

# Science

24 April 2009 | \$10

## Genomic Resources and Gene/QTL Discovery in Livestock

Jerry Taylor  
University of Missouri

FAO Intl Tech Conf on Agric Biotech in Dev Count, Guadalajara, March 2<sup>nd</sup> 2010

Livestock Decoded



## ● Available Genomic Tools

- Next generation sequencers
- High throughput SNP genotyping
- Arrays

## ● Available Populations

## ● Prospects for QTL Discovery

- Or why we need Genomic Selection!!





## ● Rapidly Evolving

- MU GAIIx (v4.0 chemistry & v1.4 base calling) yields 2.8 Gb/lane
- New platforms with enhanced capabilities

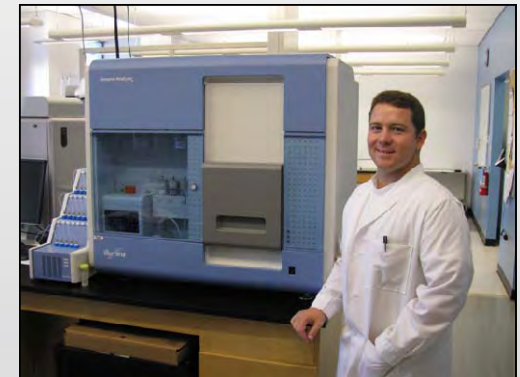
## ● Cost/Mb is decreasing

## ● Numerous Applications

- RNAseq
- Genome resequencing
  - *De novo* assembly
  - Mutation discovery
- ChIPseq
- Small non-coding RNAs
- DNA methylation (esp. Pac Biosciences)

## ● Democratizing Genomics!

- Developing countries and small labs can be genome centers!!!





# MU DNA Core Service Costs

| Library Preparation   | USD   |
|---|-------|
| Genome Sequencing Sample Preparation* (Paired-End Adapters) | \$225 |
| Digital Gene Expression (DGE)- small RNA Sample Preparation | \$445 |
| cDNA Sample Preparation*                                    | \$225 |
| ChIP-Seq Sample Preparation*                                | \$225 |
| mRNA-Seq Sample Preparation**                               | \$500 |



| Analysis                     | USD                                  | Raw Data Output/Flow Cell (Gb) |
|------------------------------|--------------------------------------|--------------------------------|
| Single Read (42 bases)       | \$828.57/lane (\$5,800/Flow Cell)    | 5-6Gb                          |
| Single Read (84 bases)       | \$1,414.29/lane (\$9,900/Flow Cell)  | 11-12Gb                        |
| Paired-End Read (2x42 bases) | \$1,657.14/lane (\$11,600/Flow Cell) | 11-12Gb                        |
| Paired-End Read (2x84 bases) | \$2157.14/lane (\$15,100/Flow Cell)  | 22-23Gb                        |

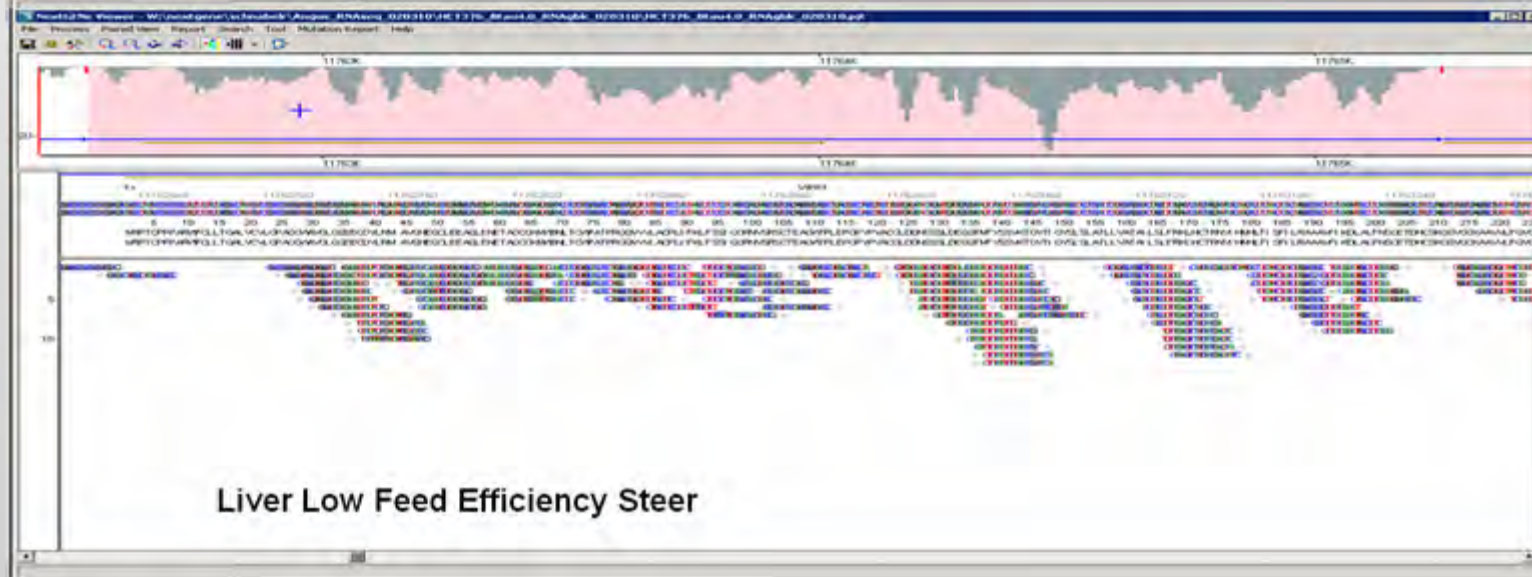
**\$3000 for 1X mammalian genome sequence coverage!!!**



