic selection?*

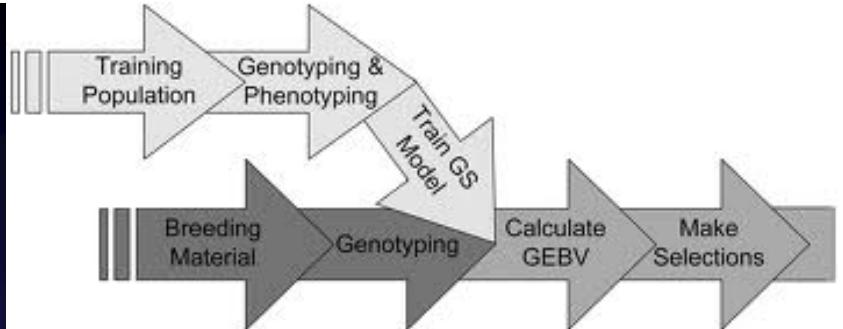


# New Way of Looking at an Old Question

- Which calf from a group of candidates is the best?



# *GS Methods & Materials*



- A genome assembly (2005)
- Validated SNP resources (2006)
- SNP Assay development (2007)
- Phenotypes for training set (1908-present)
- DNA from influential animals (1960s-present)
- Prediction algorithm (2001-2009)

