

## PHENOBANK: A PLATFORM TO FACILITATE COLLABORATION AND GENOMIC SELECTION FOR FEMALE FERTILITY IN BEEF CATTLE

K.B. Wockner<sup>1</sup>, M.S. Tahir<sup>1</sup>, J. Kour<sup>1</sup>, W.L.A. Tan<sup>1</sup>, B.S. Olasege<sup>1</sup>, L.R. Porto-Neto<sup>2</sup>, B.A. Little<sup>2</sup>, C.R. Wilson<sup>5</sup>, W. Dollemore<sup>4</sup>, A. Reverter<sup>2</sup>, B. Hayes<sup>3</sup> and M.R.S. Fortes<sup>1</sup>

<sup>1</sup>The School of Chemistry and Molecular Biosciences, the University of Queensland, Brisbane, QLD 4072, Australia <sup>2</sup>CSIRO Agriculture and Food, Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia, Brisbane, QLD, 4067, Australia <sup>3</sup>Queensland Alliance for Agriculture and Food Innovation, Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia, Brisbane, QLD, 4067, Australia <sup>4</sup>Northern Territory Department of Industry, Tourism and Trade, Katherine Research Station, 449 Stuart Highway, Katherine, NT, 0850, Australia <sup>5</sup>Agricultural Business Research Institute, University of New England, Armidale, NSW, 2351, Australia

### SUMMARY

The PhenoBank project is about creating a platform to facilitate research collaborations and genomic selection in beef cattle. The traits under investigation are female reproduction traits, measured early in life and preferably easy to measure to facilitate adoption. We now have data on more than 9000 cows from many tropical breeds, which have a record for PREG1: a binary trait for the ability to conceive, or not, after the first mating season. These cows were genotyped using SNP chip technology. Using data from 2400 Brahman cows from the PhenoBank we estimated the heritability of PREG1 and conducted a Genome Wide Association Study (GWAS). The estimated heritability was 0.17 (SE 0.03). We identified 59 suggestive SNPs (P-value  $< 9.9 \times 10^{-5}$ ) that mapped to different chromosomes. The SNP with the highest significance (P-value  $2.0 \times 10^{-7}$ ) was on BTA8. SNPs clustered on BTA21 were also significant (P-value  $1.1 \times 10^{-6}$ ). This easy to measure trait can be used for genomic selection. The associated markers need to be confirmed in further studies.

### INTRODUCTION

Female reproduction performance is a major driver of on-farm profitability and currently has immense potential for genetic improvement as many North Australian cattle enterprises have low reproductive rates. Female reproduction traits are often low in heritability and/or expensive to measure, which are hurdles for adoption (Cammack *et al.* 2009). Still, traits such as age at puberty and post-partum anoestrous interval are moderately heritable in tropical beef cattle (Johnston *et al.* 2013). This heritability makes genetic selection for female fertility traits, especially early-in-life traits, a promising strategy for improving beef production in Northern Australian herds. Genomic selection accuracy can be improved by using large phenotype datasets and identifying informative genetic markers. Once established, genomic enhanced Estimated Breeding Values (gEBVs) for easy to measure traits are more likely to be used on extensive cattle enterprises because they will reduce the need for labour intensive data collection (e.g. collecting exact birth dates).

Existing datasets with already-recorded female reproduction traits are the 'fast track' for creating the large reference population needed for robust gEBVs with the potential to transform the rate of genetic gain and the adoption of improved genetics in the beef sector, as it has done in the dairy industry (Meuwissen *et al.* 2013). By linking existing datasets, along with new data, we have created a data warehouse and platform to facilitate genomic selection: the PhenoBank. By using PhenoBank, genetic improvement of fertility traits in *Bos indicus* influenced herds can be further investigated.

### MATERIALS AND METHODS

**PhenoBank Platform.** The PhenoBank database (DB) works with the cloud-based Livestock Information Platform (LIP) developed by CSIRO and Agricultural Business Research Institute

(ABRI). LIP was designed to store and retrieve livestock phenotypes and genotypes. PhenoBank DB has been developed as a pair of customised applications with both Windows and Web user interfaces.