# RNASeq pileup over *VIPR1*



Liver High Feed Efficiency Steer



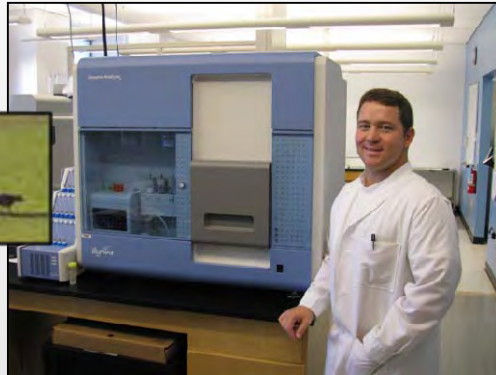
Liver Low Feed Efficiency Steer



# Public Genomic Resources

**Whole genome sequence on African, Asian, European cattle breeds will soon be publicly available**

- Whole genome sequencing of 120 animals from 12 breeds



| Library         | GAllx Lanes | Post-Filter Reads | Total Bases     | Genome Coverage<br>(2.685 Gb = 1X) | Average Read Length<br>(bp) | Unfiltered SNPs<br>and Indels |
|-----------------|-------------|-------------------|-----------------|------------------------------------|-----------------------------|-------------------------------|
| Brahman (10)    | 20          | 335,962,867       | 23,485,537,782  | 8.74                               | 69.91                       | 19,901,328                    |
| East Asian (15) | 32          | 497,458,572       | 35,220,941,134  | 13.11                              | 70.80                       | 18,662,313                    |
| Angus (3)       | 54          | 877,918,406       | 67,112,512,237  | 24.99                              | 76.45                       | 14,797,621                    |
| AN626           | 18          | 310,024,359       | 23,565,486,109  | 8.77                               | 76.01                       |                               |
| AN828           | 18          | 289,562,373       | 22,621,184,219  | 8.42                               | 78.12                       |                               |
| AN1717          | 18          | 278,331,674       | 20,925,841,909  | 7.79                               | 75.18                       |                               |
| All libraries   | 106         | 1,711,339,845     | 125,818,991,153 | 46.85                              | 73.52                       |                               |



# Advantages/Disadvantages

## ● Advantages

- Relatively inexpensive and fast
- Enormous information content
- Species independent technology
- Will make redundant many other forms of technology
  - Microarrays
  - Genotyping



## ● Disadvantages

- Suites of analysis tools only just becoming available
- Require major compute horsepower (~\$40K Servers)
  - Analysis is non-trivial - requires training/experience



- **First generation assays (~50K SNPs) available for Chicken, Pig, Sheep, Cow, Horse, Dog**
  - Under development for salmon, tilapia, catfish
- **Second generation assays (~800K SNPs) available summer 2010**
- **Utility:**
  - Genome-wide association mapping (GWAS)
  - Genomic Selection
  - Population genetics/conservation biology