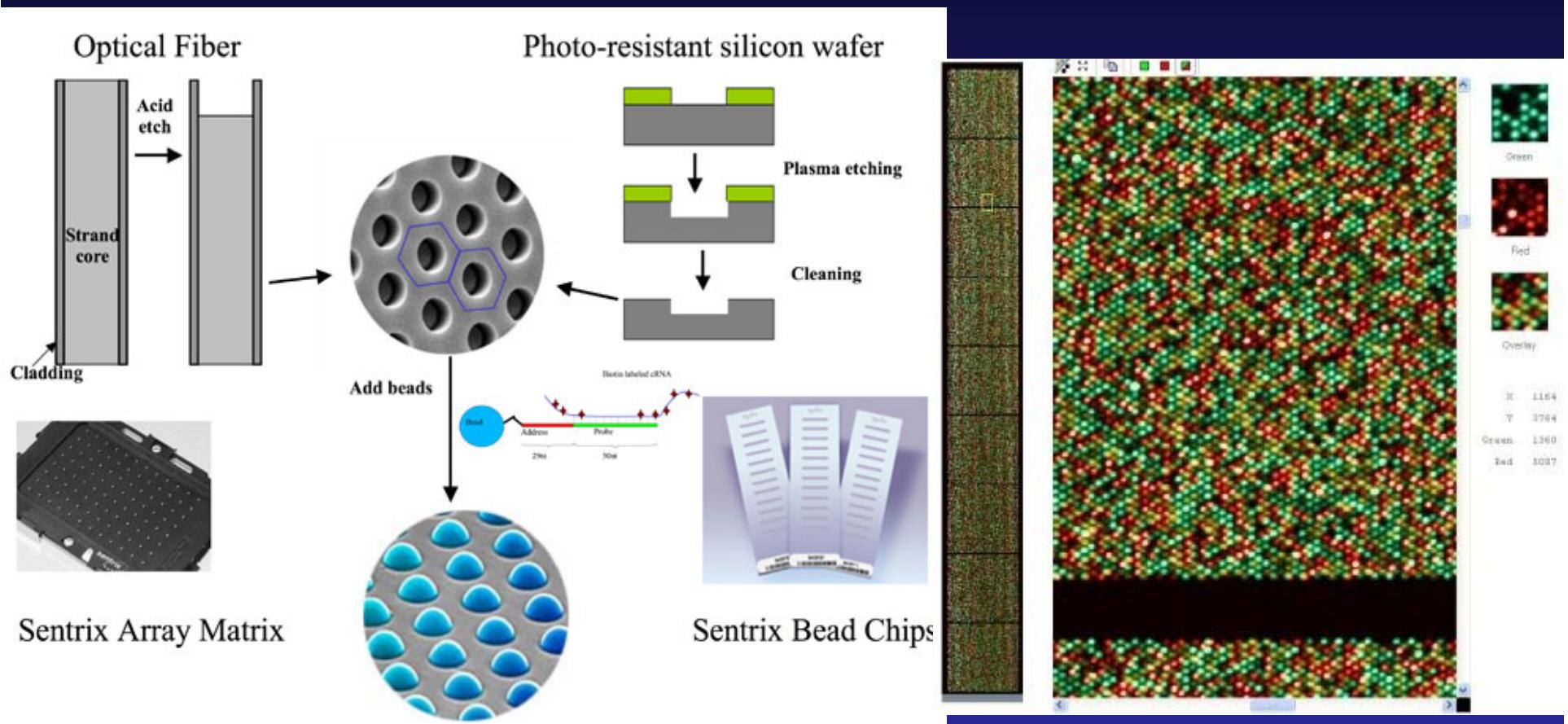


# Genotype by Sequencing

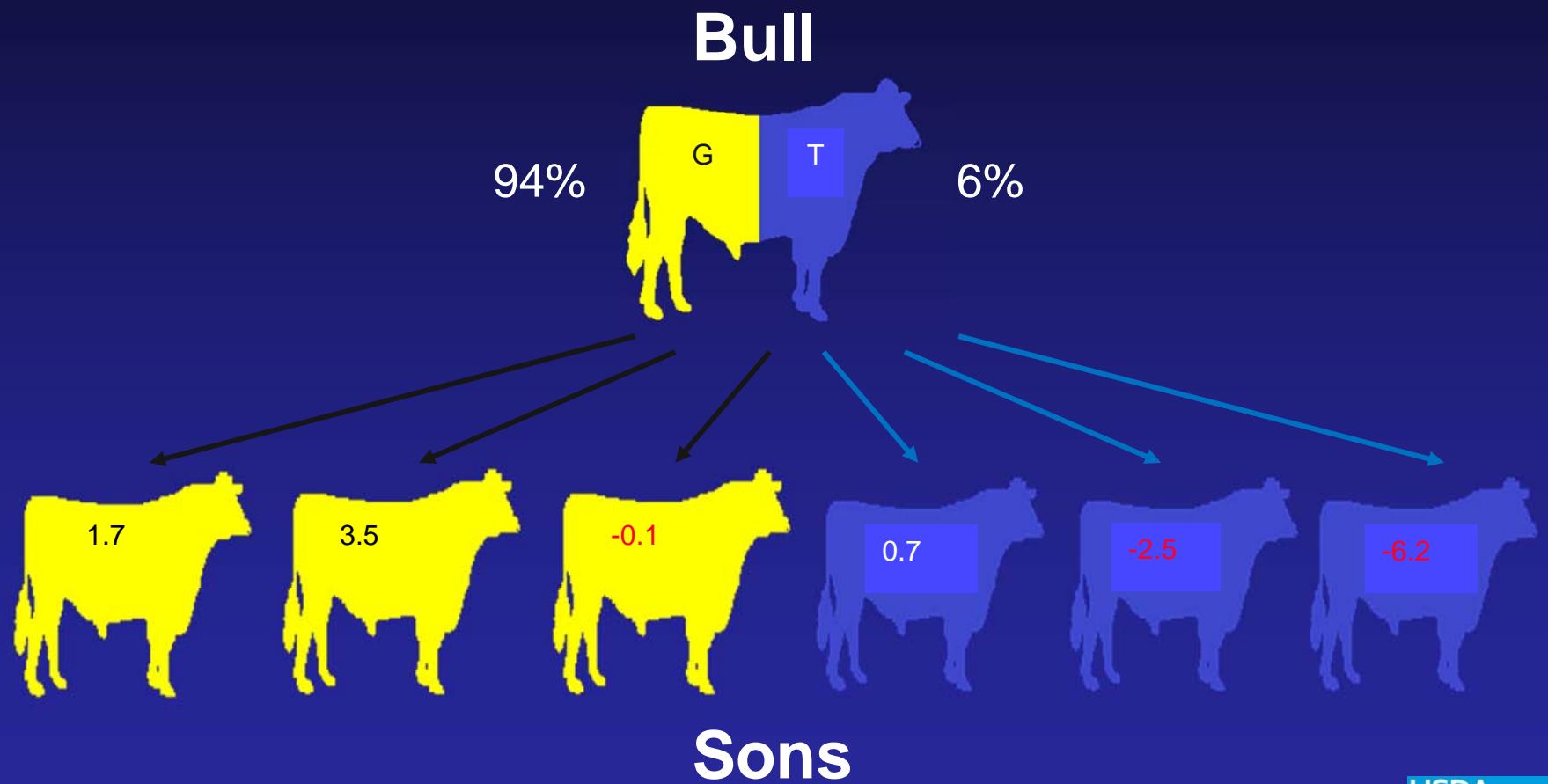


**Next generation sequencing – can now sequence more than 600 billion base pairs in 10 days**

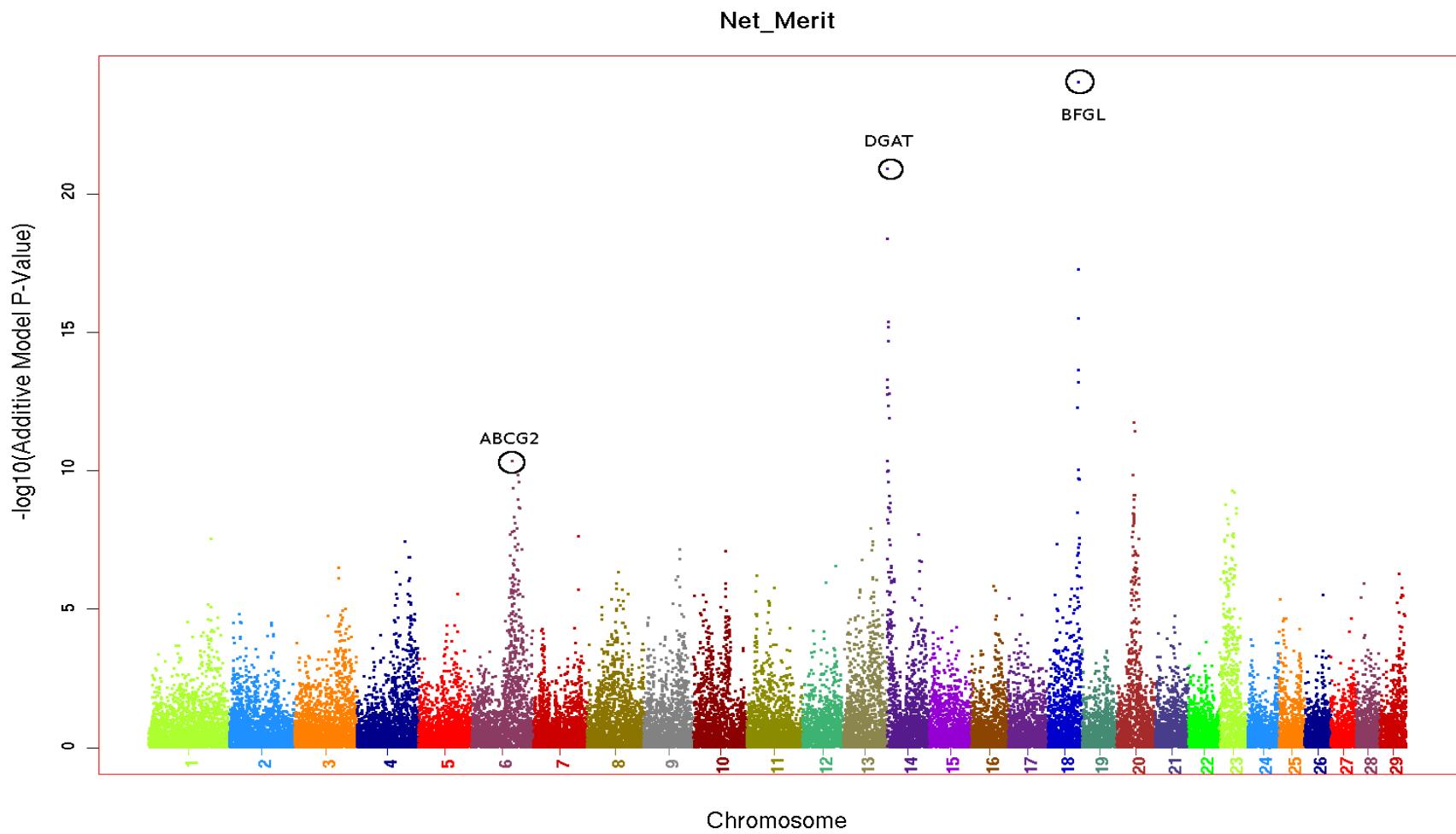
# The technology of a SNP chip



# SNP associations in Genetic Evaluations



# Times 40k!!!



# Genomics Works!

**Forbes.com** Home Page for the World's Business

U.S. EUROPE ASIA

Home Lists Business Tech Markets Personal Finance Entrepr  
Video ForbesWoman CEO Network Org Chart Wiki Mobile Portfolio Tracker Bl

Technology

**Holy Cow**

Matthew Herper, 02.05.10, 03:00 PM EST  
Forbes Asia Magazine dated February 08, 2010

Genomics has revolutionized dairy farming.



© David Deal For Forbes

## The Perfect Milk Machine: How Big Data Transformed the Dairy Industry

MAY 1 2012, 3:47 PM ET 27



Dairy scientists are the Gregor Mendels of the genomics age, developing new methods for understanding the link between genes and living things, all while quadrupling the average cow's milk production since your parents were born.



Reuters.

While there are more than 8 million Holstein dairy cows in the United States, there is exactly one bull that has been scientifically calculated to be the very best in the land. He goes by the name of Badger-Bluff Fanny Freddie.

Already, Badger-Bluff Fanny Freddie has 346 daughters who are on the books and thousands more that will be added to his progeny count when they start producing milk. This is quite a career for a young animal: He was only born in 2004.

# What is a SNP genotype worth?

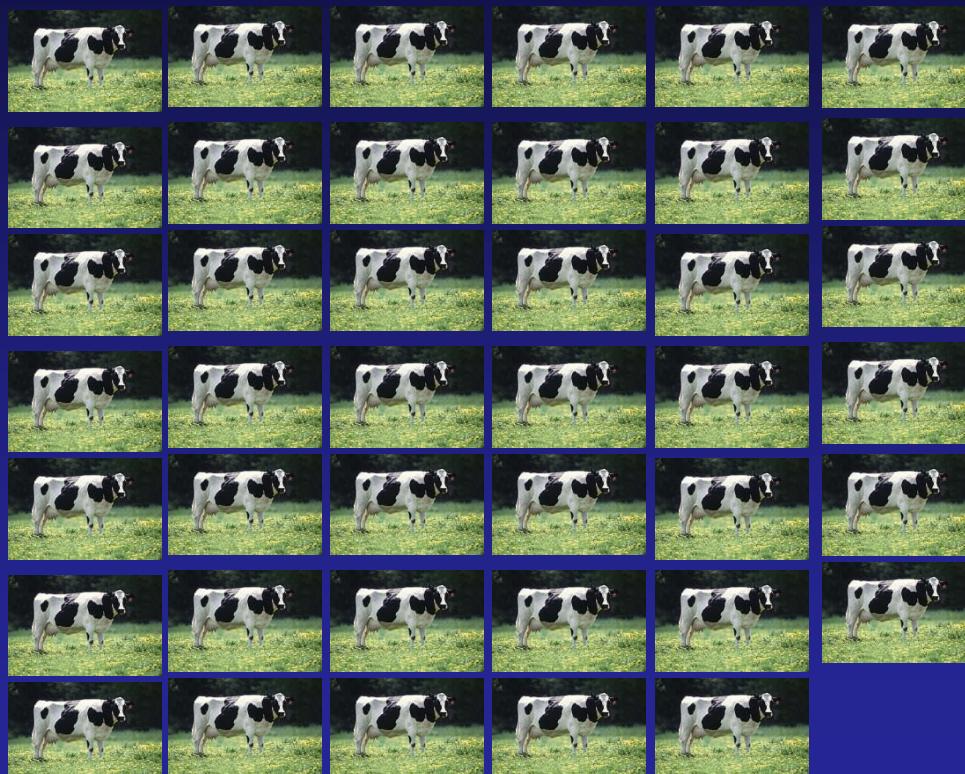


Pedigree information is equivalent to  
information on about **7** daughters



# What is a SNP genotype worth?

For milk protein ( $h^2=0.30$ ), the SNP genotype provides information equivalent to an additional 34 daughters!!!



