final report

Project code: B.NBP.0723

Prepared by: Sigrid Lehnert and Toni Reverter, CSIRO Animal, Food and Health Sciences

Marina Fortes, Queensland Alliance for Agriculture and Food Innovation, UQ

Date published: March 2013

ISBN: 9781741919875

PUBLISHED BY
Meat & Livestock Australia Limited
Locked Bag 991
NORTH SYDNEY NSW 2059

Genome-wide association study of tropical composite bulls for reproduction traits

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication

Abstract

Reproductive traits are important targets for genetic improvements programs, especially in Australia's northern beef industry. Genomic selection, based on marker panels or on genome-wide predictions, is set to become a widely-adopted tool. Building on previous investments in genomic selection for female reproductive traits in Tropical Composite and Brahman, the current study was undertaken to calculate SNP effects in a population of 1,019 Tropical Composite bulls with deep phenotypes for a range of production traits, including reproductive traits. The study was able to confirm a number of genome areas that appear to be important in reproductive and endocrine traits across breeds and, in at least one case, across both sexes. The data generated in this study has provided the foundation for the development of more accurate DNA-based selection, and for estimating its effects in both Brahman and Tropical Composite breeds in both sexes, on a large number of important production traits.

Executive summary

Reproductive traits are important targets for genetic improvements programs, especially in Australia's northern beef industry. Genomic selection for reproduction or other traits, based on genetic marker panels or on genome-wide predictions, is set to become a widely-adopted tool. The accuracies for genomic selection predictions for reproductive traits, currently available in Australia, are not very high. Higher accuracies in selection predictions will be achieved by increasing the size of datasets on which to base predictions, but also by identifying the causative mutations for the traits in question. Further, when applying DNA-based selection methods, their impact on other traits needs to be assessed so that genomic selection for reproductive traits does not result in adverse effects in other production traits.

Building on previous investments in genomic selection for reproductive traits in Tropical Composite and Brahman, the current study was undertaken to calculate SNP effects in a population of Tropical Composite bulls (n=1,019) with deep phenotypes for a range of production traits, including reproductive traits. By comparing this study with results previously obtained on linked populations of Brahman cows and bulls, and Tropical composite cows, we were able to confirm a number of genome areas that appear to be important for reproductive and endocrine traits across breeds, and in at least one case across both sexes.

For example we were able to confirm the findings from our previous study of Brahman bulls, that the X chromosome is harbouring significant associations with scrotal circumference and sperm morphology. The 25 Mb region on chromosome 14 was shown to be associated with IGF1 levels at 6 months, in agreement with our findings in Brahman cows and bulls and Tropical Composite females. These findings illustrate that it is possible to identify DNA-based predictions that will work across different cattle breeds.

The study of Tropical Composite bulls revealed clear differences between breeds in the mapping of SNP associations. For example, the level of inhibin in pre-pubertal bulls, which in Brahman bulls has been mapped to the *INHA* gene on chromosome 2, appears to map to the *INHBE* and *INHBC* genes on chromosome 5 in Tropical Composites. These two examples illustrate why most DNA-based predictions have to be calibrated for each breed, to account for the differences in genomic architecture between them.

The data generated in this study have provided the foundation for the development of more accurate DNA-based selection, and for estimating its effects in both Brahman and Tropical Composite breeds in both sexes, on a large number of important production traits. We recommend further investment in R&D on these datasets to ensure the industry derives maximum benefit from its investment in creating them. One aspect of this R&D will be the integrative analysis of the 4 reproduction populations that have now been genotyped. Another aspect will be detailed molecular biology investigations of genome regions targeting the discovery of causative mutations.

Table of Contents

1 Background

Reproductive traits are moderately heritable traits in beef cattle, which makes genetic selection for components of reproductive rate a promising strategy for beef production in northern Australian herds. The beef industry is currently investing in the development of genomic methods to facilitate more accurate selection decisions.

The Beef CRC, CSIRO and MLA have previously collaborated to conduct the first genome-wide association study (GWAS) of Brahman bull reproductive phenotypes. Brahman bulls (1,115) with measures of reproductive parameters from the Beef CRC were genotyped with 50,000 single nucleotide polymorphism (SNP) markers (Fortes et al., 2012b).

Tropical Composite cattle (male and female) enter puberty earlier and show other important differences in reproductive performance to purebred Brahman cattle. The Beef CRC team has invested the same phenotyping effort (scrotal circumferences, measures of endocrine parameters, and exhaustive semen analysis at three time points) into the Tropical Composite research population as was applied to the previously studied Brahman bulls (Corbet et al., 2009). Genotyping data will therefore allow us to establish genome-wide associations for the same reproductive traits, and to directly compare the marker associations with these traits between the males from high grade Brahman and Tropical Composite breeds.

The Beef CRC has already acquired genome data on a population of Tropical Composite females that were thoroughly characterised for reproductive phenotypes (Hawken et al., 2012). By completing the GWAS on Tropical Composite bulls, direct comparisons can now be carried out between the genomics of male and female reproductive traits in these important beef breeds.

The genotyping activity described in this project provides the data to allow future validation of markers associated with reproductive performance across breeds and sex. This analysis will increase the reliability of gene markers and prediction equations derived from this work and significantly broaden the industry applicability of the data. The data will also make it possible to investigate whether genomic selection for any male reproductive traits may have adverse effects on the reproduction rate of their female relatives.

2 Project objectives

2.1 Objective 1

Selected DNA from 1,000 Tropical Composite bulls extracted and archived in Beef CRC DNA bank

2.2 Objective 2

Acquired 50K SNP genotypes for 1,000 Tropical Composite bulls and deposited in Beef CRC database

2.3 Objective 3

Calculated marker associations for Tropical Composite bull reproductive traits

3 Methodology

3.1 Animals and traits

Cattle used in this study were from the 'Northern Breeding Project' resource population bred by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) in the tropical regions of Northern Australia (Barwick et al., 2009; Burns et al., 2013; Burrow et al., 2003; Corbet et al., 2013; Johnston et al., 2009; Prayaga et al., 2009). The stable Tropical Composite bulls used in this study were from a population of > 2,000 approximately 50% tropically adapted breeds and 50% non-tropically adapted *Bos taurus* breeds. The tropically adapted component was on average half derived from Brahman and the other half from tropically adapted Taurine breeds. The non-tropically adapted *Bos taurus* component of the Tropical Composites consisted of various combinations of Hereford, Shorthorn, Red Angus, Red Poll and Charolais.

Three hormones were measured: inhibin serum concentration (IN, ng/mL), luteinizing hormone plasma concentrations (LH, ng/mL) and serum concentration of insulin-like growth factor I (IGF1, ng/mL). Blood samples for measurement of IN and LH were taken at approximately 4 months of age and IGF1 was measured at approximately 6 month of age. Hormones and sampling time were selected for their association to puberty and testicular development, following literature evidence (see reviews Burns et al., 2011; Phillips, 2005). For IN assays, 10ml blood samples were collected by venipuncture into BD Vacutainer® serum tubes (Becton, Dickinson and Company) and allowed to clot at ambient temperature, then kept refrigerated. Within 24 hours, tubes were centrifuged at 2500rpm for 20 mins, serum was aliquoted into 5ml tubes and stored at -20°C. Samples were sent frozen to Monash University for IN assays, which followed established methodology (Phillips, 2005). The protocol for a simple measurement of gonadotrophin releasing hormone (GnRH)-stimulated LH plasma concentration was based on methods described by (Bagu et al., 2004). Briefly, bulls were divided in groups of 10 animals for better time management of GnRH (FertagylTM, Intervet Australia Pty Limited) administration. Blood samples (10ml) were collected approximately 20 minutes after the administration of GnRH, by venipuncture into lithium heparin BD Vacutainer® tubes (Becton, Dickinson and Company). Within 1 hour of sampling, tubes were centrifuged as above and 5ml of plasma were aliquot and frozen (-20°C) until assayed. Assays for LH plasma concentrations used a double-antibody radioimmunoassay (RIA) procedure, previously established (Chan et al., 2009; Hotzel et al., 1998). Blood samples for IGF1 were collected by venipuncture straight onto bloodspot collection cards supplied by PrimeGRO™, since IGF1 was measured with a commercially available enzyme-linked immunosorbent assay (ELISA), as detailed by (Moore et al., 2005).

Scrotal circumference (SC) was measured at approximately 12 months of age following industry practice (http://breedplan.une.edu.au), with a standard metal tape (Fordyce et al., 2006). Additionally, between weaning and 24 month of age 8 measurements of SC were taken for each bull, at 3 month intervals. Using these repeated measurements for individual regressions, we calculated the age of the bull when it achieved 26 cm of SC (AGE26, expressed in days). Achieving SC of 26 cm was considered a threshold for puberty in Brahman (*Bos indicus*) bulls as reported previously (Fortes et al., 2012a).

At 18 and 24 months of age, bulls with a SC of 20 cm or more were subject to electro ejaculation and semen was collected for evaluation. A sample of semen was taken into a 0.2% glutaraldehyde in phosphate buffered saline solution for preservation for an estimation of morphology. Sperm morphology of 100 spermatozoa was determined by examining a thin cover-slip preparation of semen using phase contrast microscopy (magnification at x1000). Sperm cells were classified individually according to morphology and the percentage of sperm with normal morphology (PNS) was calculated. All sperm morphology assessments were conducted by the same laboratory technician, who is accredited by the Australian Cattle Veterinarians (Fordyce et al., 2006). As gametes are necessary for fertility, the age when sperm cells first appear in the ejaculate was suggested as an indicator of puberty (Unaniam, 1997). Without enough repeated measurements to estimate the individual ages of sperm cells first appearance, we used the presence or absence of sperm cells at 18 month of age as an indicative of puberty. Although measured relatively late in life, the PNS at 24 month of age was included as a potential selection trait for young bulls because of its reported correlation with calf output (Holroyd et al., 2002).

3.2 Genotyping

The Illumina BovineSNP50 Bead Chip (Matukumalli et al., 2009) was used to genotype DNA from 1,056 Tropical Composite bulls. SNP with auto-calling rates less than 85% and SNP with minor allele frequency (MAF) less than 0.01 were excluded from later analyses. After quality control, 1,019 bulls and 48,821 SNP were retained for genome-wide association analysis.

3.3 Calculation of SNP effects

Genome-wide association studies (GWAS) were performed with the 48,821 SNP of the BovineSNP50 chip for each of the 7 traits separately. Genotype calls were coded as 0 for the homozygote of the A allele, 1 for the heterozygote, and 2 for the homozygote of the B allele. Alleles A and B were defined according to top/bottom rules from Illumina. The effect of each SNP was estimated in turn using the following mixed model:

Equation 1

$$y_{i,j} = X\beta + Zu + s_{j,k} + e_{i,j}$$

Where $y_{i,j}$ represents the vector of observations from the *i*-th bull at the *j*-th phenotype, X is the incidence matrix relating fixed effects in β with observation in $y_{i,j}$, Z is the incidence matrix relating random additive polygenic effects in u with observation in $y_{i,j}$, $s_{j,k}$ represents the additive association of the *k*-th SNP on the *j*-th phenotype, and $e_{i,j}$ is the vector of random residual effects. Fixed effects included in the model contemporary group (i.e. cohort of bulls born in the same year and raised together), month of birth, age of the dam, and the batch of the IGF1 assay. Age (in days) at the time of trait measurement was used as a linear covariate for IN, LH, SC, MOT and PNS. Solutions to the effects in the model as well as variance components were estimated using Qxpak5 (Perez-Enciso and Misztal, 2011). Qxpak5 performs a likelihood ratio test (LRT) by testing the model with versus the model without the SNP against a chi-squared distribution with 1 degree of freedom. This LRT was done one SNP at a time.

The false discovery rate (FDR) was calculated using equation 2

Equation 2

$$FDR = \frac{nP}{m}$$

where *n* represents the total number of SNP included in the study (in the present study, n = 48,821); *P* is the *P*-value threshold being used for significance and *m* is the actual number of associated SNP in the given *P*-value threshold.

The percentage of the genetic variance accounted by the *i*-th SNP was computed according to the following formula:

Equation 3

$$\%V_i = 100 \cdot \frac{2 p_i q_i a_i^2}{\sigma_g^2}$$

Where: p_i and q_i are the allele frequencies for the *i*-th SNP, a_i is the estimated additive effect of the *i*-th SNP on the phenotype in question, and σ_g^2 is the estimate of the (poly-)genetic variance for each phenotype in question.

3.4 Manhattan Plots

Manhattan Plots were compiled from GWAS results with an R script ("manhattan.r") sourced from Getting Genetics Done blog site (http://gettinggeneticsdone.blogspot.com.au/). For this task, a re-annotation of the Illumina 50K SNP chip was done in-house against the Bovine Genome UMD3.1 assembly (Childers et al., 2011).

4 Results and discussion

4.1 DNA extractions and genotyping

Out of the total population of Tropical Composite bulls available, 1,097 individuals were selected for this study by prioritising the animals that had complete data for hormone levels, scrotal circumference measures and semen assessments. Archived blood samples were located and DNA was successfully prepared from 1,056 individuals. Genotyping of 1,056 DNA samples using the Illumina 50k bovine SNP chip resulted in 1,019 genotypes that passed quality assurance criteria.

4.2 Trait statistics

Seven traits were chosen for genome-wide association (Table 1). For some traits, such as SC12, the data was very tightly clustered around the mean, whereas traits such as LH4 and MOT18 showed a wide spread of values. For a full description of the trait statistics and genetic parameter estimates, see (Burns et al., 2013)(2013) and (Corbet et al., 2013)(2013).

