GENOMIC BREEDING VALUES OF HEAT TOLERANCE IN AUSTRALIAN DAIRY CATTLE

T.T.T. Nguyen[1], P.J. Bowman[1], M. Haile-Mariam[1], J.E. Pryce[1,2], B.J. Hayes[1,2]

[1]BioSciences Research Division, Department of Economic Development, Jobs, Transport and Resources, 5 Ring Road, Bundoora, Victoria 3083, Australia
[2]La Trobe University, Bundoora, Victoria 3083, Australia

SUMMARY

In this study, we aimed to develop genomic estimated breeding values for heat tolerance in Australian dairy cattle. We combined test-day herd recording data with temperature and humidity measurements (in the form of temperature-humidity index or THI) from weather stations that were closest to the herds for test days between 2003 and 2013. Tolerance to heat stress was then estimated for each cow using random regression (intercept and slope) to model the rate of decline in production with increasing THI accumulated over the four days prior to the day of milking, for milk, fat and protein yields. The cow slopes from this model were used to define daughter trait deviations (DTD) for their sires. Data were analysed separately for Holsteins and Jerseys. The reference population for genomic prediction was 2,300 Holstein and 575 Jersey genotyped sires with DTD for response to heat stress for milk, fat and protein yield. With this reference, and using GBLUP, the range in accuracy of genomic predictions for heat tolerance across traits were 0.38 – 0.53 and 0.49 – 0.63 for 435 Holstein and 135 Jersey validation sires, respectively. When 2,191 Holstein and 1,190 Jersey cows were added in the reference populations, no substantial improvements in accuracy were observed. Genomic selection appears to be a useful tool to enable farmers to improve milk production in environments with higher heat load.

INTRODUCTION

Changes in environmental factors such as air temperature, humidity, air flow and radiation beyond the comfort zone of animals will lead to heat stress (Armstrong 1994). Heat stress in dairy cattle is an important issue as it results in reduced milk yield (Hayes *et al.* 2003), reduced fertility (Haile-Mariam *et al.* 2008) and therefore reduced profitability (St-Pierre *et al.* 2003). As the temperature in Australia is projected to continue to increase, the future of dairy farming will need measures to adapt to heat stress.

One way to address the challenge posed by heat stress is to apply management measures such as providing shade, fans and sprinklers to cows. Another approach that may have greater benefits in the long term is to select animals with better heat tolerance, as it has been demonstrated that variation in heat stress response is heritable (Hayes *et al.* 2003; Haile-Mariam *et al*. 2008). That is, production and fertility of some animals are less affected by heat stress than others and therefore these animals could be valuable candidates for selection. The genetic gain for heat tolerance will be greatest if accurate genomic estimated breeding values are available, as this will enable selection of young bulls and heifers.

In this study, we used dense DNA markers, together with field production and climate data, to develop GEBV for heat tolerance for dairy cattle in Australia.

MATERIALS AND METHODS

Hourly climate data including dry bulb temperature and relative humidity (%) were obtained from the Bureau of Meteorology (Melbourne, Australia) for all weather stations in Australia from 2003 to 2013. Average temperature-humidity index (THI) on the test day and 1, 2, 3 and 4 d before the test day were calculated following Hayes *et al.* (2009). The first-lactation test-day records between 2003 and 2013 for milk, protein and fat yield were extracted from the Australian Dairy Herd Improvement Scheme (ADHIS) database for Holstein and Jersey cows. Production records were merged with THI from the nearest weather station, or the second nearest station if it had a lower number of missing records. In total, THI from 105 weather stations were matched to production records of 1,655 Holstein and 501 Jersey dairy herds.

In our dataset, genotypes were available for a total of 2,735 Holstein and 710 Jersey sires. Illumina Bovine High-Density genotypes (777,963 SNP markers) were available for 1,620 of the Holstein sires and 125 of the Jersey sires. For all other sires, 50K (56,430 SNP) genotypes were available. After quality control and removal of non-polymorphic SNPs, 632,004 SNPs remained for animals genotyped at high density and 43,425 SNP remained for animals genotyped at the lower density. All animals genotyped at the lower density had genotypes imputed to the higher density SNP panel using BEAGLE 3 (Browning and Browning 2009).

All statistical analyses were undertaken separately for Holstein and Jersey. Mixed linear models were used to fit the data with variance components estimated using maximum likelihood in ASReml (Gilmour *et al.* 2009).

