CSIRO PUBLISHING

Animal Production Science, 2012, **52**, 143–150
http://dx.doi.org/10.1071/AN11165

Finding genes for economically important traits: Brahman cattle puberty

M. R. S. Fortes[A,B], *S. A. Lehnert*[A], *S. Bolormaa*[C], *C. Reich*[C], *G. Fordyce*[D], *N. J. Corbet*[A], *V. Whan*[A], *R. J. Hawken*[A] *and A. Reverter*[A,E]

Cooperative Research Centre for Beef Genetic Technologies.
[A]CSIRO Livestock Industries, Queensland Bioscience Precinct, St Lucia, Qld 4067, Australia.
[B]The University of Queensland, School of Veterinary Science, Gatton, Qld 4343, Australia.
[C]Biosciences Research Division, Department of Primary Industries Victoria, 1 Park Drive, Bundoora, Vic. 3083, Australia.
[D]Queensland Alliance for Agriculture and Food Innovation, Centre for Animal Science, The University of Queensland, St Lucia, Qld 4067, Australia.
[E]Corresponding author. Email: Tony.Reverter-Gomez@csiro.au

Abstract. Age at puberty is an important component of reproductive performance in beef cattle production systems. Brahman cattle are typically late-pubertal relative to *Bos taurus* cattle and so it is of economic relevance to select for early age at puberty. To assist selection and elucidate the genes underlying puberty, we performed a genome-wide association study (GWAS) using the BovineSNP50 chip (~54 000 polymorphisms) in Brahman bulls (n = 1105) and heifers (n = 843) and where the heifers were previously analysed in a different study. In a new attempt to generate unbiased estimates of single-nucleotide polymorphism (SNP) effects and proportion of variance explained by each SNP, the available data were halved on the basis of year and month of birth into a calibration and validation set. The traits that defined age at puberty were, in heifers, the age at which the first *corpus luteum* was detected (AGECL, $h^2 = 0.56 \pm 0.11$) and in bulls, the age at a scrotal circumference of 26 cm (AGE26, $h^2 = 0.78 \pm 0.10$). At puberty, heifers were on average older (751 ± 142 days) than bulls (555 ± 101 days), but AGECL and AGE26 were genetically correlated ($r = 0.20 \pm 0.10$). There were 134 SNPs associated with AGECL and 146 SNPs associated with AGE26 ($P < 0.0001$). From these SNPs, 32 (~22%) were associated ($P < 0.0001$) with both traits. These top 32 SNPs were all located on Chromosome BTA 14, between 21.95 Mb and 28.4 Mb. These results suggest that the genes located in that region of BTA 14 play a role in pubertal development in Brahman cattle. There are many annotated genes underlying this region of BTA 14 and these are the subject of current research. Further, we identified a region on Chromosome X where markers were associated (P < 1.00E–8) with AGE26, but not with AGECL. Information about specific genes and markers add value to our understanding of puberty and potentially contribute to genomic selection. Therefore, identifying these genes contributing to genetic variation in AGECL and AGE26 can assist with the selection for early onset of puberty.

Additional keywords: *Bos indicus*, *corpus luteum*, genome-wide association, scrotal circumference.

Received 5 August 2011, accepted 10 February 2012, published online 6 March 2012

Introduction

Age at puberty is an important component of cattle performance. It determines the beginning of reproductive life for any breeding animal and it influences generation interval, affecting the rate of genetic improvement and herd productivity. Selecting for early pubertal heifers and bulls is a practical approach for reducing the generation interval and potentially increasing fertility (Lesmeister *et al*. 1973; Foster 1994; Siddiqui *et al*. 2008; Johnston *et al*. 2010).

For genetic selection, puberty must be defined as a measurable and inherited trait. Puberty in heifers was defined by plasma progesterone concentration, ultrasound images of *corpus luteum* (CL) and detection of oestrous (Shamay *et al*. 2005; Romano *et al*. 2007; Johnston *et al*. 2009). Puberty in bulls was defined by scrotal circumference thresholds and was related to sperm concentration, motility and morphology (Lunstra *et al*. 1978; Bagu *et al*. 2004; Siddiqui *et al*. 2008; Corbet *et al*. 2009). Pubertal traits vary in heritability (h^2) from low to moderately high (Cammack *et al*. 2009).

