Proc. Assoc. Advmt. Anim. Breed. Genet. 22:57-60

GENOMIC ANALYSIS OF HEALTH TRAITS USING AN AUSTRALIAN GENOTYPED COW REFERENCE POPULATION

M. Abdelsayed[1], M. Haile-Mariam[1] and J.E. Pryce[1,2]

[1]Agriculture Victoria, Department of Economic Development, Jobs, Transport and Resources
[2]La Trobe University, Agribio, 5 Ring Road, Bundoora, VIC 3083, Australia

SUMMARY

In recent years there has been increasing interest internationally in estimating breeding values for traits that can reduce farm costs, such as health traits in livestock. One of the limitations in developing breeding values for health traits in Australia has been lack of data. In this study, we have estimated reliabilities of genomic breeding values for health traits when only clinical records on health disorders are used that are collected from a genomic reference population (Ginfo). Reliabilities for bulls with daughters in the reference population are 27%, and 25% for mastitis and an "all-disease" trait, respectively. For bulls with no daughters in the reference population, reliabilities are 4% and 12% for mastitis and the "all-disease" trait, respectively. In contrast, reliabilities for reproductive disorders and metabolic diseases were much lower (<15%). Mastitis and "all-diseases" have higher incidences and also higher heritability estimates than the other diseases, which is likely to be the reason for higher reliability estimates. Although estimates are still regarded as low, they are in line with expectations for a newly-recorded trait. Investigation into the improvement of reliabilities through the use of predictor traits through multi-trait analysis is the next step for this research.

INTRODUCTION

In the past, genetic selection for milk production was the main focus for the driver of dairy farm profitability. While making great genetic gains in milk production, an unfavourable relationship between production and disease resistance has become apparent (Pryce *et al*. 1997; Rauw *et al*. 1998; Koeck*et al*. 2012). Dairy cow health will continue to deteriorate if disease traits, or their predictors, are not included in breeding objectives. Healthy cows are more productive, easier to manage, require less intervention, have improved animal welfare and contribute to profitability by reducing production costs.

Health and fertility traits generally have low heritability estimates (<5%) compared to production traits (>30%) (Egger-Danner *et al*.2015). However, there is sufficient genetic variation to still make selection feasible for low heritable traits, and this has been evident in the dairy industry with the improvements made with selecting directly on fertility (Pryce *et al*. 2014).

Traits like health and fertility have large impacts on the dairy industry but sometimes data availability is low. One option is to obtain records from a dedicated reference population of genotyped cows with phenotypes of interest. This has already started in Australia with the establishment of the first 100 Genomic Information Nucleus herds (Ginfo). Ginfo was a large-scale genotyping project (103 herds and 32,386 cows) to increase the size of the Australian dairy reference population to improve the reliability of Australian genomic breeding values.

The objectives of this study were to estimate 'clinical' genomic health breeding values for the major disease traits such as mastitis, reproductive disorders, lameness, metabolic disorders and an overall "all-disease" trait using the health data collected from the Ginfo herds and secondly to determine the reliability of those estimated breeding values.

MATERIALS AND METHODS

Health data and genotypes. A total of 487,503 electronic health records were accessed from 90 (of 103) Ginfo herds. Genotypes were available on 15,632 cows that also had health records. Genotypes of 2,984 bulls with daughter health records were also obtained from DataGene.

Disease categories. The major disease traits (mastitis, reproductive disorders, lameness, and metabolic disorders) were converted into binary traits. Each disease was coded with a 0 or 1 for every cow-lactation record, where 1 corresponds to a cow having a particular disease at any time in a lactation period and 0 if it does not have that disease. For the "all-diseases" category, if a cow has any record of any disease event, it was coded 1, or otherwise 0 as healthy.

Genomic analysis. The reference dataset contained 11,458 genotyped Holstein cows (out of the total 15,632). The validation dataset contained 494 genotyped bulls, with 6,989 daughters that had health records (n = 22,276) but were not genotyped themselves, so not included in the reference set. Bulls with less than 5 daughters were excluded from the analysis.

For the estimation of genomic breeding values the following linear mixed animal model was used:

$$y = \mu + \text{HYS} + \text{Parity} + \text{MOC} + \beta_1\text{Agecalving} + \beta_2\text{Agecalving}^2 + \text{CowID} + \text{GRM} + e,$$

where y= observable health traits (binary trait 0 or 1), μ = trait mean, HYS = Herd-Year-Season contemporary group, Parity = 4 levels of parity (1, 2, 3, > 4), MOC = month of calving 1 to 12, Agecalving = age at calving from 18 months to 220 months (calving date – birth date) fitted as a covariate and 2nd order polynomial, CowID = random permanent environmental cow effect to account for repeated measures, GRM = random term for the genetic markers (SNPs), and e = random error term. The model was fitted using ASReml Version 4 (Gilmour *et al*., 2015).

Reliability of genomic prediction. Two methods were used to estimate the reliability of genomic prediction:

1. Theoretical (expected) reliability (R) $= 1 - \frac{\text{PEV}}{\sigma_g^2}$,

where, the prediction error variance (PEV) = squared standard error of the direct genetic value (DGV) for each animal in the dataset, and σ_g^2 is the additive genomic variance, obtained from the REML estimate.

