Report on the potential of genomic selection in Irish dairy cattle

Donagh Berry
Teagasc, Moorepark

February 2007

Introduction

Simply put, genomic selection as proposed by Meuwissen et al. (2001), implies selection of animals based on the sum of all haplotype effects determined from the mean of the population. Advantages of such a method are a) resources do not have to be expended on identification and validation of quantitative trait loci (QTL) or markers in population linkage disequilibrium with causative genes, b) the reliability of estimated breeding values (EBVs) is considerably higher for dams of potential young test sires and also for the young sires themselves, and c) generation intervals of all potential selection pathways may be reduced thereby increasing annual genetic gain. Because of greater reliability of EBVs for young sires, potential fluctuations in sire rankings (e.g., calving dystocia) are reduced thereby instilling greater stakeholder confidence in the genetic evaluation system. Meuwissen et al. (2001) reported that the accuracy of EBV derived from genome wide sense marker maps could be as high as 0.80 at birth for moderately heritable traits. Furthermore, because of the greater accuracy of EBVs, greater selection intensity of germplasm can be adopted which, when coupled with the higher accuracy of selection and reduced generation intervals, results in increased genetic gains and reduced costs associated with progeny testing young sires.

The objective of this report is to apply the methodology outlined by Schaeffer (2006) to Irish conditions and determine the potential benefits of genomic selection in Irish dairy cattle. Although, the example used is based on dairy cattle, the same methodology may also be applied to beef.

General methodology

For a genome of 3,000 cM, 3001 informative markers are required at distances of 1-cM. Affeymetrix (www.affymetrix.com) currently produce a bovine genome array with 23,000 single nucleotide polymorphisms (SNPs). Haplotypes need to be constructed and therefore linkage phases of the markers are required necessitating two generations of genotyped animals (i.e., sires of sons). Because the markers are SNPs and only two alleles usually exist then four different haplotypes are possible for each contiguous pair of SNPs. What is key is that enough sires are initially genotyped so that the effect of each haplotype among the selection candidates is determined. Meuwissen et al. (2001) described a series of methods for calculating the effect of haplotypes in a population.

Potential dams of young sires can then be genotyped and their breeding value estimated as the sum of the haplotype effects. By definition therefore only additive genetic effects are included in the EBV and epistatic effects are ignored although the latter may be indirectly included if two or more haplotypes, exhibiting epistatic interactions, are in linkage disequilibrium with each other.

Potential application in Ireland

In determining the potential of genomic selection to increased farmer profit through genetic selection, several assumptions have to be made. However, sensitivity analyses are facilitated in Appendix 1. In the present report genomic selection is compared to the dairy cattle progeny testing scheme currently operational in Ireland.

Current selection practices in Ireland

Genetic gain in Irish Holstein-Friesian dairy cows until recently was largely dictated by breeding programs in other countries owing to our heavy reliance on imported germplasm. However, a national breeding program, G€N€ IRELAND, was recently launched whose objective is to identify the genetically superior male calves born in Ireland; the facility is also there to identify these candidate animals prior to birth. It is envisaged that contract matings, whereby the farmer is advised to mate a particular dam to a particular high genetic merit sire, may also be undertaken.

Candidate young test sires, as identified through the national database are rejected if the dam is either not pedigree registered, less than 76 points for overall type or had low phenotypic performance in the previous lactation. Selection of the final team of young test bulls is achieved by scrutinizing the pedigree of the remaining candidate male calves. The selected male calves are then subjected to progeny testing in Irish dairy herds.

Females will be inseminated with semen from these young test bulls when the bulls are approximately 14 months of age and their progeny will be born when the bull is approximately 24 months of age. The bull's progeny will themselves calve when he is 48 months of age and thus the bull will receive his first milk production evaluation based on his producing daughters when he is approximately 50-52 months of age. However, he will be at least 60-70 months of age before he has fertility information from daughters in Ireland. The cost of progeny testing a young test sire (including farmer incentives) is assumed to be €13,000. A total of 50 young test sires were tested in 2006 implying an overall cost of €650,000. However, a target should be to increase this number to 100 young test sires annually (i.e., €1.3 m).

