Genomic Selection–A Paradigm Shift in Animal Breeding

Illumina next-generation sequencing and genotyping technologies are revolutionizing animal breeding.

Introduction

Traditional animal breeding (TAB) relies on scoring animals based on their observed physical characteristics or phenotype to determine their breeding value (BV). This process is often inexact, inefficient, and time-consuming, with most traits (e.g., meat quality, milk production) hard to assess and evident only after the animal reaches maturity, resulting in a delay in verifying breeding results. Next-generation sequencers, such as the Illumina HiSeq® and MiSeq™ systems, are enabling researchers to quickly and cost-effectively sequence livestock genomes and identify the single nucleotide polymorphisms (SNPs) associated with commercially important phenotypic traits. Genome-wide SNP panels are now available for an increasing number of livestock species, enabling breeders to cost-effectively and accurately determine a genomic estimated breeding value (gEBV), even in young animals, and rank sires and dams based on genetic merit. The resulting shift from TAB to genomic-based breeding is revolutionizing global livestock industries.

Limitations of Traditional Animal Breeding

In TAB, breeders set a breeding goal based upon desirable production traits specific for a given breed. In each breeding generation, sires and dams are phenotypically scored (using either quantitative or qualitative criteria), with the scores for each trait generating a BV that is used to rank the animals on genetic merit. Sires and dams with the most desirable physical characteristics—the highest BVs—are selected for breeding. Offspring from these matings are evaluated according to the breeding goal and those with the highest BV scores are used as breeding stock for the next generation.

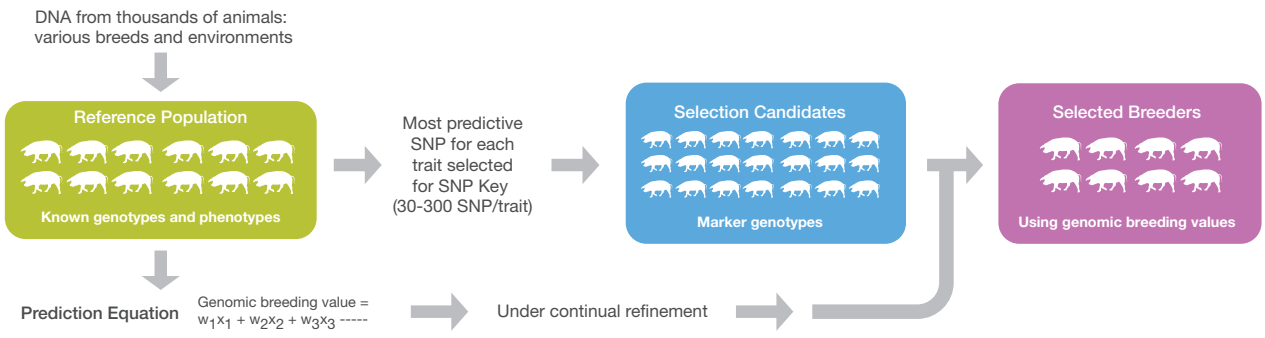
Yet, even the best phenotypic assessments cannot ensure that mating two high BV animals will result in high BV offspring. The process is limited by many variables, including:

- Limitations in understanding of most traits at the biological level due to complex inheritance (rather than simple Mendelian inheritance)
- Limitations in the number of traits that can be assessed in a breeding goal
- Decreased selection efficiency for traits that are difficult to measure (e.g., meat quality, internal parasite resistance) or have low heritability (e.g. reproduction)
- Restriction to traits that can be reproducibly and inexpensively measured in a large number of animals
- Genetic progress is slow, and a high selection pressure on key production traits may have a deleterious effect on other important traits that are not readily scored

Marker-Assisted Selection Improves Breeding Decisions

New sequencing technologies enhanced the development of genomic tools to assist in breeding decisions. Issues of cost and time limited these early sequencing efforts to reduced representations of animal genomes. Gene expression and genotyping technologies have enabled researchers to identify DNA regions, each known as a quantitative trait locus (QTL), associated with a particular phenotypic trait. QTL may involve many genes of small effect or a few genes of large effect. Linkage maps showing the relative locations of each QTL on the genome were created and then refined using linkage disequilibrium (LD) mapping (determining the correlation of these markers within a population). With the limited technologies of the day, researchers