Table 1: Descriptive statistics for each trait

Trait	Units	Description	N	Mean	SD
IN4	ng/mL	Serum levels of inhibin measured at 4 months	1,097	7.86	1.96
LH4	ng/mL	Plasma levels of luteinising hormone following GnRH challenge at 4 months	1,097	6.92	5.23
IGF6	ng/mL	IGF1 level in peripheral blood measured at 6 months	1,097	608.21	330.41
SC12	cm	Scrotal circumference at one year of age	1,097	26.52	3.19
MOT18	%	Percentage of motile sperm at 18 months of age	1,093	57.22	27.35
PNS24	%	Percentage of normal sperm at 2 years of age	1,089	73.03	20.48
AGE26	days	Age at 26 cm of SC (puberty)	1,078	400.11	76.28

4.3 Marker associations with bull reproductive traits

The genome-wide association study identified more than a thousand SNP for each trait at the significance cut-off $P \leq 0.05$ (**Error! Not a valid bookmark self-reference.**). As expected, when imposing more stringent significance thresholds, a much smaller number resulted. At the $P \leq 0.00001$ significance level no significant associations were observed for the traits LH4 and MOT18. This is probably due to the less reliable measurements available for these traits. The majority of significant SNP that were identified for SC12, PNS24 and AGE26 at $P \leq 0.00001$ were located on the X chromosome (Figure 2, Figure 3, Figure 4).

Table 2: Number of significant SNP at various P-value thresholds and for each trait

P-Value	IN4	LH4	IGF6	SC12	PNS24	AGE26	MOT18
0.05	2123	1472	1853	2367	1654	2300	1348
0.01	550	283	368	742	435	718	203
0.001	111	21	59	292	140	284	23
0.0001	33	4	18	196	76	203	1
0.00001	14	0	10	160	46	159	0

The Manhattan Plots in Figures 1 – 4 show the map positions of significant SNP with respect to bovine chromosomes. The involvement of the X chromosome in the traits SC12, AGE26 and PNS24 was immediately obvious and confirmed our findings from studying Brahman bulls from the same populations (Fortes et al., 2012a; Fortes et al., 2012b). When the map positions of the most significant SNP associations were examined in detail (Tables 6-8), it became apparent that the most significant associations on the X chromosome in the Tropical Composite study mapped to the same regions that were identified in Brahman bulls: 62 – 92 Mb for SC and AGE26, and 40

and 55 Mb for PNS24 (Fortes et al., 2012a; Fortes et al., 2012b). The *P*-values for the X chromosome SNP associations in Tropical Composite bulls were many orders of magnitude higher than the calculations we had previously obtained for Brahman bulls (Fortes et al., 2012a; Fortes et al., 2012b), (Table 6 and Table 9). The possible significance of this observation will need to be investigated further. We propose to test whether the X chromosome in Tropical Composites harbours particularly prominent signatures of selection and to investigate if *linkage disequilibrium* (LD) blocks in Tropical Composites are diverse from that observed in Brahman.

The peak in SNP associations with the IGF6 trait observed on BTA14 (Figure 3) also confirms what was found in Brahman bulls (Fortes et al., 2012b). Examining the BTA14 peak in detail (Table 5) showed that it maps to the 25 Mb region which was identified as important for Brahman age at puberty and IGF1 levels in both sexes (Fortes et al., 2012a; Hawken et al., 2012). The SNP associations between IGF6 and the BTA14 25 Mb region were more significant than those detected for the equivalent W1IGF trait in Tropical Composite cows (Hawken et al., 2012). However, our analysis detected no significant association between age at puberty in Tropical Composite bulls and markers on BTA14 (Figure 4, Table 9).

Similar to what was found in Brahman bulls, the GWAS for LH4 found very few SNP with highly significant associations to the trait. The SNP that did show association with a *P*-value ≤ 0.0001 (Table 4) mapped to different chromosomes from the ones identified in Brahman (Fortes et al., 2012b). Very few significant SNP, and none that reached $P \leq 0.0001$, were identified for MOT18 (Figure 3,Table 8).

An important contrast between the Tropical Composite and Brahman GWAS results for IN4 was the lack of an association peak on BTA2 in Tropical Composites. The Brahman study had identified the inhibin alpha gene (*INHA)* on BTA2 as a strong candidate for this trait, but this does not appear to hold true for Tropical Composites (Fortes et al., 2012b) and, Figure 1). The most significant associations for IN4 in Tropical Composites were found on BTA5 (Table 3). The genes inhibin beta C (*INHBC)* and inhibin beta E (*INHBE*) map to one of the regions with very significant SNP associations on BTA5.

It is possible to speculate that the regulation of inhibin hormone levels in the peripheral blood of prepubertal tropically adapted bulls is strongly influenced by DNA variation in or near the coding region for one of the two peptide chains that make up mature inhibin hormone. In Brahmans, the gene coding for the inhibin alpha chain may directly contribute to the levels of hormone found in blood, and in Tropical Composites, the gene coding for one of the forms of the inhibin beta chain may play that role.

The age at first scanned *corpus luteum* (AGECL) was studied as an age-at-puberty trait in Tropical Composite heifers, and showed strong associations to BTA5 (Hawken et al., 2012). In our analysis of the male age-at-puberty trait AGE26, BTA5 markers did not show any significant associations (Figure 4, Table 9).

This preliminary analysis was able to show that important differences in the genomic architecture of this small number of traits exist between breeds. The different effects of SNP in one breed context versus the other need to be taken into account when predicting the effects of genomic selection.

It is also evident that some traits are linked to common areas of the genome in both breeds. For example, the X-chromosome contains significant associations for SC12, AGE26 and PNS24 in Brahman and Tropical Composite bulls (Fortes et al 2012a, Fortes et al 2012b and Figures 2 and 3).

The 46 Mb region of BTA5 is an example of a QTL region that appears to be linked to reproduction traits in both males and females. This region harbours coding sequences for

interferon gamma (*IFNG*), as well as the immune system cytokines interleukin-22 and -26. In this study, we find strong associations with SNP ARS-BFGL-NGS-43575 and the IN4 trait in males (Table 3, – log*P* > 11). The same SNP also has a highly significant association (– log*P* > 6) with the length of post partum anoestrus (PPAI) in the female Tropical Composite population that was described by Hawken et al (2012). It is interesting to speculate how interferon gamma or interleukins could be involved both in the endocrine functions of the prepubertal testis and the re-initiation of ovarian cycles *post partum*.

Those examples of chromosome areas that may be important across breeds and for reproductive traits in both sexes point to important causative mutations underlying these traits. The data generated in this study has the potential to assist the identification of these causative mutations and to uncover the functional pathways by which DNA variation is translated into phenotype differences. This detailed knowledge of causative mutations will in future have a large impact on the accuracy and wide applicability of DNA-based selection methods.