There is a historical and increasing body of evidence reporting the genetic correlations between female and male puberty in cattle (Martin *et al*. 1992). Scrotal circumference of bulls correlates with puberty in their female relatives. Correlation estimates range from $r = -0.71$ to $r = -0.15$ (Burns *et al*. 2011). This evidence supports the indirect selection strategy for early pubertal heifers, using correlated traits measured in

bulls. Given the size of the genetic correlation between these traits, some genes and genetic markers associated with puberty in heifers are likely to be associated with puberty in bulls, and *vice versa*.

Bos indicus cattle, such as Brahman cattle, are reportedly older at puberty when compared with most *Bos taurus* breeds (Lunstra and Cundiff 2003; Lopez *et al*. 2006). It is of economic relevance to select for early age at puberty in Brahman cattle. To assist selection and identify genes underlying fertility, we performed a GWAS in a sample of bulls with a measurement for puberty, namely the age at a scrotal circumference of 26 cm (AGE26). The present study is the first report of a GWAS regarding scrotal circumference in bulls, while the GWAS on AGECL in heifers represents a re-analysis of a previously reported study (Hawken *et al*. 2012). The Brahman heifers measured for AGECL were the mothers of the bulls measured for AGE26. Heritability, genetic correlation and genes associated with AGECL and AGE26 are reported and discussed.

Materials and methods

Cattle and traits

Cattle were bred and supplied by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) as described previously (Corbet *et al*. 2009; Johnston *et al*. 2009). In brief, data from 1007 Brahman heifers and 1118 Brahman bulls were included in the current analysis. The phenotypes of the heifers have been reported previously (Johnston *et al*. 2009). The bulls included in the present experiment were the offspring of the heifers, which became breeding heifers for the Beef CRC. Their pedigree consists of over 50 grandsire families; 54 bulls that sired heifers, which were mated to 55 industry sires to produce the bulls used in the study.

The traits that defined puberty were AGECL in heifers and AGE26 in bulls. The first CL was detected by ovarian ultrasound examinations conducted every 4–6 weeks, after heifers had reached an average liveweight of 200 kg (Johnston *et al*. 2009). AGECL was estimated from annotation of the date when the first CL was observed and the date of birth. Scrotal circumference (SC) was measured with a standard metal tape (Fordyce *et al*. 2006). Between weaning and 24 months of age, eight measurements of SC were taken for each bull, at 3-month intervals. Summary statistics for the age and SC of bulls at each of the eight time points are presented in Table 1. Using these repeated measurements for individual regressions, we interpolated the age when the bull achieved 26 cm of SC (AGE26, expressed in days). Achieving SC of 26 cm was considered a threshold for puberty in *Bos indicus* bulls. The threshold was different from the previously described 28-cm threshold because *Bos indicus* typically present a more elongated scrotum and a smaller SC than do *Bos taurus* (Lunstra *et al*. 1988; McGowan *et al*. 2002; Silva *et al*. 2011). For a visual assessment of age *v.* SC in this population, see Fig. 1. A more in-depth description and quantitative analysis of all reproductive and growth phenotypes available for this population of bulls was reported elsewhere (Corbet *et al*. 2009, 2011).

Table 1. Means ± s.e. for measurements of scrotal circumference (SC) at eight time points

Time point	Age	SC
1	182.30 ± 133.64	17.02 ± 0.86
2	256.43 ± 209.77	18.03 ± 0.93
3	313.22 ± 240.97	19.68 ± 1.29
4	373.73 ± 312.85	21.23 ± 1.72
5	443.66 ± 359.56	24.55 ± 2.34
6	525.52 ± 443.19	26.38 ± 2.44
7	617.87 ± 486.12	28.03 ± 2.50
8	702.78 ± 539.66	29.88 ± 2.71

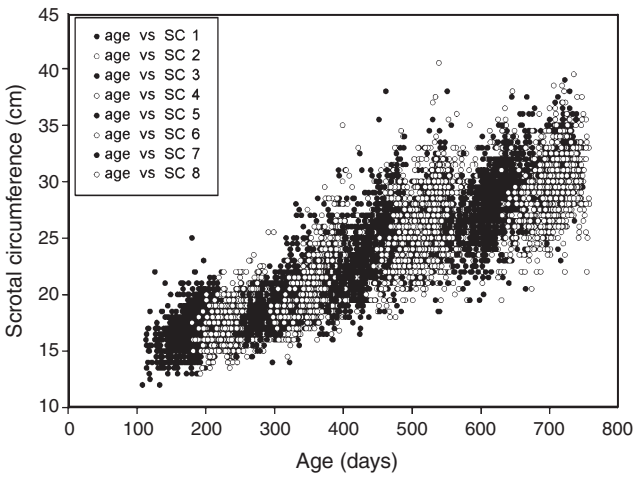

Fig. 1. Age and scrotal circumference (SC) measured across eight time points. Scatter plot of the age in days (*x*-axis) versus the SC in cm (*y*-axis), with time points represented by alternating colours (black and white).