Figure 1: Genomic Selection in Livestock Breeding



A reference population of animals is scored for key production traits and genotyped using a commercial or custom SNP array. The genotypes are represented by the variable (x) with values 0,1,2 (homozygous, heterozygous, or alternate homozygous). A prediction equation is generated combining all the marker genotypes with their effects to generate a gEBV for each animal. This prediction equation can be applied to a group of animals that have not been phenotyped, breeding values can be estimated, and the best animals selected for breeding.[2]

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA

discovered genetic markers, creating marker panels for each QTL for use in performing marker-assisted selection (MAS). Despite its limited view of the livestock genome, MAS provides breeders with a first step in leveraging genomics to achieve breeding goals[1].

Genomic Selection Tools Provide Whole-Genome View

Next-generation sequencers are transforming animal breeding, enabling cost-effective whole-genome sequencing and resequencing studies to identify genetic markers over the entire genome. In the last five years, large numbers of SNPs have been discovered in domestic animals and livestock (particularly bovine, porcine, and ovine species) by performing whole-genome association studies (WGAS). These studies can detect statistical associations between commercially important traits and one or more SNP markers, supporting the development of comprehensive marker arrays for genomic selection, a form of MAS using genetic markers covering the whole genome (Figure 1).

What's the difference between MAS and GS?

The fundamental difference between MAS and GS impacting the effectiveness of these two selection tools is 'scale'. MAS is limited in its ability to predict breeding values as it concentrates on a small number of QTLs that are tagged by markers with well-defined associations. In contrast, GS uses a dense set of markers from across the entire genome, assuring that all QTL are in LD with at least one SNP marker. This gives GS a profound advantage over TAB and two huge advantages over MAS (Table 1):

- All the genetic variance for a trait can be tracked (i.e. it is hypothesis-independent in its approach)
- By taking advantage of LD between the SNP and the QTL, the SNP allelic effect can now be estimated at the population level (no longer restricted to the family level)

Illumina and the Agrigenomic Revolution

Illumina has a proven track record in agricultural genomics, partnering with researchers and industry to apply novel genetic variation measurement technology to the sequence and assembly of a broad range of animal genomes. The results of these partnerships are genotyping tools that have significantly reduced the cost of measuring wide-scale genetic variation in livestock, including a number of consortia[3] and commercially available arrays for a range of species, including bovine, ovine, porcine, and equine.

Learn more about Illumina genomic technologies, their applications in agricultural research and animal breeding success stories, by contacting your local Illumina representative, who will work with you to understand your specific requirements. Illumina also has a dedicated global Agricultural Specialist Team with research and industry experience to assist you with your research planning.

References

1. Editorial–The genome assisted barnyard. (2009) Nat Biotechnol, 27: 487.
2. Goddard ME and Hayes BJ. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet, 10: 381–391.
3. www.illumina.com/applications/agriculture.ilmn#ag_consortia

Table 1: Comparison of TAB, MAS, and GS Breeding Approaches

Features	TAB	MAS	GS
Breeding value based on animal phenotype scores (qualitative or quantitative)	✓		
Breeding value based on a combination of genetic-genomic data and phenotype scores		✓	✓✓
Selection of complex inheritance traits		✓	✓✓
Selection of production traits without negatively impacting other desired traits		✓	✓✓
Unlimited number of traits can be assessed		✓	✓✓
Accurate measurement of low heritability traits		✓	✓✓
Reproducible and inexpensive genetic assessment of large numbers of animals		✓	✓✓
Accurate genetic selection in both genders		✓	✓✓
Reduced generation interval for BV evaluation		✓	✓✓
Increased rate of genetic improvement in the animal population		✓	✓✓
Reduced cost for progeny testing		✓	✓✓
Selection of genomic markers across the entire genome			✓
Hypothesis-independent approach to tracking of all genetic variance			✓
Estimation of marker allelic effects for a population (not restricted to the family level)			✓

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2011 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, Genetic Energy, GAIIx, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.
Pub. No. 370-2011-021 Current as of 10 October 2011

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA