NON-SYNONIMOUS POLYMORPHISM IN *HELB* IS ASSOCIATED WITH MALE AND FEMALE REPRODUCTIVE TRAITS IN CATTLE

M.R.S. Fortes[1], F.B. Almughlliq[1], L.T. Nguyen[1], L.R. Porto Neto[2], S.A. Lehnert[2]

[1]University of Queensland, School of Chemistry and Molecular Biosciences, Qld 4072, Australia.
[2]CSIRO Agriculture Flagship, Queensland Bioscience Precinct, Qld 4067, Australia.

SUMMARY

We report on the association with reproductive phenotypes of a genetic marker mapped to the *HELB* gene, in tropically adapted beef cattle. The genetic marker is a single nucleotide polymorphism (SNP) in chromosome 5 that is non-synonymous and has a predicted deleterious effect on the coded protein. Reproductive phenotypes used in these analyses were from a population of the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC). Bulls (n = 1,023) were from a mixed-breed population and represented animals not included in the original genome-wide studies (GWAS) that reported SNP associations on chromosome 5 for scrotal circumference (SC). Cows (n = 1,089) were Tropical Composites and most were from the population used in the original GWAS that described SNP associations on chromosome 5 for puberty and post-partum anoestrus interval (PPAI). Animals were genotyped with Taqman assays designed for 3 non-synonymous SNP. The results indicate that the SNP in *HELB* is significant for SC and PPAI at $P < 10^{-7}$ and for heifer puberty ($P = 0.0029$). This SNP might aid across breed phenotype prediction for reproductive phenotypes given its significance in these mixed and composite cattle with both *Bos taurus* and *Bos indicus* ancestry. Future studies should target this *HELB* SNP in other breeds and populations to confirm the associations we described.

INTRODUCTION

Discovery of functional mutations related to QTL has merit in the context of animal breeding. Functional mutations, such as non-synonymous single nucleotide polymorphisms (SNP) or frameshift SNP have potential biological meaning: they might be the causative mutation or be in strong linkage disequilibrium (LD) with the causative mutation. For these reasons, functional mutations might overcome an important issue faced by genomic selection in the beef cattle industry: prediction of crossbred performance (Jonas and de Koning, 2015). Crossbred performance is especially important in tropical regions where beef cattle herds carry a mixture of *Bos indicus* and *Bos taurus* ancestry. Functional mutations associated with reproductive traits would especially benefit *Bos indicus* influenced cattle. *Bos indicus* breeds tend to have a prolonged post-partum anoestrus period, later puberty and overall lower reproductive performance when compared to *Bos taurus* breeds (Abeygunawardena and Dematawewa, 2004). Despite lower reproductive performance, *Bos indicus* breeds and crosses are largely used in tropical regions because of their adaptive advantage in those environments (Burrow, 2012). Improvement of reproductive performance of *Bos indicus* breeds and crosses would benefit beef production.

Herein, we investigated the association between 3 non-synonymous SNP in candidate genes mapped to *Bos taurus* autosome 5 (BTA 5) and reproductive phenotypes, in tropically adapted cattle. Previous studies have reported regions of BTA 5 to be associated with male and female reproductive phenotypes in various cattle breeds (Kappes et al., 2000; Kim et al., 2009; McClure et al., 2010; Fortes et al., 2012; Hawken et al., 2012; Fortes et al., 2013a). In Tropical Composite cattle, the QTL on BTA 5 was significant for scrotal circumference (SC) and post-partum anoestrus interval (PPAI). The SNP investigated in the present study are predicted deleterious mutations in *FAU*, *INHBC* and *HELB*, candidate genes mapped to this region of BTA 5. Our objective was to search for SNP that could explain the previously identified QTL.

MATERIALS AND METHODS

Animals. Blood for DNA extraction was obtained from 1021 bulls and 1089 cows. Bulls were 113 Brahman, 741 Tropical Composite and 167 crossbred, which were from the Beef CRC populations previously described (Burns et al., 2013). Importantly, the bulls used in this study had not been genotyped for previous GWAS (Fortes et al., 2012; Fortes et al., 2013a). Cows were Tropical Composites used in previous GWAS (Hawken et al., 2012). Crossbred bulls are the product of mating Brahman to Tropical Composites. Animals studied varied in *Bos indicus* and *Bos taurus* ancestry.

Mutations and genotype assays. Non-synonymous SNP were identified within three candidate genes - *FAU*, *INHBC* and *HELB* - by mining whole genome sequences of 64 bulls of a variety of breeds including Brahman and Senepol. Potentially deleterious mutations were identified using the bioinformatics software known as Variant Effect Predictor (VEP) from Ensembl (http://www.ensembl.org/info/docs/tools/vep/index.html). Custom TaqMan assays were developed for the selected non-synonymous SNP according to TaqMan Array Design Tool (Applied Biosystems, 2010). Further, bulls were also genotyped with the 90K Illumina SNP chip.

