Translational genomics:

Progress in the adaptation of information derived from genome technologies for animal improvement

Alison Van Eenennaam

Animal Genomics and Biotechnology

Cooperative Extension Specialist

Department of Animal Science

University of California, Davis

alvaneenennaam@ucdavis.edu

(530) 752-7942

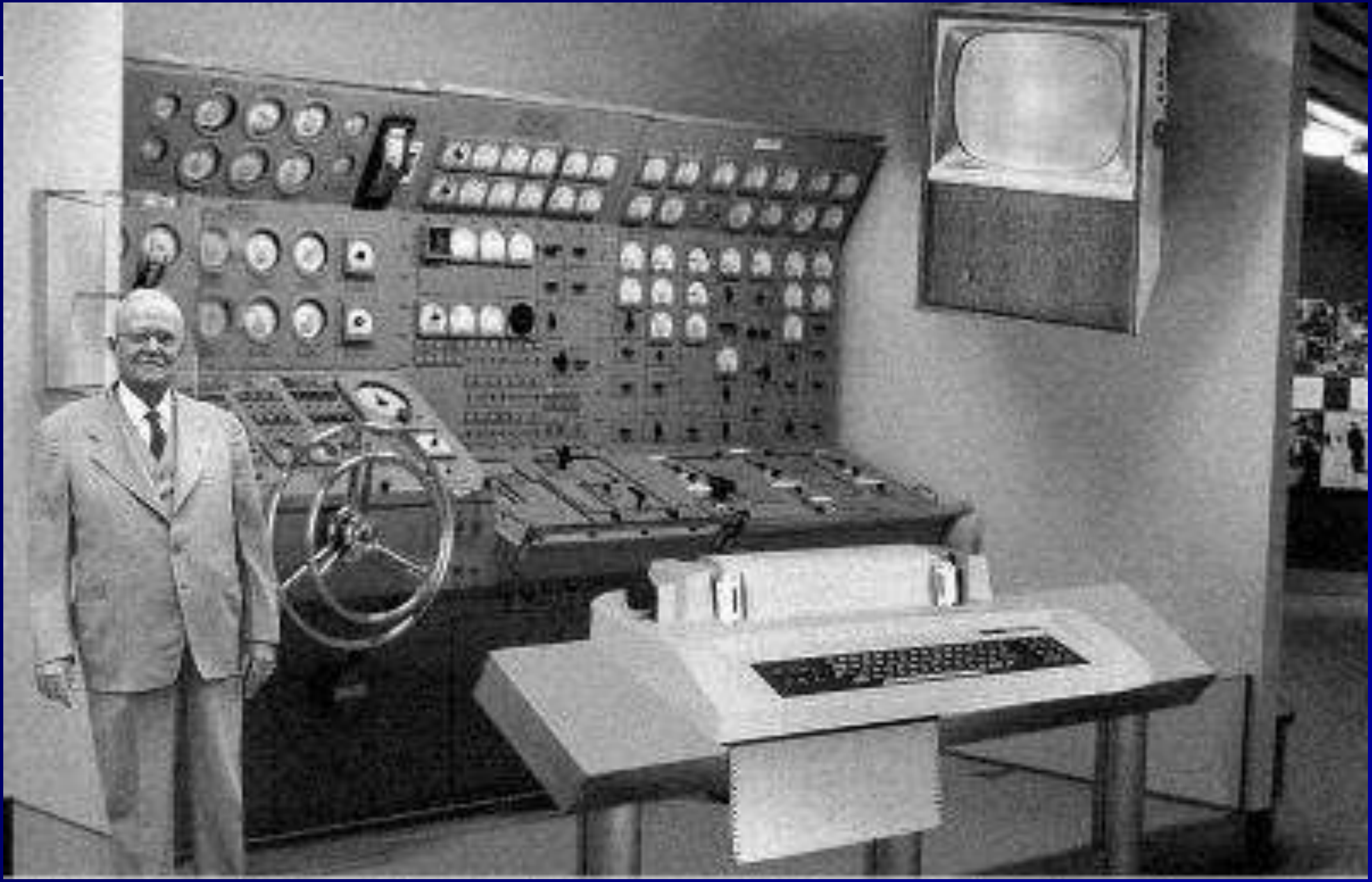
animalscience.ucdavis.edu/animalbiotech

“*Translational genomics*” is defined as the adaptation of information derived from genome technologies for animal improvement

“We believe DNA marker profiles will become widely used in livestock in the near future as the cost decreases and the benefits increase. In fact, a major research objective may be to make best use of this DNA data in commercial animal production”

Goddard, M. E., and B. J. Hayes. 2007. Genomic selection. Journal of Animal Breeding and Genetics 124: 323-330.

“1954 version of what 'home computers' might look like in 50 years time (i.e. 2004)”



Wrong Expert Predictions

I think there's a world market for about five computers.

Thomas J. Watson, chairman of the board of IBM. 1943

There is no reason anyone would want a computer in their home.

Ken Olson, president of Digital Equipment Corp. 1977

The cost for a "large" genome scan (defined as 18 chromosomes* 7 markers (i.e. 126 markers!) * $4/marker) = $504

Ben Hayes and Mike Goddard, 2003. Evaluation of marker assisted selection in pig enterprises. Livestock Production Science 81:197-211.

“what escaped their vision was that science might come up with new and different ways of commercializing and using new technologies.”

Overview

- **What is working well**
  - Parentage
  - Identification of recessive/single trait defects
  - Dairy genomic selection
- **What is not working so well**
  - Beef genomic selection
- **What does the future hold?**
  - Roadblocks to translational genomics
  - Some solutions and future prospects

Benefits of DNA-based parentage identification

- Correct pedigree errors thereby improving the rate of genetic gain
- Enables the use of multi-sire breeding pasture
  - Higher fertility
  - Elimination of sire failure
  - Tighter calving season
- Reduces the need for different breeding pastures
  - Allows for better pasture management
  - Less sorting and working of animals into different groups
- Reduces the need to disturb newborn animals
  - Labor savings so can focus on concentrate on offspring survival
  - Worker safety improvement
  - Better bonding of offspring with dam
- Enables the development of commercial-ranch genetic evaluations

McEwan, J. C. 2007 Current status and future of genomic selection. Proceedings of the New Zealand Society of Animal Production 67: 147-152.

Paternity analysis on commercial cattle ranch

- Blood collected on FTA cards from **27** herd sires and **624** calves derived from a multiple-sire pasture
- Analyzed using a 28 SNP panel in 2005

A. L. Van Eenennaam, R. L. Weaber, D. J. Drake, M. C. T. Penedo, R. L. Quaas , D. J. Garrick, E. J. Pollak. 2007. DNA-based paternity analysis and genetic evaluation in a large commercial cattle ranch setting. Journal of Animal Science. 85:3159–3169

	28 SNP Panel – 27 sires 2005 (PE=95.5%)	
One sire assigned	**175**	**23.3%**
More than one sire	**420**	**67.3%**
All excluded	**29**	**4.6%**
TOTAL	**624**	

Translational considerations for migrating to SNPs for parentage

It is likely that SNP markers will replace alternatives (i.e. microsatellites) over the next 5 or so years

- How do you switch over from microsatellites to SNPS when a lot of historical information is stored as microsatellites?
- Which SNP genotyping platform should be used and how many and which SNP markers should be included in the panel?
- What should be the number of compared loci cutoff in the case of incomplete genotyping?
- How many exclusions (as a function of number of compared loci) should be allowed to account for genotyping errors – platform dependent?
- Which sample type works best for producers to collect and genotyping entities to run?

DNA-based parentage identification has shown significant uptake by deer, cattle, and sheep breeders in New Zealand

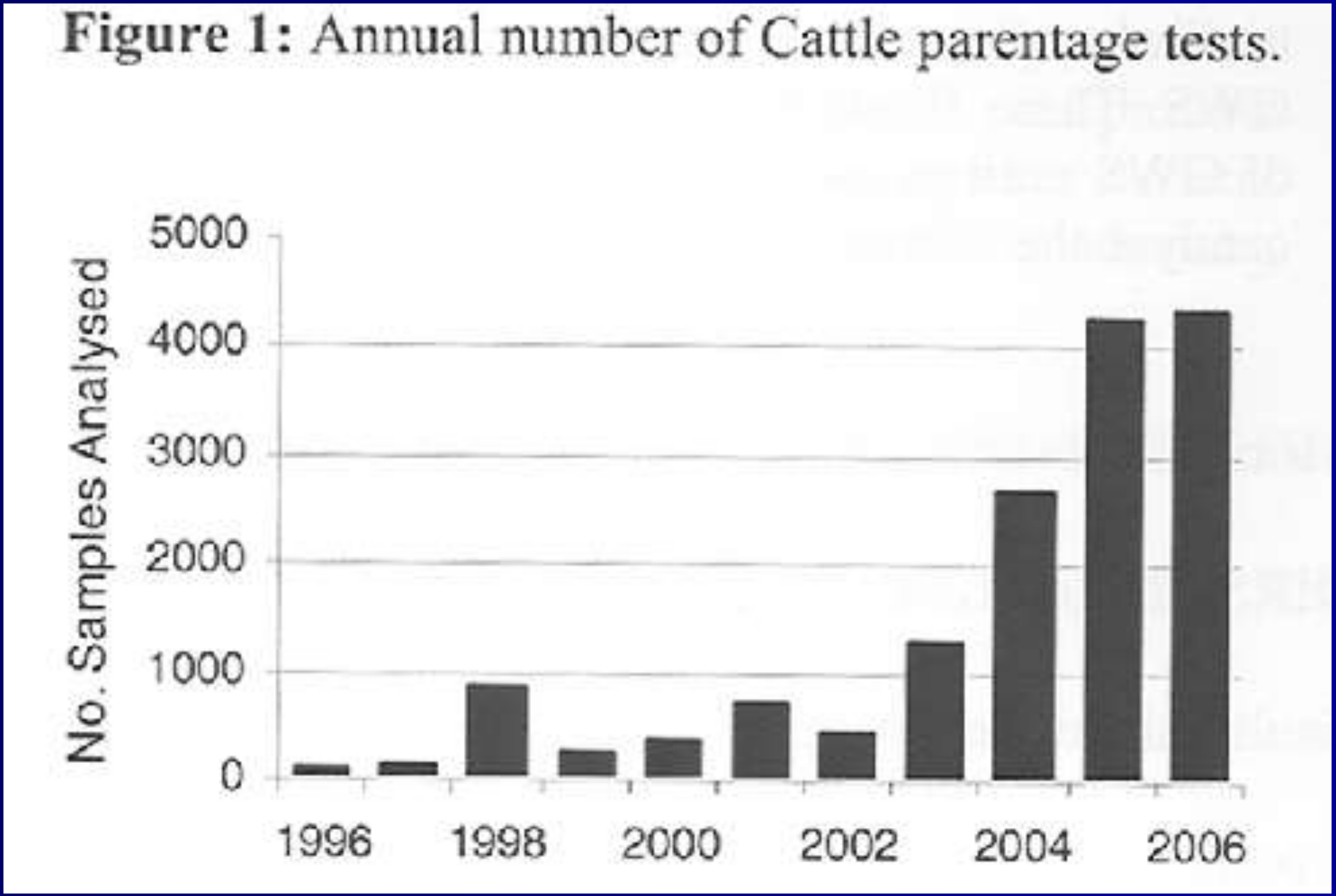


Figure 1: Annual number of Cattle parentage tests.

