llumina®



ICAR – Interbull Joint Session Use of molecular genomic tools in animal breeding

Riga, June 2nd, 2010

Chip technology for use in genomic evaluation and new developments

André Eggen Agrigenomics Specialist, Europe Illumina, Inc.



Mission

2

Innovating for the Future of Genetic Analysis



Food and feed are central to human health. Illumina is advancing the use of genetic variation-based approaches to crop and livestock markets.

Chip Technology for use in Genomic Evaluation and New Developments

The genomic era : a reality in animal breeding

The technology behind the change of paradigm

Illumina's genomic portfolio for Cattle

- Bovine HD, BovineSNP50, Bovine3K, ...

- Whole Genome Sequencing





THE AGRIGENOMIC ERA



The AgriGenomic Era ...

5

 Sequencing of livestock genomes has discovered 100s of thousands of Single Nucleotide Polymorphism (SNP) markers



The big promises of genomic selection

General idea

- With enough markers one can follow the transmission (segregation) of the entire genome and not only a set of specific regions of interest
- Parental relationship is no more used to explain similar performances in animal
- Similar performances are explain by the fact that animals are sharing identical chromosome fragments

illumına

Why is genomics so attractive ?



7

- Genetic evaluation can be performed as soon as DNA is available
- Allows accurate selection in both genders early in life
- Some direct consequences
 - Genetic progress could be doubled
 - Generation interval reduced
 - Cost for proving bulls could be reduced
 - Low heritability traits could be selected

...

illumina

Genomic Selection for Males

Age in years		0	2	6
Genomic Evaluation	Milk	62%	62%	70%
	Fert	50%	50%	50%
Traditional Evaluation	Milk	30%	35%	70%
	Fert			20%

As shown by E. Verrier, November 2009

8

Genomic Selection for Females

Age in years		0	2	3
Conomic Evaluation	Milk	62%	62%	70%
Genomic Evaluation	Fert	50%	50%	50%
Traditional Evaluation	Milk	30%	35%	45%
Traditional Evaluation	Fert			

As shown by E. Verrier, November 2009

9



THE TECHNOLOGY OF THE NEW PARADIGM



BeadArray™ Technology: Array format generation

BAD 1830 photoresist silicon wafer plasma etching cleaning

BeadChip



GoldenGate[®] Custom Genotyping





Infinium iSelect[®]HD Custom Genotyping



Consortia-developed Focused Genotyping Panels





15

ABLE 3: BOVINESNP50 BEADCHIP CONTENT VALIDATION						
BREED	SAMPLES	POLYMORPHIC LOCI*	MEAN MAF [†]	MEDIAN MAF [†]		
Angus	60	41,491	0.21	0.21		
Beefmaster	24	42,925	0.22	0.21		
Bos indicus Gir	24	23,971	0.11	0.02		
Bos indicus Nelore	21	25,814	0.11	0.02		
Brahman	25	30,284	0.13	0.08		
Brown Swiss	24	36,347	0.19	0.17		
Charolais	26	42,589	0.22	0.21		
Guernsey	21	38,632	0.19	0.17		
Hereford	32	42,992	0.20	0.23		
Holstein	64	42,730	0.22	0.22		
Jersey	28	35,976	0.18	0.14		
Limousin	45	42,821	0.22	0.22		
N'Dama	25	29,049	0.14	0.08		
Norwegian Red	21	42,782	0.22	0.21		
Piedmontese	24	42,185	0.22	0.21		
Red Angus	15	40,188	0.21	0.20		
Romagnola	24	38,830	0.20	0.19		
Santa Gertrudis	24	42,064	0.22	0.21		
Sheko	20	35,726	0.17	0.12		
Outgroup [‡]	18	11,206	0.05	0.00		
Overall	565	47,545	0.25	0.24		

BovineSNP50

SNPs also informative in many breeds



[†]Across all 54,001 loci

[‡]Bos bison, Bos gaurus, Bos grunniens, Bos javanicus, Bubalus depressicornis, and Syncerus caffer.