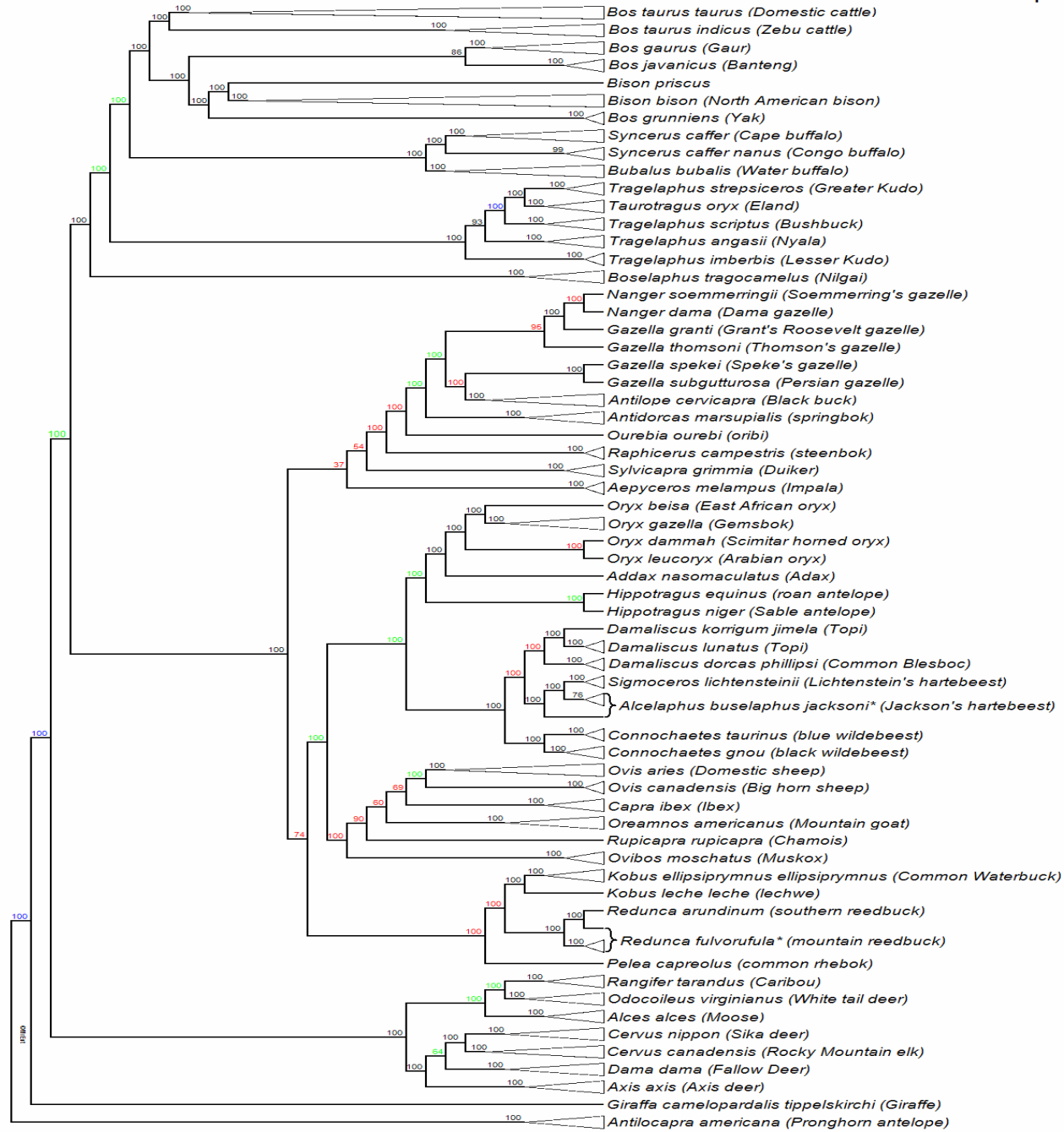
## ● Advantages

- Rapid data generation
- Highly accurate and complete data
- Work across species (30 MY evolution)
- Relatively inexpensive (research)
- Rapid research to commercialization for GS
- Data analysis very similar to BLUP genetic evaluation
- Service providers available (e.g., GeneSeek)
- Available in 96, 384, 1536, 3072, 50K, 800K

## ● Disadvantages

- Relatively expensive (commercialization)
- Available analysis tools (e.g., PLINK, ASREML)
- Contain common variants (SNP ascertained)





