

REDUCED SNP PANELS – creation, realistic expectations, and use in different livestock industries

Alison Van Eenennaam, Ph.D.

Cooperative Extension Specialist Animal Biotechnology and Genomics University of California, Davis <u>alvaneenennaam@ucdavis.edu</u> (530) 752-7942

animalscience.ucdavis.edu/animalbiotech



Animal Genomics and Biotechnology Education

Van Eenennaam 6/27/2012



What about the dog slides?







This dog belongs to one of these men....





Animal Genomics and Biotechnology Education

Van Eenennaam 6/27/2012

? Whose dog is this?







Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education



One of these cats does NOT belong to Tara McDaneld.....



Van Eenennaam 6/27/2012





Animal Biotechnology and Genomics Education

? Which one of these cats does **NOT belong to Tara McDaneld?**



Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education



Evolution of beef cattle DNA tests over the past decade

GENAISSANCE PHARMACEUTICALS	Pfizer Pfizer Animal Health	BUREKAGENOMICS
Now, know more about the genetic potential of your cattle.	igenity, illumina 🔀	ENESEEK a Neogen [®] Company
	Name/Company	# SNP
	Parentage chip	~100
	GeneMax	~100
	Igenity Profile	384
	Illumina GoldenGate Bovine3K	~3,000
😢 igenity L	Illumina Bovine LD	~ 7,000
BOVIGEN	GeneSeek Genomic Profiler	8,655
GeneSTAR [®]	Illumina BovineSNP50K/Pfizer	~50,000
1-877-BEEF-DNA • www.bovigensolutions.com	GeneSeek Genomic Profiler 80K	~80,000
	Illumina Bovine HD	~770,000

2003



Cost of commercially-available DNA tests for US beef cattle

Seedstock producers are using DNA information for pedigree verification, genetic defect testing, and genomic enhanced EBVs. Sometimes these analyses are sent to three different laboratories, and costs can be in excess of \$200 per animal. June 2012

2	Type/Purpose of DNA Test	Species	Cost (\$US)
	Microsatellite or SNP-based parentage test	Cattle	~\$ 10-25
	Genetic Defects/Single gene tests	Cattle	~\$ 15-100
P	Illumina Bovine 3K (just genotypes - no prediction equation)/Research	Cattle	~\$ 38
	Illumina Bovine LD 7K(Pfizer Animal Genetics imputation)	Dairy Cattle	~\$ 48
	Illumina Bovine 50K (just genotypes)/Research	Cattle	~\$ 80
	Illumina Bovine 50K (Pfizer Animal Genetics)	Dairy Cattle	~\$125
	Affymetrix Bovine 650K (just genotypes)/Research	Cattle	~\$200
	Illumina Bovine 770K (HD) SNP Test (just	Cattle	~\$210
5	genotypes)/Research		
	GeneMax (~100 SNP/Certified Angus Beef)	Angus	\$17
	384 SNP Angus Profile (Igenity US/AGI)/Selection	Angus	~\$ 65
	Illumina Bovine 50K (Pfizer Animal Genetics	Angus	~\$139
	US/AGI)/Selection		



Extracting DNA multiple times in different labs for different applications makes about as much sense as simultaneously paying to have access to all of the following communication devices



Van Eenennaam 6/27/2012

Animal Biotechnology and Genomics Education











Creation of reduced SNP panels – alternative approaches



1. Pick the most significant SNPs associated with the trait of interest – done on a trait by trait basis

Industry	Trait	# SNPs	Accuracy	Breed	Company
			(r _g)		
			estimate		
	Scrotal Hernia	96	0.30	Cross-bred	Genus/PIC
Swine	Finisher mortality	96	0.30	Cross-bred	Genus/PIC
	Total born	196	0.77	Cross-bred	Genus/PIC

Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education

<u>Reduced SNP panels</u>: 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.

Van Eenennaam 6/27/2012

CALIFORNI

Animal Genomics and Biotechnology Education



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



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.



