Proc. Assoc. Advmt. Anim. Breed. Genet: 24: 147-150

GENOMIC EVALUATION OF MALE FERTILITY OF AUSTRALIAN HOLSTEIN-FRIESIAN AND JERSEY BULLS

M. Haile-Mariam[1], K. Konstantinov[1], E.G Jewell[1,2], G.J Nieuwhof[1,2] and J.E. Pryce[1,3]

[1]Agriculture Victoria Research, AgriBio, 5 Ring Road, Bundoora, VIC, 3083, Australia
[2]DataGene Limited, AgriBio 5 Ring Road, Bundoora VIC 3083 Australia
[3]School of Applied Systems Biology, La Trobe University, Bundoora, VIC 3083, Australia

SUMMARY

For a long time, the artificial insemination (AI) industry has provided high-quality semen for dairy cattle breeding. With the recent trend to widely use genomically selected bulls before adequate screening of their semen, predicting bull fertility early in life has become an important area of research. In this study we used 25-day non-return rate of about 3 million Australian cows that were inseminated using semen from 5943 Holstein (H) and 1258 Jersey (J) bulls that had high density SNP data (HD), to estimate the proportion of variance explained by SNP data and assess the accuracy of prediction of validation bulls. The proportion of variance explained by SNP data was about 1.2% in Jersey and 0.6% in Holstein bulls. The mean bull solution for both breeds was near zero (-0.05% for H and 0.43% for J). The standard deviation of the bull solutions of was 2.36% in H and 3.30% in J bulls. For both H and J bulls, the difference between the best and worst bulls was about 18% units. Genomic prediction (GP) accuracies were estimated using 5-fold cross validation and varied from 0.20 to 0.25 in H bulls and 0.08 to 0.36 in J bulls. For H bulls the GP accuracy for young bulls were lower (0.13) than average accuracies calculated from 5-fold cross validation. In the case of J bulls, the accuracy for young bulls were the same (0.22) as the average accuracy from 5-fold cross validation. The results show that despite the low heritability, GP of male fertility in Australian H and J breeds is possible and could be used for monitoring and making early decisions to avoid the use of semen from extremely poor fertility bulls.

INTRODUCTION

Genetic improvement programs in dairy cattle have focused on female fertility but ignored male fertility assuming the artificial insemination (AI) industry is able to properly screen and standardize the quality of semen before it is widely distributed. Most studies have not found significant genetic difference in outcomes of insemination among bulls used for mating, possibly because of screening on semen parameters (e.g., Carrick *et al.* 2000; Kuhn and Hutchison 2008). As a result, bull fertility is a phenotypic evaluation used to rank bulls on AI success. Nevertheless, there is evidence that AI success varies among bulls and information on bull non-return rate (NRR) following insemination could be useful for improving overall herd fertility (e.g., Abdollahi-Arpanahi *et al.* 2017). The economic impact of even a small difference in semen fertility between bulls could be large because a single bull is mated to thousands of cows and the benefit of using bulls with good semen fertility is immediate and has a direct effect on the overall herd fertility.

With the recent shift in the dairy industry towards fast tracking of young genomically selected bulls for intensive use before adequate screening, exploring causes of variation in bull fertility early has become an emerging area of research (Taylor *et al*. 2018). The renewed interest to assess the extent of genetic variation in male fertility is partly due to the opportunity to carry out genomic-enabled screening of bulls before they are extensively used for semen collection (Abdollahi-Arpanahi *et al.* 2017; Rezende *et al*. 2019). The main aim of this study was to examine if the use of genomic evaluations can provide an opportunity for early culling of bulls based on 25-day non-return rate (success or failure of insemination outcomes) of their mates. For this study we used genotype and phenotype data of 5934 Holstein (H) and 1258 Jersey (J) bulls that mated to

about 3 million cows. Accuracy of genomic predictions (GP) for both breeds were tested using a 5-fold cross validation and by predicting direct genomic values (DGVs) for younger bulls.

MATERIALS AND METHODS

Phenotype data. Detailed description of the phenotype data used for this study is given by Carrick *et al.* (2000) and Haile-Mariam and Pryce (2021). Briefly the outcome of each insemination of AI bulls, called non-return rate (NRR), is derived by coding each insemination as successful (1) or failed (0) based on a minimum of interval of at least 25-days after insemination. In the first instance, any insemination performed at least 25-days before the end of the AI period was coded as successful and was changed to failed if it is followed by another insemination or mating at least 10 days after the previous insemination. Currently these data are used for calculating semen fertility values (SFV) of bulls by DataGene (https://datagene.com.au). In total there were 10941 bulls with 3.8 million inseminations between 1995 and 2020 in 3289 herds in Australia. AIs involving H and J bulls that mated to all breeds of cows (predominantly H and J, respectively) were selected for this study. The number of H and J bulls with phenotype and genotype data are given in Table 1.

