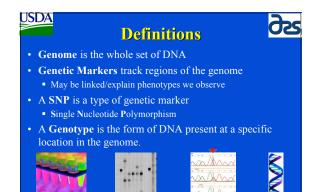


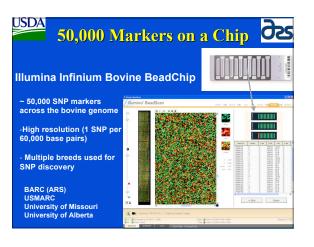
Overview

Introduction

USDA

- 2,000 Bull Project
- Training Data and Process
- Validation Data and Process
- Industry Application
- Conclusions
- Technical Issues





USDA

Whole Genome Selection (WGS)

- Use genotypes of thousands of genetic markers, like SNP, to predict breeding values (EPDs)
 Like marker sets for traits being offered now but mide much much being benefits.
- with much more density
- Genetic differences in DNA that cause phenotypic differences likely close to many markers
- Accounts appropriately for ambiguity regarding which SNP influence which traits
- Allows SNP with smaller effects on target traits to be used effectively. In theory, this will allow WGS to account for a greater percentage of genetic variation

USDA

ටින

Problem:

محو

ටින

- Whole-genome selection seems to be big leap technology that needs to either be implemented in a big way or not at all.
- Early adopters are likely to receive little benefit until a critical mass of influential and extensively phenotyped animals is tested and the data is analyzed jointly through NCE.



USDA ටින **2000 Bull Project** • Breed associations provided semen for DNA on influential sires • USMARC ran the 50K SNP chip on those 2,000 sires • USMARC provides extensively phenotyped animals for use as training data set

USDA

Objectives

25

d 23

- Extend genetic predictions from USMARC phenotypes to industry bulls
 - EPDs for traits not typically reported (e.g. feed efficiency) developed and delivered to breed associations (they decide use)
 Hopefully leads to development and use of commercial genotyping of additional animals with the 50K chip
- Validate the effectiveness of WGS using EPDs from the 2000 bulls relative to USMARC data on common traits (e.g., weaning weight)
 Improve accuracy of EPDs for common traits
- Determine to what extent training data must be of the same breed as in which WGS will be applied

Number	0Í	Sires	Sampled	<u>5</u>
 Angus Hereford Simmental Red Angus Gelbvieh Limousin Charolais Shorthorn 	402 317 253		Brangus Beefmaster Maine-Anjou Brahman Chiangus Santa Gertrudis Salers Braunvieh	68 64 59 53 47 43 42 27 026

USDA

Sire Sampling Criteria

- Breed associations responsible for selecting sires
 - Current candidates for selection (transfer genetic evaluations for new traits directly into genetic improvement)
 - High accuracy (verify the process)
 - Influential within their breeds (many progeny)
 - Sample each breed broadly

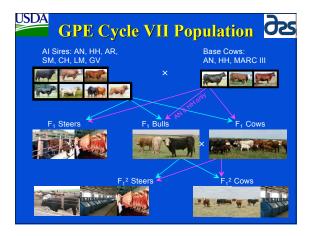
USDA

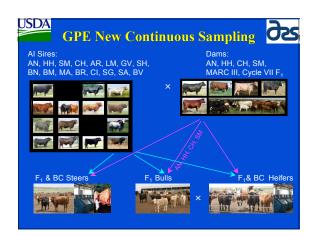
Current Status

- Have received semen on all of the 2,000 bulls.
- 50K genotyping is essentially complete and has generated about 250 million genotypes.
 - Research at USMARC previous to the 50K chip generated about 1.7 million genotypes at a substantially higher cost.
- Data analysis is underway.

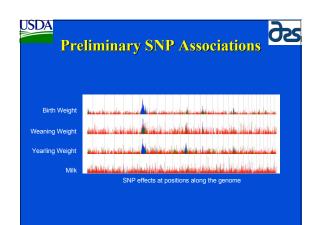
<u>ds</u>







		50 1	ΚΊ	rai	inir	lg l	Dat	8		đ
	EPD	Weights	Feed Efficiency	Carcass	Meat Ouality	Meat Comp.	Female Fertility	Mature Maint.	Longevity	Male Fertility
				GPE (Cycle V	П				
AI Sires	150									
F ₁ Bulls		73								73
F1 Steers		568		568	568	568				
F ₁ Cows		642					641	362	641	
F ₁ ² Steers		1313	1306	1249	1306	1220				
F ₁ ² Cows		710	691				702	Fut.		
		G	PE - Ne	ew Cor	ıtinuou	is Sam	pling			
AI Sires	135									
F ₁ Bulls		59								59
F ₁ Steers		269	257	252						
F ₁ Heifers		353	345				Fut.			
Total	285	3987	2599	2069	1874	1788	1343	362	641	132