Creation of reduced SNP panels – 2 alternative approaches

- 1. Pick the most significant SNPs associated with the trait of interest done on a trait by trait basis
- 2. Impute up to a high density genotype and use this information to predict genetic merit on any trait that has been recorded in the training population

Industry	Trait	# SNPs	Accuracy	Breed	Company
			(r _g) estimate		
			cotiniate		
	Scrotal Hernia	96	0.30	Cross-bred	Genus/PIC
Swine	Finisher mortality	96	0.30	Cross-bred	Genus/PIC
	Total born	196	0.77	Cross-bred	Genus/PIC
	Body Weight	384 (being	0.58	Broiler	Aviagen Ltd.
Poultry	Hen house production	used for 41K	0.60	Broiler	Aviagen Ltd.
		imputation)			



Imputation is a method of dealing with missing genotypes by filling in values based on a reference population



a	Study sample	
	AAA	. A . A

Reference haplotypes

CGAGATCTCCTTCTTCTGTGC CGAGATCTCCCGACCTCATGG CCAAGCTCTTTTCTTCTGTGC CGAGACTCTCCCGACCTTATGC CGAGACTCTCCCGACCTCATGG CGAGATCTCCCGACCTCATGG CGAGACTCTTCCCGACCTTGTGC CGAGACTCTTCCGACCTCGTGC CGAGACTCTTCCGACCTCGTGC

C Study sample

cgagAtctcccgAcctcAtgg
cgaaGctcttttCtttcAtgg

Reference haplotypes

CGGCCCCCGGCAATTTTTT CGAGATCTCCCGACCTCATGG CCAAGCTCTTTCTTCTGTGC CGAGACTCTCCGACCTTATGC TGGGATCTCCCGACCTCATGG CGAGATCTCCCGACCTCATGG CGAGACTCTTCCCGACCTGTGC CGAGACTCTTCCGACCTCGTGC CGAGCTCTTTCTTTCTTCGTGC



b Study sample

Annu. Rev. Genom. Human Genet. 2009.10:387-406.

Animal Biotechnology and Genomics Education



Accuracy of lifetime net merit MBV based on progeny test data for all SNP and selected or equally spaced reduced SNP sets in a testing set of 1,398 Holstein bulls (Weigel *et al.*, 2009)

No. SNP	Method of SNP Selection	Accuracy (r)
300	Largest Effects	0.428
300	Equally Spaced	0.253
500	Largest Effects	0.485
500	Equally Spaced	0.333
750	Largest Effects	0.519
750	Equally Spaced	0.435
1,000	Largest Effects	0.537
1,000	Equally Spaced	0.422
1,250	Largest Effects	0.554
1,250	Equally Spaced	0.477
1,500	Largest Effects	0.559
1,500	Equally Spaced	0.518
2,000	Largest Effects	0.567
2,000	Equally Spaced	0.539
32,518	All Available	0.612

Weigel K.A., de los Campos G., Gonzalez-Recio O., Naya H., Wu X.L., Long N., Rosa G.J.M. and Gianola D. (2009) J. Dairy Sci. 92: 5248.

- 1. Calving ease direct
- 2. Birth weight
- 3. Weaning weight
- 4. Yearling weight
- 5. Yearling height
- 6. Mature weight
- 7. Mature height
- 8. Dry matter intake
- 9. Residual feed intake
- 10. Scrotal circumference 11. Docility
- 12. Calving ease maternal
- 13. Milking ability
- 14. Carcass weight
- 15. Fat thickness
- 16. Ribeye area
- 17. Marbling score
- 18. Tenderness

Lead Today with 50K





Pfizer Animal Health Animal Genetics 50K SNP chip assays 50,000 SNPs spread throughout genome



The Power of the IGENITY[®] profile for Angus 384 SNP

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





American Angus Association performs weekly evaluations with genomic data – recently updated to include new traits