A random regression model was used to derive individual cow sensitivity to changes in THI of milk, fat and protein yields (i.e. the slope of the regression, or cow slope): $y_{ijl} = \mu + HTD_i + YS_j + \sum_{n=1}^{3} A_n X_n + \sum_{n=1}^{8} D_n Z_n + \sum_{n=0}^{l} P_{ln} W_n + e_{jli}$ (model 1), where y_{ijl} is yield of milk in litres, fat in kg x100 or protein in kg x 100 from the i^{th} herd test day, j^{th} year season of calving, and l^{th} cow in her first lactation; μ is the intercept, HTD_i is the effect of the i^{th} herd test day; YS_j is the effect of the j^{th} year season of calving; X_n is the n^{th}-order orthogonal polynomial corresponding to age on day of test; A_n is a fixed regression coefficient of milk/fat/protein yield on age at test; Z_n is the n^{th}-order orthogonal polynomial corresponding to days in milk (DIM) at test; D_n is a fixed regression coefficient of milk/fat/protein yield on DIM; P_{ln} is a random regression coefficient on THI for the l^{th} cow; W_n is either the intercept or slope solution for heat load index for cows; and e_{ijl} is the vector of residual effects. In this random regression model, all THI values below 60 were set to 60 (Hayes *et al.* 2009).

The effects of the sires (sire slope) to sensitivity of milk, protein and fat yield of cows to changes in THI were obtained using the following model: $y_i = \mu + Sire_i + e$ (model 2), where, y_i is a vector of slope value for a daughter of the i^{th} sire obtained from model 1, $Sire_i$ is the effect of the i^{th} sire on cow slope ~ $N(0, \sigma_s^2)$, e is the vector of residuals ~ $N(0, \boldsymbol{I}\sigma_e^2)$ where $\boldsymbol{I}$ is identity matrix and σ_e^2 is residual variance.

Proportion of additive variance in cow slope was calculated as 4 times of sire variance divided by total variance obtained from model 2.

To assess the accuracy of using genomic breeding values to predict heat tolerance, in each breed, the sires were split into a reference and a validation population. These populations were split by age, with sires born before 2005 included in the reference population, and sires born in or after 2005 placed in the validation population for Holsteins; sires born before 2004 were included the reference population, and sires born in or after 2004 were placed in the validation population for Jerseys. Sires that are paternal half-sibs were placed in either the reference set or the validation set. The genomic breeding values (GEBV) of the validation sires (the phenotype of which were not included in the analysis), were estimated by GBLUP using model: $y = \mu + Zg + e$ (model 3), where y is a vector of sire slope values (solutions for model 2), μ is the intercept, Z is a design matrix relating records to genomic breeding value of animals, g is a vector of genomic breeding values ~ $N(0, \boldsymbol{GRM}\sigma_g^2)$, where $\boldsymbol{GRM}$ is the genomic relationship matrix, σ_g^2 is the additive genetic variance captured by SNP, and e is a vector of random residuals ~ $N(0,\ \boldsymbol{W}\sigma_e^2)$, where $\boldsymbol{W}$ is identity matrix and ${\sigma_e}^2$ is residual variance. The genomic relationship matrix (***GRM***) was constructed amongst all genotyped individuals following Yang *et al.* (2010).

Proportion of additive variance of sire slope that is explained by SNP was calculated as additive variance divided by total variance obtained from model 3.

In some analyses the genotyped cows were used in the reference population as well as reference sires, and in this case cows that were daughters of validation bulls were excluded from the analyses. A similar model to model 3 was fitted to the reference data, but in this case the difference in residual variances for bull and cow phenotypes were taken into account by constructing the diagonal matrix $\boldsymbol{W}$ as $g(1/w_i)$, where w_i is the weighting coefficient for the i^{th} animal. Weighting coefficient was calculated differently for bulls and cows following Garrick *et al.* (2009), as follows:

$$w_i(bulls) = \frac{1-h^2}{ch^2+\frac{4-h^2}{p}} \quad ; \quad w_i(cows) = \frac{1-h^2}{ch^2+\left[\frac{1+(n-1)t}{n}-h^2\right]}$$

where $h^2$is the heritability of heat tolerance, c is the proportion of variance not explained by SNP (c =0.2), d is the effective number of daughters, n is the number of repeat records (i.e. lactations), r is the number of records per cow, and t is the repeatability of the trait (average repeatability for cow slopes in relation to milk, fat and protein was 0.34 for Holsteins and 0.44 for Jerseys in the current datasets).

The accuracy of genomic prediction was taken as the correlation of the genomic estimated breeding values, and the slopes for the validation sires (from model 2), divided by square root of the proportion of variance of sire slope explained by SNP obtained from model 3.