# Why genomics works in dairy cattle

- Extensive historical data available
- Well-developed genetic evaluation program
- Widespread use of artificial insemination bulls
  - Historic DNA source
  - Large half-sib families
  - Progeny test programs – accurate genetic merit
- High value animals, justify genotyping expense
- Long generation interval associated with data collection

**GeneSeek® Genomic Profiler™**

**One Sample, One Simple Comprehensive Solution**

### GeneSeek® Genomic Profiler™ for Dairy Cattle

The custom GeneSeek Genomic Profiler (GGP) low density BeadChip utilizes Illumina Infinium® chemistry and features nearly 10,000 SNPs for accurate, high density BeadChip imputation and also includes disease and performance trait SNPs.

**The GGP BeadChip offers a comprehensive parentage, disease and trait profile:**

- Approximately 8,500 SNPs (including the original 7,000 Illumina LD SNPs) for increased Imputation efficiency.
- All commonly utilized USDA parentage SNPs.
- Approximately 400 SNPs to enable conversion of ISAG microsatellite parentage data.
- Several disease and trait markers that are included below.

<ul style="list-style-type: none"> <li>◦ Beta Casein A/B</li> <li>◦ Beta Lactoglobulin</li> <li>◦ Black/Red Coat Color (346, 358, 373)</li> <li>◦ BLAD</li> <li>◦ Chondrodysplasia</li> <li>◦ Citrullinemia</li> <li>◦ DUMPS</li> <li>◦ Dun Color</li> <li>◦ Factor XI</li> <li>◦ Holstein haplotypes impacting fertility HH1, HH2, HH3</li> </ul>	<ul style="list-style-type: none"> <li>◦ Jersey fertility haplotype I</li> <li>◦ Kappa Casein I</li> <li>◦ Kappa Casein II</li> <li>◦ Silver Color Dilutor</li> <li>◦ Y chromosome infertility</li> </ul>
<ul style="list-style-type: none"> <li>◦ CVM*</li> <li>◦ Beta Casein A2*</li> <li>◦ Brachyspina*</li> <li>◦ Hypothrichosis_KRT71 (Rat tail)*</li> </ul>	<small>*additional royalty fees will be applied for these optional licensed SNPs</small>

GeneSeek, a Neogen Company, was founded in 1998 and has developed into a comprehensive agricultural biotechnology service provider. GeneSeek provides comprehensive genomics solutions for research and development, and commercial applications.

Contact GeneSeek for more details  
[genseekinfo@neogen.com](mailto:genseekinfo@neogen.com).



Acknowledgement: GeneSeek gratefully acknowledges the USDA (BARC, APL and BFL labs; MARC, Genetics and Breeding Research; NADC, Ruminant Diseases and Immunology Research Unit) for its contribution of SNPs and assistance in chip design in the development of GeneSeek Genomic Profiler.

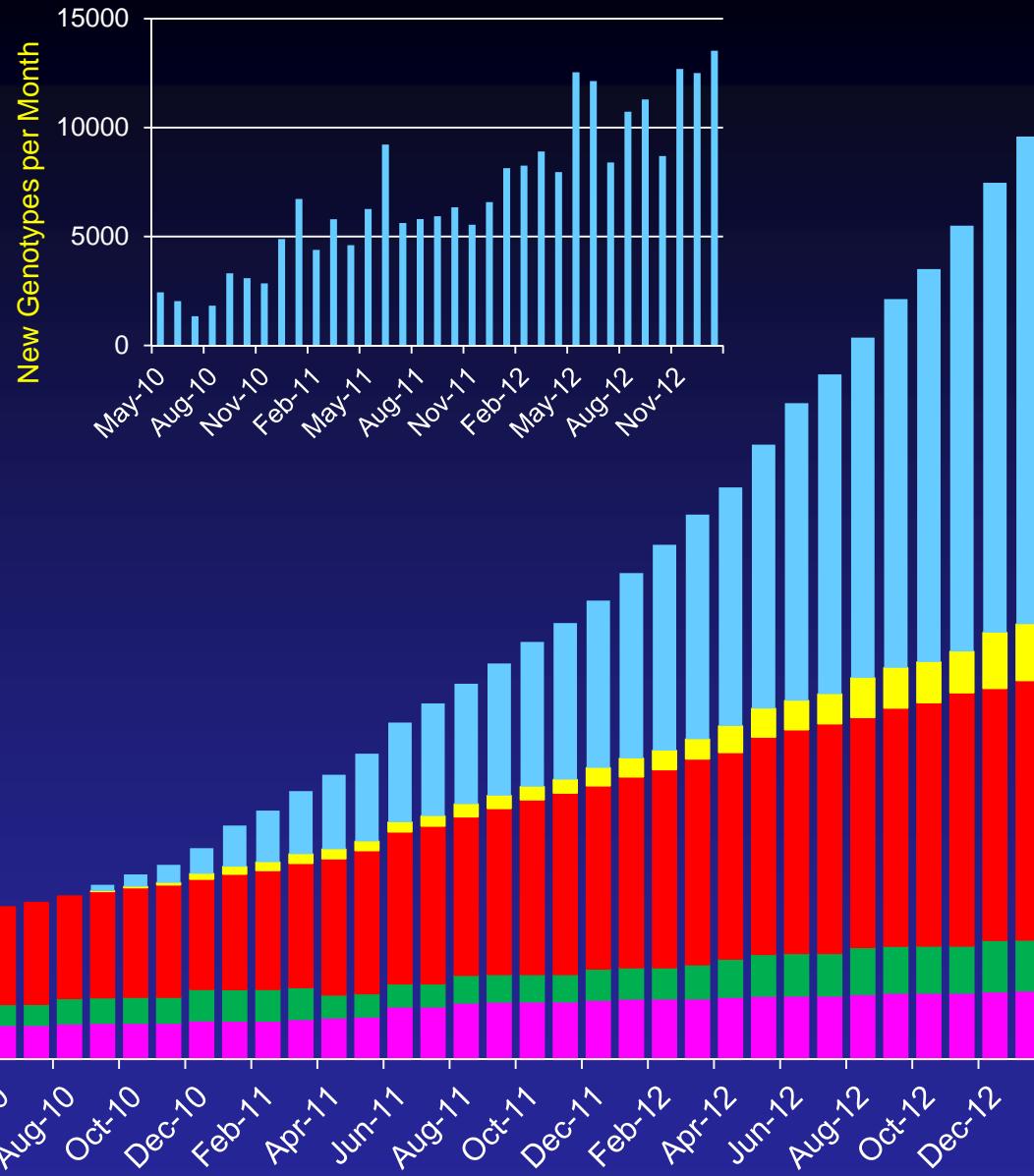


[d.Sonstegard@ars.usda.gov](mailto:d.Sonstegard@ars.usda.gov)