Genotypes

The Illumina BovineSNP50 array (Van Tassell *et al*. 2008; Matukumalli *et al*. 2009) was used to genotype samples according to the manufacturer's protocols (Illumina Inc., San Diego, CA). Positions for each SNP were based on the UMD3 assembly of the bovine genome sequence (available from http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/btauUMD3/, verified 14 December 2011). Genotypes for 843 heifers had been generated previously using the BovineSNP50 v1 array and were reported in two recent GWAS that analysed growth traits and reproduction traits, including AGECL (Bolormaa *et al*. 2011; Hawken *et al*. 2012).

In the present study, 1118 bulls were genotyped using the Illumina BovineSNP50 v2 array. Quality control was similar to the previous study, with repeat samples included and the Bead Studio software from Illumina (www.illumina.com, verified 28 February 2012) used to determine the genotype calls. Animals with call rates <98% were excluded, resulting in 1105 bulls being retained for analyses. SNPs with auto-calling rates <85% and SNPs with a minor allele frequency <0.01 were excluded from further analyses.

For the current study, missing genotypes for bulls and heifers were imputed using the BEAGLE 3.2 program (Browning and Browning 2010). Imputation and quality-control procedures

yielded a total of 43 821 SNPs used in the genome-wide association analysis. Genotype calls were coded as 0 for the homozygote of the first allele (A), 1 for the heterozygote, and 2 for the homozygote of the second allele (B). Alleles A and B were defined as per top and bottom rules from Illumina (http://www.illumina.com/documents/products/technotes/technote_topbot.pdf, verified 17 February 2012). Imputation resulted in better coverage of the X chromosome, adding new data to the previous heifer study on AGECL (Hawken *et al*. 2012). This improved X chromosome coverage for the heifers and all genotypes obtained for bulls are reported here for the first time.

Genetic correlation and heritability

Genetic correlation and h^2 (for AGE26) were estimated fitting the two traits in a bivariate analysis using a mixed model and three generation pedigree relationships. For AGECL, fixed effects included in the model and h^2 have been reported previously (Johnston *et al*. 2009); here, we estimated the h^2 *de novo* in the bivariate analysis, using the same fixed effects. Fixed effects included in the model for AGE26 were contemporary group (bulls born in the same year and location – or cohort – and post-weaning location) and birth month. Solutions to the effects in the model as well as variance components were estimated using VCE v.6 software (Groeneveld and García-Cortés; http://vce.tzv.fal.de, verified 20 February 2012).

Genome-wide association study

Genome-wide association studies (GWAS) were performed for AGECL and AGE26 separately and by using two approaches. First, GWAS was performed with the full datasets for bulls (1105) and heifers (843) and, subsequently, each population was split into calibration and a validation datasets, for providing additional evidence of SNP association and estimated effects. The calibration and validation datasets were generated according to birth date, dividing the populations into a younger half (~50%) and an older half (~50%). For AGECL, the calibration dataset was formed by heifers born until November 2001 ($n = 480$) and older half was born from December 2001 onwards ($n = 363$). For AGE26, the calibration dataset comprised those bulls born until December 2005 ($n = 565$) and the bulls in the validation dataset were born from January 2006 onwards ($n = 550$). Using birth date to define calibration and validation datasets is a tested approach (Hayes *et al*. 2010). A GWAS for AGECL using the full dataset has been reported previously (Hawken *et al*. 2012), but the results reserving 363 heifers for a validation exercise are reported only here. No previous study exists to relate the GWAS for AGE26 reported here.