Some of the SNP effects we were able to detect in this study were large (for example the X chromosome-linked SNP associations), and carried very large significance values. This means that there may be promise in exploring some low-cost DNA diagnostic tools for ranking bulls in situations where pedigree and other genetic improvement technology is not currently applied. It is possible to envisage simple DNA diagnostic tests (similar to the poll gene test) in situations where one region of the genome seems to explain most of the variation: chromosome X regions for SC, AGE26 and PNS.

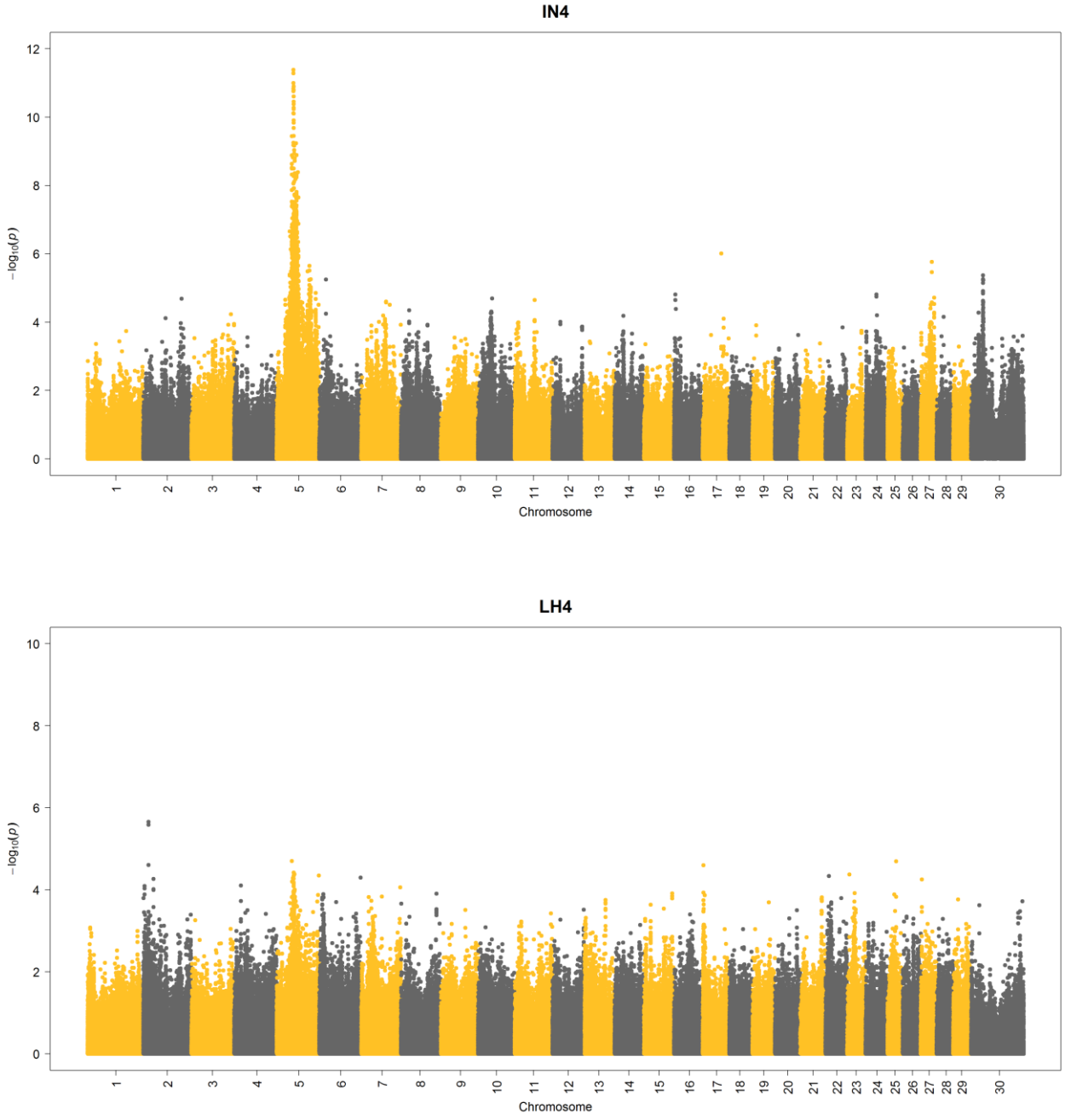

Figure 1: Manhattan Plots of SNP association with IN4 and LH4 in Tropical Composite bulls

Chromosome numbers are shown along the X-axis,"30" denotes the X-chromosome, and "31" denotes SNP that could not be mapped to the current version of the bovine genome (Btau UMD v3.0). Significance values are plotted along the Y axis. To allow all GWAS results to be plotted on the same scale, we used an arbitrary cut-off of "10" for the –log*(P)* in the y-axis. For actual values in –log*(P)* greater than 10 refer to Tables 3 and 4.