Genomic selection in Ireland

Analysis of the genetic evaluation database on Holstein-Friesian sires revealed that 45 grandsires have at least 10 sire sons with more than 100 daughter records for milk production in Ireland; the number of sire sons per grandsire from this sample vary from 54 (Valiant) to 9. However, semen for DNA extraction may not be available for some of the more prominent and older grandsires; DNA of these sires is currently being procured. Having a large number of unrelated families is an advantage as it will allow from a greater diversity in haplotype permutations. Nonetheless, we may assume that we can genotype 700 sires (also with genotypic information on their grandsires). This is somewhat lower that what might be necessary to estimate all potential haplotype effects.

A genotyping cost of €350 per animal is assumed. Daughter yield deviations of the sires will be used to estimate the additive effect of each haplotype using methodology suggested by Meuwissen et al. (2001). It is of utmost importance that effects for the vast majority (preferably all) of the alternative combinations of haplotypes are estimated accurately.

Two options then exist: a two stage or one stage approach.

1) *Two stage approach.* In this approach potential dams of sires are genotyped and their breeding values estimated based on the sum of their respective haplotype effects. Genotypic information may also be used to select dams of young test sires in future years. When these test sires are born they are also genotyped, their breeding values estimated, and the genetically superior selected for widespread use. Assumptions include starting with a potential 1000 dams of which 500 are selected as sire dams. This results in approximately 250 male calves of which 20 are chosen based on superior genotypes while simultaneously having minimal effect on future genetic diversity. The genotypes of the dams can be used for producing more young test sires in future generations as well as identifying female calves with potential merit as sire mothers in subsequent generations. Cost of this scheme is €437,500 excluding the cost of estimating the haplotype effect.
2) *One stage approach.* Similar to the two stage approach except that dams are not preselected and therefore 1000 male calves are genotyped and the top 20 (with cognizance taken of relatedness between them) retained for widespread use. The advantage of the one stage approach is that it can potentially reduce the number (and cost) of genotyping although the accuracy of dam selection will be lower. The cost of this scheme is €350,000 excluding the cost of estimating the haplotype effect

Overall cost and impact on genetic gain

Table 1 summarises the genetic gain currently achievable using progeny testing in Ireland across the four alternative selection pathways assuming a heritability of 0.35 (e.g., milk yield). Hence, genetic gain under the current scheme is expected to be 0.199 genetic standard deviations per year (i.e., 4.725 / 23.75). The annual cost is €650,000 and thus the cost per genetic standard deviation change is €3.27m.

Table 1. *Current values for the alternative selection pathways using progeny testing*

Selection pathway	Percentage selected	Selection intensity	Accuracy of selection	Generation interval	Intensity*accuracy
Sires of sires	1%	2.66	0.92	8.15	2.45
Dams of sires	20%	1.4	0.55	3.94	0.77
Sires of dams	1%	2.66	0.53	7.63	1.41
Dams of dams	90%	0.20	0.50	4.03	0.10
Σ				23.75	4.725

Table 2 summarises the genetic gain achievable using the two stage approached previously outlined with genomic selection. It is assumed that the proportion of the population selected as parents of subsequent generations does not change. However, the accuracy of selection of sires of sires is lower while that of dams of sires and sires of

dams increase; the effect on the accuracy of selecting dams of dams does not change. The lower accuracy of the sire to sire pathway is because in this example young sires with information only on their genotypes are used as sires of subsequent generations. The accuracy of selection of 0.75 for the three main pathways may be conservative since Meuwissen et al. (2001) reported accuracy of select as higher as 0.85 using genome wide marker maps.