2. Empirical (observed) reliability using cross-validation

$$= r(\text{DGV}, \text{DTD})^2$$

Cross-validation was performed by predicting DGVs for the 494 genotyped bulls that had daughters with health records but were not genotyped. Reliability was then estimated as a simple Pearson's squared correlation between the direct genomic breeding value (DGV) and the corrected phenotypes (residuals) which were used to calculate the daughter trait deviations (DTD) for each bull. The reliability was adjusted by dividing it by the average reliability of DTDs (h^2*average effective number of daughters for the genotyped bulls) (Haile-Mariam *et al*., 2012).

RESULTS AND DISCUSSION

A summary of the number of records used in the genomic analysis for each health trait is reported in Table 1 for Holsteins.

Table 1. Summary of the number of cow-lactations, cases of disease (n) recorded for each health trait (MAST = mastitis, REPRO = reproductive disorders, LAME = lameness, METAB = metabolic diseases, ALL DIS = "all-diseases") and heritability estimates ($\hat{h}^2$ ±standard errors) for Holsteins using all parity records

Traits	n	$\hat{h}^2$ ±S.E
Cow-Lac	33,000	
MAST	3,735	0.03 ± 0.004
REPRO	2,498	0.01 ± 0.002
LAME	248	0.00 ± 0.00
METAB	241	0.002 ± 0.002
ALL DIS	6,085	0.02 ± 0.004

Mastitis and the all disease category had the largest number of records followed by reproduction, lameness and metabolic disorder categories. The same patterns were also evident with the reliabilities of genomic predictions with the highest being mastitis and the all disease category, followed by reproductive and metabolic disorders (Table 2).

Table 2. Average expected reliabilities (R) of genomic breeding values for cows and bulls with daughters in the reference dataset and bulls in the validation dataset (V) and Cross-validation accuracy and reliability (r^2) for each health trait (MAST = mastitis, REPRO = reproductive disorders, LAME = lameness, METAB = metabolic diseases, ALL DIS = all diseases)

	Expected Reliability			Cross-Validation	
Traits	Bulls*	Cows	Bulls_V^	Accuracy	r^2
MAST	0.33	0.23	0.18	0.12	0.04
REPRO	0.15	0.09	0.05	0.02	0.004
METAB	0.04	0.01	-0.01	-0.01	0.003
ALLDIS	0.31	0.20	0.16	0.18	0.12

*Bulls with daughters in the reference set (n= 948); ^Bulls with no daughters in the reference set (n= 494)

The prediction error variance and cross-validation methods produce similar reliability estimates. The reliabilities are low but are comparatively higher for mastitis and the all disease category (Table 2). Bulls generally had higher reliabilities than cows, due to bulls having greater than 5 daughters in the data.

The lower reliability for metabolic disease is associated with fewer records in comparison to mastitis and the all disease trait. Further, mastitis and the "all-diseases" trait had higher heritabilities and incidences than the other disease traits (Table 2), possibly an indication of why their reliabilities are higher. There is still potential for improving these traits' reliabilities by two

means: 1) by including DTDs in the reference set for the 948 genotyped bulls, and 2) by incorporating predictor traits, for example inclusion of both mastitis and SCC data is expected to improve the reliabilities of GEBVs for mastitis.

The reliability estimate for lameness was unsatisfactory to report (R=0) due to the low number of records associated with this trait, and zero heritability. However, there may be merit in recording different types of lameness (e.g. laminitis, etc.) and developing new ways of recording, such as using a phone app. We expect that collection of more data and distinguishing between types of lameness may help to develop genomic breeding values for this trait.

CONCLUSIONS

Overall the results from this study are in line with expected reliabilities for new traits with comparatively small amounts of data and provide a good foundation for further improvement of reliabilities for health traits. It is encouraging that reasonable reliabilities were achieved for diseases such as mastitis and the all disease trait. Having more health event data being identified and made available to the dairy industry, and further investigation in combining predictor traits, will assist in providing genomic breeding values with greater reliability.

REFERENCES

Egger-Danner C., Cole J. B.,Pryce J. E., GenglerN., HeringstadB.,Bradley A. and StockK. F.(2015) *Anim.* **9**: 207.

Gilmour A. R., Gogel B. J.,Cullis B.R., Welham S. J. and Thompson R. (2015). ASReml User Guide Release 4.1 Structural Specification, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Haile-Mariam M.,Nieuwhof G. J., Beard K.T., Konstatinov K.V.and Hayes B. J. (2012) *J. Anim. Breed. Genet.* **130**:31.

Koeck A., Miglior F., Kelton D. F. and Schenkel F. S. (2012) *J. Dairy Sci*. **95**: 4099.

Pryce J.E., Veerkamp R.F., Thompson R., Hill W.G. and Simm G. (1997) *Anim. Sci*. **65**: 353.

Pryce J.E., Woolaston R., Berry D.P., Wall E., Winters M., Butler R. and Shaffer M. (2014) World Trends in Dairy Cow Fertility. *Proc. 10th World Congress Genet. App.Livest. Prod.***154**

RauwW. M., Kanis E.,Noordhuizen-Stassen N. and GrommersF. J. (1998) *Livest. Prod. Sci.***56**:15.