Analysis. Single SNP regression was applied for genotyped animals using a mixed model analysis of variance with the SNP & Variation Suite software (Release 8.3.0, Golden Helix, Inc.). The mixed model can be described with the equation: $y_i = X\beta + Z\mu + S_j a_j + e_i$; where yi represents the phenotypic measurement for the i^{th} animal, X is the incidence matrix relating fixed effects (contemporary group and breed) in β with observations in y, Z is the incidence matrix relating to random additive polygenic effects of animal in μ with observations in y and Sj is the observed animal genotype for the j^{th} SNP (coded as 0, 1 or 2 to represent the number of copies of the B allele), aj is the estimated SNP effect, and ei is the random residual effect. Age was fitted as a covariate for SC and PPAI. Same model for candidate SNP and GWAS.

RESULTS AND DISCUSSION

Previous QTL mapping and GWAS reported the importance of BTA 5 for cattle reproduction. A QTL for ovulation rate was mapped to 40 cM in the mixed-breed population known as MARC herd (Kappes et al., 2000). In Angus, a QTL for SC was reported at 13 Mb, at 104 Mb and at 127 cM (McClure et al., 2010). Associations with twinning rate for SNP between 55 and 75 Mb were reported in Holstein cattle (Kim et al., 2009). Hawken et.al (2012) carried GWAS in Brahman and Tropical Composites cows and found SNP associated with female reproductive traits in Tropical Composites located on BTA 5. The largest concentration of SNP associated with PPAI was on 44.0 to 44.3 Mb, 58.2 Mb and 113.6 Mb. Age of puberty was also reported to be associated with BTA 5 in 2 positions: 28.7 Mb and 96 Mb (Hawken et al., 2012). In Tropical Composite bulls, significant SNP association for levels of inhibin hormone were located between 42 and 61 Mb (Fortes et al., 2013a). Subsequently, significant SNP in BTA 5 were found for levels of inhibin and insulin-like growth hormone (IGF1) in Brahman and Tropical Composite bulls (Fortes et al., 2013b). Taken together the evidence points to one or more QTL on BTA 5 that could be important for reproductive physiology in many cattle breeds, with various *Bos indicus* and *Bos taurus* ancestry. In the current study, we tested 3 SNP in candidate genes located between 47 and 56 Mb of BTA 5, because these locations were important for tropically adapted cattle in the literature.

We have identified a SNP in *HELB* associated with SC and PPAI (Table 1, $P < 10^{-7}$). An association with heifer puberty, defined by the age at the first corpus luteum, was also noted ($P = 0.0029$). Breeding cows with the favourable allele could reduce PPAI in 30 days and puberty in 18 days. This SNP explained more than 2% of the genetic variance for SC and PPAI. The SNP in

FAU and *INHBC* failed to associate with studied phenotypes. Out of the 3 tested SNP, the one in *HELB* is the only mutation likely to be in high LD with the causative mutation in BTA 5. We considered the *HELB* SNP a potential functional mutation because it is non-synonymous and has predicted deleterious effect. It is a G to A substitution that causes a T to M residual change in the protein, with a SIFT score 41, deemed deleterious (0.04). The physiological effect of this protein change in SC, PPAI and AGECL is unknown and merits investigation. The causative mutation(s) in BTA 5 could be affecting the endocrine pathways that control both male and female reproduction; so that it is associated with SC, AGECL and PPAI. Previously, we observed an association between this QTL in BTA 5 and levels of IGF1 and inhibin (Fortes et al., 2013a). Links between this QTL and reproductive hormones should be explored.

Table 1. Estimated significance and effect of non-synonymous polymorphisms mapped to candidate genes on chromosome 5.*

	Phenotype	***Gene***	***P-Value***	***Effect***	***% Variance***
Bulls	SC18	*FAU*	$2.82x10^{-1}$	0.3507	0.12
		HELB	**$5.23x10^{-7}$**	**0.7018**	**2.50**
		INHBC	$7.86\ x10^{-1}$	-0.0368	0.01
	SC24	*FAU*	$5.77\ x10^{-1}$	0.1724	0.03
		HELB	**$1.94x10^{-11}$**	**0.8949**	**4.40**
		INHBC	$4.22x10^{-1}$	-0.1044	0.06
Cows	AGECL	*FAU*	$7.67x10^{-1}$	3.9005	0.01
		HELB	**$2.90x10^{-3}$**	**18.3899**	**0.95**
		INHBC	$8.47x10^{-1}$	1.1082	0.00
	PPAI	*FAU*	$7.68x10^{-1}$	3.4762	0.01
		HELB	**$2.15x10^{-8}$**	**30.4043**	**3.87**
		INHBC	$8.53x10^{-1}$	-0.9703	0.00

*Genes and values highlighted in bold represent significant associations for reproductive traits: scrotal circumference at 18 and 24 months of age (SC18 and SC24, cm), age at first corpus luteum (AGECL, days) and post-partum anoestrus interval (PPAI, days). The SNP ***Effect*** is provided in the same measuring unit at the trait and ***% Variance*** is the percentage of the additive genetic variance explained by each SNP.

