GENOMIC PREDICTIONS FOR FERTILITY TRAITS IN TROPICAL BEEF CATTLE FROM A MULTI-BREED, CROSSBRED AND COMPOSITE REFERENCE POPULATION

B.J. Hayes, G. Fordyce and S. Landmark

Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD, 4067 Australia

SUMMARY

Cow fertility is a major driver of profitability in Northern beef herds. Cow fertility has been difficult to select for, and the availability of genomic estimated breeding values (GEBV) would enable more rapid gains to be made. Ideally GEBV would be from a multi-breed genomic evaluation, given the wide range of breeds, composites and crossbreds used in Northern Australia. With this ultimate goal in mind, 14,552 heifers in 54 herds across Northern Australia were genotyped and phenotyped for CLscore (presence or absence of a corpus luteum at approximately 600 days), a proxy trait for age at puberty, a trait in turn correlated with cow lifetime productivity. Genomic heritabilities estimated from the data set were 0.32, 0.42, 0.22 and 0.25 for weight, hip height, body condition score and CLscore respectively. The accuracy of GEBV in nine validation herds (where accuracy was the correlation of GEBV for CLscore and the actual CLscore for the heifers within a herd, representing a wide range of breed composition were 0.30, 0.50, 0.25 and 0.40 for weight, hip height, body condition score and CLscore respectively. For CLscore, this accuracy suggests gains for fertility could be made through selection on GEBV. The data set analysed here represents approximately half the data that will be collected in the Northern Genomics Project.

INTRODUCTION

Cow fertility is a key driver of productivity and profitability of beef production in northern Australia (Taylor and Rudder 1986; Fordyce 2012; Johnstone *et al.* 2014). Genomic estimated breeding values (GEBV) for cow fertility would enable more rapid genetic gains for these traits. However, accurate genomic evaluations for low heritability traits such as fertility require large reference populations (e.g. Goddard and Hayes 2009), with thousands of cows measured for both the traits of interest and genotyped for genome wide markers. Assembling such large reference populations for each breed used in Northern Australia is likely to be challenging, as the cattle population consists of many breeds, crossbreds and composites. Cattle populations include high proportion *Bos indicus* breeds (e.g. Brahman), stabilised composites (e.g. Droughtmaster and Santa Gertrudis), adapted *Bos taurus* breeds, and many composites. An alternative to constructing reference populations within each breed is to use multi-breed genomic evaluations, where the reference set includes cows from across Northern Australia.

Here we test the accuracy of GEBV for fertility and other traits from using such a reference set, in this case consisting of 14,552 heifers (reference and validation) from 54 commercial properties across Northern Australia.

MATERIALS AND METHODS

Animals and Phenotypes. Fifty-four collaborator herds from across Northern Australia are participating in the Northern Genomics project. The data set includes crossbred and, in some cases, purebred Angus, Belmont Red, Brahman, Charolais, Droughtmaster, Shorthorn, Limousin, Santa Gertrudis, Boran and Wagyu heifers. The fertility trait measured on the heifers to date is cycling or

not cycling by approximately 600 days (CLscore) assessed by ovarian scanning, as described by (Corbet *et al.* 2018). To maximise genetic variation, the trait is actually measured when an estimated 50% of heifers are pubertal, ie, at 1.0-2.5 years of age. As an alternative to CLscore, CLrate was also measured, where 1 = Acyclic, 2 = Dominant follicle 10mm or less, 3 = Dominant follicle greater than 10mm, 4 = Corpus luteum is present, and 5= Cow is pregnant (Burns *et al*. 2016). Weight, body condition score, hip height, fly lesions, and tick scores were also collected at the time of scanning. Tail hairs have currently been taken from all heifers for genotyping.

Genotypes. All heifers were genotyped with the 35K tropBeef SNP array by Neogen, Australasia. Genotypes were imputed up to 728,785 SNP (Bovine HD array) using the Fimpute software (Sargolzaei, *et al.* 2014), and a panel of 3,140 cattle of relevant breeds genotyped for the Bovine HD array.