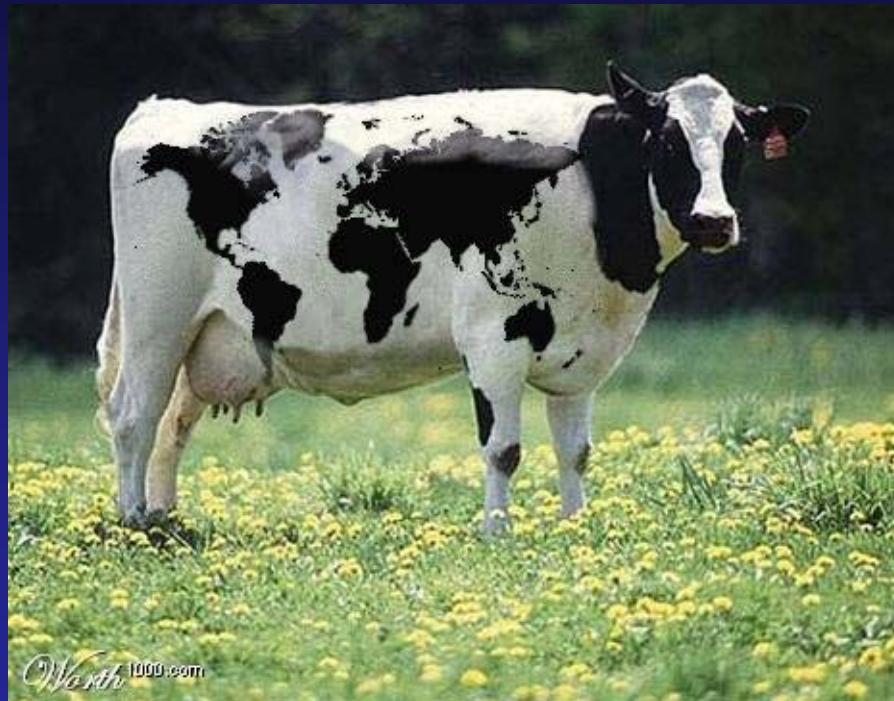
**Better - cheaper tools**



# Genotypes



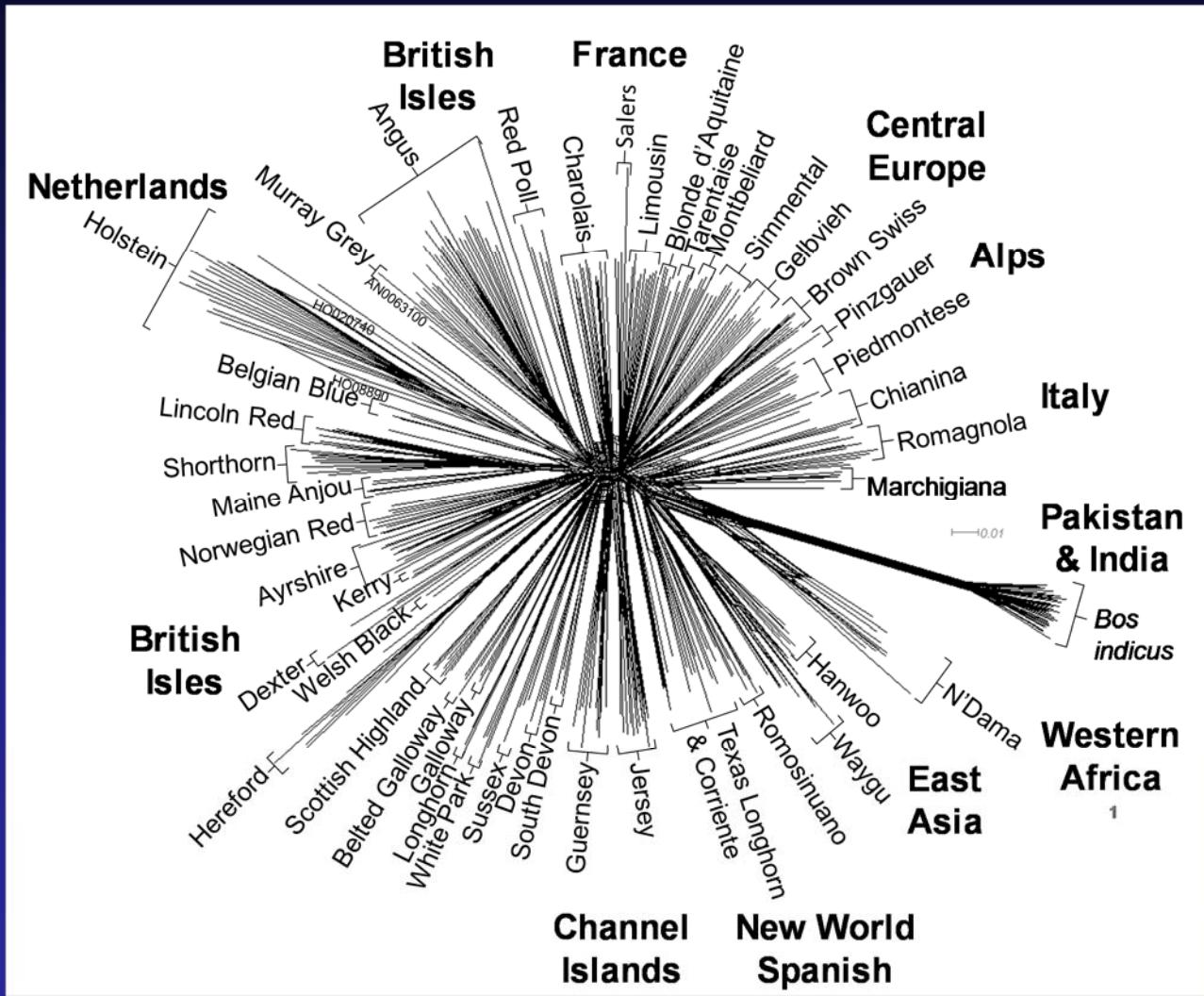
# Now What?!



# Other challenges in cattle genomics?

- Select animals for bioconservation
- Improve indigenous breeds
- Apply genome selection in beef cattle
- Improve fertility across breed types
- Find traits related to tropical adaptation

# *Can decode the origin of breeds...*



# Seleção Animal Genoma

37.5% Golias  
12.6% Karvadi  
<0.9% taurine



FUTURO



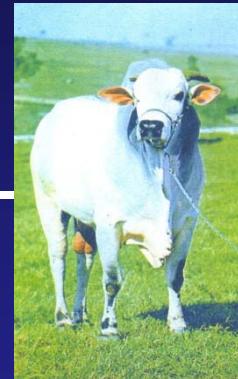
Fauji SC



Papaina IV



Golias – imp.  
Ganta – imp.



Ediri – ½ Golias

Papaina

- Selecionados touros com 0.16 het indice BovineSNP50 & mtDNA derivados *Bt-indicus*
- Data used for Illumina and Affy Bovine HD designs