The effect of each SNP was estimated three times, as follows: first, using phenotypic data from the full dataset; second, using phenotypic data from the calibration dataset; and finally, using phenotypic data from the validation dataset. In all three instances, the effect of each SNP was estimated using the mixed model of Eqn 1, as follows:

$$y_{ij} = \mathrm{X}\beta + \mathrm{Z}u + \mathrm{S}_k s_{jk} + e_{ij}, \quad (1)$$

where y_{ij} represents the vector of observations from the ith heifer or bull at the jth trait ($j = 1$ and 2 for AGECL and AGE26, respectively); X is the incidence matrix relating fixed effects in β with observations in y_{ij}; Z is the incidence matrix relating random additive polygenic effects in u with observations in y_{ij}; S_k in the vector of genotypes for the kth SNP across all animals; s_{jk} represents the additive association of the kth SNP on the jth trait; and e_{ij} is the vector of random residual effects. Fixed effects included in the model were the same as described for the models used to estimate the genetic correlations and h^2. Tests for the significance of SNP association were conducted using Qxpak5 (Perez-Enciso and Misztal 2004, 2011). Qxpak5 performs a likelihood ratio test, testing the model with SNP *versus* the model without the SNP, against a chi-square distribution with 1 degree of freedom and this was carried out one SNP at a time. Chromosome X genotypes of males and females were not analysed together, because the traits were analysed separately.

False discovery rates were estimated using Eqn 2, as follows:

$$FDR = \frac{nP}{k}, \quad (2)$$

where n represents the total number of SNPs included in the study (in the present study, $n = 43\,821$), P is the P-value threshold being used and k is the actual number of associated SNPs in the given P-value threshold.

The percentage of the genetic variance accounted for by the ith SNP was computed according to the formulae of Eqn 3, as follows:

$$\%\mathrm{V}_i = 100 \times \frac{2p_i q_i a_i^2}{\sigma_g^2}, \quad (3)$$

where p_i and q_i are the allele frequencies for the ith SNP estimated across the entire population, a_i is the estimated additive effect of the ith SNP on the trait in question (AGECL or AGE26), and σ_g^2 is the REML estimate of the (poly-)genetic variance for the trait in question.

Results

Descriptive statistics, h^2 and associated errors for AGECL and AGE26 are presented in Table 2. At puberty, heifers were on average older (751 ± 142 days) than bulls (555 ± 101 days). The heritability estimated for puberty in heifers was lower (AGECL, $h^2 = 0.56 \pm 0.11$) than that in bulls (AGE26, $h^2 = 0.78 \pm 0.10$). The traits AGECL and AGE26 were genetically correlated, with $r = 0.20 \pm 0.10$ estimated using REML methods.

Table 2. Summary statistics and heritability estimates for age at the first *corpus luteum* (AGECL) and age at the scrotal circumference of 26 cm (AGE26)

Std, standard deviation; AGECL phenotype is reported in Johnston *et al*. (2009)

Parameter	AGECL	AGE26
N (animals)	1007	1118
Mean	750.60	555.06
Std	142.14	100.52
Minimum	394.00	284.00
Maximum	1211.00	1174.00
h^2 ± s.e.	0.56 ± 0.11	0.78 ± 0.10

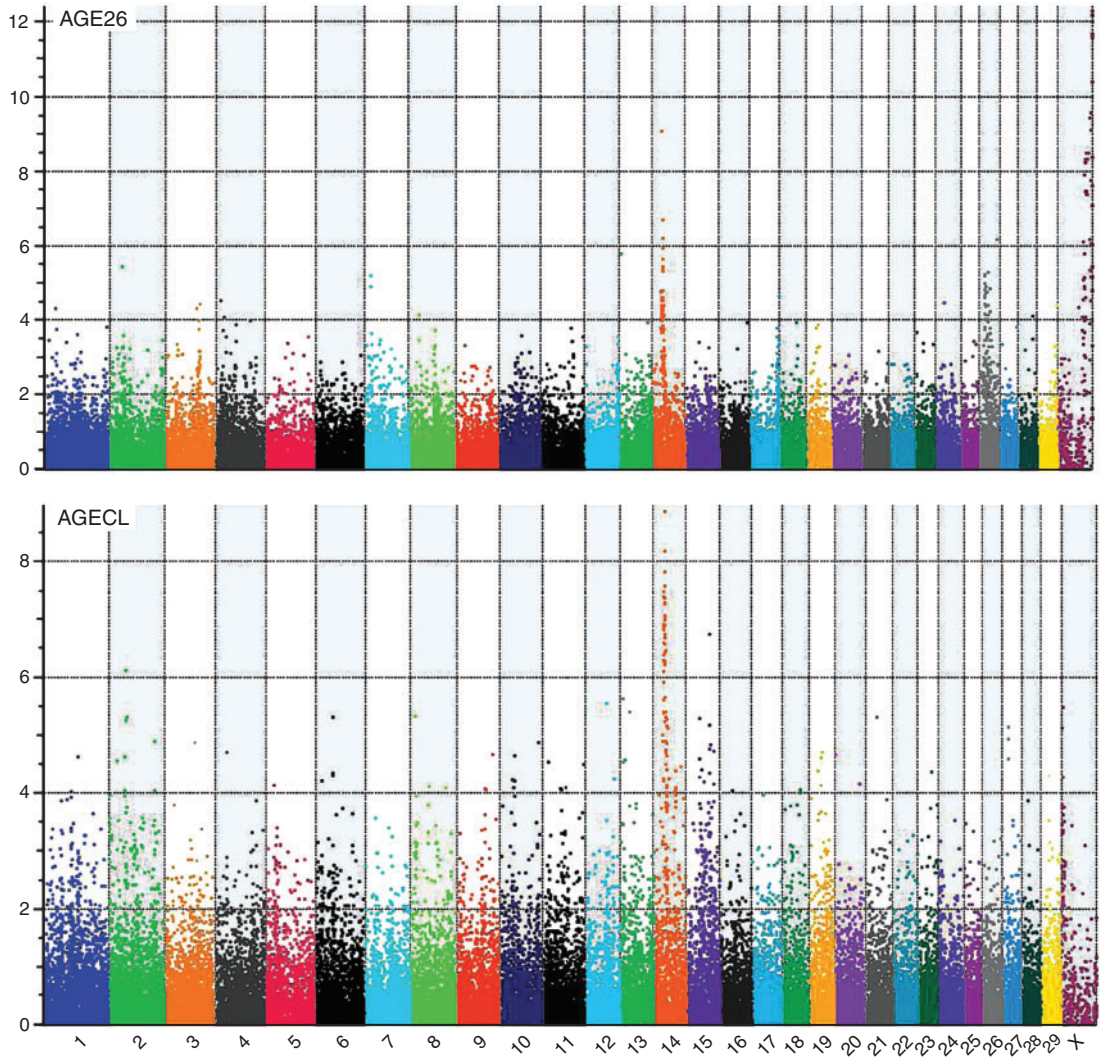

Fig. 2. Manhattan plots for age at a scrotal circumference of 26 cm (AGE26) and age at the first *corpus luteum* (AGECL). Plots represent the associations between genome-wide singe-nucleotide polymorphisms (SNPs) and AGE26, as well as, SNPs and AGECL. In both plots, chromosomal positions are in the *x*-axis and –log (*P*-values) are in the *y*-axis.

Figure 2 provides an overview of GWAS results for the full datasets presented as Manhattan plots (i.e. significance on the *y*-axis *v.* genome map position on the *x*-axis) for AGE26 and AGECL. Two regions showed clear association peaks in the GWAS of AGE26, one on the X chromosome and another on BTA14. The GWAS of AGECL also showed a peak on BTA14 for the same region where SNPs were significant for AGE26. There were 134 SNPs associated with AGECL and 146 SNPs associated with AGE26 ($P < 0.0001$, false discovery rate = 0.03). Of these SNPs, 32 (~22%) were associated with both traits ($P < 0.0001$, Table 3). These 32 SNPs that were in common for both traits were all located on BTA14, between 21.95 Mb and 28.4 Mb. The most significant SNP for AGECL was Hapmap23509-BTC-073113 at Position 27 198 715 of BTA14 ($P = 1.36E{-}09$). The most significant SNP for AGE26 was BTA-30242-no-rs at Position 85 495 447 of Chromosome X ($P = 5.35E{-}13$). The effects, *P*-values and proportion of genetic variance explained for each of the 32 SNPs associated ($P < 0.0001$) with both AGECL and AGE26 are presented in Table 4. The effects, *P*-values and proportion of the variance explained for the SNPs underlying the association peak located on X chromosome are also reported (Table 5).