Crawford, A. M., R. M. Anderson, and K. M. McEwan. 2007. Uptake of DNA testing by the livestock industries of New Zealand. Proceedings of the New Zealand Society of Animal Production 67:168-174

Industry adoption looking at New Zealand as an example

- >200,000 dairy parentage tests out of a herd of about 4.6M cows. ~10% of the commercial tier. Most of these tests are SNP based.
- 20% of the ram, and 30% of the deer breeding industry (majority of stag breeders); mostly microsatellite-based parentage tests
- Emerging area in NZ is parentage testing in aquaculture species

- One of the reasons for the widespread adoption of this technology is the development of an integrated ID and collection system
- This is especially important for lower value animals such as sheep
- If DNA samples are already being collected for parentage verification or as part of a national animal identification scheme, then other DNA technologies can be introduced cost-effectively.

McEwan, J. C. 2007 Current status and future of genomic selection. Proceedings of the New Zealand Society of Animal Production 67: 147-152.; and pers. comm.

A key issue in commercial situations is ease of DNA sampling, tracking and quality of resultant DNA

Reconcile of hair and meat samples (based on 427 records)

- 427 meat samples (3K genotyped) and hair samples (99 SNP parentage genotyped)
  - – 149 had no exclusions with hair (35%)
  - – 185 had 1 exclusion (43%)
  - – 31 had 2 exclusions (7%)
  - – 10 had 3 exclusions (2%)
  - – 31 had 4-10 exclusions (7%)
  - – 21 > 10 exclusions (5%)

Genetic Abnormalities

Images from an article by David S. Buchanan, Department of Animal Sciences, North Dakota State University

http://www.ag.ndsu.edu/williamscountyextension/livestock/genetic-defects-in-cattle

Compare dwarfism response in the 50s to the response to curly calf (AM)

A Triple Compounded

EILEENMERE

Look at the Blood Line Wealth concentrated in one animal. "SHORT SNORTER"

An early '50's advertisement that superimposed a measuring stick in the picture of this bull who was nick-named "Short Snorter."

Based upon his height and age, he was less than a frame score 1.

Image from https://www.msu.edu/~ritchieh/historical/shortsnorter.jpg

Curly calf – Arthrogryposis multiplex

- From a scientific standpoint, AM is the complete deletion of a segment of DNA that encompasses two different genes
- One of these genes is expressed at a crucial time in the development of nerve and muscle tissue. The mutation results in no protein being produced from this gene and therefore it is unable to carry out its normal function so homozygotes show phenotype
- Dr. David Stefan of the University of Nebraska and Dr. Jon Beever of the University of Illinois worked to develop a genetic test from September – October, 2008

From September 8 – November 3, 2008 identified genetic problem, developed test, and released carrier status of 736 bulls!

- In the 10 months following the release of the test, the AAA posted the results of tests for AM on about 90,000 cattle.
- These AM test costs less than $30.
- Of these, almost 5,000 bulls and more than 13,000 heifers have tested as carriers of AM. **That leaves more than 22,000 bulls and more than 50,000 heifers which tested as free of AM.**

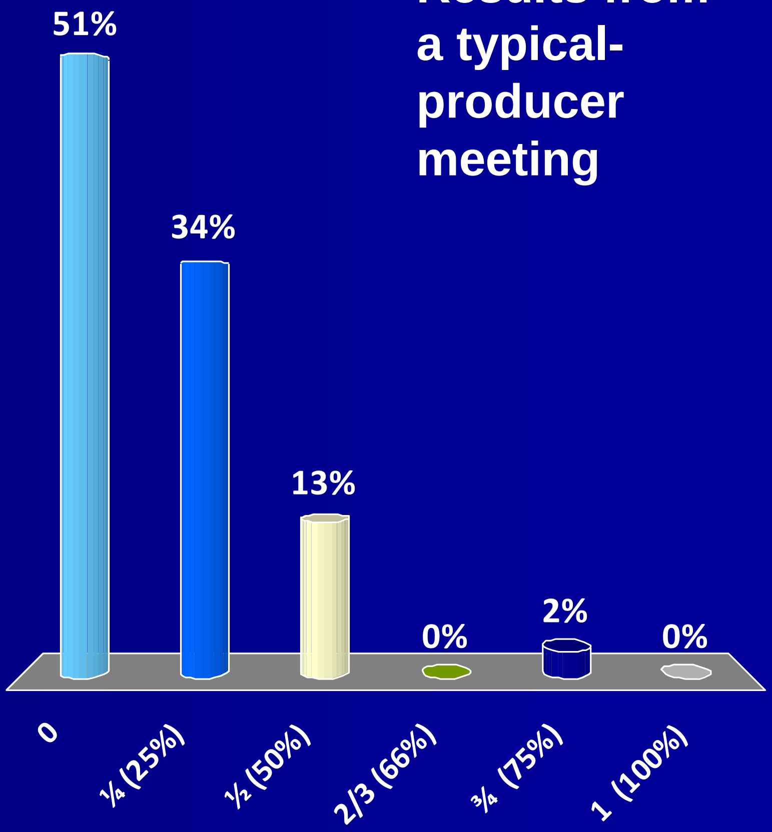
From: Buchanan, D.S. Genetic Defects in Cattle.
http://www.ag.ndsu.edu/williamscountyextension/livestock/genetic-defects-in-cattle

Early extension education about dwarfism explaining carriers and inheritance

Image from Special CollectionsUniversity Libraries, Virginia Tech: http://spec.lib.vt.edu/imagebase/agextension/boxseven/screen/AGR3618.jpg

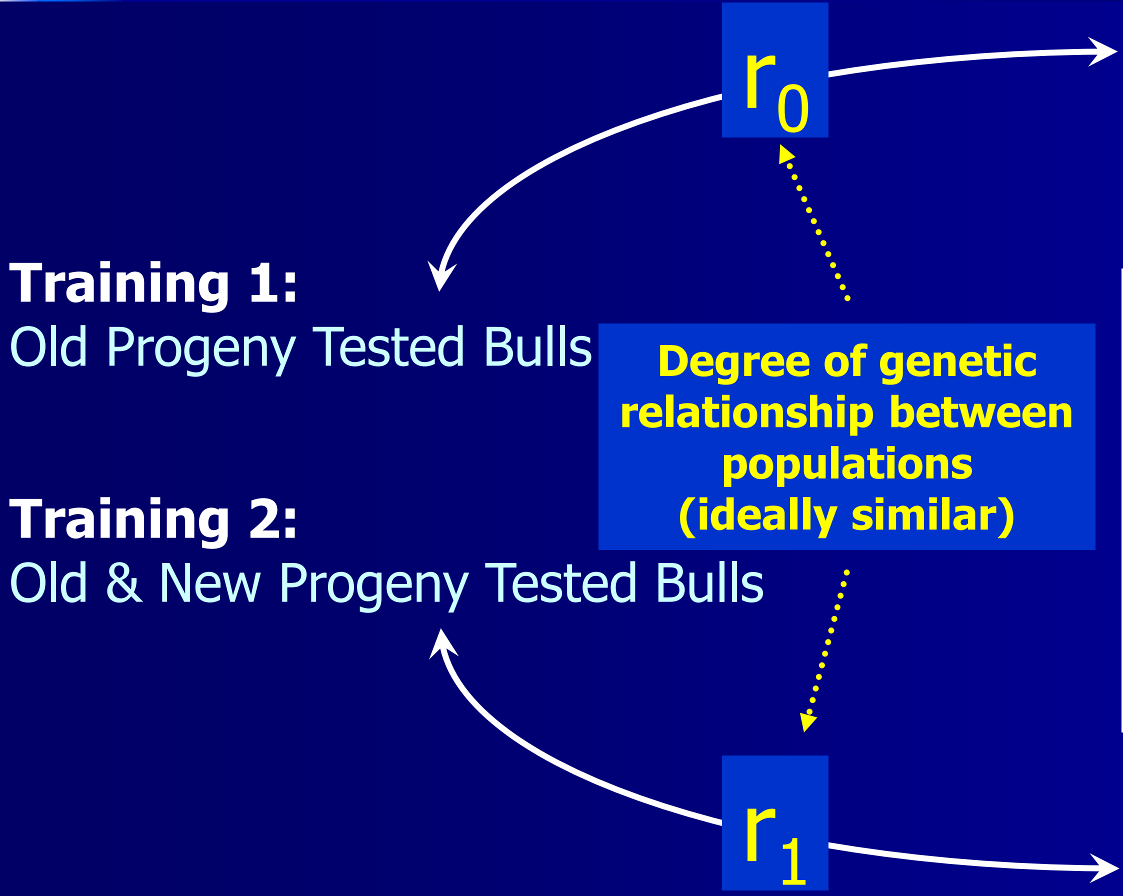
If you breed a curly calf carrier cow (AMC) to an curly calf free bull (AMF), what is the chance that the offspring will be stillborn as a result of being curly calf?

1. 0
2. ¼ (25%)
3. ½ (50%)
4. 2/3 (66%)
5. ¾ (75%)
6. 1 (100%)



Results from a typical-producer meeting

Moving onto Genomic Selection



r_0

Validation:
New Progeny
Tested Bulls

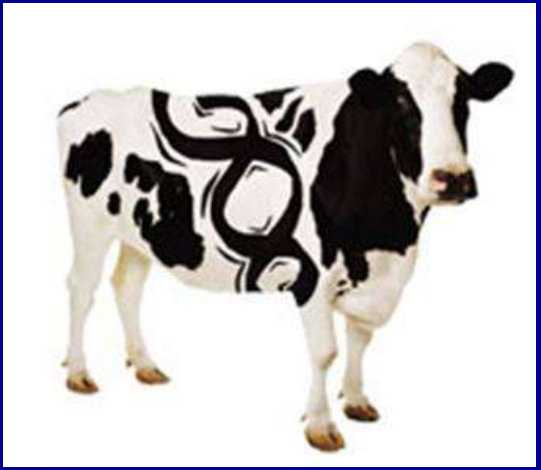
Training 1:
Old Progeny Tested Bulls

Degree of genetic relationship between populations (ideally similar)

Training 2:
Old & New Progeny Tested Bulls

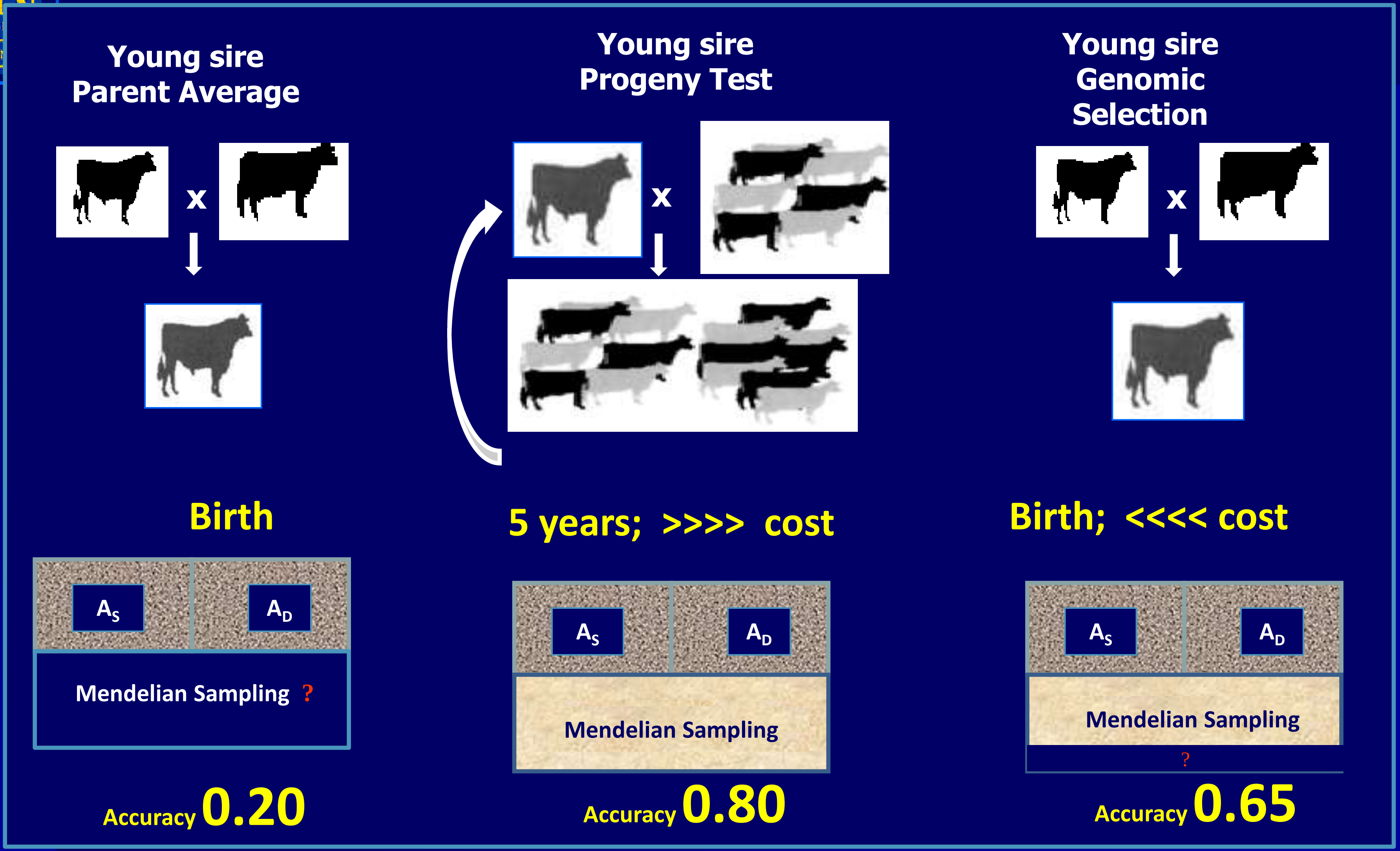
r_1

Application:
New Sire
Candidates



Slide courtesy of Marc Thallman, US MARC

Breeding value prediction in Dairy Sires



Young sire
Parent Average

x

Birth

A_S A_D

Mendelian Sampling ?

Accuracy **0.20**

Young sire
Progeny Test

x

5 years; >>>> cost

A_S A_D

Mendelian Sampling

Accuracy **0.80**

Young sire
Genomic
Selection

x

Birth; <<<< cost

A_S A_D

Mendelian Sampling

?

Accuracy **0.65**

Genomic selection can help breeders identify animals with superior breeding values at a young age

ΔG = ***intensity of selection X***

accuracy of selection X

genetic variation in the population /

generation interval)

Dairy industry suited to WGS

- High use of AI
- Clear selection goal
- One breed used extensively
- Large number of high accuracy A.I. sires for training
- Extensive, uniform collection of data on traits
- Central evaluation (AIPL) receiving genotypes
- Obvious way to increase rate of genetic gain
- AI companies funding the genotyping because they get a clear cost savings in terms of young sire program

Translational Questions for other animal industries

- How many phenotypic records are required in the initial experiment estimating the effect of chromosome segments?
- How many markers are needed– 50K, 800K, whole genome?
- How does the relationship between the training population and the selection candidate affect accuracy?
- How often do chromosome segment effects need to be re-estimated?
- Do predictions work across breeds?
- What is the value generated by the increased accuracy?
- Does this technology change optimal breeding program design?

Accuracy of the prediction equation proportional to:

$$\frac{Th^2}{N_eL}$$

T: total number of animals in the training population

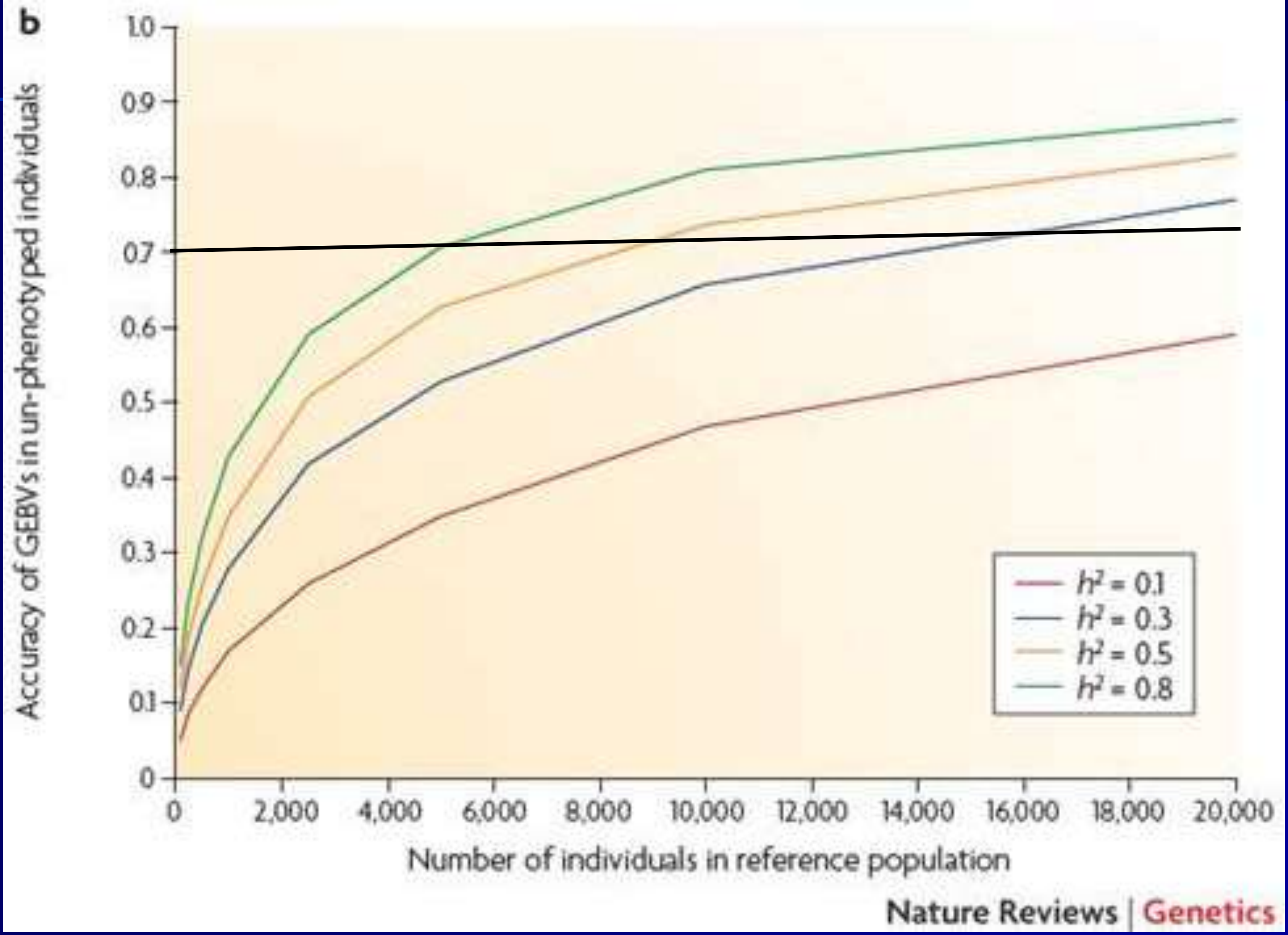
h^2: heritability of the trait

L : length of chromosomes (in Morgans)

N_e: effective population size

Also influenced by trait architecture, number of markers, availability of economically-relevant phenotypes, and relationship between animals in the training and target population

Effect of number of animals on accuracy of prediction equation (for a N_e of 100)