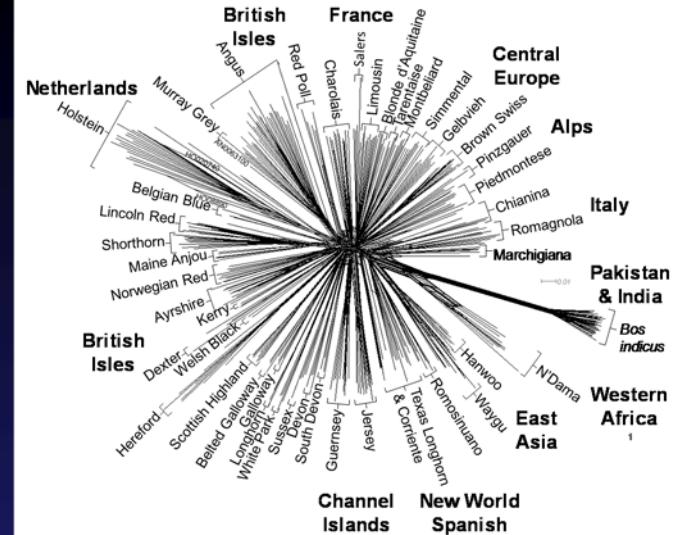
unesp Araçatuba



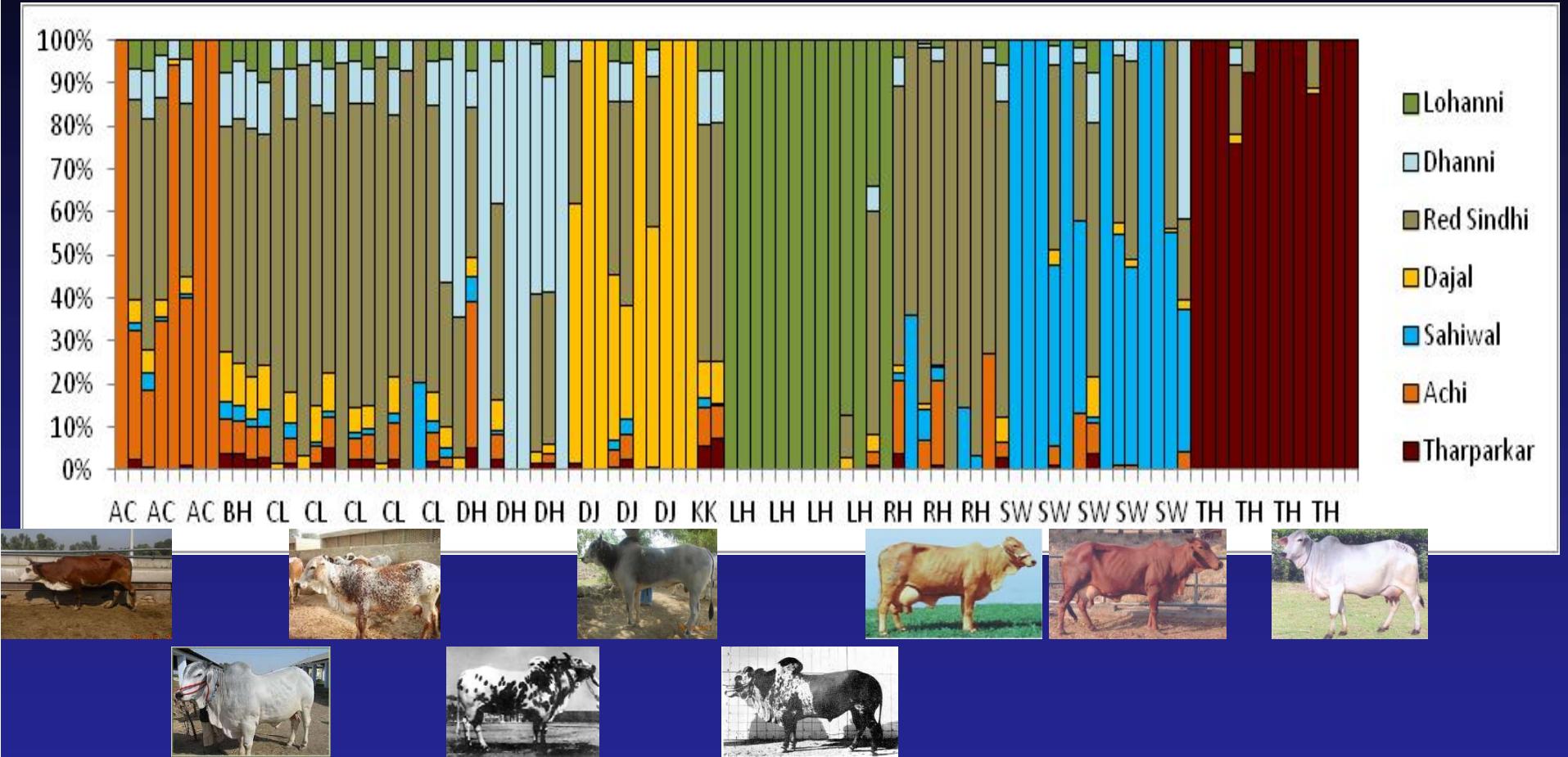
UNIVERSITY OF MARYLAND USDA 1862-2012 150 YEARS

# BovineHD SNP Discovery & Mining

- 49 taurine animals (88X – 7 breeds)  
vs. 21 zebu animals (40X - 4 breeds)
- Focus on identifying SNP involved in subspecies differences – makes a high density SNP chip
- **>45 million SNPs discovered – validation rate 97%**
  - Taurus only : 17,227,326
  - Indicus only: 14,777,958
  - Common : ~ 17 million SNP

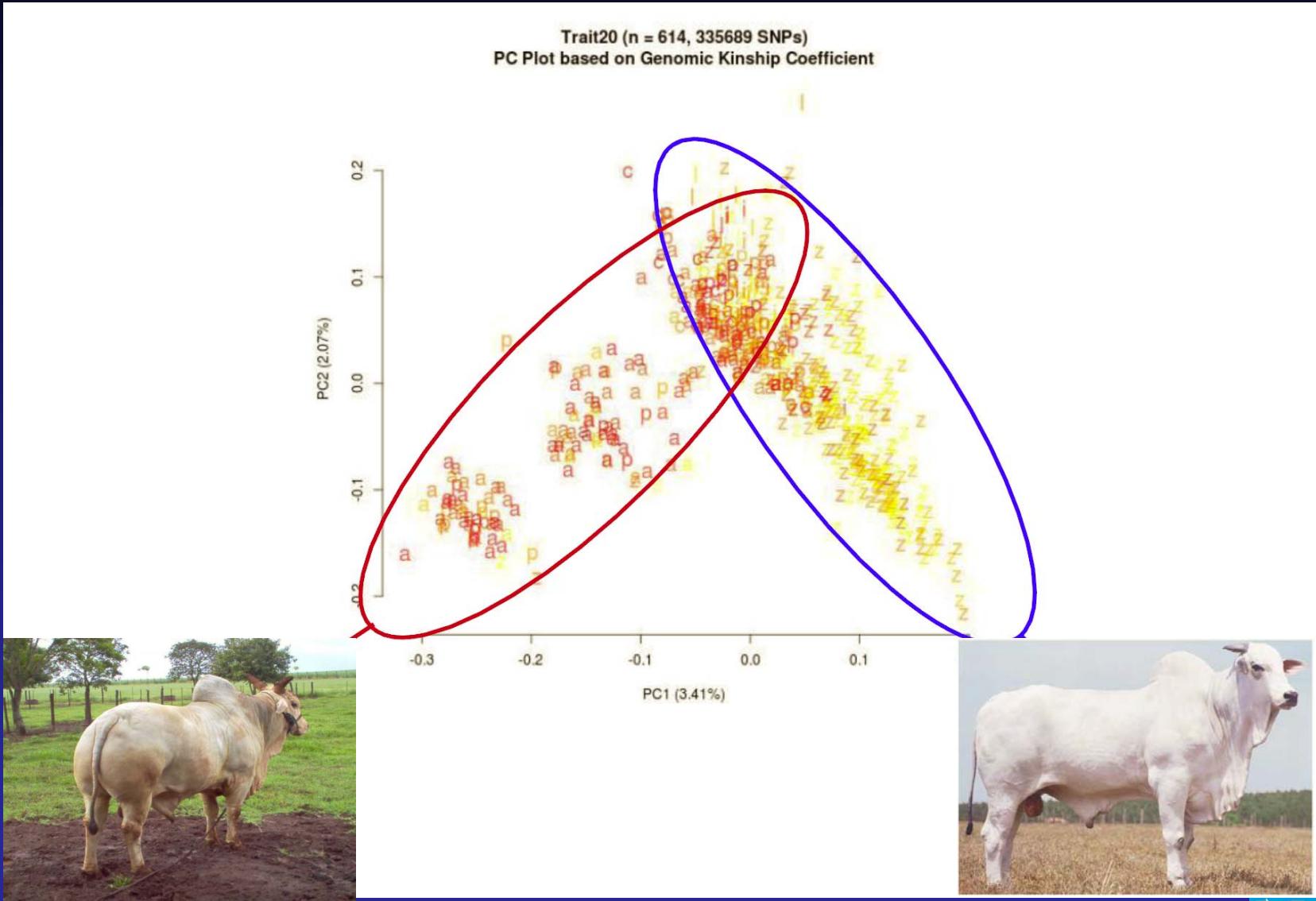


# Diversity of Pakistani Cattle



Proves Bovine high density SNP assays can inform bio-conservation and future research decisions

# Differences in selection in Brazil





# Heat-tolerance in Tropical Cattle



## SLICK hair-coat:

Dominant trait - very short, sleek hair-coat found in tropically adapted cattle demonstrating thermo-tolerance

Move the trait into other breeds to help improve tropical production

# Origins of SLICK?

- Composite Breed of the St. Croix Virgin Island



Senepol



Red Poll



N'Dama



Criollo

# Continued analysis: *Ancestry modeling*

## Slick hair-coated breeds



Senepol  
St. Croix



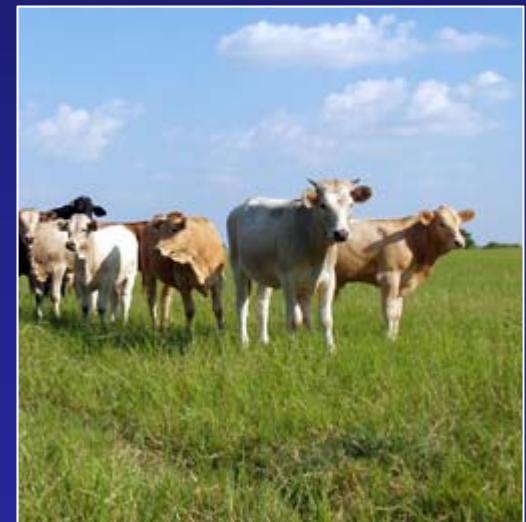
Romosinuano  
Venezuela



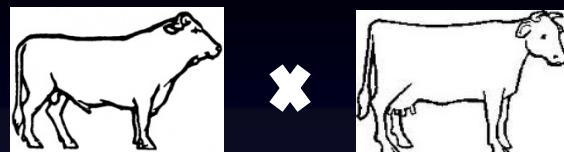
Carora  
Venezuela

## Common link

“Criollo”- a group of cattle descended from Spanish stock imported to the Americas