**PhenoBank Data.** The existing phenotype and genotype data contributed to PhenoBank was sourced from the Cooperative Research Centre for Beef Genetics Technologies (Beef CRC), the Northern Territory Department of Industry, Tourism and Trade (NT DITT) breeding herd and the Kamilaroi herd investigated in a CSIRO-led project. Genotypes of the Beef CRC cows were from the Bovine50K v.1 chip (Hawken *et al.* 2012). The NT DITT cows were genotyped in the project with the GGP Bovine50K SNP chip (NEOGEN Inc.), while Kamilaroi cows were genotyped with the GGP TropBeef 35K SNP chip (NEOGEN Inc.). Additionally, we have received phenotypic and genotypic data from 10 industry herds. To contribute to PhenoBank, producers provided the mating outcomes from the first two breeding seasons in a cow's life and its DNA sample for genotyping. New cows for PhenoBank are being genotyped with the GGP TropBeef 35K SNP chip (NEOGEN Inc.).

**Analysis.** We selected, curated and combined data of 2400 Brahman cows for which we defined new phenotypes. These records were sourced from the Beef CRC, NT DITT and Kamilaroi data contributions, and all have an early-in-life, easy-to-measure record for PREG1: a binary trait for the ability to conceive, or not, after the first mating season, outlined in Table 1.

**Table 1. Scoring criteria of early reproductive traits in Brahman heifers**

No.	Trait	Score	Scoring Criteria
1	PREG1	1	Not pregnant as a result of the first mating opportunity (n = 600)
		2	Pregnant as a result of the first mating opportunity (n = 1719)

A reference panel of 546 Brahman animals were genotyped with the BovineHD (770K) SNP chip (NEOGEN) and used to impute genotypes from the medium-density SNP panels. A combination of Eagle v2.4.1 (Loh *et al.* 2016) and Minimac3 (Das *et al.* 2018) were used for imputation. The combined genotypes dataset was passed through final quality control (SNPs with a call rate < 0.9 and multiple allele frequency < 0.05 were discarded) to get over 500,000 SNPs for all cows included in this study.

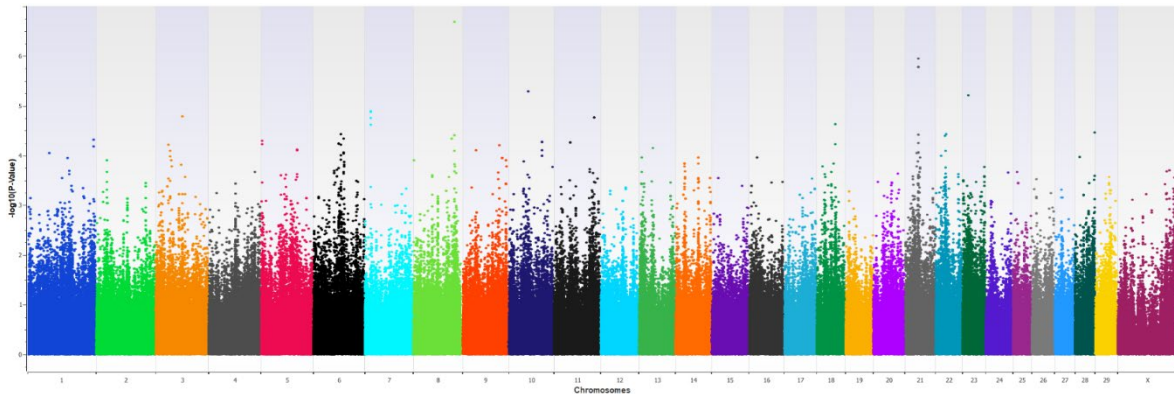
Each phenotype dataset had particular fixed effects to account for contemporary group effects. Contemporary groups were defined by farm location (animals raised together in the same farm) and by birth year and month, which inform the cow cohort (year) and the birth month class (Aug to Nov = Class A; Dec to April = Class B). For the Beef CRC dataset, farm, cow cohort and birth month class were used as fixed effects. For the NT DITT dataset, cow cohort and birth month class were used as fixed effects. For the Kamilaroi dataset, cow cohort was used as a fixed effect. After adjusting for fixed effects, the three datasets were combined to make a single dataset for pooled analyses. Genome wide association studies (GWAS) were conducted for the combined PREG1 dataset using *SNP & Variation Suite v8.8* (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com).