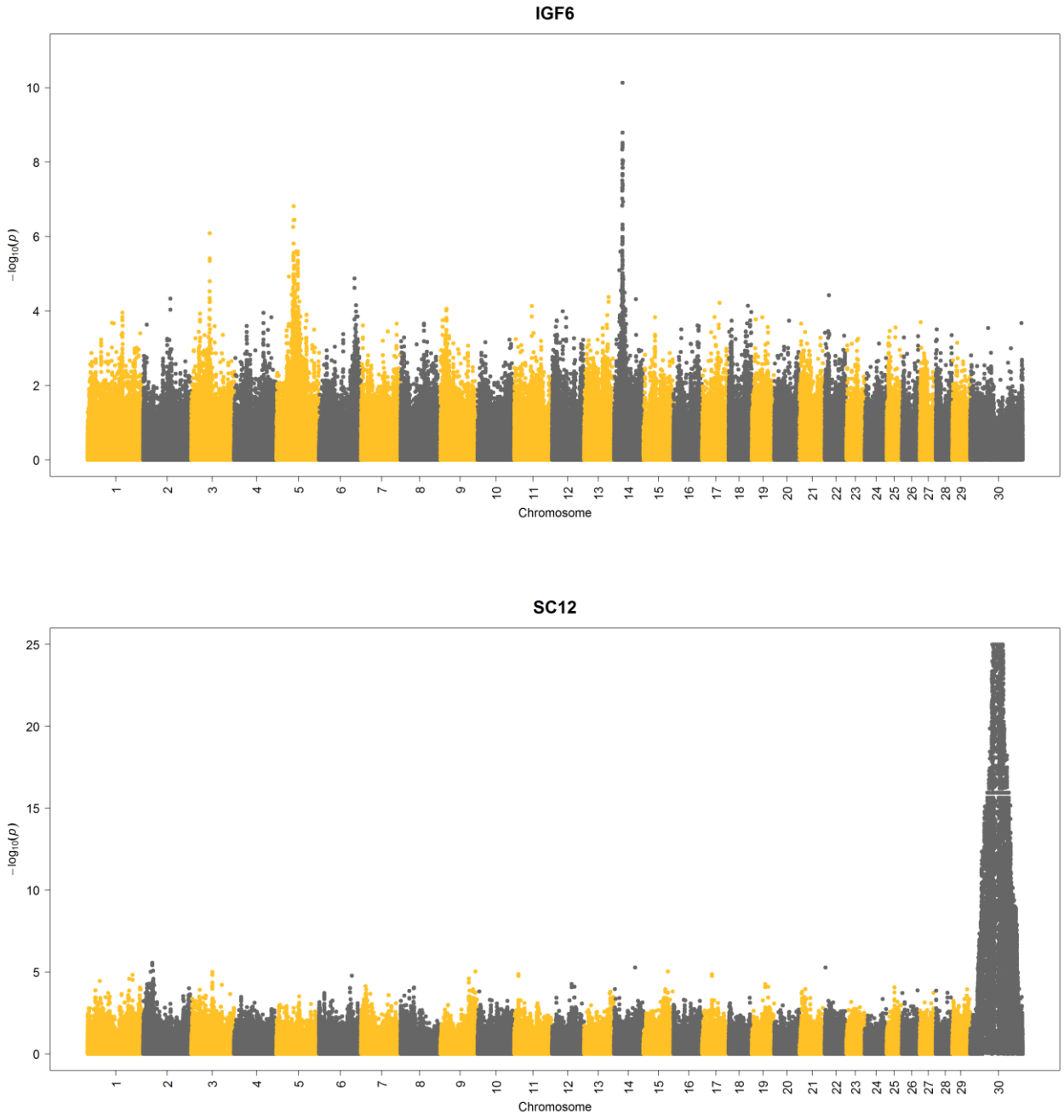

Figure 2: Manhattan Plots of SNP association with IGF6 and SC12 in Tropical Composite bulls

Chromosome numbers are shown along the X-axis,"30" denotes the X-chromosome, and "31" denotes SNP that could not be mapped to the current version of the bovine genome (Btau UMD v3.0). Significance values are plotted along the Y axis. To allow all GWAS results to be plotted on the same scale, we used an arbitrary cut-off of "10" for the –log*(P)* in the y-axis. For actual values in –log*(P)* greater than 10 refer to Tables 5 and 6.

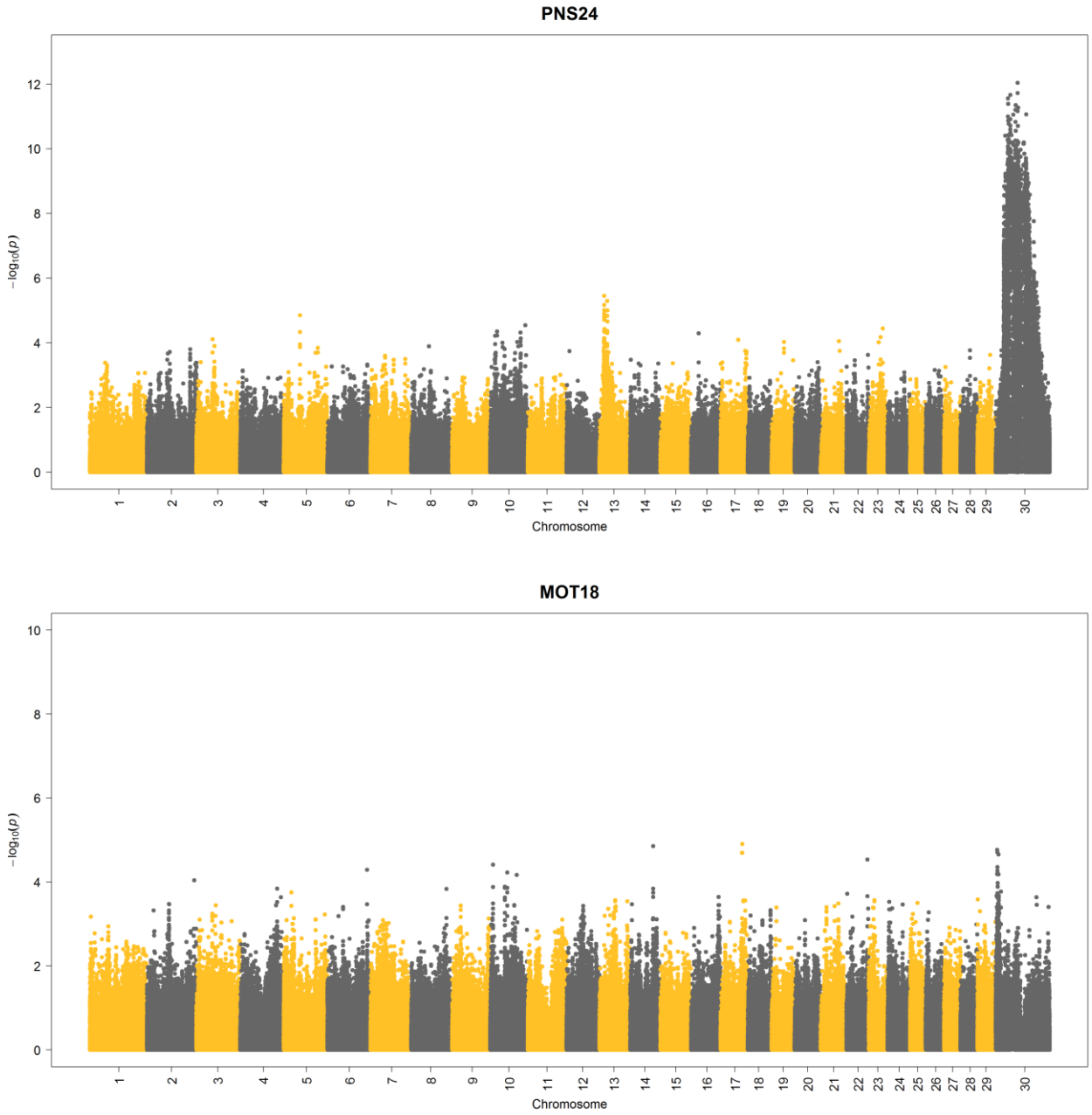

Figure 3: Manhattan Plots of SNP association with PNS24 and MOT18 in Tropical Composite bulls

Chromosome numbers are shown along the X-axis,"30" denotes the X-chromosome, and "31" denotes SNP that could not be mapped to the current version of the bovine genome (Btau UMD v3.0). Significance values are plotted along the Y axis. To allow all GWAS results to be plotted on the same scale, we used an arbitrary cut-off of "10" for the –log*(P)* in the y-axis. For actual values in –log*(P)* greater than 10 refer to Tables 7 and 8.