Table 3. Number of single-nucleotide polymorphisms (SNPs) associated with age at the first *corpus luteum* (AGECL), age at the scrotal circumference of 26 cm (AGE26) and both traits at various *P*-value thresholds

FDR, false discovery rate

P-value	AGECL		AGE26		Both
	SNPs	FDR	SNPs	FDR	
$P < 0.05$	4924	0.44	3865	0.57	505
$P < 0.01$	1629	0.27	1241	0.35	119
$P < 0.001$	410	0.11	315	0.14	49
$P < 0.0001$	134	0.03	146	0.03	32
$P < 0.00001$	61	0.01	74	0.01	9

Table 4. Effect, *P*-values and proportion of the variance explained for single-nucleotide polymorphisms (SNPs) associated ($P < 0.0001$) with both age at the first *corpus luteum* (AGECL) and age at the scrotal circumference of 26 cm (AGE26) and located in chromosome 14

SNP positions and distance to nearest gene are reported in base pairs and are based on the UMD3 assembly of the bovine genome sequence. Most significant SNPs in terms of *P*-values are reported in bold. Note that SNP BTA-97369-no-rs was the only SNP not associated with both calibration and validation dataset for neither AGE26 nor AGECL (see Supplementary Material Tables S1, S2). MAF, minor allele frequency

SNP	Chromosome 14		AGECL			AGE26			Gene	Distance
	Position	MAF	Effect	*P*	%Var.	Effect	*P*	%Var.		
Hapmap27177-BTA-147405	21 954 567	0.40	24.52	9.83E-06	8.22	16.86	1.79E-05	2.17	CSF2RA	62 207
Hapmap32552-BTA-129045	22 323 719	0.42	27.89	8.00E-07	10.79	16.80	1.66E-05	2.19	SNTG1	0
BTB-01252028	22 544 496	0.45	–28.34	4.10E-07	11.34	–16.48	4.71E-05	2.14	PCMTD1	124 865
ARS-BFGL-BAC-12159	22 587 081	0.50	–27.55	4.96E-07	10.80	–16.97	2.47E-05	2.29	PCMTD1	82 280
Hapmap31202-BTA-162588	23 988 778	0.44	–27.01	1.20E-06	10.25	–16.74	2.96E-05	2.20	RP1	1415
ARS-BFGL-NGS-104268	**24 057 354**	**0.42**	**–28.3**	**1.20E-07**	**11.16**	**–25.00**	**8.45E-10**	**4.86**	RP1	58 016
BTB-01532239	24 437 778	0.48	–29.43	3.24E-08	12.31	–15.86	6.30E-05	2.00	XKR4	0
BTB-01530778	24 482 969	0.45	–27.06	2.57E-07	10.33	–16.39	3.96E-05	2.12	XKR4	0
BTB-01530788	24 524 205	0.48	–29.33	4.13E-08	12.23	–15.61	8.21E-05	1.94	XKR4	0
BTB-01530836	24 573 257	0.48	29.33	4.13E-08	12.23	15.89	6.10E-05	2.00	XKR4	0
BTB-00557532	24 643 266	0.49	–29.58	3.32E-08	12.45	–15.91	5.95E-05	2.01	XKR4	32 308
Hapmap41234-BTA-34285	25 107 556	0.47	–30.12	1.45E-08	12.85	–17.69	4.39E-06	2.48	LOC526726	0
BTB-02056709	25 175 950	0.48	28.47	5.51E-08	11.53	18.10	4.72E-06	2.60	SDR16C6	0
Hapmap46986-BTA-34282	25 307 116	0.50	–28.40	5.09E-08	11.48	–18.73	1.19E-06	2.79	PENK	84 125
BTB-01779799	25 351 733	0.44	–28.87	1.01E-07	11.71	–16.15	3.01E-05	2.05	PENK	128 742
ARS-BFGL-NGS-529	25 638 580	0.44	–29.06	1.04E-07	11.82	–17.98	3.65E-06	2.53	IMPAD1	77 701
ARS-BFGL-NGS-111395	25 731 992	0.48	28.35	8.71E-08	11.43	16.83	1.65E-05	2.25	IMPAD1	171 113
BTA-97369-no-rs	25 887 784	0.44	–24.14	5.73E-06	8.16	–16.06	7.68E-05	2.02	FAM110B	162 458
Hapmap24966-BTC-054594	26 080 801	0.49	–29.67	6.43E-09	12.52	–15.76	9.54E-05	1.97	FAM110B	0
Hapmap32434-BTC-011497	26 473 490	0.40	–27.53	2.09E-07	10.39	–17.06	2.49E-05	2.23	NSMAF	0
ARS-BFGL-NGS-35159	26 508 236	0.41	27.97	1.27E-07	10.78	16.54	4.41E-05	2.11	NSMAF	12 525
Hapmap28828-BTC-011250	26 713 734	0.40	–27.19	3.70E-07	10.08	–16.97	2.90E-05	2.19	TOX	0
Hapmap30932-BTC-011225	26 766 010	0.40	–26.69	5.31E-07	9.71	–16.60	4.25E-05	2.10	TOX	0
Hapmap25761-BTC-065280	26 949 215	0.44	26.78	1.96E-07	10.07	18.53	2.32E-06	2.69	TOX	7489
BTB-01280026	27 035 971	0.40	27.28	2.63E-07	10.17	17.02	2.73E-05	2.21	TOX	94 245
Hapmap27934-BTC-065223	27 155 254	0.40	27.69	1.84E-07	10.47	16.31	5.58E-05	2.03	TOX	213 528
Hapmap23509-BTC-073113	**27 198 715**	**0.46**	**–31.20**	**1.36E-09**	**13.79**	**–17.80**	**4.32E-06**	**2.49**	TOX	256 989
Hapmap23060-BTC-072978	27 360 366	0.40	29.11	4.22E-08	11.54	21.53	1.95E-07	3.53	CA8	277 183
Hapmap26621-BTC-072953	27 380 992	0.39	–28.75	8.60E-08	11.16	–20.56	6.53E-07	3.19	CA8	256 557
BTB-00560182	28 247 205	0.37	24.69	4.87E-06	8.08	17.28	3.96E-05	2.21	CHD7	75 924
BTB-02008412	28 300 924	0.43	–22.06	1.96E-05	6.79	–16.29	7.73E-05	2.07	CHD7	129 643
BTB-00561430	28 443 928	0.38	–28.39	3.36E-07	10.85	–16.35	7.13E-05	2.01	RLBP1 L1	94 261