## RESULTS AND DISCUSSION

We have created a data platform to facilitate collaborations and genomic selection in beef cattle research. PhenoBank now contains data on more than 9000 cows. PhenoBank presents an opportunity for researchers and producers to engage and collaborate. Datasets can be uploaded, merged, stored securely and shared when desired. The traits under investigation are female reproduction traits, measured early in life and preferably easy to measure to facilitate adoption by beef producers. The accuracy of imputation for CRC, Kamilaroi and NT DITT cows' datasets was

0.95, 0.93 and 0.92, respectively. Linear mixed model analysis of PREG1 resulted in an estimated heritability of 0.17 (SE 0.03). This compares to a heritability of 0.18 for PREG1 after fixed-time artificial insemination in another study on Brahman heifers (Porto-Neto *et al.* 2015).

GWAS for PREG1 identified 59 suggestive SNPs ( $P\text{-value} < 9.9 \times 10^{-5}$ ) that mapped to different chromosomes. The SNP with the highest significance ( $P\text{-value} 2.0 \times 10^{-7}$ ) was on BTA8. SNPs clustered on BTA21 were also significant ( $P\text{-value} 1.1 \times 10^{-6}$ ) as shown in Figure 1.



**Figure 1. Manhattan plot of PREG1: genome-wide association results**

## CONCLUSIONS

By re-defining and merging data from previous projects with new data, we are expanding the number of samples available for research use. Using a sample of available PhenoBank data, we were able to show that PREG1 has a heritable component and we identified some potential genomic markers for this trait. Next, we will analyse the rebreeding ability for these cows, as we have the outcomes of the second mating season for most of them.

Our next step for PhenoBank is the imputation of all uploaded genotypes to sequence level data and continued analyses investigating heifer and cow fertility. We will continue to collaborate with North Australian producers and upload their data contributions to PhenoBank to create a world class digital infrastructure for beef cattle genomics. Our goal is to contribute to the sustainability and profitability of Australia's beef industry.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding received from Meat and Livestock Australia (MLA) for the Female Fertility PhenoBank project (L.GEN.1710) to undertake this work. We also acknowledge all our collaborators for their data contributions to PhenoBank.

## REFERENCES

- Cammack K. M., Thomas M. G., and Enns R. M. (2009) *Prof. Anim. Sci.* **25**:517.
- Das S., Forer L., Schönherr S., Sidore C., Locke A.E., Kwong A., Vrieze S.I., Chew E.Y., Levy S., McGue M., Schlessinger D., Stambolian D., Loh P-R., Iacono W.G., Swaroop A., Scott L.J., Cucca F., Kronenberg F., Boehnke M., Abecasis G.R., and Fuchsberger C. (2016) *Nat Genet* **2016 48**: 1284.
- Hawken R.J., Zhang Y.D., Fortes M.R.S., Collis E., Barris W.C., Corbet N.J., Williams P.J., Fordyce G., Holroyd R.G., and Walkley J.R.W. (2012) *J Anim Sci* **2012 90**:1398.

- 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.
- Loh P-R., Danecek P., Palamara P.F., Fuchsberger C., Reshef Y.A., Finucane H.K., Schoenherr S., Forer L., McCarthy S., Abecasis G.R., Durbin R., and Price A.L. (2016) *Nature* **48**: 1443.
- Meuwissen, T., Hayes B., and Goddard, M. (2013). *Annu Rev Anim Biosci* **1**: 221.
- NEOGEN Corporation. (2021) GeneSeek® Genomic Profiler™. Accessed July 30, 2021. <https://www.neogen.com/categories/genotyping-arrays>.
- Porto-Neto L.R., Edwards S., Fortes M.R.S., Lehnert S.A., Reverter A., McGowan M. (2015). *J Anim Sci* **93**: 5119.
- SNP & Variation Suite™ (Version 8.8) [Software]. Bozeman, MT: Golden Helix, Inc. Available from <http://www.goldenhelix.com>.