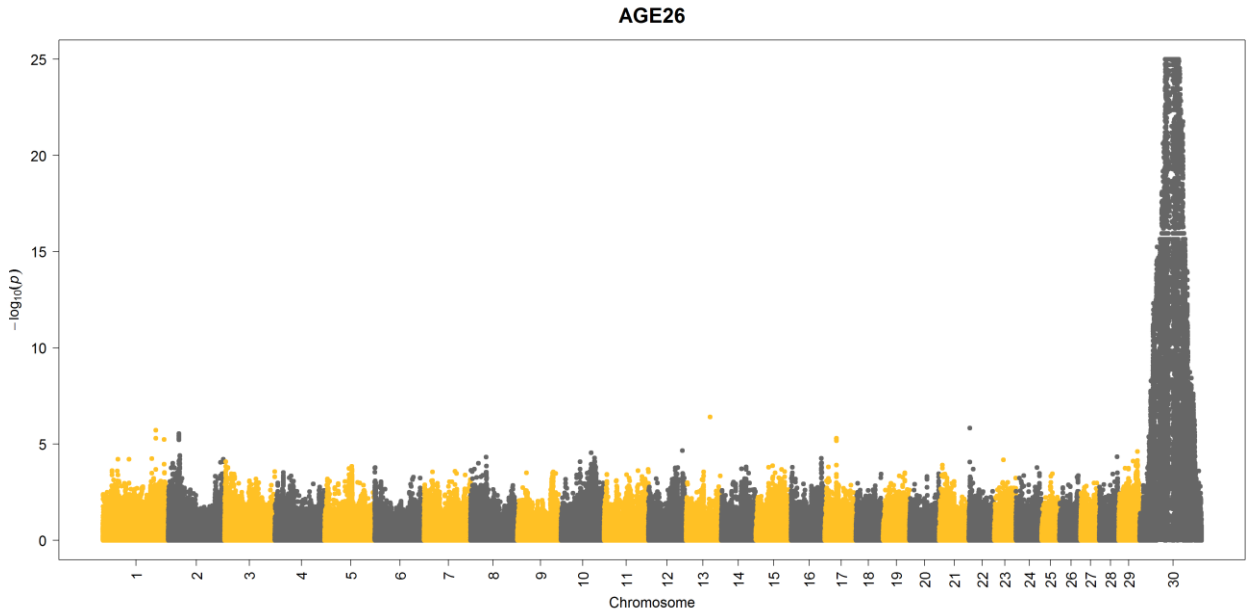

Figure 4: Manhattan Plot of SNP association with AGE26 in Tropical Composite bulls

Chromosome numbers are shown along the X-axis," 30" denotes the X-chromosome, and 31 denotes SNP that could not be mapped to the current version of the bovine genome (Btau UMD v3.0). Significance values are plotted along the Y axis. To allow all GWAS results to be plotted on the same scale, we used an arbitrary cut-off of "10" for the –log*(P)* in the y-axis. For actual values in –log*(P)* greater than 10 refer to Table 9.

Table 3: Details of SNP markers with most significant associations with IN4 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
5	ARS-BFGL-NGS-43575	0.235	0.636	11.42	45,971,062	IFNG IL22 IL26 LOC514916
5	Hapmap26409-BTA-143127	0.330	0.457	7.03	55,803,789	ARHGAP9 AVIL B4GALNT1 CTDSP2 CYP27B1 DCTN2 DDIT3 DTX3 FAM119B GEFT GGAP2 GLI1 INHBC INHBE KIF5A MARCH9 MARS METTL1 OS9 PIP4K2C R3HDM2 SLC26A10 TSFM TSPAN31
24	BTB-01258307	0.462	-0.327	4.59	31,086,312	KCTD1 PSMA8 SS18 TAF4B
27	ARS-USMARC-Parent-EF141102-rs29015783	0.420	0.356	4.75	37,513,923	C27H8orf40 CHRNA6 CHRNB3 CSGALNACT1 FNTA HGSNAT HOOK3 INTS10 RNF170 SGK196 SLC20A2 THAP1

Table 4: Details of SNP markers with most significant associations with LH4 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
5	BTA-11044-rs29016809	0.339	0.864	4.55	48,080,259	GRIP1 HELB IRAK3 MSRB3 TMBIM4
6	UA-IFASA-3742	0.371	0.812	4.28	117,106,304	LDB2
29	ARS-BFGL-NGS-34318	0.454	-0.86	4.09	41,329,056	AHNAK ASRGL1 B3GAT3 BEST1 BSCL2 C29H11orf10 DAGLA EEF1G EML3 FADS1 FADS2 FADS3 FEN1 GANAB GNG3 HNRPUL2 INCENP INTS5 MTA2 NXF1 POLR2G RAB3IL1 ROM1 SCGB1A1 SCGB1D2 STX5 TAF6L TMEM179B TTC9C UBXN1

Table 5: Details of SNP markers with most significant associations with IGF6 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
3	Hapmap31706-BTA-122950	0.242	32.762	5.54	52,065,209	CDC7 TGFBR3 ZNF644
5	Hapmap30258-BTA-143119	0.434	-31.874	5.57	56,661,587	ARHGAP9 ATP5B B4GALNT1 BAZ2A DCTN2 DDIT3 GEFT GLI1 HSD17B6 INHBC KIF5A MARS MYO1A NAB2 NACA PIP4K2C PRIM1 PTGES3 R3HDM2 RDH16 SDR9C7 SHMT2 SLC26A10 STAC3 STAT6 TAC3 TMEM194A ZBTB39
6	ARS-BFGL-NGS-116512	0.354	-26.034	3.96	105,050,270	CRMP1 EVC EVC2 JAKMIP1 PPP2R2C STK32B
14	Hapmap46986-BTA-34282	0.293	38.657	8.42	25,307,116	CHCHD7 PLAG1 IMPAD1 LYN PENK RPS20 SDR16C6

Table 6: Details of SNP markers with most significant associations with SC12 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
2	BTB-01084842	0.443	-0.557	4.35	15,021,561	ITGA4 SSFA2
3	BTB-00130211	0.414	-0.568	4.91	59,835,074	CTBS DNASE2B EDG7 GNG5 MCOLN2 PRKACB RPF1 SSX2IP UOX
X[A]	Hapmap24353-BTA-19502	0.222	1.056	28.89	72,447,651	BRWD3 CHM CYLC1 LRCH2 NSBP1 PLS3 POU3F4 SH3BGRL ZNF711

[A]For the X chromosome the interval was widened to 2.5 Mb on either side of the SNP.

Table 7: Details of SNP markers with most significant associations with PNS 24 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
8	ARS-BFGL-NGS-78015	0.385	0.039	4.02	49,550,836	ALDH1A1 ANXA1 TMC1
10	ARS-BFGL-NGS-85148	0.433	-0.041	4.52	97,043,822	MIR2293
X[A]	Hapmap30571-BTA-153233	0.230	-0.045	9.43	47,984,536	DIAPH2 RPL7

[A]For the X chromosome the interval was widened to 2.5 Mb on either side of the SNP.