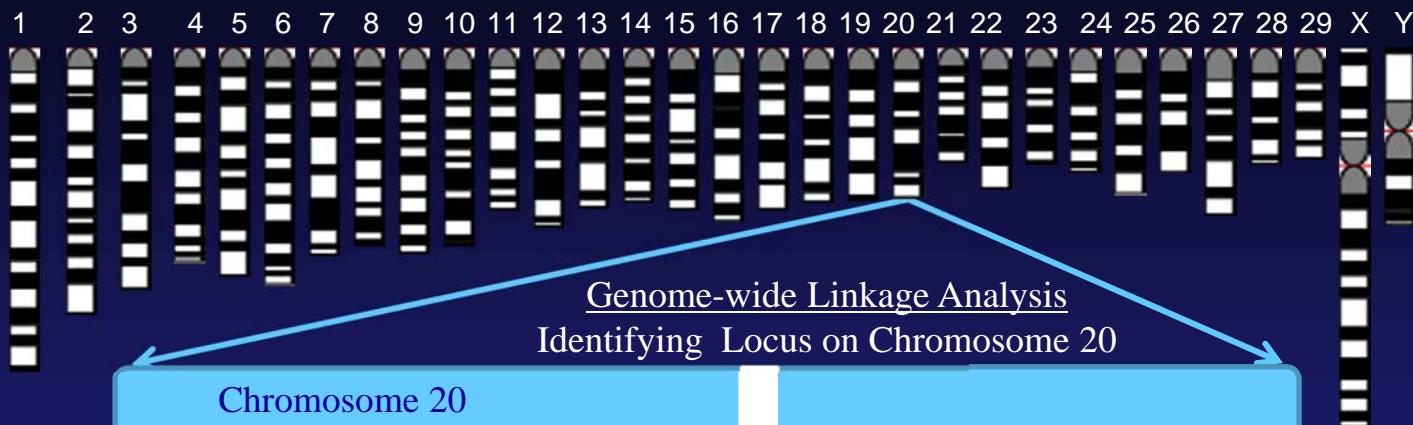


2003  
Olson *et al.*

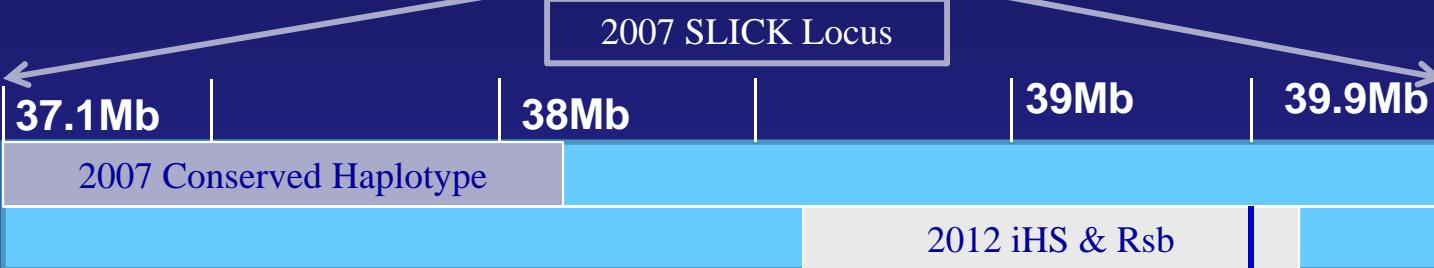


Breeding Trials: Dominant inheritance of a major gene responsible for SLICK phenotype

