GENOMIC PREDICTION OF METABOLIC PROFILES IN DAIRY COWS

T.D.W. Luke[1,2], T.T.T. Nguyen[1], S. Rochfort[1,2], W.J. Wales[3], C. Richardson[1,2], M. Abdelsayed[4] and J.E. Pryce[1,2]

[1]Agriculture Victoria Research, AgriBio, Centre for AgriBioscience, Bundoora, Victoria, 3083 Australia
[2]School of Applied Systems Biology, La Trobe University, Bundoora, Victoria, 3083 Australia
[3]Agriculture Victoria Research, Ellinbank Centre, 1301 Hazeldean Rd., Ellinbank, Victoria, 3820 Australia
[4]Datagene Ltd., AgriBio, Centre for AgriBioscience, Bundoora, Victoria, 3083 Australia

SUMMARY

Improving animal health and resilience is an increasingly important breeding objective for all livestock industries. In this study we estimated genetic parameters of serum metabolic profiles in early lactation dairy cows. A single serum sample was taken from 1,393 cows, located on 14 farms in south eastern Australia, within 30 days after calving. Sera were analysed for biomarkers of energy balance (β hydroxybutyrate (BHBA) and non-esterified fatty acids (NEFA)), mineral status (Ca and Mg), protein nutrition (urea and albumin) and immune status (globulins and albumin to globulin ratio (A:G)). After editing, 47,162 single nucleotide polymorphism marker genotypes were used for estimating genomic heritabilities and breeding values (gEBV) for these traits in ASReml. Heritabilities were low for BHBA, NEFA, Ca, Mg and urea (0.09, 0.18, 0.07, 0.19 and 0.18, respectively), and moderate to high for albumin, globulins and A:G (0.27, 0.46 and 0.41, respectively). The accuracy of genomic predictions was assessed by (1) calculating empirical accuracy using 5-fold cross validation, and (2) calculating theoretical accuracy using the prediction error variance obtained from ASReml. Empirical accuracies ranged from 0.20 to 0.40, being higher for traits with higher heritabilities. Theoretical accuracies were higher than respective empirical accuracies (0.31 – 0.51), but the results of the 2 methods were in excellent agreement ($R^2 = 0.89$). While increasing the size of the reference population should theoretically improve accuracies, our results suggest that genomic prediction may allow identification of healthier cows that are less susceptible to diseases in early lactation.

INTRODUCTION

Most disease events affecting dairy cows occur in the first 30 days after calving (LeBlanc *et al.* 2006) and many of these diseases are associated with metabolic disorders such as ketosis and hypocalcaemia (Ospina *et al.* 2010). While heritability estimates of metabolic disorders are generally low (Uribe *et al.* 1995), sufficient genetic variance exists to suggest that improvements in metabolic stability can be achieved through genetic selection.

One way of assessing the metabolic health of cattle is serum metabolic profiling, which employs well-established epidemiological associations between the concentrations of several metabolites in serum, and the presence of both subclinical and clinical metabolic disorders (Payne *et al*. 1970). These metabolites include those associated with energy balance (BHBA and NEFA), mineral status (Ca and Mg), protein nutrition (urea and albumin) and immune status (globulins and albumin to globulin ratio). While extremely valuable, these phenotypes are costly and invasive to collect, making them impractical for traditional large-scale genetic evaluations. Genomic selection offers exciting potential for achieving genetic improvement in such difficult to measure and lowly heritable traits, by using data obtained from relatively small genotyped reference populations with high quality phenotypic data.

The objectives of this study were to (1) estimate the genetic parameters of serum biomarkers of

health in early lactation dairy cows using data collected from a genotyped female reference population, and (2) estimate the accuracy of genomic predictions of serum biomarker concentrations. If sufficiently accurate, genomic selection for metabolic stability offers the potential to provide permanent and incremental improvements in dairy cow health and welfare, thereby increasing farm profitability and sustainability.

MATERIALS AND METHODS

Phenotypes. A single serum sample was taken from of 1,393 Holstein-Friesian cows from 14 farms in south eastern Australia between August 2017 and October 2018, according to the protocol described in Luke *et al.* (2019). All animals had been calved 30 days or less at the time of sampling. Sera were analysed for biomarkers of energy balance (BHBA and NEFA), mineral status (Ca and Mg), protein nutrition (urea and albumin) and immune status (globulins) by Regional Laboratory Services (Benalla, Victoria, Australia). Descriptive statistics of phenotypes are shown in Table 1.

Genotypes. Genotypes for the 1,393 animals used in this study were provided by DataGene Ltd. (Victoria, Australia). After editing, 47,162 single nucleotide polymorphism (**SNP**) markers were available for genomic analyses. A genomic relationship matrix (**GRM**) was constructed according to Yang *et al.* (2010).