Table 1. The structure of Holstein and Jersey data used for genomic analyses

Reference set	Holstein bulls	Jersey bulls
No. of records	2114529	300560
No. of bulls with data	4654	1057
Year of birth of bulls	1990-2014	1990-2012
Mean NRR (%)	51.77(49.97)[A]	55.95(49.65)[A]
No. of inseminations per bull	449(10-43221)[B]	285(10-14147)[B]
Validation set		
No. of records	234401	61493
No. of bulls with data	799	201
Year of birth of bulls	2015-2019	2013-2019
Mean NRR (%)	49.69(49.99)[A]	52.96(49.91)[A]
No. of inseminations per bull	293(10-3313)[B]	309(10-4497)[B]

[A]Standard deviation; [B]Range in number of inseminations per bull.

Genotype data. Most bulls were genotyped using 50K SNP chips from various commercial providers, while about a quarter had HD genotypes. The first stage of the imputation was to a standard 50K SNP chip for all bulls followed by imputation to HD. Imputation of all 50K genotypes to HD was implemented using Fimpute v3 (Sargolzaei *et al.* 2014) with a reference set (RS) of 2700 HD genotypes. All 50K variants that passed quality control but did not overlap the HD set were then added back into the final imputed set which included the combined HD and 50K SNP sets. The 720,521 SNP set used for this study are located on all 30 chromosomes including the pseudo-autosomal region of the X Chromosome (Nguyen et al. 2021). The SNP data were used to create genomic relationship matrix (GRM) following Yang *et al.* (2011) separately for H and J bulls applying a minor allele frequency of 0.01 and 0.05 for H and J, respectively. To test if a joint RS of H and J bulls is beneficial, a third GRM using genotyped data of both breeds was also constructed.

Statistical analyses. This study used NRR coded as 100 (for successful) and 0 (for failed) as the response variable to evaluate male fertility compared to studies in the literature (Abdollahi-Arpanahi *et al.* 2017; Rezende *et al.* 2020) that used summarized bull solutions (e.g., sire conception rate or SFVs). The use of the raw NRR data jointly with important fixed and random effects and the GRM of bulls is expected to capture more of the variance and increase the accuracy

GP of bulls. Data analyses were carried out assuming a linear animal model using ASReml (Gilmour *et al*. 2015). Details of the fixed and random effects that were fitted are described by Haile-Mariam and Pryce (2021). Briefly a contemporary group effect that included herd-year-AI technician, mating number, cow breed, month of insemination, data processing centre, age of cow and bull at insemination, days in milk at insemination and days from insemination to the end of the AI period were fitted. The random effects fitted were the permanent environmental effect for the cow and the GRM for the bulls with insemination data. First, we used the genotype and phenotype data of all H and J bulls to quantify the proportion of variance captured by GRM. Then accuracies of GP were tested into 2 ways: Firstly, in a 5-fold cross validation scheme where the data were split into 5 parts of approximately equal size, by allocating the offspring of each sire to one of the 5 datasets. In this approach no bull in the validation set had paternal half sibs in the RS. This analysis was performed 5 times using each dataset in turn as a validation and the other 4 sets as the reference. Secondly, validation using young bulls (forward prediction) where bulls born after 2014 were used as a validation set and those born between 1990 and 2014 were used as RS in H. For J, bulls born after 2012 were included in validation set because the number born after 2014 were fewer (see Table 1). In both cases validation bulls were included in the GRM but had missing phenotypes when calculating their DGVs. Accuracy of prediction is calculated as the correlation between corrected phenotype (for effects considered in the model described above) and DGVs for bulls with at least 100 inseminations.

RESULTS AND DISCUSSION

The mean NRR for both H and J bulls used in the reference and validation set are shown in Table 1. The mean NRR are lowest in H validation bulls and highest in J reference bulls. The proportion of variance explained by the GRM was lower in H (0.6%) than in J bulls (1.2%). In both cases the permanent environmental effect of the cow accounted for 3% of the total variance and more than 95% of the variation was not accounted for by the model. Despite this, the bull solutions for both breeds show considerable variation. The mean bull solutions for all Holstein bulls were close to zero (-0.05%) with a standard deviation (SD) of 2.36%. In the case of Jersey bulls, the mean was 0.43% with SD of 3.30%. The bull solutions for both breeds show an approximate normal distribution (-9.0 to +9.0%) with few extremely poor fertility bulls. There were 9 H and 7 J bulls with solutions of below -9.0%.