Statistical Analysis. We first estimated breed proportions of each heifer for each of the 12 breeds known to be in the data set (using the 35K array data only). Previously, a separate large data set consisting of only purebred cattle was used to estimate SNP effects for breed composition. A GBLUP model was fitted, where the phenotype was 1 if the animal was of that breed and 0 if not (Dodds *et al*. 2014). The effects of each SNP for the proportion of each breed was then derived by back-solving for the SNP effects (Yang *et al*. 2011), and the resulting prediction equations for each breed were used to estimate breed proportions in the heifers. Then the model fitted to the CLscore, CLrate, height and weight data was

$$y = \mu + cohort + year + het + breedprop + animal + \text{error}$$

where **y** is a vector of trait records (CLscore,CLrate,weight, hip height or body condition score, μ is the population mean, **cohort** is the property+yeardrop+paddock that the heifers were in prior to mustering for trait recording, year is the year of recording, **het** is the heterozygosity of each heifer as measured by the proportion of markers that were heterozygous (to capture heterosis effects), fitted as a liner effect **breedprop** is a series of 12 covariates (11 breeds and *Bos indicus* content), measuring the proportion of each breed in the heifers as described above, and **animal** is a vector of random effects $\sim N(0, \mathbf{G}\sigma_g^2)$, with G the genomic relationship matrix among all heifers (Yang *et al*. 2011) and σ_g^2 the genetic variance captured by the SNP markers, and **error** is a vector of random deviations $\sim N(0, \mathbf{I}\sigma_e^2)$. Variance components were estimated in GCTA (Yang *et al*. 2011), and the heritability of the traits (actually the proportion of phenotypic variance captured by the SNP) was estimated as $h^2=\widehat{\sigma_g^2}/(\widehat{\sigma_g^2}+\widehat{\sigma_e^2})$.

The accuracy of GEBV was evaluated by dropping out 9 herds at random (but these 9 herds had to have at least two-year cohorts in the data set). The breed composition within the 9 herds (2,205 heifers) ranged from purebred Brahman to crossbreds of *Bos taurus* breeds. There were 12,347 heifers in the reference population. GEBV were predicted for the heifers in the 9 excluded herds, then the GEBV were correlated with the actual phenotypes (adjusted for fixed effects) of the heifers within each herd. This correlation was divided by the square root of the heritability of the trait to get the accuracy of genomic prediction. Accuracy was calculated either dropping out all of the data from the 9 herds, or just the last year drop. The latter approach was taken to assess the improvement in accuracy when a herd has some data in the reference set.

Ultimately for multi-breed evaluations, head to head comparisons of breeds in the same herd/environment are necessary. We assessed how many head to head comparisons as $\sum_i^n X_i' X_i$ where for each of *n* cohorts in the data set, X_i is a matrix of breed proportions in cohort *i*, of dimensions number of heifers in the cohort x number of breeds (12).

RESULTS AND DISCUSSION

The heritability of the traits estimated from the genomic data was moderate (for CLscore, CLrate and body condition score), and higher for weight and hip height (Table 1).

Table 1. Trait genomic heritabilities and standard errors

Trait	Heritability	Standard error
Weight	0.32	0.02
Hip height	0.43	0.02
Body condition score	0.22	0.02
CLscore	0.25	0.01
CLrate	0.22	0.01

Heritabilities were consistent with previous estimates for these traits in tropical beef cattle data sets derived from pedigree (eg Corbet *et al*. 2018). Accuracies of GEBV in the 9 validation herds were moderate (Figure 1). Accuracies of GEBV were higher when the validation herds had a cohort in the dataset.