Genetic parameters. Variance components were estimated for each trait using univariate linear mixed animal models in ASReml (Gilmour *et al.* 2015). In matrix notation, the model used was $\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$ (Model 1), where **y** is a vector of metabolite concentrations (BHBA, NEFA, Ca, Mg, urea, albumin, globulins), **b** is a vector of fixed effects of DIM, herd, parity, age and sample collection date, **u** is a vector of random genetic effects, and **e** is a vector of the random residual effects; and **X** and **Z** are incidence matrices for **b** and **u** respectively. It is assumed that $\text{var}(\mathbf{u}) = \mathbf{GRM}\,\boldsymbol{\sigma}_u^2$, and $\text{var}(\mathbf{e}) = \mathbf{I}\boldsymbol{\sigma}_e^2$. Estimated variance components were then used to calculate the genomic heritability of each biomarker.

Genomic predictions. Genomic estimated breeding values (**GEBV**) were predicted using genomic best linear unbiased prediction (**gBLUP**), using variance components estimated from the univariate model (Model 1). The accuracy of genomic predictions was assessed in 2 ways. Firstly, empirical accuracy was calculated using 5-fold cross validation. This involved randomly dividing the reference population into 5 equally sized groups or folds. Data from 1 fold (approximately 20% of the reference population) were set aside as a validation set, and data from the remaining 4 folds (approximately 80% of the reference population) formed the training set for model development. The resulting model was then used to predict GEBVs for animals in the validation set. This was repeated 5 times so that all animals appeared in the testing set once. Empirical accuracy was then calculated as the Pearson's correlation between the predicted GEBVs and actual phenotype values, corrected for the fixed effects described in Model 1. Predicted accuracies of the true breeding values were calculated by dividing the empirical accuracies by the square root of the heritability of the trait. Secondly, theoretical accuracy was calculated as

$$\mathrm{r}_i = \sqrt{1 - \frac{SE_i^2}{\sigma_g^2\, GRM_{ii}}}$$

where SE_i SE_i is the standard error of GEBV of individual *i*, and $\sigma_g^2\sigma_g^2$ is the genetic variance of each trait estimated from Model 1, adjusted for inbreeding by multiplying by the corresponding diagonal elements in the GRM for each individual (GRM_{ii}).

RESULTS AND DISCUSSION

Estimated heritabilities for all traits, obtained from Model 1, are shown in Table 1. Heritability estimates were low for serum BHBA, NEFA, Ca, Mg and urea at 0.09 0.18, 0.07 0.19 and 0.18, respectively. Heritabilities of albumin, globulins and A:G were higher at 0.27, 0.46 and 0.41, respectively. Standard errors for all heritabilities were low (0.04 - 0.06).

Heritability estimates were consistent with the literature for NEFA (Oikonomou *et al.* 2008), Mg (Tsiamadis *et al.* 2016), albumin, globulins and A:G (Cecchinato *et al.* 2018). We could find no reports of the heritability of serum urea concentration in the literature, however our results are consistent with the reported heritability of milk urea nitrogen (Mitchell *et al.* 2005), the concentration of which is linearly correlated with serum urea.

The heritability of serum BHBA in our dataset was 0.09 ± 0.04, which is in excellent agreement with the findings of Weigel *et al.* (2017) (0.093 ± 0.045), slightly lower than those of van der Drift *et al.* (2012) (0.17 ± 0.06), and considerably lower than those of Oikonomou *et al.* (2008) and Cecchinato *et al.* (2018) (0.40 ± 0.06 and 0.37 ± 0.14, respectively). Oikonomou *et al.* (2008) demonstrated that the heritability of BHBA concentration is highest in immediately post calving and decreases rapidly over the first 7 weeks of lactation. In our study only 209 cows were in the first week of lactation at the time of sampling, and we expect that adding more data from animals in this highest risk period could improve heritabilities. Similarly, the heritability of Ca in our dataset was 0.07 ± 0.04, significantly lower than reported by Tsiamadis *et al.* (2016) who found that the heritability of serum Ca at days 1, 2, 4 and 8 post-partum ranged from 0.23 (± 0.02) to 0.32 (± 0.03). Serum Ca concentrations drop in the 12 to 24 hours immediately post-calving before rapidly returning to normal physiological levels once homeostatic mechanism are restored, and it is likely that our low heritability estimate is the result of having sampled only 14 cows in this period of highest phenotypic variability. These results demonstrate the importance of careful trait definition when investigating the genetic parameters of health traits in the transition period.