2007  
Mariasegaram *et al.*

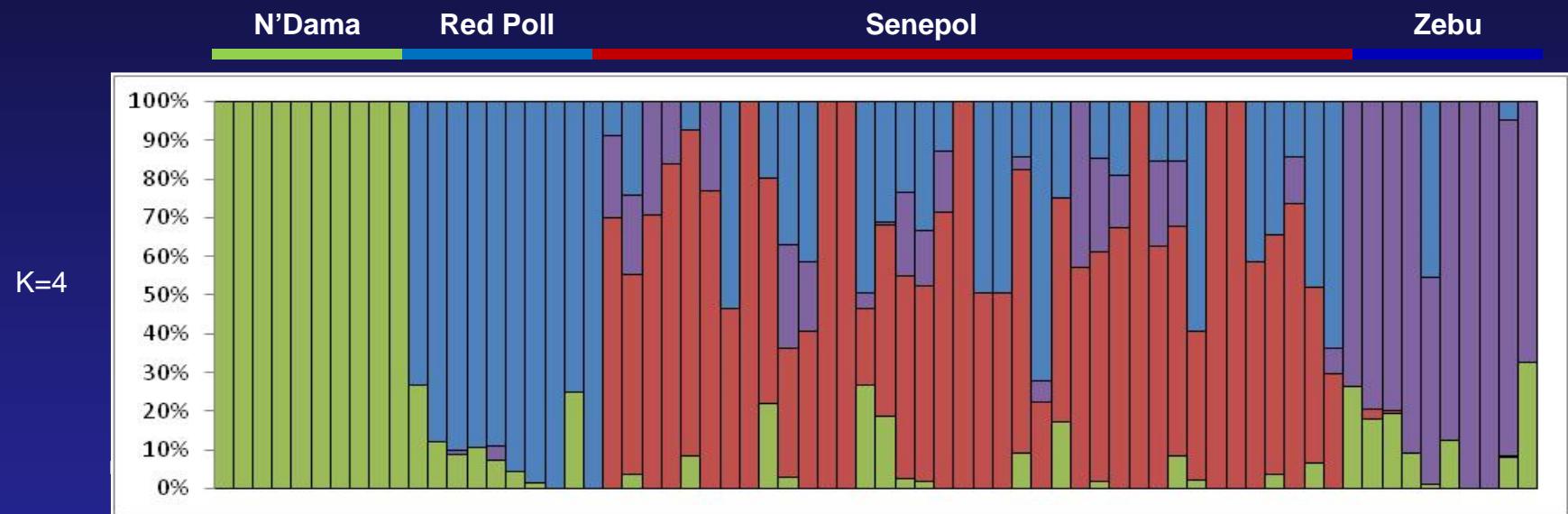


2007  
Mariasegaram *et al.*  
2012  
Flori *et al.*

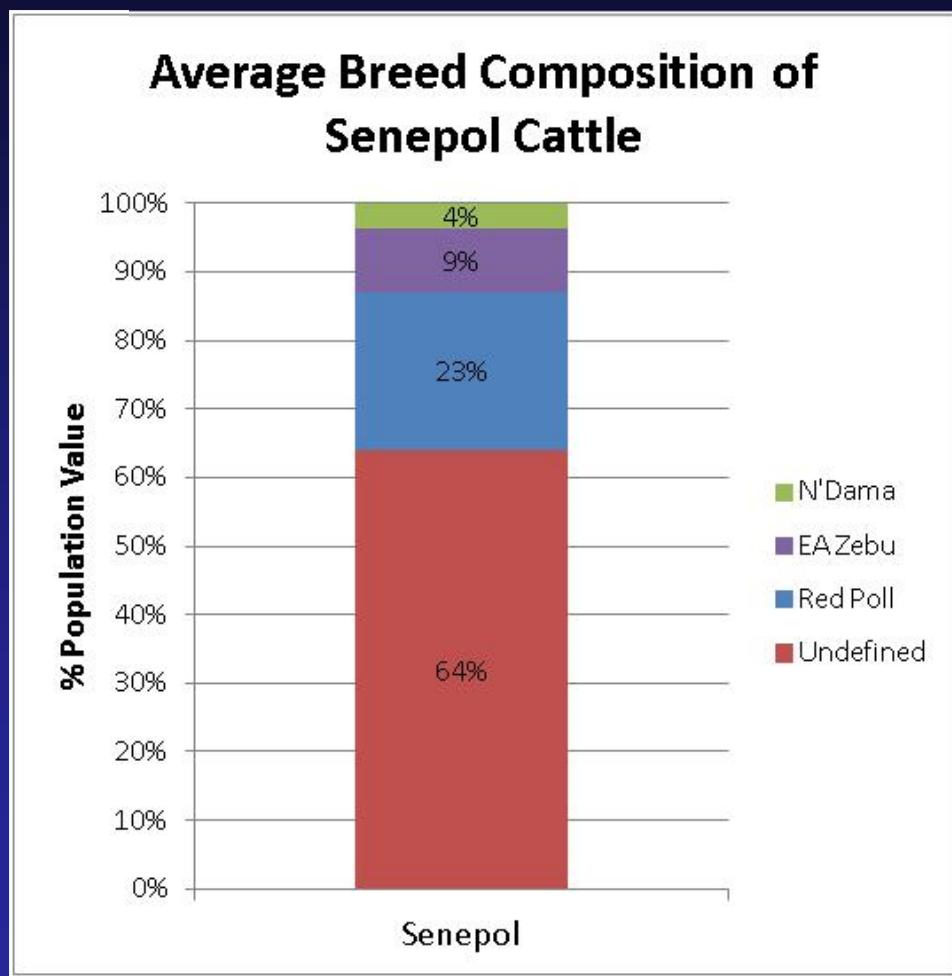


Improve Diagnostic & Identify Causative Mutation

# Genomic verification of Senepol ancestry



# Genomic verification of Senepol ancestry



N'Dama



East  
African  
Zebu



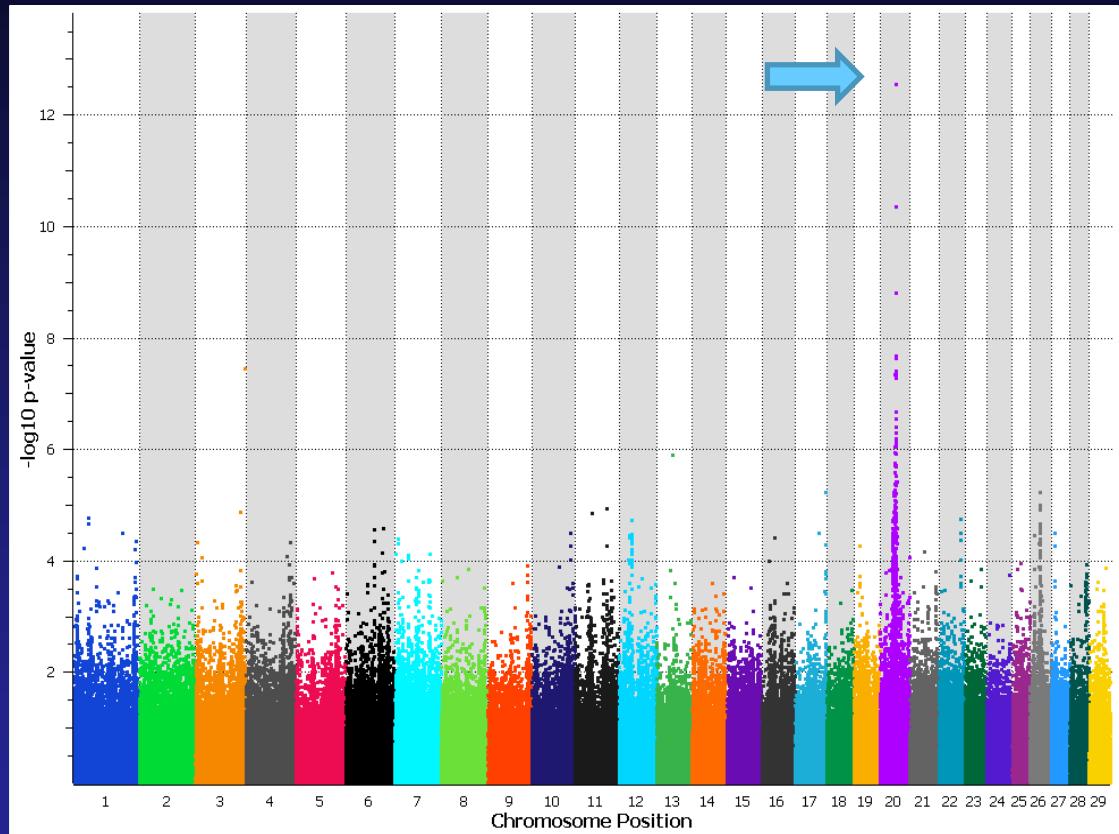
Red Poll



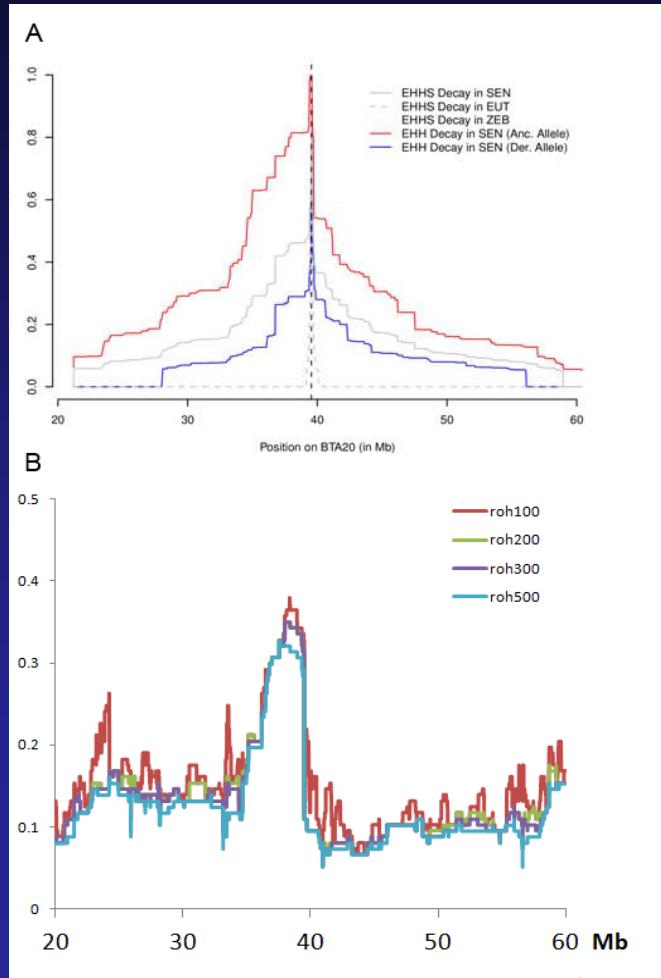
# GWAS of SLICK hair-coat

✓ Use ancestral breeds to balance SLICK cases and controls

- 639,663 SNPs
- Cases = 70
  - Senepol- 36
  - Senepol x Angus-3
  - Senepol/Angus x Angus-1
  - Romosinuano/Angus x Angus-1
  - Romosinuano x Angus-11
  - Holstein x Senepol- 7
  - Carora- 10
- Controls = 64
  - Senepol-2
  - Red Pole-10
  - Angus-10
  - Senepol x Angus-1
  - Senepol/Angus x Angus-1
  - N'Dama-10
  - Holstein- 7
  - Brown Swiss- 10
  - Zebu- 10



# Slick Mapping



(Flori et al. 2012)

**Footprint of Selection** (Flori et al. 2012)

50K iHS in farm Senepol

**Identifying causative variation**

HD SNP / Sequencing in Senepol, Carora, Senepol x Holstein (Kim et al. in progress)

(Kim et al. in progress)

2003  
Olson *et al.*

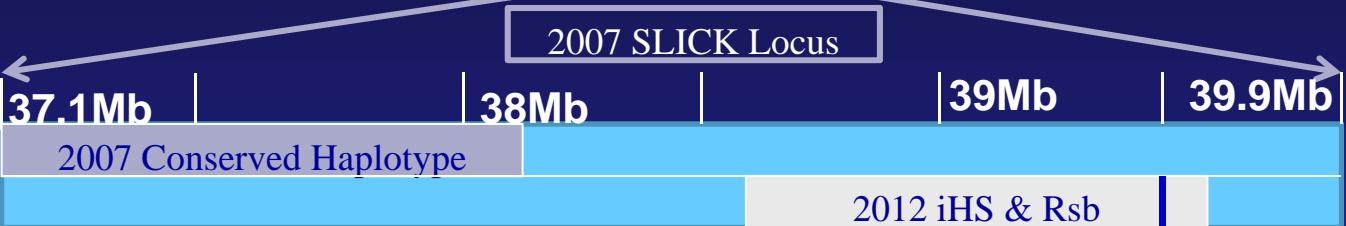


Breeding Trials: Dominant inheritance of a major gene responsible for SLICK phenotype