A Numbers of Significant S for Birth Weight	и ду
# Animals (N)	2,578
# SNP analyzed	44,602
# SNP Significant (#Sig) at P<0.0001	308
# SNP Expected Sig. (#Exp) at P<0.0001under H ₀	4.4602
False Discovery Rate (FDR)	0.014

Ľ

	eaner	thar	ı Exp	ected	by (Than	ce 👘
			0.0001		• P<		
Trait		#Sig	#Exp	FDR	#Sig	#Exp	FDF
Birth Wt.	2578	308	4.5	0.014	638	45	0.07
Weaning Wt.	2569	103	4.5	0.043	273		0.10
Milk	2569	79	4.5	0.056	376		0.12
Yearling Wt.	2540	283	4.5	0.016	649		0.0
Fat Thickness	1693	25	4.4	0.177	115	44	0.38
Ribeye Area	1693	20	4.4	0.221	120	44	0.3
Marbling	1693	11	4.4	0.401	113	44	0.39
Carcass Wt.	1693	42	4.4	0.105	166	44	0.2
Shear Force	1667		4.4	0.736	48	44	0.92
Res. Feed Intake	1192	15	4.4	0.294	116	44	0.38
Flight time	1188		4.4	0.883	51	44	0.8

Prediction of Genetic Merit

029

020

- Prediction of EPD of cattle closely related to the training data is quite good.
- Right now, prediction of EPD of AI bulls that did not sire progeny in the training data is not good enough.
- Currently, results are mixed regarding how well EPD of breeds not included in training data can be predicted.

USDA

Prediction of Genetic Merit

- Have tried two basic statistical approaches thus far.
- Many other approaches are yet to be tried.

Experience from Human Genetics Human geneticists use higher density chips (about 500K to 1,000K SNP) in whole genome association (WGA) to try to explain the quantitative variation in disease risk. They are finding that WGA has been more challenging and explained less of the genetic variation than initially expected.

- A potential explanation is that much of the variation is due to rare variants (Goldstein, 2009, New. Engl. J. Med. 360:1696).
 - This suggests that associations may be more breedspecific, and perhaps more family-specific than we had hoped.

USDA

Experience from Dairy Cattle Breeding

- WGS being used to select young bulls to progeny test. Initial results look promising.
- However, the training data are essentially the entire Holstein breed.
 - Therefore, the target population for application is very closely related to the training data and to phenotypes.
 - This should work well for traits for which phenotypes are abundant.
 - However, it seems likely that tests for traits for which phenotypes are very limited (e.g., feed efficiency) will be difficult in dairy cattle.

USDA

USDA

D

02

Initial Emphasis will be on Weight Traits

- Many more animals with 50K chip data than any other set of traits.
 - Set of traits for which we are likely to have a success the soonest.
- Best set of traits with which to explore how well DNA tests will work across a variety of breeds, including breeds that were not included in the discovery data sets.
- Will be able to validate against EPDs of industry AI sires.

USDA

Continued Development of Training Data

- We are currently sampling about 80-100 sires per year in our Germplasm Evaluation Project.
 - Don't yet have funding to run 50K chip over calves from Fall 2008 forward.
- Have cow capacity to sample about 50% more sires, but it would require a substantial funding increase to cover expenses of increased AI and phenotype collection.

USDA

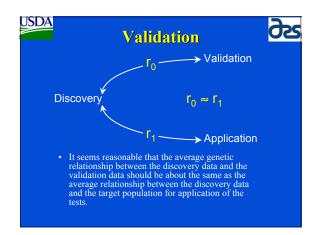
ටින

Validation

<u>ds</u>

d

- It is relatively easy to explain a lot of the genetic variation within discovery data.
- It is much more difficult to explain much of the variation within validation data that is completely independent of the discovery data.



Validation Resources 2,000 Bull Project International Collaboration Australia: Beef CRC Canada: University of Guelph University of Alberta U.S.: New Mexico State University

USDA

Industry Application

- We will provide molecular breeding values (**MBV**) on the 2,000 industry AI sires, when we have developed sufficient confidence in them.
- The initial focus is on weight traits, for which high accuracy EPD are available.
- · Carcass traits will likely be next.
- The ultimate goal is MBV for more difficult traits, such as feed efficiency.

USDA

ටින

Conclusions

- Industry sire sampling is complete.
- Laboratory work is complete.
- Statistical analysis is underway.
- Preliminary results suggest that more complicated statistical analysis will be required.
 - Consequently, results will take longer than initially hoped.

Contributors Exceed Associations Am. Angus Assn. Am. Chainina Assn. Am. Chainina Assn. Am. Hereford Assn. Am. Intl. Charolais Assn. Am. Salers Assn. Am. Shorthorn Assn. Am. Shorthorn Assn. Am. Shorthorn Assn. Am. Shorthorn Assn. Beefmaster Brdrs. United Braunvieh Assn. of Am. Intl. Brangus Brdrs. Assn. N. Am. Limousin Fndn. Red Angus Assn. of Am. Santa Gertrudis Brdrs. Intl.