

Thoughts and Concerns about Genomics

L. R. Schaeffer

*Centre for Genetic Improvement of Livestock
Department of Animal & Poultry Science
University of Guelph
Guelph, ON Canada*

The pace of genomic advances in humans and cattle is not slowing down. The dairy cattle industries around the world have grabbed hold of the new technology and have been racing to find better ways to evaluate dairy sires and cows. The purpose of this meeting is to find ways to combine different genomic solutions for international comparisons of bulls. However, we should not think that SNP chips are here to stay. Already it is possible to sequence the entire genome within 24 hours (for a price). The computing industry is already writing software to zip through the genome to make comparisons between different individuals, and to find genes. Thus, in a few years we will likely have chips to genotype individuals for every gene, and we will be in the business of estimating the effects of every allele of those genes. We may even estimate epistatic and epigenetic effects. The era of gene assisted selection will be upon us, selecting animals with the best combinations of alleles. This is the basis for my following comments.

Inbreeding

Katarzyna Stachowicz recently completed her PhD thesis on inbreeding and genetic diversity, in which she did a simulation of gene assisted selection versus traditional animal model selection on EBVs. A scenario with 30 chromosomes, 50 biallelic QTLs, 1000 females and 10 males, selection was followed for 6 generations. She compared selection procedures by the degree of identity by descent for each location on the chromosomes. With selection on EBVs, the degree of IBD was at about 0.2 throughout the 30 chromosomes. Every locus had equal selection pressure on it during the selection. With gene assisted selection, the average degree of IBD was lower around 0.1 throughout the 30 chromosomes, but at the locations of the QTLs, the degree of IBD was greater at 0.2 or higher, depending on the size of the QTL effect and the starting allele frequency. Thus, there was differential IBD rather than uniform IBD.

The **A** matrix of additive relationships only represents the average relationships between relatives, and would be entirely unsuitable for use in gene assisted selection. Should relationships be computed only for the QTLs? An inbreeding coefficient no longer conveys the proper message because it is based on

pedigree only. An inbreeding coefficient based on QTL genotypes would not be useful because we want a higher frequency of the good alleles of QTLs. How do we indicate the degree of homozygosity in the bad QTLs? The bad alleles could be selected against with gene assisted selection. My point here is that the concepts about inbreeding will need to be revised, and the models for genetic evaluation of individuals may become more biased due to using inappropriate relationship matrices.

Similarly, if the QTLs are becoming more highly IBD, then do we need to worry about conserving genetic variability? Should we make sure that all alleles survive or can we let some alleles disappear from our populations? Interbull should have someone monitoring genetic diversity in the dairy cattle populations, if it has the genotypes.

Genetic Evaluation

Most statistical analyses are built on the premise of having observed random samples of individuals from a population. Typical genetic evaluation models assume that animals are random variables from a population with null mean and a particular covariance matrix. If gene assisted selection is applied to dairy

cattle, then when we have a group of sons of a particular sire, the ones that will be progeny tested are those having the 'better' alleles for QTLs. The other sons will not have any progeny. The selected sons will not be a random sample of possible sons of that sire. Thus, the sons should not be used to evaluate their sire, and their sire perhaps should not be used to evaluate the sons. Hopefully, the sons will have a random sample of daughters from the dams to which they are bred. The use of the **A** matrix in genetic evaluation needs to be reconsidered, firstly because of the inbreeding issues previously discussed, and secondly because sons will not be random samples of their sire.

The animal model is becoming old technology and may suffer biases due to pre-selection of animals based on genes or markers. Therefore, instead of evaluating individuals, national genetic evaluations should be more about estimating the effects of alleles of all QTLs, and their interactions. Thus, instead of evaluating 2.5 million animals we should evaluate 25,000 QTLs (assume an average of 10 alleles per gene gives only 250,000 allelic effects), as either fixed or random effects (see Gianola's paper about that). This would avoid the non-random sampling of sons and inbreeding problems. Every animal would need to be sequenced and QTL alleles determined. Gene interactions could also be included if the models are not linear models. Given an animal's genotypes then an estimate of genetic merit could be derived. Thus, I agree totally with the proposal of Mike Goddard, and I agree with Daniel Gianola about thinking of models as abstractions, not necessarily linear models.

Conclusions

SNPs will be replaced by actual QTLs and their alleles in the near future through complete genome sequencing. All dairy animals should be sequenced and phenotyped.

Due to non-random sampling of sons of sires, genetic evaluation models should be modified to estimate allelic effects of QTLs and their interactions rather than estimated breeding values of individual animals by traditional animal models. This would eliminate the need for genetic relationship matrices.

The concepts around inbreeding coefficients need to be revised when gene assisted selection is applied. Identity by descent will not be homogenous through the genome, but will have spikes of high IBD at the locations of the QTLs that are being favourably selected.

A concern, when gene assisted selection is applied, will be maintaining genetic diversity and variability within populations, and to avoid loss of alleles. Detecting new mutations may also be part of this activity.

The future, in general, is unknown, but full of opportunities. Thank you.