| Species  | Number of Diploid Chromosomes |
|--|-------------------------------|
| <i>Bos taurus taurus</i> (Domestic cattle)                     | 60                            |
| <i>Bos taurus indicus</i> (Zebu cattle)                        | 60                            |
| <i>Bos gaurus</i> (Gaur)                                       | 58                            |
| <i>Bos javanicus</i> (Banteng)                                 | 60                            |
| <i>Bison priscus</i>   |                               |
| <i>Bison bison</i> (North American bison)                      | 60                            |
| <i>Bos grunniens</i> (Yak)                                     | 60                            |
| <i>Syncerus caffer</i> (Cape buffalo)                          | 52                            |
| <i>Syncerus caffer nanus</i> (Congo buffalo)                   | 54                            |
| <i>Bubalus bubalis</i> (Water buffalo)                         | 50                            |
| <i>Tragelaphus strepsiceros</i> (Greater Kudu)                 | 32 Female, 31 Male            |
| <i>Taurotragus oryx</i> (Eland)                                | 32 Female, 31 Male            |
| <i>Tragelaphus scriptus</i> (Bushbuck)                         | 34 Female, 33 Male            |
| <i>Tragelaphus angasii</i> (Nyala)                             | 56 Female, 55 Male            |
| <i>Tragelaphus imberbis</i> (Lesser Kudu)                      |                               |
| <i>Boselaphus tragocamelus</i> (Nilgai)                        | 46                            |
| <i>Nanger soemmerringii</i> (Soemmerring's gazelle)            | 36                            |
| <i>Nanger dama</i> (Dama gazelle)                              | 38-40                         |
| <i>Gazella granti</i> (Grant's Roosevelt gazelle)              | 30 Female, 31 Male            |
| <i>Gazella thomsoni</i> (Thomson's gazelle)                    | 58                            |
| <i>Gazella spekei</i> (Speke's gazelle)                        | 32 Female, 33 Male            |
| <i>Gazella subgutturosa</i> (Persian gazelle)                  | 32                            |
| <i>Antilope cervicapra</i> (Black buck)                        | 30-32 Female, 31-33 Male      |
| <i>Antidorcas marsupialis</i> (springbok)                      | 56                            |
| <i>Ourebia ourebi</i> (oribi)                                  |                               |
| <i>Raphicerus campestris</i> (steenbok)                        | 30                            |
| <i>Sylvicapra grimmia</i> (Duiker)                             |                               |
| <i>Aepyceros melampus</i> (Impala)                             | 58-60                         |
| <i>Oryx beisa</i> (East African oryx)                          | 58                            |
| <i>Oryx gazella</i> (Gemsbok)                                  | 56                            |
| <i>Oryx dammah</i> (Scimitar horned oryx)                      | 56-58                         |
| <i>Oryx leucoryx</i> (Arabian oryx)                            | 57-58                         |
| <i>Addax nasomaculatus</i> (Adax)                              |                               |
| <i>Hippotragus equinus</i> (roan antelope)                     | 60                            |
| <i>Hippotragus niger</i> (Sable antelope)                      | 60                            |
| <i>Damaliscus korrigum jimela</i> (Topi)                       |                               |
| <i>Damaliscus lunatus</i> (Topi)                               | 36                            |
| <i>Damaliscus dorcas phillipsi</i> (Common Blesboc)            | 38                            |
| <i>Sigmoceros lichtensteinii</i> (Lichtenstein's hartebeest)   |                               |
| <i>Alcelaphus buselaphus jacksoni</i> * (Jackson's hartebeest) | 40                            |
| <i>Connochaetes taurinus</i> (blue wildebeest)                 | 58                            |
| <i>Connochaetes gnou</i> (black wildebeest)                    | 58                            |
| <i>Ovis aries</i> (Domestic sheep)                             | 54                            |
| <i>Ovis canadensis</i> (Big horn sheep)                        |                               |
| <i>Capra ibex</i> (Ibex)                                       | 60                            |
| <i>Oreamnos americanus</i> (Mountain goat)                     | 42                            |
| <i>Rupicapra rupicapra</i> (Chamois)                           | 58                            |
| <i>Ovibos moschatus</i> (Muskox)                               | 48                            |
| <i>Kobus ellipsiprymnus ellipsiprymnus</i> (Common Waterbuck)  | 50                            |
| <i>Kobus leche leche</i> (lechwe)                              | 48                            |
| <i>Redunca arundinum</i> (southern reedbuck)                   |                               |
| <i>Redunca fulvorufula</i> * (mountain reedbuck)               | 56-57                         |
| <i>Pelea capreolus</i> (common rhebok)                         |                               |
| <i>Rangifer tarandus</i> (Caribou)                             | 70                            |
| <i>Odocoileus virginianus</i> (White tail deer)                | 70                            |
| <i>Alces alces</i> (Moose)                                     | 70                            |
| <i>Cervus nippon</i> (Sika deer)                               | 68                            |
| <i>Cervus canadensis</i> (Rocky Mountain elk)                  |                               |
| <i>Dama dama</i> (Fallow Deer)                                 | 68                            |
| <i>Axis axis</i> (Axis deer)                                   | 66                            |
| <i>Giraffa camelopardalis tippelskirchi</i> (Giraffe)          | 30                            |
| <i>Antilocapra americana</i> (Pronghorn antelope)              | 58                            |

**Bovidae**



Bovinae



Antilopinae



Cephalophus  
Alcelaphinae



Caprinae  
Reduncinae

**Cervidae**



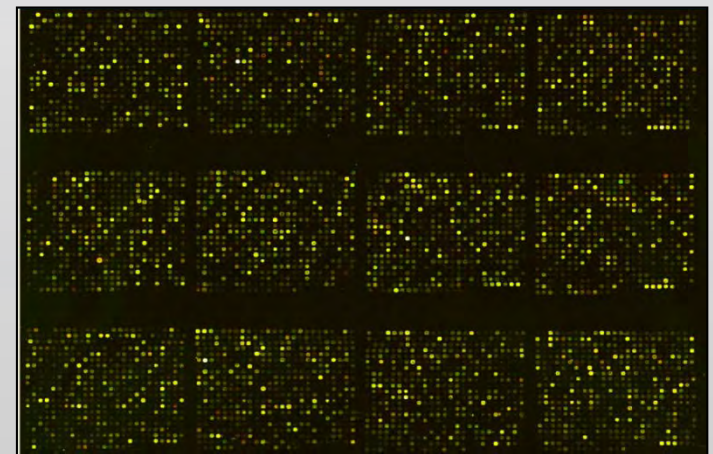
**Giraffidae  
Antilocapridae**

When compared to supermatrix analysis of Marcot (1)

- Novel Relationship
- Improvement in bootstrap support >10%
- Improvement in bootstrap support >5%



- **Genome-wide oligo microarrays available for chicken, cattle, sheep, pigs**
- **Availably commercially (Agilent, Illumina, Affymetrix) or through species consortia (BOM: [bovineoligo.org](http://bovineoligo.org); SPAM: [pigoligoarray.org](http://pigoligoarray.org))**
- **CNV, Sequence Capture (e.g., exon, target regions), CpG arrays by custom design (e.g., Nimblegen, Agilent)**





## ● Advantages

- Relatively inexpensive (\$55 per array + RNA extraction and labeling kits etc = ~\$250 per sample) but require replication!!!!
- Analysis tools readily available (e.g., R open source)
- Available for many species

## ● Disadvantages

- Nontrivial chemistry
- Non exhaustive (not all genes are queried)
- Probes only for one exon (alternate splicing, promoter and 3' UTR utilization missed)
- Not available for all species