RESULTS AND DISCUSSION

Table 1. Correlations between sire slopes and GEBV, and accuracies of genomic estimated breeding values using 632,004 SNP panel for heat tolerance on milk, fat and protein yield using GBLUP

Breed	Reference	Validation	Production traits affected by heat stress	Correlation	Accuracy
Holstein	2,300 sires	435 sires	Milk yield	0.26	0.46
			Fat yield	0.22	0.38
			Protein yield	0.26	0.47
	2,300 sires + 2,191 cows	435 sires	Milk yield	0.27	0.48
			Fat yield	0.22	0.38
			Protein yield	0.29	0.53
Jersey	575 sires	135 sires	Milk yield	0.34	0.49
			Fat yield	0.40	0.60
			Protein yield	0.42	0.63
	575 sires + 1,190 cows	135 sires	Milk yield	0.39	0.56
			Fat yield	0.39	0.60
			Protein yield	0.43	0.64

There was substantial test-day milk, fat and protein yield variation in the datasets. In all dairy farming regions of Australia, cows experienced some degree of heat stress, which was variable among regions. For example, in Queensland the median of daily temperatures and **THI** were 20.7 and 67.3, respectively. Northern Victoria experienced a wide spectrum of weather conditions, with average daily temperatures ranging from -5°C to 38°C (median of 13°C), and **THI** ranging from 33.8 to 81.0 (median 56.6).

The proportion of additive variance in cow slope in relation to milk, fat and protein yields were 0.14, 0.16 and 0.13 in Holsteins and 0.27, 0.26 and 0.21 in Jerseys, respectively. The corresponding proportions of variance explained by SNP in sire slope were 0.32, 0.34 and 0.30 in Holsteins and 0.49, 0.44 and 0.45 in Jerseys. This confirms that selection for heat tolerance is possible and could be particularly effective in the geographical regions with high heat load. Correlations of sire slopes among milk, fat and protein yield ranged from 0.46 – 0.86 for Holsteins and 0.60 – 0.86 for Jerseys.

Genetic correlations between heat tolerance and production traits in cows were negative. For example, in Holsteins the correlations between heat tolerance with milk, fat and protein yields were -0.38, -0.40 and -0.54, respectively. This confirms the antagonistic relationships between heat tolerance and production traits reported in previous studies (Ravagnolo and Misztal 2000).

Correlations between sire slopes and GEBV, and accuracies of genomic prediction are presented in Table 1. When only genotyped sires were included in the reference population, accuracies of genomic prediction ranged from 0.38 to 0.53 in Holsteins and 0.49 to 0.63 in Jerseys. When the reference set was expanded to include genotyped cows, the accuracies of genomic predictions showed a slight increase in some cases but not all.

The accuracies of genomic predictions for heat tolerance we have reported indicate that genomic selection offers a promising tool to predict heat tolerance for individual animals based on their genotypes. This will enable farmers to improve the milk production at higher heat load conditions of their herd over time through selection decisions.

ACKNOWLEDGEMENTS

The authors thank the Department of Agriculture of Australia (Canberra, Australia) and Dairy Futures CRC for funding this project and the Australia Dairy Herd Improvement Scheme (Melbourne, Australia) for providing production data.

REFERENCES

Armstrong D. V. (1994). *J. Dairy Sci.* **77**: 2044.

Bernabucci U., Biffani S., Buggiotti L., Vitali A., Lacetera N. and Nardone A. (2014). *J. Dairy Sci.* **97**: 471.

Browning B. L. and Browning S. R. (2009). *Am. J. Hum. Genet.* **84**: 210.

Garrick D. J., Taylor J. F. and Fernando R. L. (2009). *Genet. Select. Evol.* **41**:1.

Gilmour, A. R., Gogel B. J., Cullis B. R. and Thompson R. (2009). VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Haile-Mariam, M., Carrick M. J. and Goddard M. E. (2008) *J. Dairy Sci.* **91**: 4840.

Hayes, B. J., Bowman P. J., Chamberlain A. J., Savin K., van Tassell C. P., Sonstegard T. S. and Goddard M. E. (2009) *PLOS One* **4**: 1.

Hayes, B. J., Carrick M., Bowman P. and Goddard M. E. (2003) *J. Dairy Sci.* **86**: 3736.

Ravagnolo, O. and Misztal I. (2000) *J Dairy Sci* **83**: 2126.

St-Pierre, N. R., Cobanov B. and Schnitkey G. (2003) *J. Dairy Sci.* **86**: 52.

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. Genet.* **42**:565.