Table 8: Details of SNP markers with most significant associations with MOT18 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
13	Hapmap23870-BTA-128381	0.462	-4.64	3.57	43,427,963	ABHD12 ACSS1 ANKRD16 APMAP ASB13 CST7 ENTPD6 GDI2 NET1 PYGB UCN3 VSX1
X[A]	BTA-30238-no-rs	0.122	3.873	3.24	5,808,198	ATP1B4 CUL4B GRIA3 LAMP2 ZBTB33

[A]For the X chromosome the interval was widened to 1 Mb on either side of the SNP.

Table 9: Details of SNP markers with most significant associations with AGE26 (one SNP represents QTL region)

Chr	SNP	MAF	Effect	-Log(P)	BP	Genes in ± 0.5 Mb region
2	Hapmap25643-BTA-47070	0.372	-13.125	4.20	28,012,303	LASS6 STK39
16	ARS-BFGL-NGS-110639	0.430	13.836	4.18	73,287,916	ATF3 INTS7 LPGAT1 NEK2 NENF PPP2R5A TMEM206
X[A]	Hapmap24353-BTA-19502	0.222	-29.719	33.55	72,447,651	BRWD3 CHM CYLC1 LRCH2 NSBP1 PLS3 POU3F4 SH3BGRL ZNF711

[A]For the X chromosome the interval was widened to 2.5 Mb on either side of the SNP.

5 Success in achieving objectives

5.1 Objective 1

Selected DNA from 1,000 Tropical Composite bulls extracted and archived in Beef CRC DNA bank

This objective was achieved. DNA from 1,056 individuals was prepared for genotyping and an aliquot archived in the Beef CRC DNA bank (see section 4.1).

5.2 Objective 2

Acquired 50K SNP genotypes for 1,000 Tropical Composite bulls

This objective was achieved. 50K SNP genotypes for 1,019 bulls passed QC and were securely archived at CSIRO (see section 4.1).

5.3 Objective 3

Calculated marker associations for Tropical Composite bull reproductive traits

This objective was achieved. SNP associations for 7 reproductive traits were calculated, as reported in section 4.3 of this report.

6 Impact on meat and livestock industry – Now and in five years time

This study was not intended to achieve an immediate impact on the meat and livestock industry. This R&D has produced a valuable dataset that lays the foundations for integrative analysis of genomics data for the two breeds (Brahman and Tropical Composite), two sexes, and a large number of traits (reproduction and others). The 7 reproduction traits that have been the subject of this report are only a small portion of the traits that have been measured on these bulls. The availability of genotypic data for this extensively phenotyped population means that any genomic selection equations or other DNA-based selection technology can now be tested for their predicted impact on unrelated phenotypes, in these two important beef industry breeds, and both sexes. In particular, any production trait trade-offs from DNA-based selection for reproductive performance in males or females, can now be calculated using the complete data set.

The study has also opened a window on so far undiscovered genome regions that are likely to harbour causative DNA variation for some important endocrine and reproduction traits. The discovery of these regions will contribute to more accurate genomic predictions, as well as the development of low-cost DNA diagnostics that may be applied in commercial animals. By making DNA-based section more accurate and more affordable, the rate of genetic progress will be accelerated across the industry.

7 Conclusions and recommendations

7.1 Conclusions

This study has contributed important information on the genomic architecture on several bull reproductive traits. Genotype data and analysis of Tropical Composite bulls from this study together with the results obtained from a group of Brahman bulls raised and phenotyped at the same time enabled comparisons between the SNP associations found in each breed.

In particular, two regions on the X-chromosome appear to be involved in the traits SC12, PNS24 and AGE26 in both Brahmans and Tropical Composites. A region on chromosome 14 that has been shown to be significantly linked to IGF1 levels, as well as a range of other growth traits, again shows up in this study, as linked to IGF1.

Important differences between the two breeds include the finding that the IN4 trait, which in Brahman bulls has been mapped to the *INHA* gene on chromosome 2, appears to map to the inhibin beta genes (*INHBC* and *INHBE*) on chromosome 5 in Tropical Composites.

The profound breed differences detected here have important implications for the Australian beef industry and the multiple breeds it employs. The development and application of DNA-based predictions to the northern beef industry needs to take into account the genomic architecture of the trait for each breed and use to its advantage situations when diverse breeds yield similar results (i.e. chromosome X and 14).

The high quality phenotyping data assembled by the Beef CRC teams are a world-leading dataset, particularly in the area of genomics of reproduction in tropically-adapted beef breeds. The collaborative efforts of the Beef CRC team have provided outstanding opportunities to translate international advances in genomics technology into real progress towards genetic solutions for the problem of reproductive inefficiencies of the northern beef herd.

7.2 Recommendations

7.2.1 Peer-reviewed publication of the results

The results presented here need to me more thoroughly explored and evaluated and submitted for peer review in an internationally-recognised journal.

7.2.2 Conduct integrative analysis of entire “reproduction” dataset

The 1,019 Tropical Composite bull genotypes complete a set of genotypes that consists of almost 1,000 Tropical composite cows, an equivalent number of Brahman cows, and 1,000 Brahman bulls, all from the same populations. These data have recently been complemented by relevant RNASeq data from Brangus heifers (acquired via our international collaborative network), and will be joined by RNASeq data from Brahman heifers (in progress by our QAAFI collaborators), and large amounts of genotypic data on commercial Brahman and Tropical Composite cows collected as part of other studies.

This data set offers substantial opportunities to discover genomic links between production and reproductive traits, and how to work with DNA-based predictions in a multi-breed context such as the northern Australian beef industry. The project is well suited to a postdoctoral project that would deliver outcomes to industry, fulfil an important training and succession planning need and open new horizons for genomics science solutions for livestock industries.

7.2.3 Invest in detailed studies of chromosome regions with large effect to uncover causative mutations

The opportunity to identify the nature of the DNA variations that are linked to endocrine and fertility traits such as inhibin levels, scrotal circumference, percent normal sperm and age at puberty, which this study has created, needs to be followed up with targeted molecular biology studies. Recently completed whole genome sequences and RNASeq datasets are available for this, but need to be complemented by the sequencing of testis tissue libraries and more targeted sequencing of key individuals. The identification of causal mutations is important for more accurate genomic predictions, for the development of lower-cost targeted DNA assays and for the emergence of new insights into the physiological regulation of the traits in question, which may lead to new biotechnologies.

Higher R&D students (Master or PhD) and/or postdoctoral level projects would be able to progress these issues.

