GENETIC AND GENOMIC EVALUATION NOTES

CHAPTER 01

FRAMEWORK, CONCEPTS, AND GENERAL BACKGROUND

Framework

Genetic evaluation of individual animals in a breeding population, selection of the best individuals for traits of importance in some sense (adaptation, production, economic), followed by mating of selected individuals to generate the next generation of animals.

Definition of terms

Population: group of interbreeding animals that live in some geographical area. A population is connected through the presence of common animals across geographical, management, and time categories. Usually these categories are combined to create comparison or contemporary groups. For example, a contemporary group in dairy cattle could be formed by all animals born in some herd, year, and season, whereas in beef cattle it could be all animals within the same herd, year, season, management type, and sex.

Phenotype: is a characteristic or trait that we can observe and measure. For example: body weight and amount of muscle and fat in beef cattle, and milk yield and amount of milk protein and fat in dairy cattle.

Trait: usual name given by geneticists to a characteristic of interest (adaptation, production, economic). A trait is determined by the combined effects of all the genes associated with it, the internal environment within an animal, the external environment surrounding the animal, and interactions among these genetic and environmental effects.

Gene: a sequence of DNA that is transcribed into RNA that either codes for a protein or else has a different specific function (e.g., transfer RNA, ribosomal RNA, small nuclear RNA, guide RNA, micro RNA). Genes are the units of inheritance that are transmitted from parents to progeny, thus they are the biological basis of quantitative genetic analysis.

Allele: a particular DNA sequence of a gene. Genes can have different sequences, some producing different forms of a protein or RNA molecule. Each one of these sequences of a gene is called an allele. Alleles are the basis of genetic variation among animals.

Single Nucleotide Polymorphism (SNP): Differences among animals or pairs of chromosomes within animals for a single nucleotide (Adenine, Cytosine, Thiamine, Guanine).

Genotype: is the set of genes that determine a phenotypic trait within an animal.

Genome: all the genetic material contained in the cells of an individual. In domestic animals, it can refer to all the nuclear DNA and mitochondrial DNA.

Locus: location of a gene in a chromosome.

Quantitative Trait Locus (QTL): sequence of DNA that is associated with a phenotypic trait. It may represent one gene or a series of closely linked genes.

Haplotype: A set of alleles within a chromosome that tend to be transmitted together. It also refers to a set of SNP located at several loci that tend to be transmitted together.

Linkage Disequilibrium: combinations of alleles at 2 or more loci are inherited in higher proportion than by chance alone (loci may be in the same or different chromosomes).

Marker gene: is an identifiable sequence of DNA that is biologically linked or statistically associated with a QTL of unknown location, thus it can be used as a pointer in quantitative genomic analyses.

Genetic Evaluation: prediction of the genetic value of an animal for a trait using phenotypic measurements. Usual assumption: infinite number of genes, each having a very small effect. A predicted genetic value for an animal is a single number for a particular trait or a single number predicting an animal's worth for a linear combination of traits.

Genomic Evaluation: prediction of the genomic value of an animal for a single trait or multiple traits using phenotypic measurements and genotypes for a large number of markers spread throughout the genome. These markers are assumed to be in linkage disequilibrium with one or more quantitative trait loci (QTL) affecting the trait or traits of interest. A genomic model predicts the genomic value of marker locus effects (e.g., single nucleotide polymorphisms or SNP). Subsequently, the genomic value of an animal is predicted as a weighted sum of the gene content at each locus (number of "second" alleles) in an animal times the predicted genomic value of these "second" alleles at each marker locus.

Selection: some animals leave more descendants for the next generation than the average number of progeny per parent in a population. In natural populations this may be due to fertility, survival, behavior, and better ability to cope with environmental conditions. In artificial populations, this is primarily due to human decision on what animals will become parents of the next generation based on some genetic evaluation system.

General Background for Genetic and Genomic Evaluation

Although the initial concept of genomic evaluation (Meuwissen et al., 2001) represented a new paradigm for the genetic evaluation of animals, there has been a convergence of research and development aimed at the estimation of single markers or QTL of large effects on a trait (e.g., *dairy* *cattle:* DGAT1 on milk fat, milk protein, and milk yield; casein on milk protein; *beef cattle:* myostatin or GDF8 on meat weight, fat depth, intramuscular fat, mu-calpain, calpastatin) and genomic evaluation based on tens or hundreds of thousands of markers evenly distributed throughout the genome.

Private companies (e.g., Pfizer Animal Health, Igenity) developed molecular evaluation of animals for small number of markers. The number of markers used in these test panels increased since they became available in the mid-2000. For example, the GeneSTAR panel from Pfizer originally had 11 genes originally and subsequently increased to 56 genes associated with feed efficiency, marbling, and tenderness in cattle (Pfizer Animal Health, 2009). On the other hand, research work on genomic chips has been directed towards the development of chips with increasingly larger number of markers spread throughout the genome (e.g., 250K, 500K, 770K, …) that also contain sets of markers with either known or suspected association with purported QTL.

Modeling Approaches for Genetic and Genomic Evaluation

Assumptions

1) Traits determined by genes at a large number of loci (quantitative traits).
2) Markers are in some level of linkage disequilibrium with QTL.
3) Animals have either phenotypic and(or) genotypic information.
4) Direct genetic effects only (i.e., no maternal genetic effects).
5) Single-trait models only.

Models

1) Genetic Evaluation Model

2) **Polygenic-Genomic Models with Complete Genotype Information**
   a. **Polygenic -Genomic Marker Model with a single fixed marker effect**
   b. **Polygenic –Genomic Marker Model with a small number of fixed marker effects**
   c. **Polygenic -Genomic Marker Model with a small number of random marker effects**
   d. **Polygenic-Genomic Marker Model with a large number of random marker effects**
   e. **Polygenic-Genomic Animal Model with a large number of random marker effects**
3) **Polygenic-Genomic Models with Incomplete Genotype Information**
   a. **Polygenic-Genomic Marker Model with actual and predicted genotypes**
   b. **Polygenic-Genomic Animal Model with actual and predicted genotypes**
   c. **Polygenic-Genomic Animal Model with phenotypic, pedigree, and genotypic information**

References

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2009. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluations of Holstein final score. J. Dairy Sci. 93:743-752.

Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual purpose Belgian Blue cattle. Animal 1:21-28.

Gianola, D., Campos, G. de los, Hill, W. G., Manfredi, E., Fernando, R., 2009. Additive genetic variability and the Bayesian alphabet. Genetics 183, 347-363.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656-4663.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Nkrumah, J. D., C. Li, J. Yu, C. Hansen, D. H. Keisler, and S. S. Moore. 2005. Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feed intake, feeding behavior, and measures of carcass merit. J. Anim. Sci. 83:20-28.

Pfizer Animal Health, Animal Genetics. 2009. GeneSTAR MVPs – Molecular Value Predictions for beef feed efficiency, marbling, and tenderness. Technical Summary, March 2009. http://www.pfizeranimalgenetics.com/sites/PAG/Documents/GeneSTAR_MVP_TechSummary_NA.pdf.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414-4423.