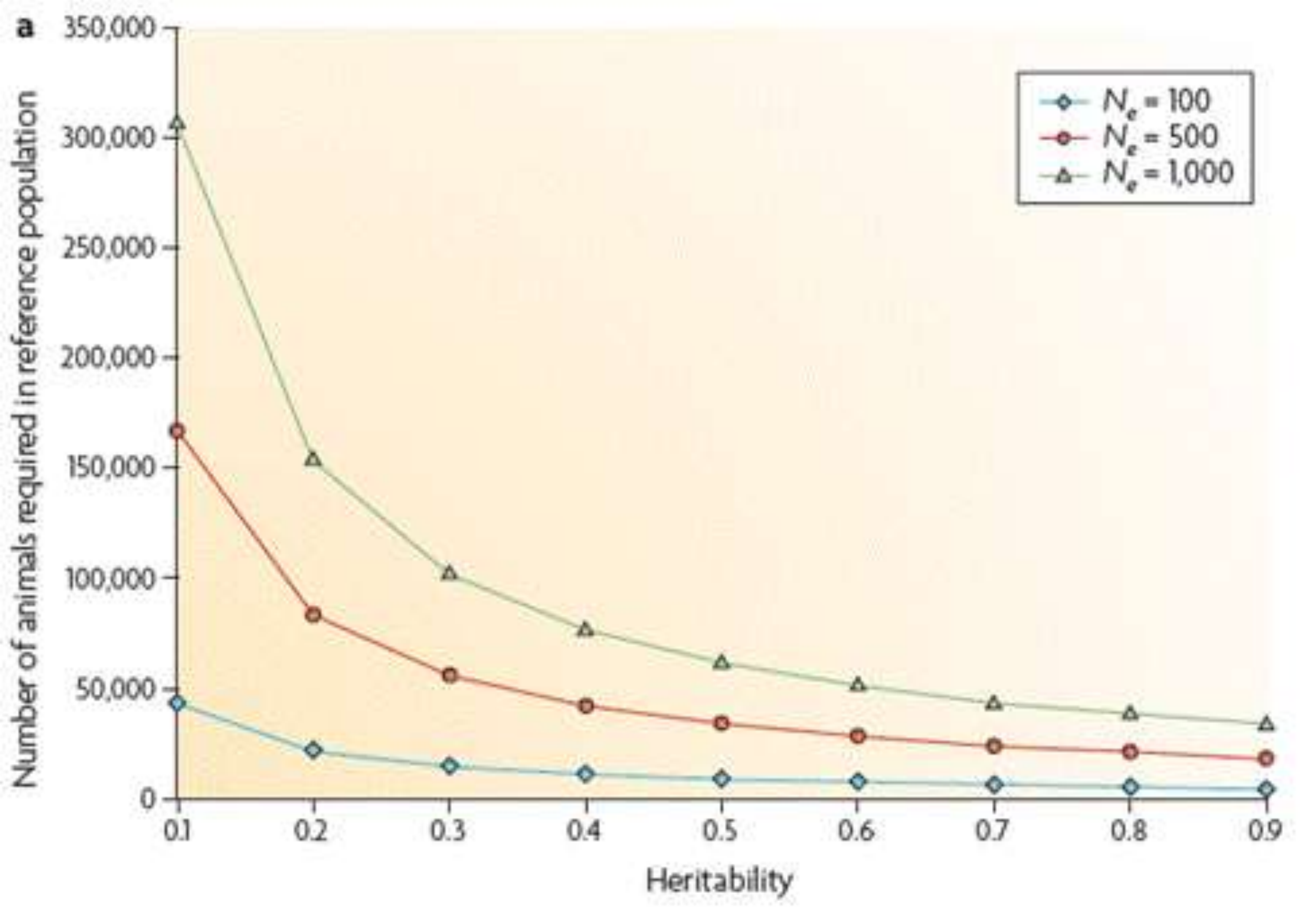
Goddard, M. E., and B. J. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10: 381-391.

Effective population size estimates for cattle

Breed	**N_e**	**Breed**	**N_e**
Angus	**136**	Brown Swiss	**61**
Charolais	**110**	Guernsey	**76**
Hereford	**97**	Holstein	**99**
Limousin	**174**	Jersey	**73**
Red Angus	**85**	Norwegian Red	**106**
Brahman	**115**	Gir	**133**
Nelore	**86**		
Beef Master	**106**	Merino (sheep)	**~ Big (> 100)**
Santa Gertrudis	**107**		**Ben Hayes (pers. comm.)**

Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. 2009 The Bovine HapMap Consortium. Science 3245: 528-532. Supporting Online Material. Table S1. http://www.sciencemag.org/content/suppl/2009/04/22/324.5926.528.DC1

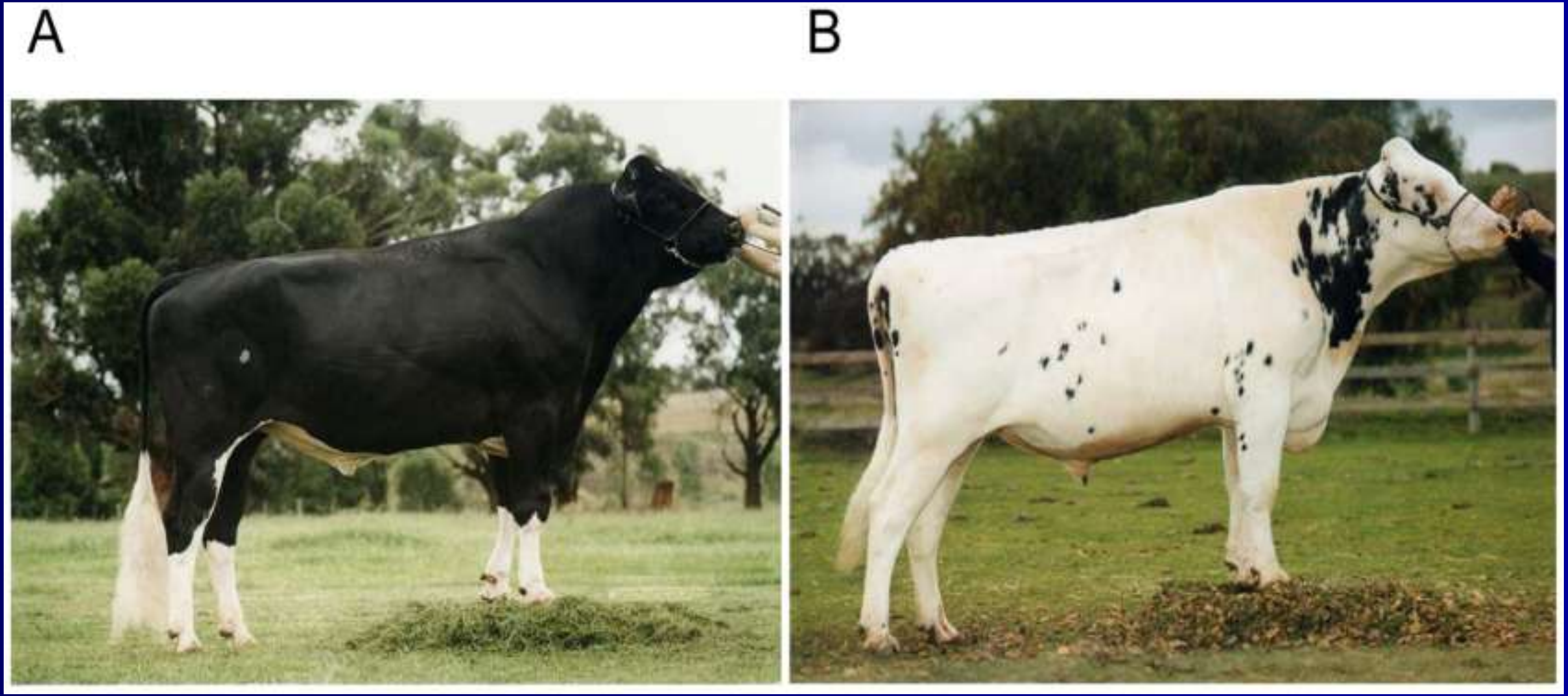
Effect of population size and heritability on the number of animals required in the training population (for an accuracy of 0.7)



Goddard, M. E., and B. J. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10: 381-391.

There is also an effect of trait architecture

The accuracy of predicting genetic values is higher for traits with a proportion of large effects (e.g. proportion black and fat percentage) than for a trait with no loci of large effect (e.g. overall type), provided the method of analysis takes advantage of the distribution of loci effects.

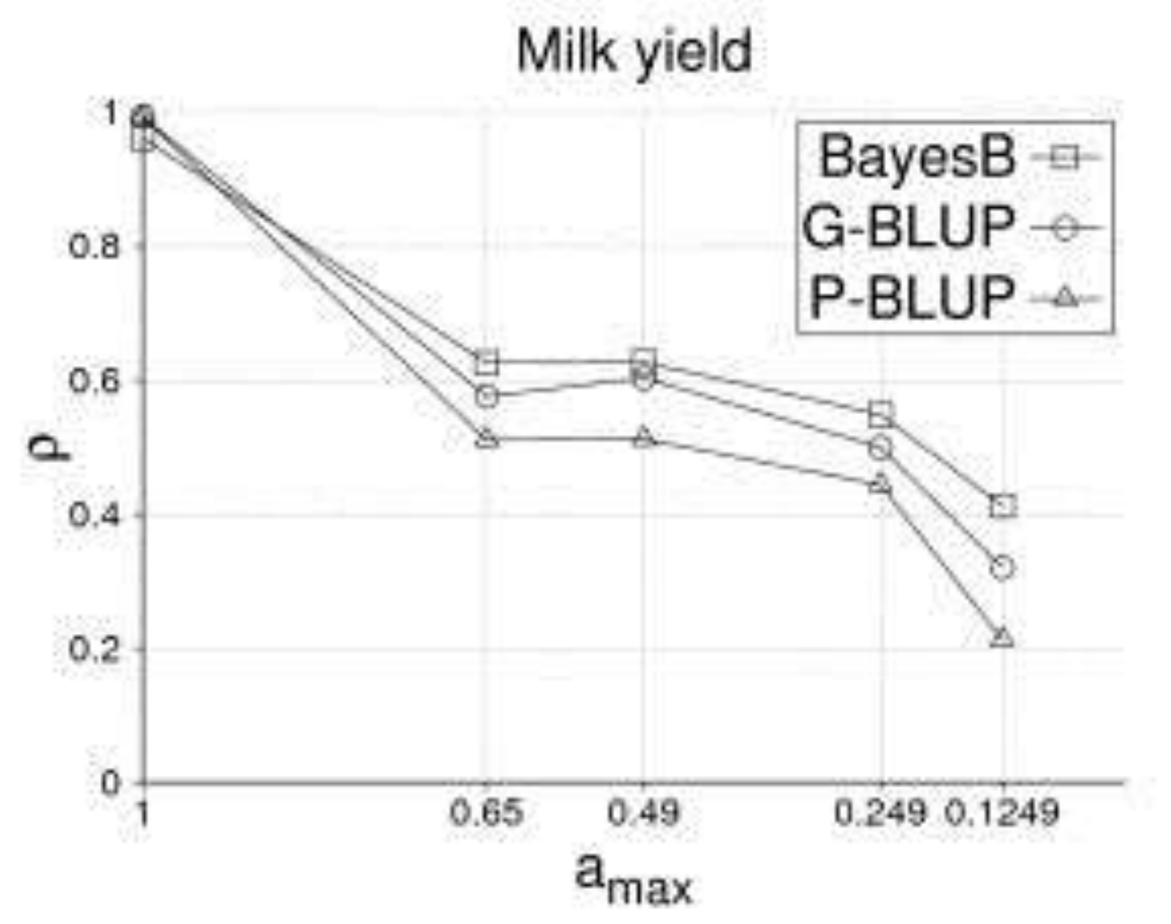


Hayes, B. J., J. Pryce, A. J. Chamberlain, P. J. Bowman, and M. E. Goddard. 2010. Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits. Plos Genet 6

Maybe R.A. Fisher was onto something?

If a nearly infinitesimal model is correct as seems to be the case for most quantitative traits; then large sample sizes will be needed to achieve high accuracy

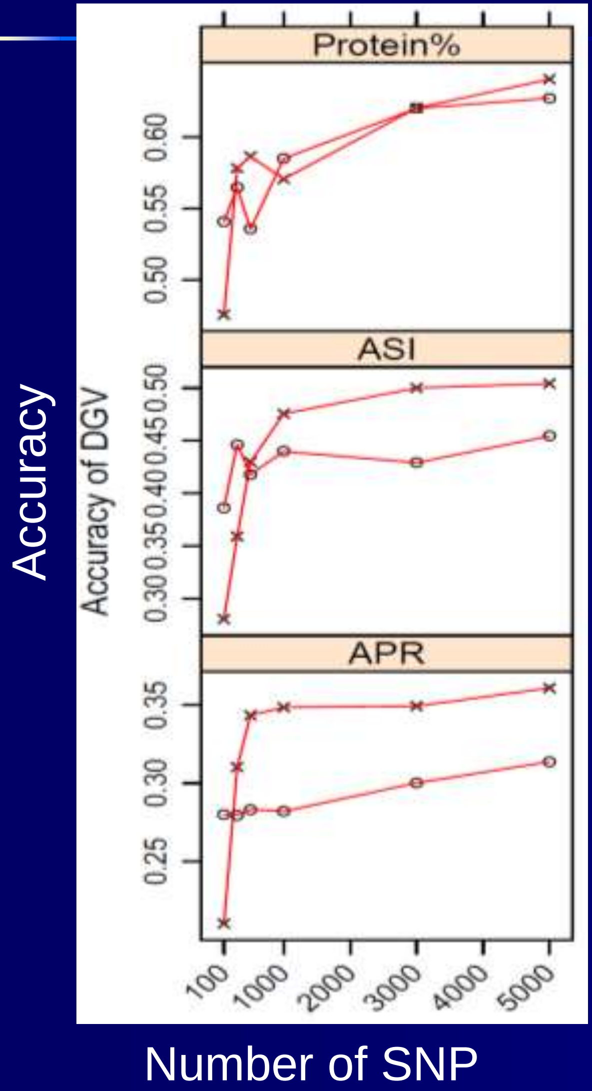
Markers can predict family relationships between animals, independently of LD between the markers and QTL (i.e. due to family relationships or linkage)



Additive-genetic relationships between training and validation animals was found to be a good indicator of accuracy

Habier, D., J. Tetens, F.-R. Seefried, P. Lichtner, and G. Thaller. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genetics Selection Evolution 42: Article No.: 5

Reduced SNP panels: Accuracy of direct genomic value (DGV) of dairy bulls using subsets of 5,000 or less of best SNP for each trait



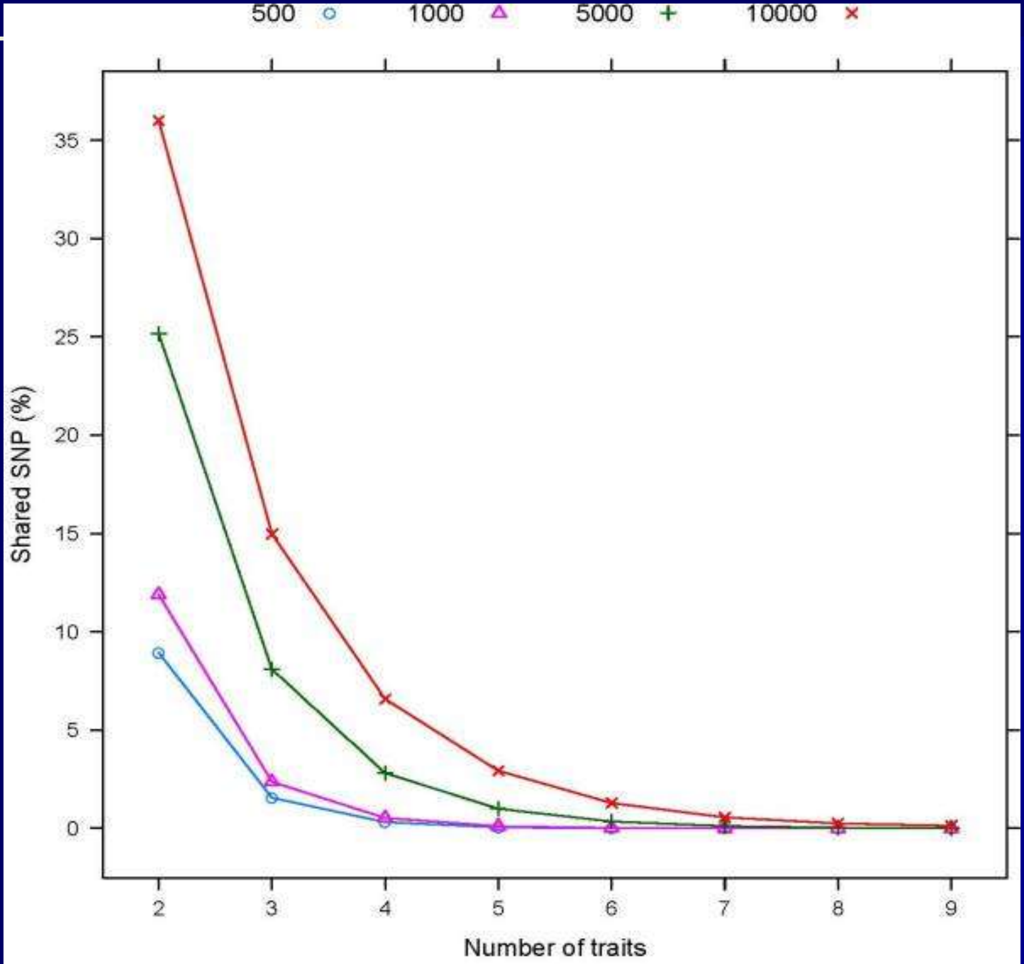
Traits :
Protein %
ASI (Australian Selection Index)
APR (Australian Profit Rank)

Predictions based on <1,000 SNP panels were very sensitive to the selection method and tended to be low accuracy

Moser, G., M. S. Khatkar, B. J. Hayes, and H. W. Raadsma. 2010. Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP markers. Genetics Selection Evolution 42.

Reduced SNP panels: Percentage of the highest ranked SNP that are shared between sets of traits*

for subsets including 500, 1,000, 5,000 or 10,000 SNP



* **Dairy traits included**:

1. Protein
2. Protein %
3. Survival
4. Fat %
5. Milk
6. Overall Type
7. APR (Australian Profit Rank)
8. ASI (Australian Selection Index)
9. Fat

Few SNPs were in common between the trait-specific subsets

Moser, G., M. S. Khatkar, B. J. Hayes, and H. W. Raadsma. 2010. Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP markers. Genetics Selection Evolution 42.

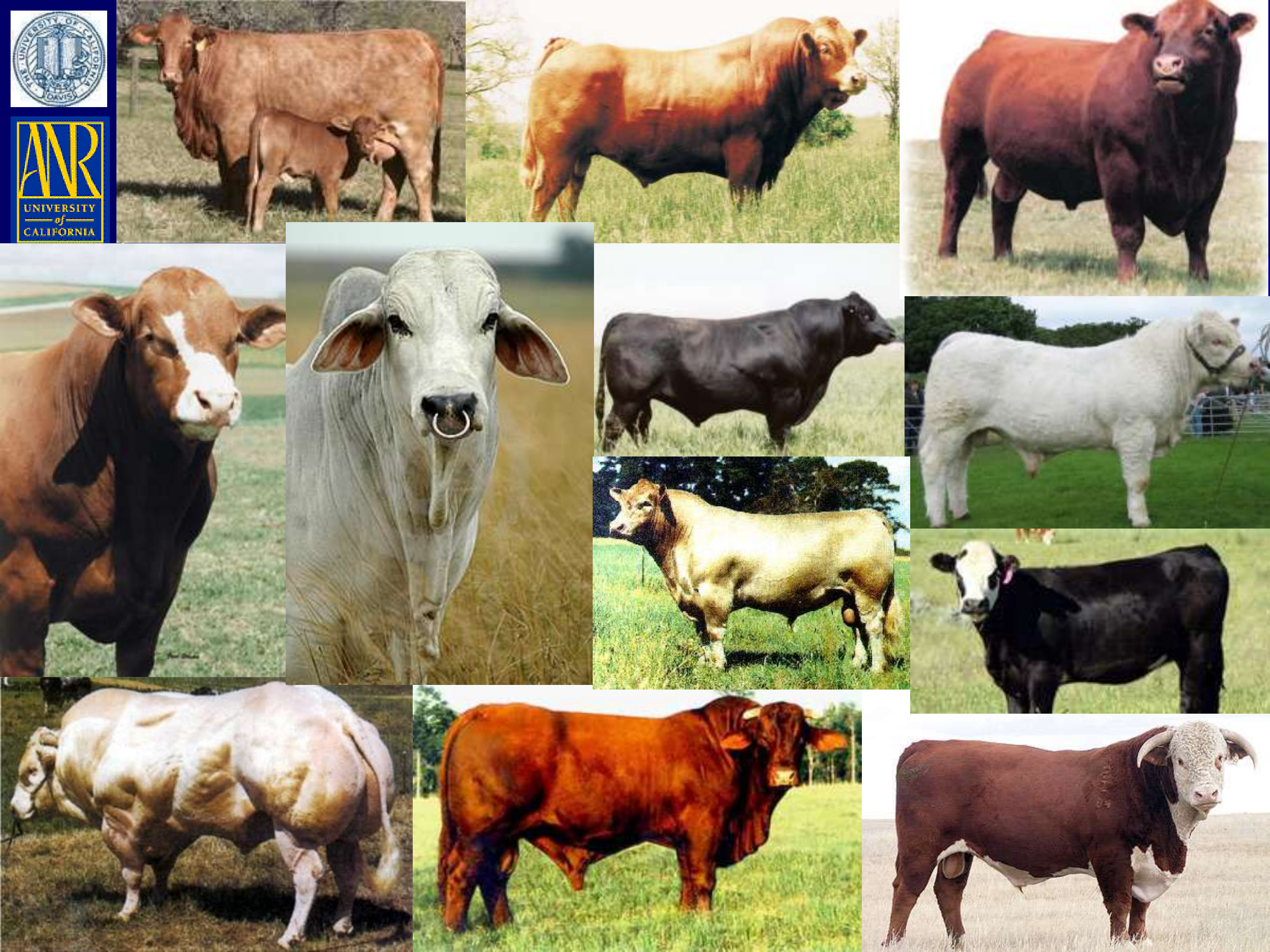
In general accuracy is higher when:

- Small effective population size so small number of chromosome segments to track
- Small number of QTL effecting the trait so there is a marker associated with every QTL
- High density of makers
- Trait is highly heritable
- A large number of animals and high-quality phenotypic records available for training
- There are genetic relationships (linkage) between training and selection candidates

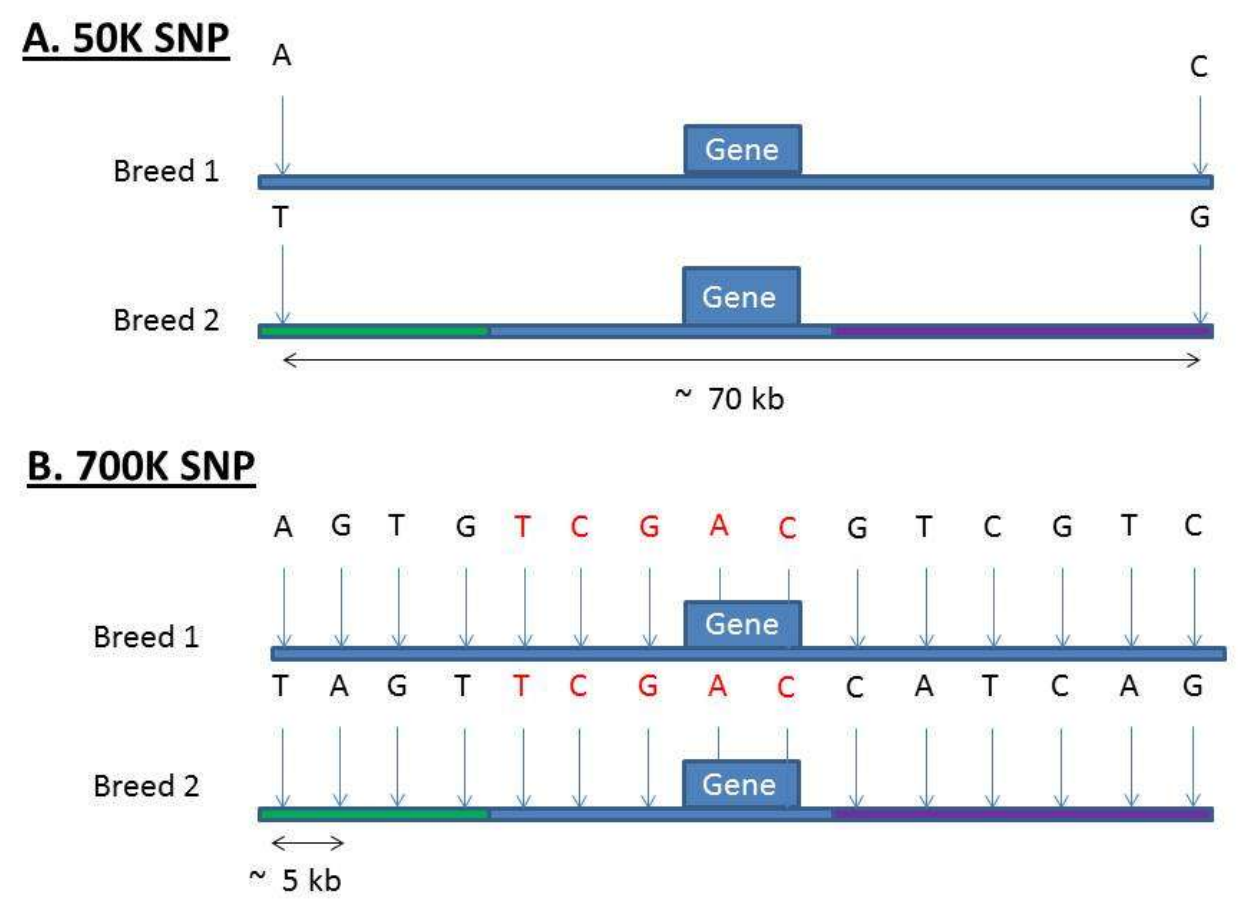
Other relevant interesting findings with practical implications

- If markers are picking up family relationships (linkage), then the accuracy of marker-based selection will decay over generations within breed
- Prediction equations derived in one breed do not predict accurate GEBVs when applied to other breeds
- To find markers that are in LD with QTL across diverged breeds, such as Holstein, Jersey, and Angus, will require high density SNPs (>300,000 markers)
- Combining breeds into one large multi-breed reference population gives reasonable accuracies in purebreds
- Few of the "best" markers for one trait are common to another

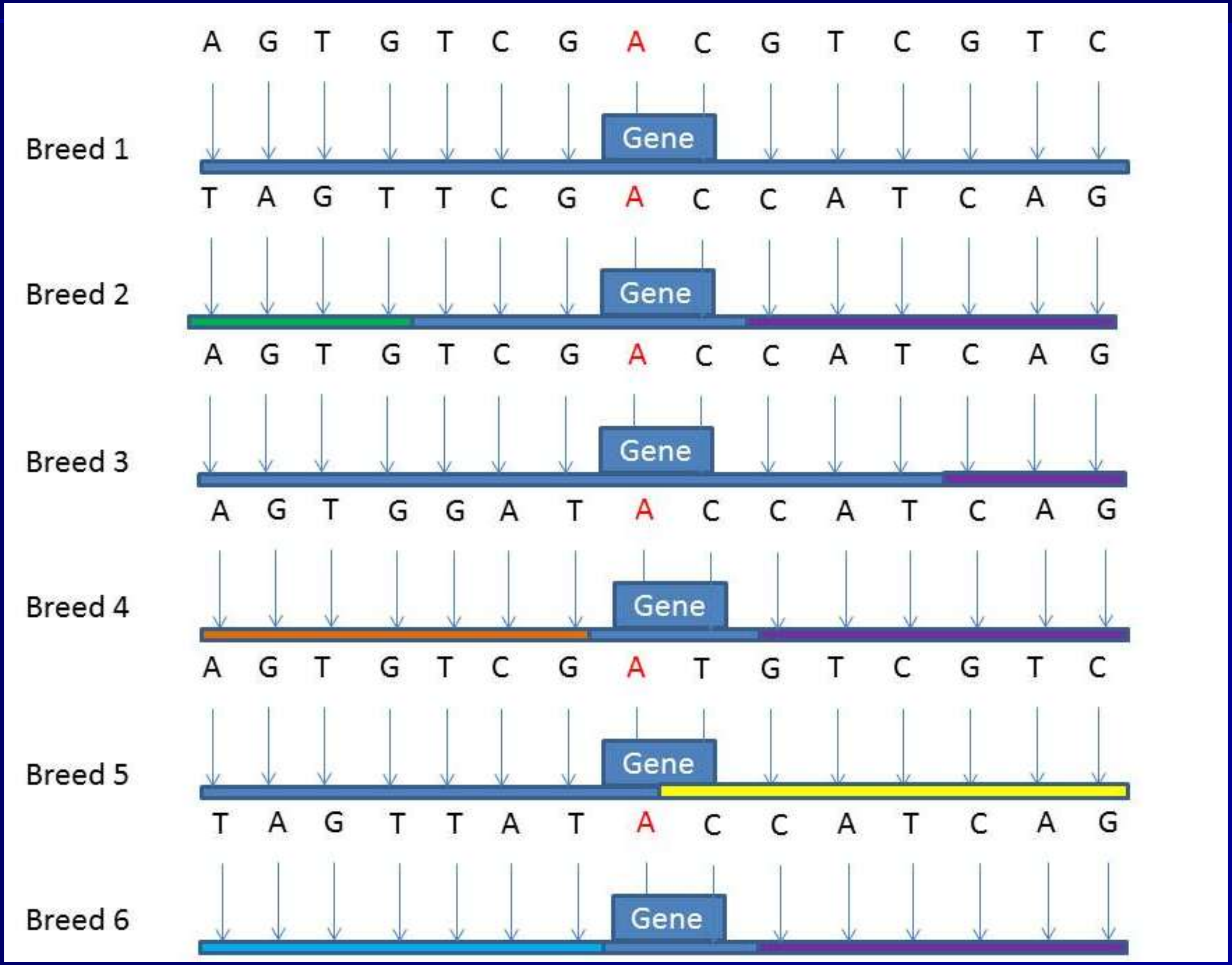
UNIVERSITY of CALIFORNIA



Marker location relative to the gene of interest in two breeds when using the (A) 50K SNP chip assay (markers spaced at ~ 70 kb intervals), or (B) the high density 700 K SNP chip assay (markers spaced at ~ 5 kb intervals)



High density panels offer the opportunity to accelerate discovery of the causal mutations underlying genetic variation – especially if combined with full sequence data on key ancestors



The Beef Cattle Industry

- Little use of AI
- Relatively few high accuracy sires for training
- Multiple competing selection goals – cow/calf, feedlot, processor – little data sharing between sectors
- Few/no records on many economically-relevant traits
- Many breeds, some small with limited resources
- Crossbreeding is important
- No centralized “national” cattle evaluation

UNIVERSITY of CALIFORNIA

What commercial products are out there for beef cattle producers?

GeneSTAR®

SureTRAK

SireTRACE

GeneSTAR®

Pfizer Pfizer Animal Genetics

igenity™

ImmGen.

GENETICVISIONS

genmark®

BIOGENETIC SERVICES, INC.

GENESEEK

Molecular Solutions for Breeding and Genetics

ANGUSURE™

GENAISSANCE PHARMACEUTICALS

MMI GENOMICS, INC.

BREED TRU™

DNA VERIFIED

METAMORPHIX, INC.

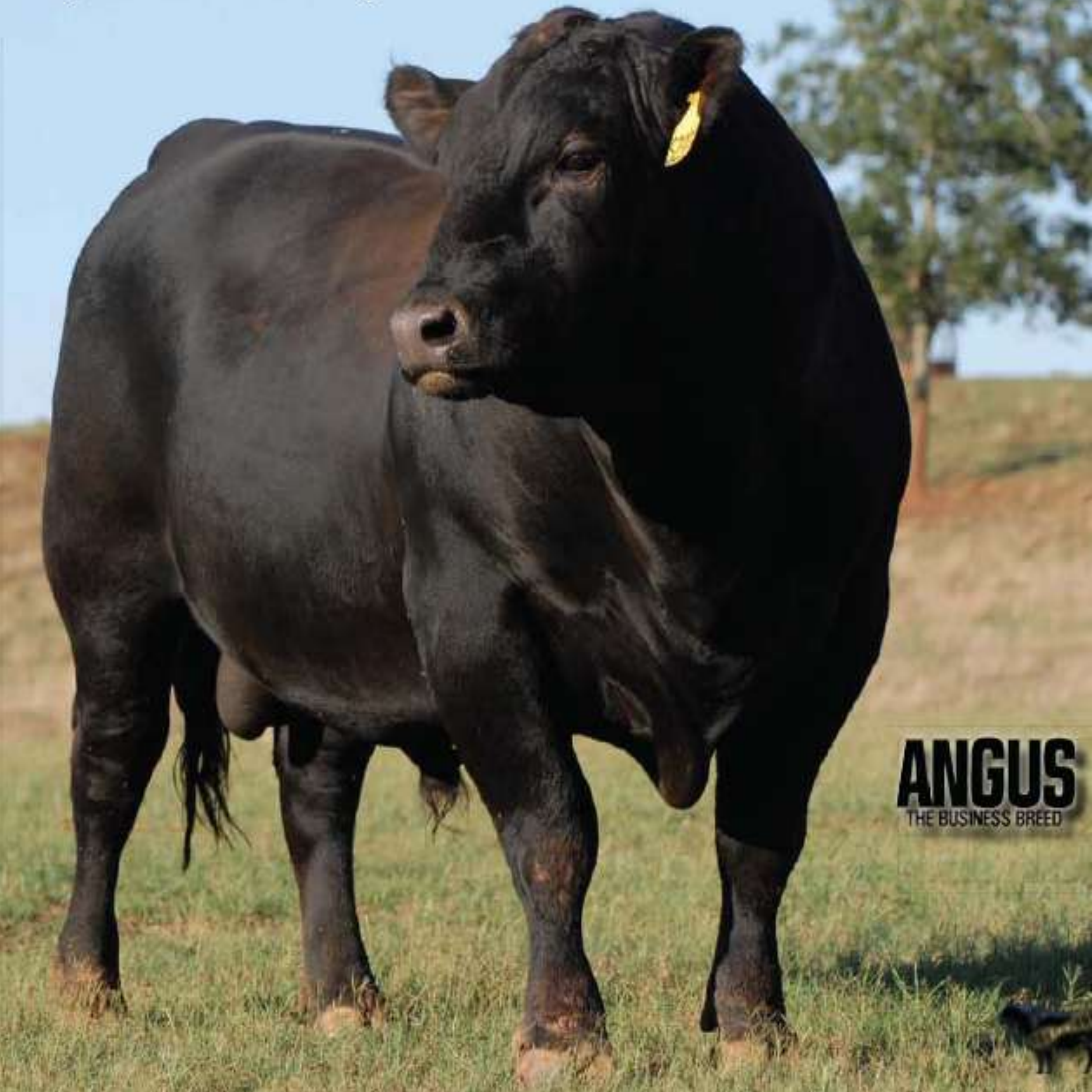
The Power of the IGENITY® profile for Angus

The American Angus Association® through its subsidiary, Angus Genetics Inc.® (AGI), has a vision to provide Angus breeders with the most advanced solutions to their genetic selection and management needs.

Genomic-enhanced Expected Progeny Differences (EPDs) can now be calculated for your animals using the highly predictable American Angus Association database along with IGENITY® profile results to provide a more thorough characterization of economically important traits and improved accuracy on young animals.

Using the IGENITY profile for Angus, breeders receive comprehensive genomic results for multiple, economically important traits.

1. **Dry Matter Intake**
2. **Birth Weight**
3. **Mature Height**
4. **Mature Weight**
5. **Milk**
6. **Scrotal Circumference**
7. **Weaning Weight**
8. **Yearling Weight**
9. **Marbling**
10. **Ribeye Area**
11. **Fat Thickness**
12. **Carcass Weight**
13. **Tenderness**
14. **Percent Choice (quality grade)**
15. **Heifer Pregnancy**
16. **Maternal Calving Ease**
17. **Direct Calving Ease**
18. **Docility**
19. **Average Daily Gain**
20. **Feed Efficiency**
21. **Yearling Height**



ANGUS
THE BUSINESS BREED

Genetic correlations (r) between carcass traits and IGENITY® Angus Profile molecular breeding values (384 reduced-SNP panel) in Angus cattle

TRAIT	Trait heritability	Accuracy (r)	% Genetic Variation (r^2)
Carcass weight	0.39	.54	29%
Backfat thickness	0.36	.50	25%
Ribeye area	0.40	.58	34%
Marbling score	0.37	.65	42%

MacNeil, M. D., S. L. Northcutt, R. D. Schnabel, D. J. Garrick, B.W. Woodward, and J. F. Taylor. 2010. Genetic correlations between carcass traits and molecular breeding values in Angus cattle. 9th World Cong. Genet. Appl. Livest. Prod. Leipzig, Germany. August, 2010.

Lead Today with 50K

1. Birth weight
2. Weaning weight
3. Weaning maternal (milk)
4. Calving ease direct
5. Calving ease maternal
6. Marbling
7. Backfat thickness
8. Ribeye area
9. Carcass weight
10. Tenderness
11. Postweaning average daily gain
12. Daily feed intake
13. Feed efficiency (net feed intake)

Pfizer Animal Health
Animal Genetics

50K SNP chip assays 50,000 SNPs spread throughout genome

Trait	h^2	Pfizer 50K HD in Angus	
		Number of animals in training population[1]	% Genetic variation (r^2)
			Predicted from LD
Average Daily Gain	0.28	1254	7%
Net Feed Intake	0.39	1254	10%
Dry matter intake	0.39	1254	10%
Tenderness	0.37	1445	11%
Calving Ease (Direct)	0.1	1188	2%
Birth weight	0.31	1169	7%
Weaning Weight	0.25	1192	5%
Calving ease (maternal)	0.1	1177	2%
Milking Ability	0.25	1187	5%
Carcass weight	0.39	1100	9%
Backfat thickness	0.36	1097	8%
Ribeye area	0.4	1114	10%
Marbling score	0.37	1143	9%

[1] Pfizer Animal Genetics. 2010. Technical Summary. http://www.pfizeranimalgenetics.com/sites/PAG/Documents/50K%20Tech%20Summary.pdf

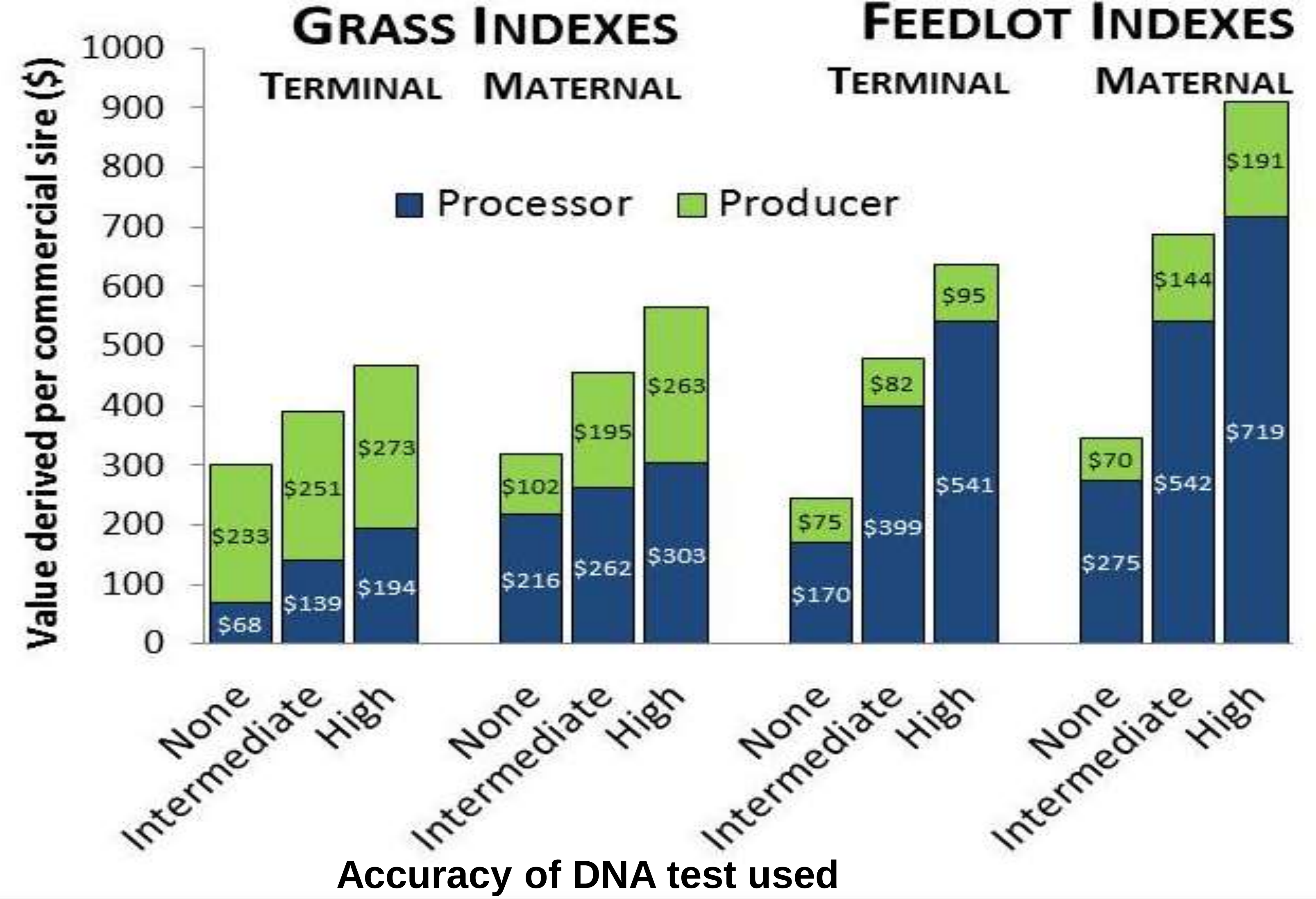
[2] Animal Genetics and Breeding Unit (AGBU). 2010. Evaluation of Pfizer Animal Genetics HD 50K MVP Calibration. http://agbu.une.edu.au/pdf/Pfizer_50K_September%202010.pdf

Approx. cost of commercial tests

(estimates only!! - derived from web-sites or personal experience - not official quotes!!)

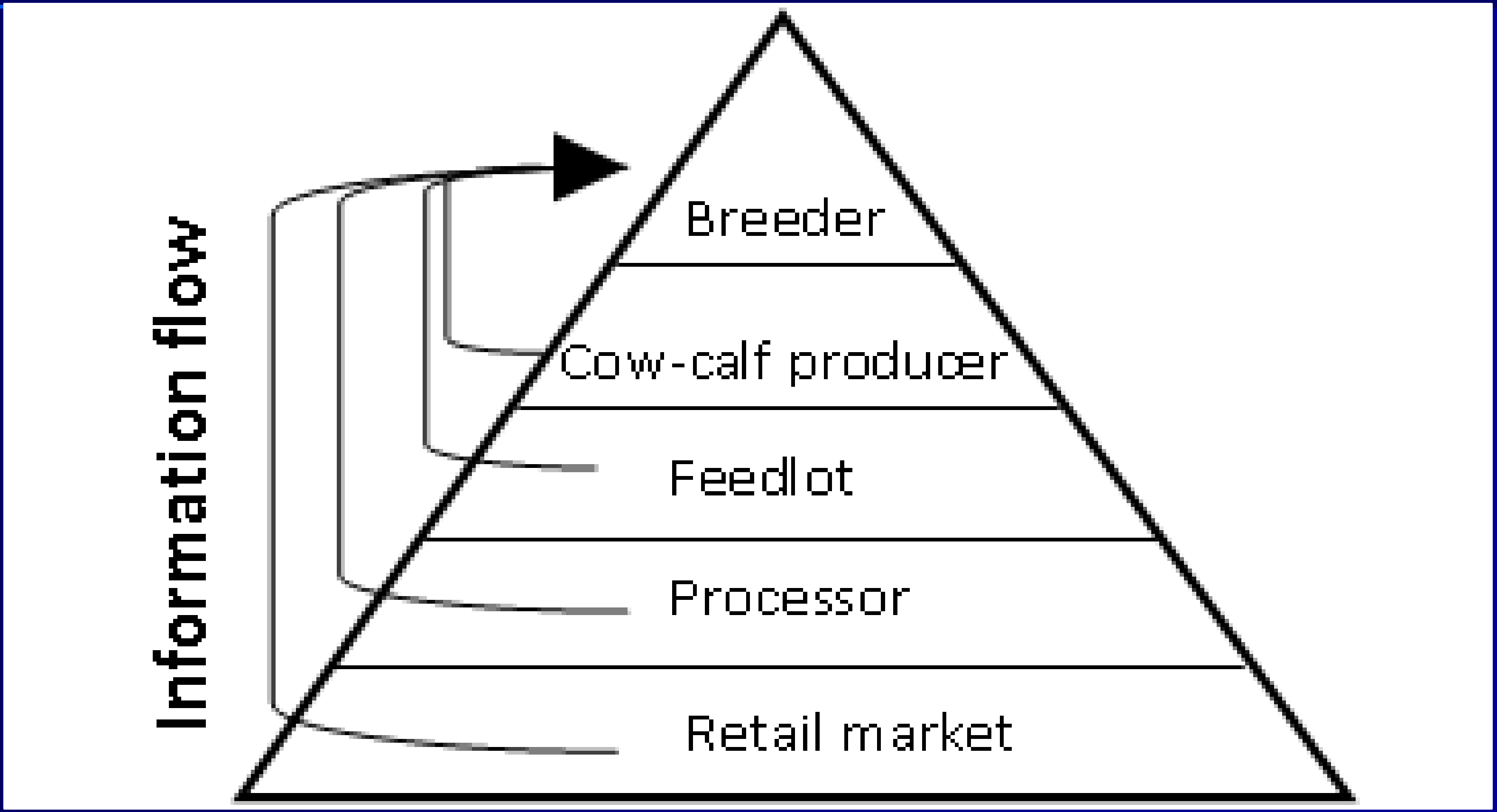
Test	Species	Cost ($US)
Parentage	Cattle	**$ 13-25**
Genetic Defects	Cattle	**$ 15-150**
3K (just the genotypes)	Cattle	**$ 38**
50K (just the genotypes)	Cattle	**$150**
800K (just the genotypes)	Cattle	**$340**
384 Angus Profile (Igenity US/AGI)	Beef Cattle	**$ 65**
3K (Pfizer US)	Dairy Cattle	**$ 45**
50K (Pfizer US/AGI)	Beef Cattle	**$139**
50K (Holstein Ass.)	Dairy Cattle	**$150**
800K (Holstein Ass.)	Dairy Cattle	**$365**
50K (Pfizer NZ)	Sheep	**$756 (NZ$990)**

Industry breakdown of ΔG value derived from increased accuracy from genomic selection



Van Eenennaam, A. L., J.H. van der Werf, and M.E. Goddard. 2011. The economics of using DNA markers for beef bull selection in the seedstock sector. Journal of Animal Science. 89 (2) *In press*.

The beef industry needs to share data and profit between sectors to most benefit from genomic selection



McEwan, J. C. 2007 Current status and future of genomic selection. Proceedings of the New Zealand Society of Animal Production 67: 147-152.

CONCLUSION: Ramifications of genomic selection

- The benefits of genomic selection are best captured in well-structured industries that are already making significant genetic progress
- May encourage more vertical integration to collect phenotypes to enable predictions for ERTs for all sectors
- May see genetic evaluations developed for novel traits – if large enough populations can be amassed and data shared
- May see breeds/counties start to share data – especially with HD chips and whole genome sequencing
- Will beef follow the pig/poultry model of vertically-integrated breeding companies owning all sectors?

USDA

United States
Department of
Agriculture

National Institute
of Food and
Agriculture

"This project is supported by National Research Initiative Grant no. 2009-55205-05057 from the USDA National Institute of Food and Agriculture."

Come to Melbourne, Australia !!!
2-5 May, 2011

Date Claimer

Applied Genomics for Sustainable Livestock Breeding

2011

2-5 May 2011
The Sebel Albert Park
Melbourne

www.smogenomics.org

genomics-conf@jkconnections.com.au


35
CITY CIRCLE
City Circle
728
Luna Park

Questions?

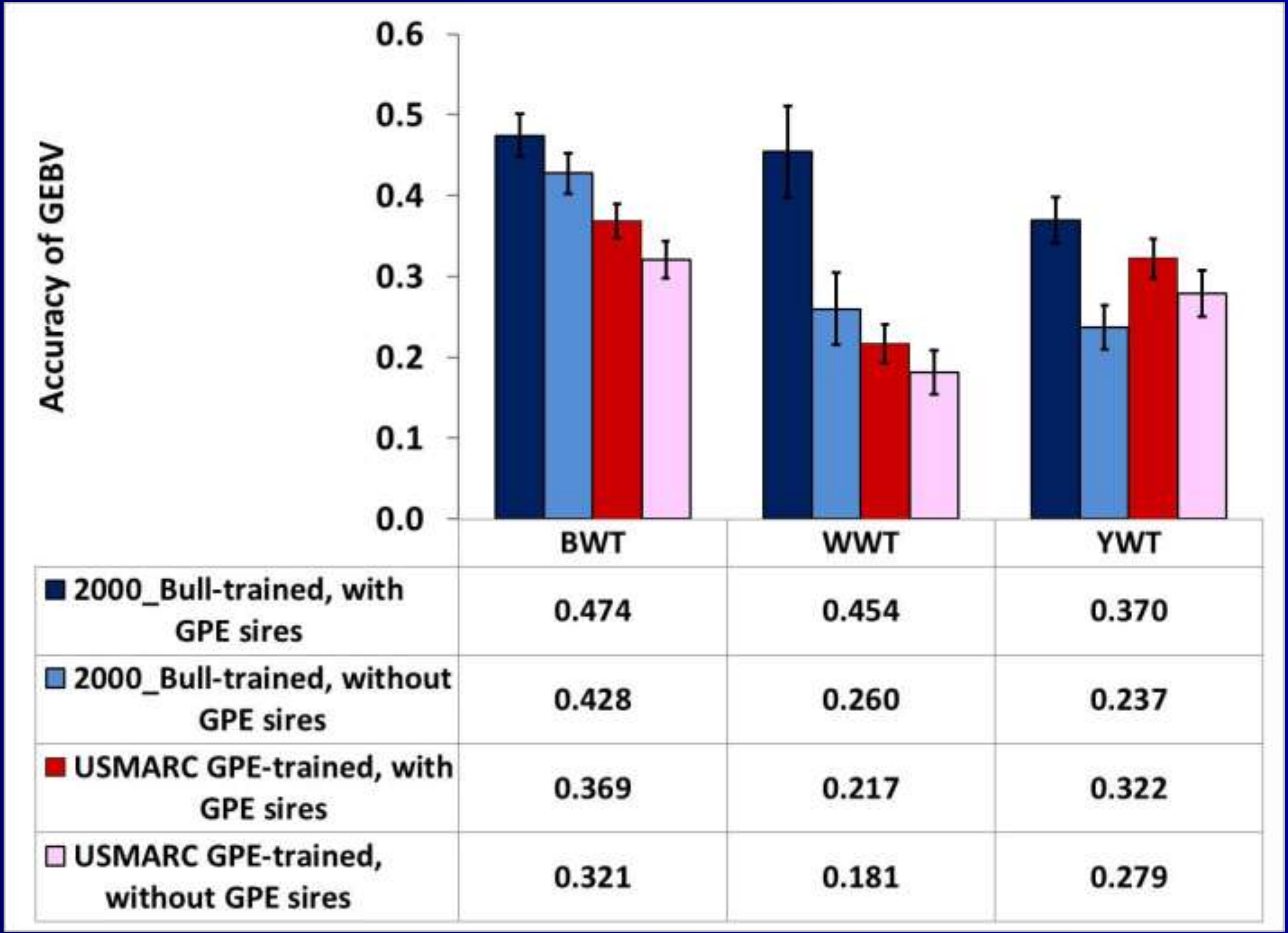
Value of improved selection response for beef seedstock sector due to DNA-test increase in index accuracy

Variable	Unit	Accuracy of DNA test used	GRASS INDEX		FEEDLOT INDEX	
			Terminal	Maternal	Terminal	Maternal
Improvement in selection response	**%**	**Intermediate**	**29**	**46**	**94**	**95**
		High	**54**	**81**	**157**	**158**
Increased value derived from ΔG in commercial sires	$/ DNA test	Intermediate	45	69	118	170
		High	83	124	196	282
Increased value derived from ΔG in stud sires	$/ DNA test	Intermediate	160	203	421	506
		High	297	366	701	836
Total value per test to seedstock operator	**$/ DNA test**	**Intermediate**	**$ 204**	**$ 272**	**$ 539**	**$ 676**
		High	**$ 380**	**$ 490**	**$ 897**	**$1119**

Van Eenennaam, A. L., J.H. van der Werf, and M.E. Goddard. 2011. The economics of using DNA markers for beef bull selection in the seedstock sector. Journal of Animal Science. 89 (2) *In press*.

Genomic Selection In Beef Cattle: Training And Validation In Multibreed Populations

Kristina Weber P514: Monday morning poster session



	BWT	WWT	YWT
■ 2000_Bull-trained, with GPE sires	0.474	0.454	0.370
■ 2000_Bull-trained, without GPE sires	0.428	0.260	0.237
■ USMARC GPE-trained, with GPE sires	0.369	0.217	0.322
□ USMARC GPE-trained, without GPE sires	0.321	0.181	0.279