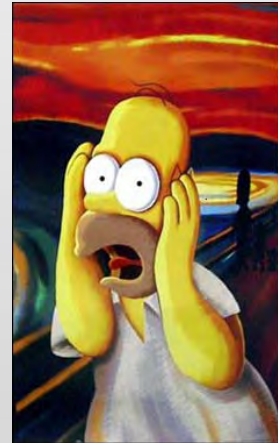
## ● All Arrays

- NextGen sequencing will make most arrays redundant



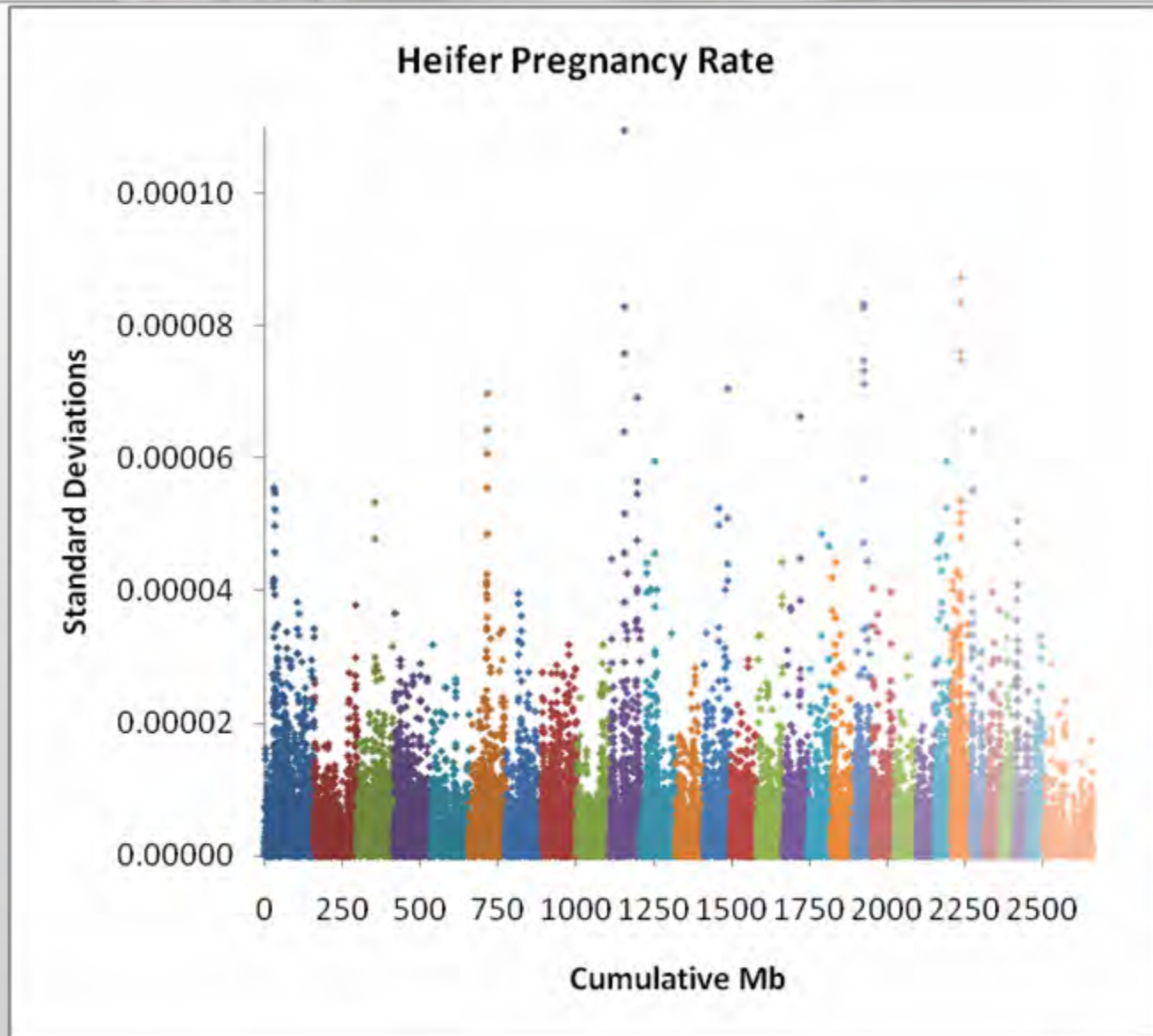
# Available Populations

- **The greatest limitation to all genomic technologies previous and imminent is that we do not have populations of animals with high-quality DNA, tissue samples and phenotypes...**
- **In the US**
  - The only population with adequate numbers of animals to build accurate GS prediction models is Holsteins
  - Angus 2,700 AI bulls with EPDs
  - Angus 3,500 steers with growth and carcass composition
  - Angus 700 steers with feed intake
  - Limousin 2,200 registered males and females with EPDs
  - Angus (660), Limousin (285), Charolais (702), Hereford (1192) , Simmental (521) with meat tenderness
- **Worldwide priority is assembly of populations and phenotypes**