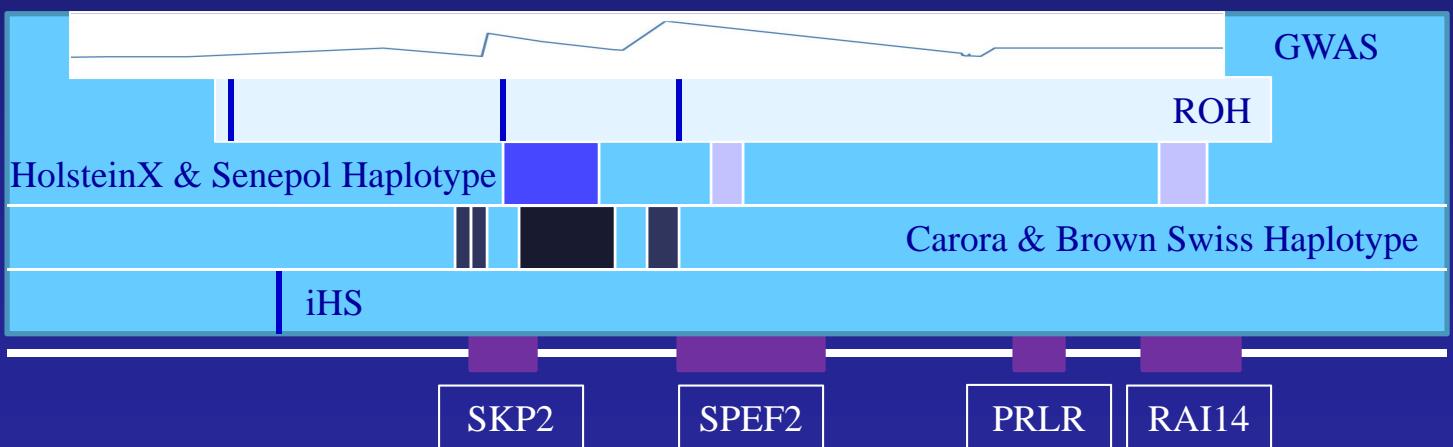
2007  
Mariasegaram *et al.*



2007  
Mariasegaram *et al.*  
2012  
Flori *et al.*



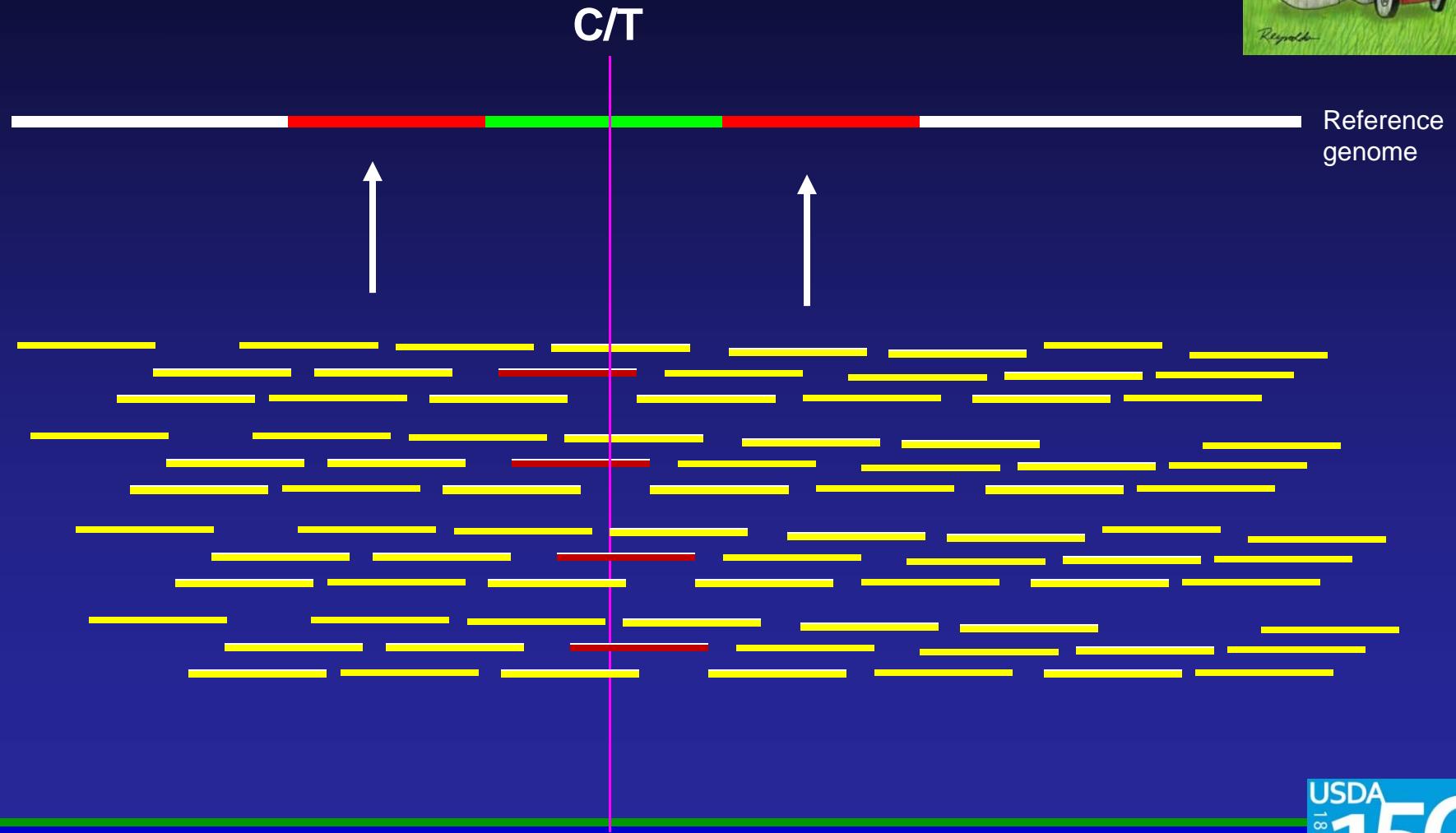
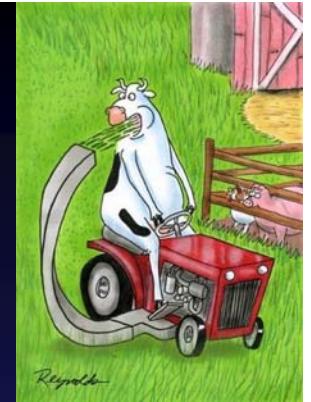
2012  
Kim & Huson *et al.*



# Genome sequence analysis

- **Exome analysis: CoFactor Genomics**
  - 2 Senepol: 1 homozygous, 1 heterozygous
  - 1 Hereford (Dominette)
  - 13 Holstein
  - 7 Brown Swiss
- **Whole-genome sequence analysis: USDA, 5x coverage**
  - 2 Senepol
  - 3 Holstein x Senepol cross individuals
  - 7 Angus
  - 17 Holstein

# Bioinformatics of NGS SNP Discovery



# Current Status - Summary

- Tested 51 new SNP in the SLICK locus region
- 32 of them were accurate in predicting SLICK in Senepol cattle
  - One exception
- Highly predictive in SLICK Holsteins
  - 4 unresolved discrepancies – SLICK phenotype correct?
- None of the SNP were predictive in Carora
  - SLICK phenotypes correct??
- Is there a second SLICK mutation?

# Future Outlook

- Genotype and Sequence Criollo cattle
- Test Genesseek SNP Panel on Australian composite animals
- Deliver the Results to producers as soon as possible
  
- Thanks for your attentions and hospitality!

# Acknowledgements

- **BFGL (USDA-ARS)**
  - Curt Van Tassel
  - George Liu
  - Steve Schroeder
  - Eui-Soo Kim
  - Matthew McClure
  - Heather Huson
  - Adriana Garcia
- **AIPL (USDA-ARS)**
  - Derek Bickhart
- Robert Godfrey – UVI
- Jose Fernando Garcia Lab – UNESP
- Aleksey Zimin – UMD-CBCB
- Jerry Taylor Lab – Univ. MO
- Peter Hanson/Serdal Dikmen – U. Florida
- Tim Olson – Retired
- Art Martinez – PRR
- Red Poll Association
- Senepol Association
- Rita Rizzi – Milan
- Chad Chase – USMARC
- Alta Genetics
- Neogen-Geneseeek
- DNA Genotek