The relevance of the SNP in *HELB* to previously described QTL was tested by fitting this SNP as a fixed effect in the GWAS (Figure 1). When comparing the *P-values* between two models, we observe that the *P-values* are reduced from 10^{-8} to 10^{-6} for SC18, and from 10^{-13} to 10^{-10} for SC24. However, the strong association signal from common chip SNP was still present after fitting the *HELB* SNP. This result is different from expectations for causative mutations. When the causative mutation in the *PLAG1* gene was fitted the QTL on BTA 14 practically disappeared from GWAS (Fortes et al., 2013c). This result suggests that more in-depth molecular characterisation is required. It is unlikely that the SNP in *HELB* is the only causative mutation associated to reproductive traits in this genomic region.

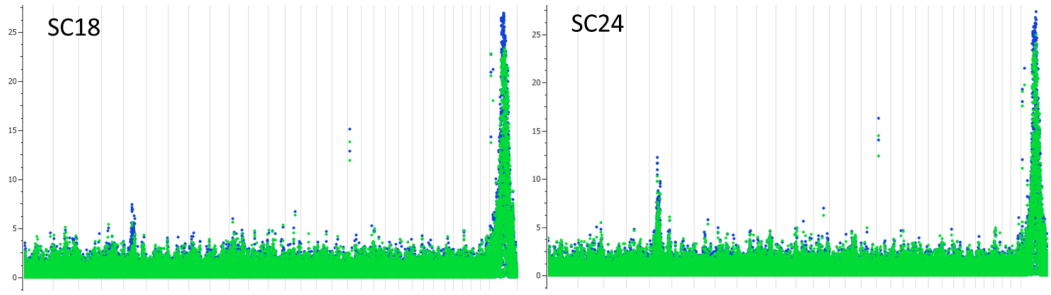

Figure 1. Significance of the associations with SC18 and SC24 for SNP across the genome. *P-values* are on y-axis and genomic positions are on x-axis. Note the reduction in *P*-values from the blue dots (animal model) to the green dots (animal model with *HELB* SNP fitted as a fixed effect).

CONCLUSION

The identification of a SNP associated with SC in a mixed population of bulls and with PPAI and AGECL in Tropical Composite cows might aid across breed phenotype predictions. Future studies should target this *HELB* SNP in other breeds and populations to confirm associations. For female traits, this SNP needs to be validated independently from the original GWAS.

ACKNOWLEDGEMENTS

The authors acknowledge all collaborators of the Beef CRC. Meat and Livestock Australia provided support (project B.NBP.0786). Fortes is supported by UQ Postdoctoral Fellowship.

REFERENCES

Abeygunawardena H. and Dematawewa C.M. (2004) *Anim. Reprod. Sci.* **83**: 373-387.

Burns B.M., Corbet N.J., Corbet D.H., Crisp J.M., Venus B.K., Johnston D.J., Li Y., McGowan M.R. and Holroyd R.G. (2013) *Anim. Prod. Sc.* **53**: 87-100.

Burrow H.M. (2012) *Anim.* **6**: 729-740.

Fortes M.R.S., Reverter A., Hawken R.J., Bolormaa S. and Lehnert S.A. (2012) *Bio. Reprod.* **87**: 1-8.

Fortes M.R.S., Reverter A., Kelly M., McCulloch R. and Lehnert S.A. (2013a) *Androl.* 1: 644-650.

Fortes M.R.S. Reverter A., Porto Neto L.R., Kelly M., Moore S.S. and Lehnert S.A. (2013b) *Proc. Ass. Adv. Anim. Breed. Genet.* **20**: 389-392.

Fortes M.R.S., Kemper K., Sasazaki S., Reverter A., Pryce J.E., Barendse W., Bunch R., McCulloch R., Harrison B., Bolormaa S., Zhang Y.D., Hawken R.J., Goddard M.E. and Lehnert S.A. (2013c) *Anim. Genet.* **44**: 636-647.

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., Walkley J.R.W., Barendse W., Johnston D.J., Prayaga K.C., Tier B., Reverter A. and Lehnert S.A. (2012) *J. Anim. Sci.* **90**: 1398-1410.

Jonas E., and D.J. de Koning. (2015) *Front. Genet.* **6**: 49.

Kappes S.M., Bennett G.L., Keele J.W., Echternkamp S.E., Gregory K.E. and Thallman R.M. (2000) *J. Anim. Sci.* **78**: 3053-3059.

Kim E.S., Berger P.J. and Kirkpatrick B.W. (2009) *J. Anim. Sci.* **87**: 835-843.

McClure M.C., Morsci N.S., Schnabel R.D., Kim J.W., Yao P., Rolf M.M., McKay S.D., Gregg SJ, Chapple R.H., Northcutt S.L. and Taylor J.F. (2010) *Anim. Genet.* **41**: 597-607.