Table 2. *Values for the alternative selection pathways using genomic selection*

Selection pathway	Percentage selected	Selection intensity	Accuracy of selection	Generation interval	Intensity*accuracy
Sires of sires	1%	2.66	0.75	1.75	1.995
Dams of sires	20%	1.40	0.75	1.75	1.05
Sires of dams	1%	2.66	0.75	2.00	1.995
Dams of dams	90%	0.20	0.50	4.03	0.098
Σ				9.53	5.138

Genetic gain using genomic selection is expected to be 0.539 genetic standard deviations per year (i.e., 5.138 / 9.53). The initial cost of this scheme attributable to genotyping is €682,500 with an annual cost thereafter of €437,500. The greater cost in the first year is associated with genotyping sufficient sires to estimate the haplotype effects. It is not known how often these haplotype effects will have to be estimated, however, because of mutations and recombination, haplotype effects will have to be systematically estimated. Assuming the 20 selected young sires remain in service for 2 years results in an increased annual cost of €280,000 (i..e., €7,000*40; the full cost of proving a bull is not included since farmer incentives will not have to be paid due to the high accuracy of the young test sires). Hence the cost per genetic standard deviation change in the first year is €1.59 m. Annual costs thereafter will be lower.

Concluding remarks

This is a relatively new methodology and has not been extensively tested although Holland genetics have allegedly selected their first sire using this methodology. Ireland currently do not have the expertise or physical resources to undertake this project, but it is very likely that if we do not embark on such an undertaking we will be left behind. Future research will have to focus on methods of minimising loss of genetic diversity in subsequent generations using this technology. However, the increased accuracy of selection on the dams, using the two stage approach, may increase the likelihood of introducing new germplasm as well as reducing potential genetic bottlenecks from aggressive selection on sire families. Schaeffer (2006) suggests a “heterozygosity index” based on the genotypes of the dams. This may be incorporated into optimal contribution-type algorithms to maximize genetic gain while minimsing or constraining future inbreeding. Estimation of breeding values using genomic markers is similar to current traditional methods with the exception that the random animal effects are replaced by random haplotype effects and the numerator matrix is not required.

Recommendations

- This proposal should be discussed amongst the main bodies in animal breeding in Ireland and if deemed appropriate this proposal should be peer reviewed prior to any concrete recommendations to the industry.
- This project should be part funded by the exchequer with the deficit met by the dairy industry
- It is likely that the cost of genotyping may be reduced if large quantities are undertaken
- There may be potential to collaborate on this project (e.g., share information on genotypes of different sires) as well as statistical methods to analyze the data and develop breeding programs to maximize the benefit attainable from this technology.

References

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.

Schaeffer, L.R., 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123:218-223.

Appendix 1. Sensitivity analysis. Figures highlighted in yellow contain formulae. To update highlight and press F9

Progeny testing scheme

Cost to progeny test sire	13,000
Number of sires to progeny test	50

Genomic selection

Cost per genotype	350
Sire families	37
Average number of sires per family	22
Males to genotype	814
Dams to genotype	1000
Bull calves to genotype	200
Initial genotyping cost	€704,900
Lay-off annual cost of proven sire	7,000

Traditional progeny testing

Selection pathway	Percentage selected	Selection intensity	Accuracy of selection	Generation interval	Intensity*accuracy
Sires of sires	1%	2.66	0.92	8.15	2.4472
Dams of sires	20%	1.4	0.55	3.94	0.77
Sires of dams	1%	2.66	0.53	7.63	1.4098
Dams of dams	90%	0.196	0.5	4.03	0.098
TOTAL				23.75	4.725
Cost per genetic standard deviation change					€3,267,196

Genomic selection

Selection pathway	Percentage selected	Selection intensity	Accuracy of selection	Generation interval	Intensity*accuracy
Sires of sires	1%	2.66	0.75	1.75	1.995
Dams of sires	20%	1.40	0.75	1.75	1.05
Sires of dams	1%	2.66	0.75	2	1.995
Dams of dams	90%	0.20	0.5	4.03	0.098
TOTAL				9.53	5.138
Cost per genetic standard deviation change					€1,587,454