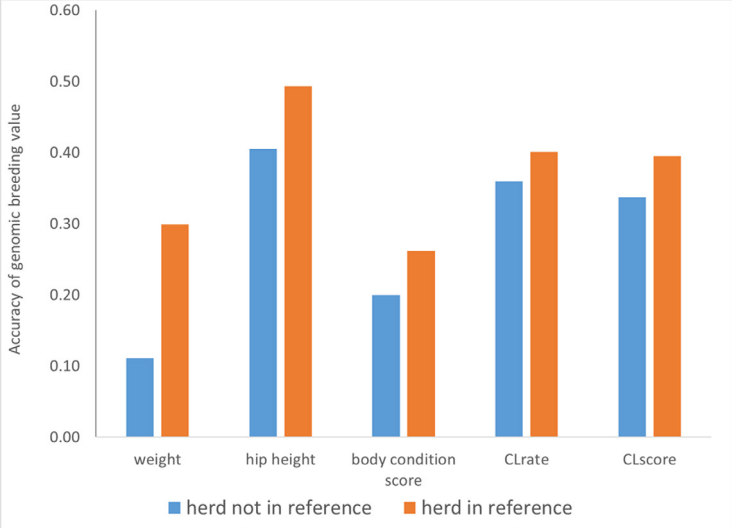

Figure 1. Accuracy of GEBV in nine validation herds. The data for these herds was either completely removed from the reference population ("herd not in reference") or the last cohort of heifer data was removed from the reference and used as the validation set ("herd in reference)

The number of head to head comparisons possible from the data set, which enables estimates of breed effect, reasonable for Angus versus Brahman, Brahman versus Droughtmaster and Brahman versus Santa Getrudis, but was lower for other breed combinations, Table 2. This suggests the data set will contribute to multi-breed genomic evaluations for many, but not all breeds used in Northern Australia.

CONCLUSIONS

The results of this preliminary study, both in terms of genomic heritabilities, and accuracy of GEBV are promising. Heritability's of the traits measured on the 14,552 heifers phenotypes and genotyped to date are consistent with heritability previously reported for tropical beef cattle, based on pedigree and data. Accuracies of GEBV, including for the fertility traits CLscore and CLrate were moderate, but of sufficient magnitude to suggest genetic gains could be made by selecting for GEBV for these

traits. The utility of GEBV are enhanced by the fact that they work to some extent across breeds (the validation set included herds with *Bos indicus*, *Bos indicus* x *Bos taurus* and *Bos taurus* cattle) The heifer data here represents approximately half the data that will be collected in the Northern Genomics Project (which aims to genotype and phenotype 30,000 heifers from the 54 collaborating herds). Additional traits will include heifer rebreed success and follow up pregnancy tests for a number of years. Given the results reported here, the complete data set should enable reasonably accurate GEBV for several fertility traits related to cow lifetime productivity, especially when this data is combined with other data sets, for example in BREEDPLAN.

Table 2. Number of head to head breed comparisons in the data set, where each cell represents the number of genomes for a breed being compared to the number of genomes of the other breed. Empty cells indicate no comparisons for that breed combination

	Angus	Belmont Red	Brahman	Charolais	Droughtmaster	Hereford	Limousin	Santa Gertrudis	Shorthorn	Wagyu
Angus										
Belmont Red	11									
Brahman	315	37								
Charolais	26		165							
Droughtmaster	120	17	603	42						
Hereford	52		94	16	59					
Limousin	25		87	14	31	15				
Santa Gertrudis	116	12	311	34	144	37	31			
Shorthorn	36		79	12	58	16	12	45		
Wagyu	12		40		27			19		
Boran			51		13					

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the participation and efforts of the 54 Northern herds in the Northern Genomics Project, and assistance from various DAF staff at various stages of the project. We are grateful for funding from the MLA Donor Company and University of Queensland through project P.PSH.P0833.

REFERENCES

Burns, B.M., (2016) Queensland Government Smart Futures Research Partnerships Program (2012-2015). University of Queensland.

Corbet N.J., Allen J.M., Laing A.R., Fordyce G., McGowan M.R., and Burns B.M. (2017) *Anim. Prod. Sci*. **58**: 1735.

Dodds K.G., Auvray B., Newman S.A. and McEwan J.C. (2014) *BMC Genetics* **16**: 15.

Fordyce G (2012) Beef CRC exit Report..

Johnston D.J., Barwick S.A., Fordyce G., Holroyd R.G., Williams P.J., Corbet N.J. and Grant T. (2013) *Anim. Prod. Sci*. **54**: 1.

Sargolzaei M., Chesnais J.P. and Schenkel F.S. (2014) *BMC Genomics*. **15**:478.

Taylor, W.J., and T. H. Rudder. (1986). Proceedings of the Australian Society of Animal Production 16:379-382.

Yang J, Lee SH, Goddard ME, Visscher PM (2011) Amer. Jour. Human. Genet. **88**:76-82.