	Igenity	Pfizer
Calving ease (CED)	\checkmark	\checkmark
Growth (BW WW YW Milk)	\checkmark	\checkmark
Residual Average Daily Gain (RADG)	\checkmark	\checkmark
Docility (DOC)	\checkmark	\checkmark
Yearling Scrotal/Height (SC,YH)	\checkmark	\checkmark
Mature Weight (MW)	\checkmark	\checkmark
Carcass (CWT MARB RIB FAT)	\checkmark	\checkmark

http://www.angus.org/AGI/GenomicChoice11102011.pdf (updated 11/18/2011)

Animal Biotechnology and Genomics Education

	384 SNP	50K SNP
Genetic Correlation (r)/(r²%)	Igenity	Pfizer
Calving Ease Direct	.47 (22%)	.33 (11%)
Birth Weight	.57 (32%)	.51 (26%)
Weaning Weight	.45 (20%)	.52 (27%)
Yearling Weight	.34 (12%)	.64 (41%)
Dry Matter Intake (component of RADG)	.45 (20%)	.65 (42%)
Yearling Height	.38 (14%)	.63 (40%)
Yearling Scrotal	.35 (12%)	.65 (42%)
Docility	.29 (.08%)	.60 (36%)
Milk	.24 (06%)	.32 (10%)
Mature Weight	.53 (28%)	.58 (34%)
Mature Height	.56 (31%)	.56 (31%)
Carcass Weight	.54 (29%)	.48 (23%)
Carcass Marbling	.65 (42%)	.57 (32%)
Carcass Rib	.58 (34%)	.60 (36%)
Carcass Fat	.50 (25%)	.56 (31%)

UNIVERSITY of CALIFORNIA

http://www.angus.org/AGI/GenomicChoice11102011.pdf (updated 11/18/2011)



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.

Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education



Independent Assessment of Commercial Tests for Beef Cattle Production Traits

The objective of this study was to estimate the genetic correlation between MBVs from various sources and target traits based on ranch genetic evaluations of herd bulls sourced from the Angus seedstock sector.

Dissertation project of UC Davis Ph.D. student Kristina Weber

Weber, K. L., D.J. Drake, J. F. Taylor, D.J. Garrick, L.A. Kuehn, R.M. Thallman, R.D. Schnabel, W.M. Snelling, E.J. Pollak, A.L. Van Eenennaam. 2012. The accuracies of DNA-based genetic merit prediction equations derived from Angus- and multibreed beef cattle training populations. Journal of Animal Science. *In press*. Validation Population: American Angus Assoc.

Training Population: DNA Companies

Target Population: Commercial Beef Industry

Animal Genomics and Biotechnology Education

Van Eenennaam 6/27/2012



The accuracy for herd bulls on commercial ranches for WW, HCW, RE and MS was similar to that observed in AAA data, although standard errors were large as N was small



IGENITY MBV (AAA) IGENITY MBV (Est.) Pfizer MVP (AAA) Pfizer MVP (Est.)

Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education



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



Additivegenetic relationships between training and validation animals was found to be a good indicator of accuracy (r)

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



Practical implication of markers picking up family relationships

- Accuracy of marker-based selection will decay over generations within breed as the relationship between the training population and the evaluation population becomes more distant.
- This might not be an issue for seedstock breeders as elite seedstock typically provide the next generation of selection candidates and so selection candidates will most likely be closely related to the training population.
- Practically this means that SNP effects will have to be reestimated frequently to include data from each generation of selection candidates, although this may create logistical complications for genetic evaluation entities, especially if they do not have access to both the phenotypes and the genotypes or if additional costly phenotyping is required.









4665 Innovation Drive, Suite 120 . Lincoln, NE 68521 402/435-0665 · Fax: 402/435-0664 geneseekinfo@neogen.com · www.neogen.com

GeneSeek Genomic

One Sample, One Simple Comprehensive Solution

GeneSeek[®] Genomic Profiler™ for Dairy Cattle

The custom GeneSeek Genomic Profiler (GGP) low density Bead-Chip 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.
- Beta Casein A/B
- Beta Lactoglobulin
- Black/Red Coat Color
- (346, 358, 373) BLAD
- Chondrodvsplasia
- Citrullinemia
- DUMPS.
- Dun Color
- Factor XI
- Holstein haplotypes impacting fertility
- HH1, HH2, HH3

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.

Acknowledgement: GeneSeek gratefully acknowledges the USDA (BARC APL and BRL labs; MARC, Genetics and Breeding Research; NADC, Puminant Ш umina i 💽 Diseases and immunology Research Unit) for its contribution of SNPs and CERTIFIED assistance in chip design in the development of GeneSeek Genomic Profiler.



GeneSeek **Genomic Profiler** 8,655 SNPs

- 6,909 LD and other • SNP for improved imputation to 50K
- SNP for proprietary • single-gene tests for recessive conditions including genetic defects
- Detection of • haplotypes that affect fertility in dairy cattle
- Imputation of • microsatellite alleles to facilitate parentage validation

Animal Biotechnology and Genomics Education

Van Eenennaam 6/27/2012

- Jersey fertility haplotype1
- Silver Color Dilutor
- CVM*
- Beta Casein A2*
- Brachvspina*
- Hypotrichosis KRT71 (Rat tail)*
- "additional royalty fees will be applied for these optional licensed tests

Kappa Casein I Kappa Casein II and in case

-

THURSDAY.

Y chromosome infertility

? How much would you pay for a DNA test that could be used for multiple purposes – such as imputation, parentage, recessive conditions....

- 1. \$0
- 2. **\$1-5**
- з. **\$5-10**
- 4. \$10-20
- 5. **\$20-30**
- 6. **\$30-50**
- 7. >\$50



Van Eenennaam 6/27/2012

Animal Genomics and Biotechnology Education



Ideally cattle would be genotyped ONCE early in life and genotypes shared with downstream production sectors to derive the maximum value from the fixed DNA collection and extraction costs

Cattle industry Sector

Nucleus seedstock/AI bulls

Seedstock/bull multiplier

Registered females and stock bulls for commercial sector

Commercial cattle – Markerassisted management (MAM), replacement heifer selection Feedlot cattle purchasing, sorting and marker-assisted management (MAM) Traceability for voluntary labelling e.g. Angus beef

Traceability for disease outbreak/contaminated meat

Type of DNA product // DNA information access required

Full genome sequence

HD 770 K genotype

50K genotype + parentage + beef single gene traits/recessives

Imputation LD chip + parentage + beef single gene traits/recessives

Access genotypes from supplier (subset of LD imputation chip).

Access genotypes from supplier (subset of LD imputation chip).

Access genotypes from supplier (subset of LD imputation chip).



Conclusions

- Reduced SNP panels for imputation to higher density genotypes enable genetic prediction for multiple traits
 Decreased genotyping costs may soon make panels with fewer than several thousand SNPs obsolete
 - One test may provide information for different uses (genetic prediction, parentage, recessive testing) thereby delivering more value to offset the costs of DNA collection and genotyping



Troubled times come to UC Davis November 2011



Animal Biotechnology and Genomics Education

? Who would be the person least likely to participate in the "Occupy Davis" movement last Fall?







Thanks for inviting me!





National Colorado State University-Cornell University-University of Georgia-Iowa State University Beef Cattle Evaluation Consortium

Thanks to funding source: National Research Initiative competitive grant no. 2009-55205-05057 ("Integrating DNA information into beef cattle production systems") from the USDA National Institute of Food and Agriculture Animal Genome Program.



United States Department of Agriculture

National Institute of Food and Agriculture

Van Eenennaam 6/27/2012

Animal Biotechnology and Genomics Education