The current trend in animal breeding (especially Dairy Cattle)

Make use of the genomic information by genometry genetic/genomic value of animals !

 Marker Assisted Breeding is chan productivity

Towards Genomic Seleg

🦲 Cost savings iŋ

A very large

an animal have been forecast.

he structure of the breeding industry

Illumina BovineSNP50 has become the standart whole-genome genotyping tool for research and industry in Cattle

welfare and

THE BOVINE GENOMIC PORTFOLIO AT ILLUMINA

The current Bovine Genomic Portfolio



18

BovineSNP50 transition to a 24-sample format

From BovineSNP50 v1 to v2

Illumina BovineSNP50 has become the standart whole-genome genotyping tool for research and industry in Cattle

BovineSNP50 has been resynthesized on the Bov 2µm 24-sample BeadChip Format

- Faster sample throughput
- Faster scan times per sample
- Availability of 24-sample tip guide for more accurate loading

Due to the re-synthesis of beadpool onto 2µm beads,

the content of v2 is slightly different than v1

- 54,609 loci are on BovineSNP50v2
- 52,340 loci are common to v1 and v2
- 1661 loci on v1 are not on v2
- 2269 loci are new on v2
- The concordance of genotypes between v1 and v2 is 99.72%





BovineHD

The Next Generation High Density Bovine BeadChip

Whole genome research array to complement the current BovineSNP50 for the identification of agriculturally important genes and for a deeper implementation of genomic selection in cattle

- Developed in collaboration
 - USDA-ARS
 - Pfizer Animal Genetics
 - University of Missouri
 - UNCEIA, INRA
 - Other high-profile agricultural organizations

BovineHD : Groups, institutions and companies involved in the development

USA

- 🔶 USDA
- George Mason University
- University of Missouri
- University of Illinois
- 🦲 Pfizer
- Brasil
- UNESP, Aracatuba
- 🥌 Embrapa, Brasil
- France
- UNCEIA-INRA
- Netherlands
- CRV

8

Republic of Korea

- Chungbuk National University
- Yeungnam University, Gyeongsan
- Danemark
- University of Aarhus, Danemark

Germany

Technical University of Munich

Italy

- University of Piacenza
- University of Milano

UK

Roslin Institute

Assay Design Results

All of BovineSNP50 SNPs will be attempted

- ~800,000 bead types after 60,800 beads of BovineSNP50 positioned
- 795,000 SNP positioned on BTA1-29,X
- 5,000 beads represent unknown contigs, BTA Y, and mitochondrial SNP

Breed groups used:

- Holstein, Angus, Nelore, Taurine dairy, Taurine beef, Indicine, tropically adapted Taurine
- 🤍 852,645 total gaps
 - 850,816 <20kb
 - 1795 >20kb, < 100kb</p>
 - 34 > 100 kb

22

C. Van Tassell, PAG, January 2010



Intervalls between adjacent SNPs under 20,000 bp

C. Van Tassell, PAG, January 2010



Bovine 3K panel : 3072 SNPs (GoldenGate Technology)

- Designed in collaboration with the USDA and Dairy Cattle Breed Associations
- Enables broader access of cattle herds to genetic tests for the evaluation of net merit
- Content derived from BovineSNP50 and designed to accurately impute back to BovineSNP50
- Includes > 100 parentage SNPs (Heaton's panel, USDA)
- Less sensitive to DNA quality (esp. Sheared DNA)



In Silico Genome Re-Sequencing

Use Low-Density SNP Chip



Modified from Hayes, 2009

25

In Silico Genome Re-Sequencing

Use Low-Density SNP Chip to infer Sequence data



Modified from Hayes, 2009

Selection of SNPs (for low-plex genotyping)

Dense genotypic data available for

- Both parents
- Only sires
- Selection of SNPs
 - Magnitude of the estimated marker effects
 - Expected contributions to the genetic variance
 - In previous generation
 - In the next generation
 - Evenly spacing on chromosome
- Some concerns about the accuracy of Genomic Evaluation
 - Across populations
 - Over generations