The accuracy of GP from the 5-fold cross validation are similar in both breeds despite the larger reference size of the H breed. The accuracy values for H bulls are lower than those reported by Abdollahi-Arpanahi *et al.* (2017) who used 7447 bulls with sire conception rate in the USA. Part of the reason for the difference could be the response variable used and the way the data were analysed in both studies. The difference in the RS between the two studies may also have contributed to the lower GP accuracy of the current study. For J bulls our estimates are slightly lower than those for J bulls from the USA (0.28-0.29) which was based on about 1500 bulls (Rezende *et al.* 2019). Interestingly for Australian J bulls, a bivariate model that used sire conception rate from the USA and SFV from Australia resulted in accuracy of 0.24 (Rezende *et al.* 2020), which is similar to our result in Table 3. The analyses by Rezende et al. (2020) used about half of J bulls used in the current study and about 1500 bulls from the USA.

To the best of our knowledge the accuracy of GP for young bulls for male fertility is not available in the literature. GP accuracy for young H bulls is lower than that the average from 5-fold cross validation (Table 3). This could be because the young bulls in H are less related to the RS set due to the fast turn-over of bulls in the post genomic era. Furthermore, the lower proportion of genetic variance explained by GRM and the higher genetic diversity of all H bulls relative to J bulls may have contributed to lower accuracy of prediction for the young bulls. Possibly also changes to the level of screening on semen parameters after the introduction of genomic selection

may have contributed to low accuracy (Taylor *et al.* 2018). The use of joint H and J RS gave similar accuracy for young bulls (Table 3) suggesting a potential to have a single step genomic evaluation by including both genotyped and ungenotyped bulls of both breeds. This is appealing for the Australian dairy industry as the current evaluation for SFV uses data of all breeds.

Table 2. Variance component estimates for semen fertility value and proportion of variance explained by the different random effects in Holstein and Jersey bulls

Random effects	Holstein bulls		Jersey bulls	
	Variance	Proportion of total	Variance	Proportion of total
GRM	13.60±0.68	0.006±0.000	27.74±2.24	0.012±0.001
PE of cows	70.08±1.49	0.031±0.001	73.05±4.10	0.033±0.002
Residual	2190.35±2.39	0.963±0.001	2139.93±6.37	0.955±0.002

Table 3. Accuracy of genomic prediction for validation bulls for semen fertility value in Holstein and Jersey bulls with at least 100 inseminations

Breed	Five-fold cross validation		Validation in young bulls		
	No.	Accuracy	No.	Breed specific reference	Joint reference
Holstein	717-898	0.197-0.252(0.220)	482	0.128	0.123
Jersey	100-176	0.078-0.357(0.221)	126	0.219	0.239

CONCLUSIONS

The results of this study show that prediction of DGVs for H and J bulls using raw insemination data is feasible. At this stage the accuracies of GP particularly for young bulls are low. Nevertheless, there is a potential to use these results for monitoring and making early decisions to avoid using semen from extremely poor fertility bulls.

ACKNOWLEDGEMENTS

We thank DataGene Ltd. for access to the data. Drs I. M. Macleod, Bolormaa S. and T. Nguyen performed the quality control and imputation of the genotype data. The research was funded by DairyBio (Dairy Australia, Agriculture Research Victoria and Gardiner Foundation) (Melbourne, Australia).

REFERENCES

Abdollahi-Arpanahi R., Morota G. and Peñagaricano F. (2017) *J. Dairy Sci*. **100**: 9656.

Carrick M., Goddard M.E. and Bowman P.J. (2000) Pilot system for routine collation of non-return data for bulls. The University of Queensland, Australia pp. 9-24.

Haile-Mariam M. and Pryce J.E. (2021) *J. Dairy Sci*. In press.

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2015) ASReml User Guide Release 4.0. VSN International Ltd., Hemel Hempstead, UK.

Kuhn M.T. and J.L. Hutchison. (2008) *J. Dairy Sci*. **91**: 2481.

Nguyen T.V., Bolormaa S., Reich C. M., Chamberlain A. J., Medley A., Schrooten C., Daetwyler H. D. and MacLeod I. M. (2021) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **24**:(in press).

Rezende F.M., Nani J. P. and Peñagaricano F. (2019) *J. Dairy Sci*. **102**: 3230.

Rezende F.M., Haile-Mariam M., Pryce J.E. and Peñagaricano F. (2020) *J. Dairy Sci*. **103**: 11618.

Sargolzaei M, Chesnais J.P. and Schenkel F.S. (2014) *BMC genomics* **15(1)**: 478.

Taylor J.F., Schnabel R.D. and Sutovsky P. (2018) *Animal* **12(S1):** S172.

Yang J., Lee S.H., Goddard M.E. and Visscher P.M. (2011) *Am. J. Hum. Genet*. **88**: 76.