8 Acknowledgements

We need to acknowledge the efforts of a large number of collaborators in the CRC for Beef Genetic Technology that have made this study possible. Dr Richard Holroyd, Dr Jim Walkley, Dr Brian Burns and Dr David Johnston were instrumental in designing and resourcing the phenotype collections. Northern Pastoral Group, the Queensland government (DEEDI) and CSIRO collaborated to collect phenotypes. The authors wish to thank the many colleagues from the Beef CRC who contributed towards data collection and analysis for this project, in particular Ms Bronwyn Venus, Mr Nick Corbet, Ms Debra Corbet, Mr Matt Woolcott and Mr Paul Williams. Dr Rachel Hawken laid the foundations for the current study by initiating a genome-wide association study of 1,000 Brahman bulls from the same resource population. Mr Russell McCulloch, Ms Rama Addepalli, and Dr Emily Piper ensured efficient and reliable DNA preparations and genotyping.

9 Bibliography

Bagu, E.T., Madgwick, S., Duggavathi, R., Bartlewski, P.M., Barrett, D.M., Huchkowsky, S., Cook, S.J., Rawlings, N.C., 2004. Effects of treatment with LH or FSH from 4 to 8 weeks of age on the attainment of puberty in bull calves. Theriogenology 62, 861-873.

Barwick, S.A., Johnston, D.J., Burrow, H.M., Holroyd, R.G., Fordyce, G., Wolcott, M.L., Sim, W.D., Sullivan, M.T., 2009. Genetics of heifer performance in 'wet' and 'dry' seasons and their relationships with steer performance in two tropical beef genotypes. Animal Production Science 49, 367-382.

Burns, B., Gazzola, C., Holroyd, R., Crisp, J., McGowan, M., 2011. Male reproductive traits and their relationship to reproductive traits in their female progeny: a systematic review. Reproduction in Domestic Animals 46, 534-553.

Burns, B.M., Corbet, N.J., Corbet, D.H., Crisp, J.M., Venus, B.K., Johnston, D.J., Li, Y., McGowan, M.R., Holroyd, R.G., 2013. Male traits and herd reproductive capability in tropical beef cattle. 1. Experimental design and animal measures. Animal Production Science 53, 87-100.

Burrow, H.M., Johnston, D.J., Barwick, S.A., Holroyd, R.G., Barendse, W., Thompson, J.M., Griffith, G.R., Sullivan, M., 2003. Relationships between carcass and beef quality and components of herd profitability in Northern Australia. 50 years of DNA: Proceedings of the Fifteenth Conference, Association for the Advancement of Animal Breeding and Genetics, Melbourne, Australia, 7-11 July 2003, 359-362.

Chan, E.K., Hawken, R., Reverter, A., 2009. The combined effect of SNP-marker and phenotype attributes in genome-wide association studies. Anim Genet 40, 149-156.

Childers, C.P., Reese, J.T., Sundaram, J.P., Vile, D.C., Dickens, C.M., Childs, K.L., Salih, H., Bennett, A.K., Hagen, D.E., Adelson, D.L., Elsik, C.G., 2011. Bovine Genome Database: integrated tools for genome annotation and discovery. Nucleic Acids Res 39, D830-834.

Corbet, N.J., Burns, B.M., Corbet, D.H., Johnston, D.J., Crisp, J.M., McGowan, M., Prayaga, K.C., Venus, B.K., Holroyd, R.G., 2009. Genetic variation in growth, hormonal and seminal traits of young tropically adapted bulls. Proceedings of the Association for the Advancement of Animal Breeding and Genetics 18, 121-124.

Corbet, N.J., Burns, B.M., Johnston, D.J., Wolcott, M.L., Corbet, D.H., Venus, B.K., Li, Y., McGowan, M.R., Holroyd, R.G., 2013. Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits. Animal Production Science 53, 101-113.

Fordyce, G., Entwistle, K., Norman, S., Perry, V., Gardiner, B., Fordyce, P., 2006. Standardising bull breeding soundness evaluations and reporting in Australia. Theriogenology 66, 1140-1148.

Fortes, M.R.S., Lehnert, S.A., Bolormaa, S., Reich, C., Fordyce, G., Corbet, N.J., Whan, V., Hawken, R.J., Reverter, A., 2012a. Finding genes for economically important traits: Brahman cattle puberty. Animal Production Science 52, 143-150.

Fortes, M.R.S., Reverter, A., Hawken, R.J., Bolormaa, S., Lehnert, S.A., 2012b. Candidate Genes Associated with Testicular Development, Sperm Quality, and Hormone Levels of Inhibin, Luteinizing Hormone, and Insulin-Like Growth Factor 1 in Brahman Bulls. Biology of Reproduction 87, 58, 51-58.

Hawken, R.J., Zhang, Y.D., Fortes, M.R., Collis, E., Barris, W.C., Corbet, N.J., Williams, P.J., Fordyce, G., Holroyd, R.G., Walkley, J.R., Barendse, W., Johnston, D.J., Prayaga, K.C., Tier, B., Reverter, A., Lehnert, S.A., 2012. Genome-wide association studies of female reproduction in tropically adapted beef cattle. J Anim Sci 90, 1398-1410.

Holroyd, R.G., Doogan, W., De Faveri, J., Fordyce, G., McGowan, M.R., Bertram, J.D., Vankan, D.M., Fitzpatrick, L.A., Jayawardhana, G.A., Miller, R.G., 2002. Bull selection and use in northern Australia. 4. Calf output and predictors of fertility of bulls in multiple-sire herds. Animal Reproduction Science 71, 67-79.

Hotzel, M.J., Markey, C.M., Walkden-Brown, S.W., Blackberry, M.A., Martin, G.B., 1998. Morphometric and endocrine analyses of the effects of nutrition on the testis of mature Merino rams. Journal of Reproduction and Fertility 113, 217-230.

Johnston, D.J., Barwick, S.A., Corbet, N.J., Fordyce, G., Holroyd, R.G., Williams, P.J., Burrow, H.M., 2009. Genetics of heifer puberty in two tropical beef genotypes in northern Australia and associations with heifer- and steer-production traits. Animal Production Science 49, 399-412.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P., Sonstegard, T.S., Van Tassell, C.P., 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS One 4, e5350.

Moore, K.L., Johnston, D.J., Graser, H.U., Herd, R., 2005. Genetic and phenotypic relationships between insulin-like growth factor-I (IGF-I) and net feed intake, fat, and growth traits in Angus beef cattle. Australian Journal of Agricultural Research 56, 211-218.

Perez-Enciso, M., Misztal, I., 2011. Qxpak.5: Old mixed model solutions for new genomics problems. BMC Bioinformatics 12.

Phillips, D.J., 2005. Activins, inhibins and follistatins in the large domestic species. Domestic Animal Endocrinology 28, 1-16.

Prayaga, K.C., Corbet, N.J., Johnston, D.J., Wolcott, M.L., Fordyce, G., Burrow, H.M., 2009. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. Animal Production Science 49, 413-425.

Unaniam, M.M., 1997. Precocidade sexual em machos da raca Nelore. Nelore ano VI, 66-68.