Time consuming (computing involved)

Whole Genome (re-)Sequencing

Research

Open Access

Whole genome sequencing of a single Bos taurus animal for single nucleotide polymorphism discovery

Sebastian H Eck¤*, Anna Benet-Pagès¤*, Krzysztof Flisikowski†, Thomas Meitinger*‡, Ruedi Fries† and Tim M Strom*‡

Addresses: "Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr., 85764 Neulerberg, Germany. "Lehrstuhl für Tierzucht, Technische Universität München, Hochfeldweg, 83354 Freising-Weichenstephan, Germany. "Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Trogerstr., 81675 München, Germany.

× These authors contributed equally to this work.

Correspondence: Tim M Strom. Email: TimStrom@helmholtz-muenchen.de

Published: 6 August 2009

Gename Biology 2009, 10:R82 (doi:10.1186/gb-2009-10-8-r82)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2009/10/8/R82

© 2009 Eck et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creative.commons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: 21 April 2009 Revised: 22 June 2009

Accepted: 6 August 2009

Abstract

Background: The majority of the 2 million bovine single nucleotide polymorphisms (SNPs) currently available in dbSNP have been identified in a single breed, Hereford cattle, during the bovine genome project. In an attempt to evaluate the variance of a second breed, we have produced a whole genome sequence at low coverage of a single Flectivieth bull.

Results: We generated 24 gigabases of sequence, mainly using 36-bp paired-end reads, resulting in an average 7.4-fold sequence depth. This coverage was sufficient to identify 2.44 million SNPs,

82% of which were previously unknown, and 15,000 genotypes of the same animal, generated on a 50 k oligo of 74% and 30% for homozygous and heterozygous SN determined by comparison with genotypes determin approximately 1.1%. We further determined the allele f and 48 Braumieh bulls. 95% of the SNPs were polymor of 24.5% and with 83% of the SNPs having a minor allel

Conclusions: This work provides the first single catt The chosen approach - low to medium coverage re-sec SNPs to the currently publicly available SNP resour construction of high density oligonucleotide arrays in studies. - 24 Gb of sequence
- 7.4-fold coverage
- 2.44 millions SNPs

- 115'000 indels

HiSeq2000 launched in January 2010

200 Gb per run or over 2 bovine

genomes at a quality allowing the

detection of single base pair

differences (to the SNP level)

25 Gb/day; 7-8 days for 2 x 100 bp



Using Whole Genome Sequence Data in Genetic

Evaluation (Meuwissen & Goddard, 2010)

Accuracies of prediction of genetic Not possible for a cost point of view. values increased by more than 40% The costs may be reduced relative to the use of dense ~30K SNP substantially by the whole-genome chips. sequencing of a limited subset of the Predictions of genetic values individual. remained These results suggest that with a combination of genome training a the sequence data, large sample sizes and a statistical method generati that detects the polymorphisms that are informative, Similar a high accuracy is attainable and training auto come same generation.

CONCLUSIONS



Optimization of Genomic Breeding



Genomic Tools



Optimization of Genomic Breeding *Different combinations of genomic tools*



Imputation needed ; Imputation Error ; Accuracy of GE

Optimization of Genomic Breeding *Different combinations of genomic tools*



Imputation needed ; Error due to Imputation reduced ; Accuracy of GE

Optimization of Genomic Breeding *Different combinations of genomic tools*



No imputation ; no biais ; individuals to be integrated in future reference population

Towards « The genome-assisted Barnyard » Nature Biotechnology 27, 487 (2009)

Livestock genomes are providing

Hugely valuable biological information

- Immediate benefits to the way livestock breeders go about their business

In contrast to the slow translation of human genome information into medicine, animal genomics is likely to have a rapid and tangible impact on Agriculture »

NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 6 JUNE 2009

Economical optimization of genomic breeding :

Genomic evaluation is one important step but not the only one

Price (DNA + Genomic Tool + Analysis)

Information & Accuracy(GenomicInformation)



Discover. Validate. Screen



A new evolution in agriculture genetic analysis.