Table 1. Number of samples (n), phenotypic means (μ) and standard deviations (σ), estimated genomic heritabilities (± standard errors), empirical reliabilities, and theoretical reliabilities of serum metabolic profiles

Phenotype	n	μ	σ	h^2	r_e	r_t
BHBA	1393	0.48	0.22	0.09 ± 0.04	0.29	0.34
NEFA	1393	0.55	0.33	0.18 ± 0.05	0.36	0.41
Ca	1327	2.31	0.18	0.07 ± 0.04	0.20	0.31
Mg	1294	0.98	0.14	0.19 ± 0.06	0.28	0.41
Urea	1393	5.24	0.17	0.18 ± 0.05	0.30	0.41
Albumin	1294	32.8	2.95	0.27 ± 0.06	0.38	0.44
Globulin	1294	38.4	6.04	0.46 ± 0.06	0.40	0.51
A:G	1294	0.88	0.17	0.41 ± 0.06	0.40	0.49

Accuracies of genomic predictions resulting from univariate models are shown in Table 1. Empirical accuracies of the true breeding values were low to moderate (0.20 and 0.40), with more heritable traits having higher prediction accuracies. Theoretical accuracies, calculated from the standard errors estimated from Model 1, were higher than respective empirical accuracies, but the results of the 2 methods were in excellent agreement ($R^2 = 0.89$). Although low, our results are consistent with a small female reference population and low to moderate trait reliabilities (Gonzalez-Recio *et al.*

2014). We expect that increasing the size of the reference population and refining trait definitions to maximise heritabilities should improve genomic prediction accuracies. Given the cost and logistical challenges of blood sampling large numbers of cows, one method for dramatically increasing the number of phenotypes may be to use mid-infrared spectroscopy of milk to predict serum biomarker concentrations. Other high throughput metabolomic methods such as nuclear magnetic reasonance spectroscopy may also offer potential for the discovery of novel biomarkers of health in milk and serum, which could help to further improve the genomic prediction accuracies.

CONCLUSIONS

Our results show that genetic variance exists in the concentration of biomarkers of energy balance, protein nutrition, micromineral status and immune status in early lactation dairy cows. Genomic prediction accuracies were low, and while increasing the size of the reference population should theoretically improve accuracies, our results suggest that genomic prediction may allow identification of healthier cows that are less susceptible to diseases in early lactation.

ACKNOWLEDGEMENTS

The authors thank the farmers involved and acknowledge DairyBio (Agriculture Victoria, Dairy Australia and Gardiner Foundation) for funding this study.

REFERENCES

Cecchinato A., Bobbo T., Ruegg P.L., Gallo L., Bittante G. and Pegolo S. (2018) *J Dairy Sci*. **101**: 11108.

Gilmour A.R., Gogel B.J., Cullis W.S.J. and R. Thompson. (2015) ASReml. 4.1 ed. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Gonzalez-Recio O., Coffey M. P. and Pryce J. E. (2014) *J Dairy Sci*. **97**:7905.

LeBlanc S.J., Lissemore K.D., Kelton D.F., Duffield T.F. and Leslie K.E. (2006) *J Dairy Sci*. **89**: 1267.

Luke T.D.W., Rochfort S., Wales W.J., Bonfatti V., Marett L. and Pryce J. E. (2019) *J Dairy Sci*. **102**: 1747.

Mitchell R.G., Rogers G.W., Dechow C.D., Vallimont J.E., Cooper J.B., Sander-Nielsen U., and Clay J.S. (2005) *J Dairy Sci*. **88**: 4434.

Oikonomou G., Valergakis G.E., Arsenos G., Roubies N., and Banos G. (2008) *J Dairy Sci.* **91**: 2814.

Ospina P., Nydam D., Stokol T., and Overton T. (2010) *J Dairy Sci.* **93**: 1596.

Payne J., Dew S.M., Manston R., and Faulks M. (1970) *Vet. Rec*. **87**: 150.

Tsiamadis V., Banos G., Panousis N., Kritsepi-Konstantinou M., Arsenos G. and Valergakis G.E. (2016) *J Dairy Sci*. **99**: 5535.

Uribe H.A., Kennedy B.W., Martin S.W. and Kelton D.F. (1995) *J Dairy Sci*. **78**: 421.

van der Drift S.G., van Hulzen K.J., Teweldemedhn T.G., Jorritsma R., Nielen M. and Heuven H.C. (2012) *J Dairy Sci.* **95**: 6781.

Weigel K.A., Pralle R.S., Adams H., Cho K., Do C., and White H.M. (2017) *J Anim. Breed. Genet.* **134**: 275.

Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., Madden P.A., Heath A.C., Martin N.G., Montgomery G.W., Goddard M.E., and Visscher P.M. (2010) *Nat. Gen.* **42**: 565.