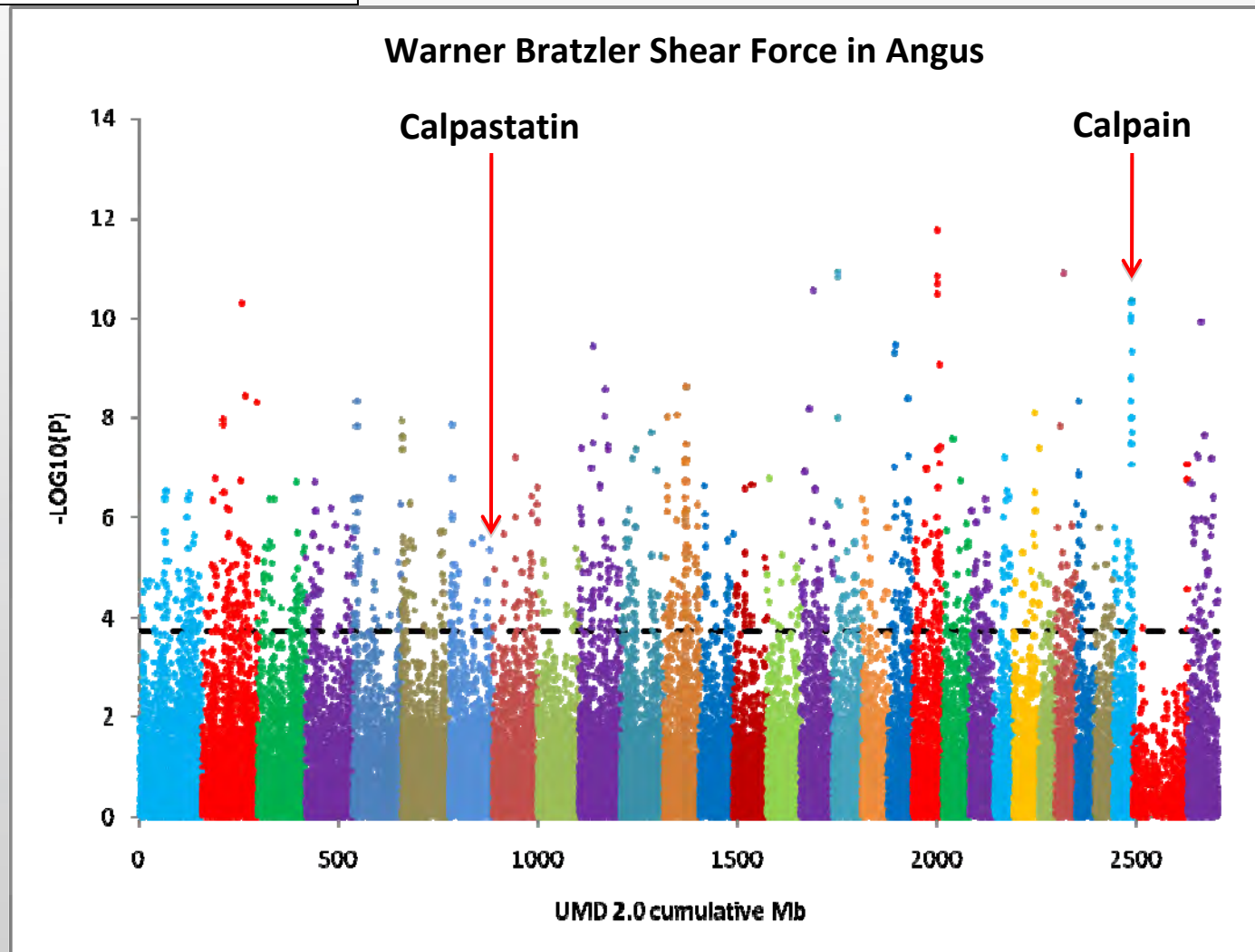


# Too Many Genes for MAS!!!



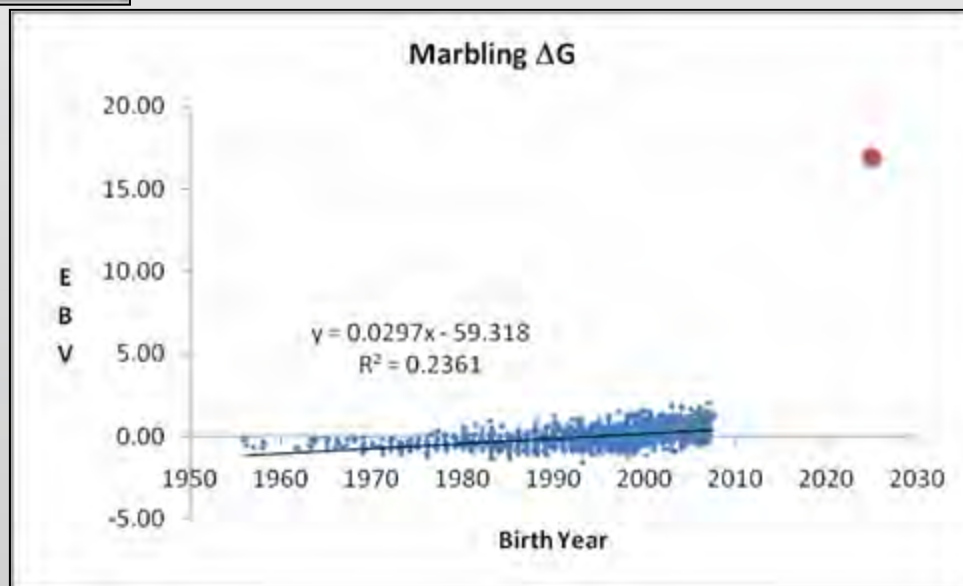
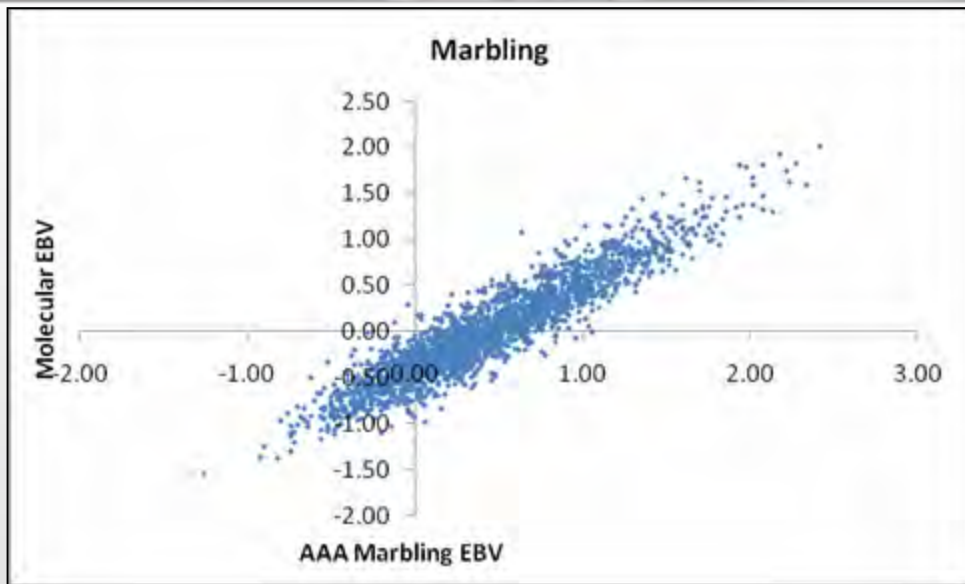
# Use Experimental Populations to Produce EBVs for an Entire Breed

| Breed     | No. Sires | Sire Identified | Sire Unknown | Total |
|-----------|-----------|-----------------|--------------|-------|
|           |           | Steers          | Steers       |       |
| Angus     | 20        | 520             | 140          | 660   |
| Hereford  | 29        | 566             | 626          | 1192  |
| Limousin  | 23        | 285             | 0            | 285   |
| Charolais | 18        | 350             | 352          | 702   |
| Simmental | 24        | 374             | 147          | 521   |
| Total     | 114       | 2095            | 1265         | 3360  |





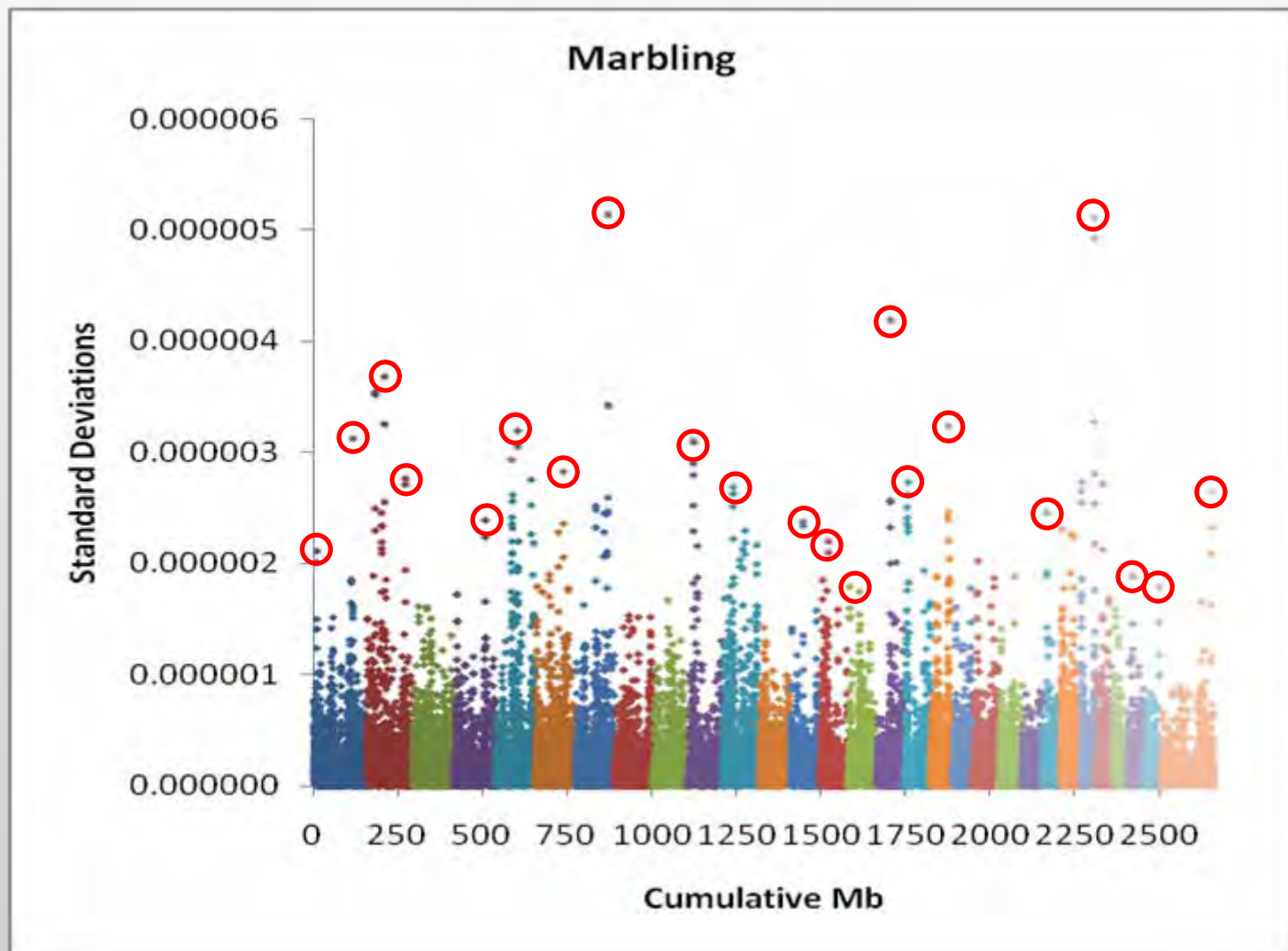
# Genomic Selection!







# Genomic Selection with Reduced SNP Sets



The End