The nearest annotated genes and their distance to the reported SNPs were determined using the UMD3.0 bovine genome assembly (Tables 4, 5). However, to consider only the genes nearest to each SNP could be misleading. It is preferable to consider all genes mapped to the association regions. In total, 23 genes mapped to the association region in Chromosome 14, including *ATP6V1H*, *CA8*, *CHD7*, *CSF2RA*, *FAM110B*, *IMPAD1*, *NPBWR1*, *NSMAF*, *OPRK1*, *PCMTD1*, *PENK*, *POLR2K*, *RAB2A*, *RB1CC1*, *RGS20*, *RLBP1 L1*, *RP1*, *SDR16C6*, *SNTG1*, *SOX17*, *TGS1*, *TOX* and *XKR4*. Also, 46 genes mapped to the association region in Chromosome X, including *ABCB7*, *AKAP4*, *APOL*, *AR*, *ARR3*, *BRWD3*, *CACNA1F*, *CCDC120*, *CHIC1*, *CHM*, *CLCN5*, *CXCR3*, *CYLC1*, *DGAT2 L3*, *DGAT2 L6*, *EDA*, *EFNB1*, *FGF16*, *FOXP3*, *GPKOW*, *H11CXORF26*, *KIF4A*, *LRCH2*, *MGC140080*, *MGC152340*, *MIR374A*, *MIR374B*, *NHSL2*, *PAGE4*, *PJA1*, *PLS3*, *POU3F4*, *RBM10*, *SLC9A7*, *SNX12*, *STARD8*, *SUV39H1*, *TAF1*, *TAF9B*, *TIMM17B*, *TMEM28*, *UPRT*, *USP11*, *USP27X*, *ZDHHC15* and *ZNF182*.

Manhattan plots for the calibration and validation datasets are provided in the Supplementary Material Fig. S1. There were 111 SNPs that were associated ($P < 0.01$) with AGE26 in both calibration and validation datasets (Supplementary Material Table S1). There were 98 SNPs that were associated ($P < 0.01$) with AGECL in both calibration and validation datasets (Supplementary Material Table S2). Chromosome X harboured more SNPs associated with AGE26, in both calibration and validation sets, than did any other chromosome (49%, or 55 of 111 SNPs), followed by BTA 14 (20%, or 23 of 111 SNPs). Most (58%, or 57 of 98 SNPs) of the SNPs associated with AGECL in both calibration and validation sets were located at BTA 14. As a result of our calibration and validation exercise, the same regions and candidate genes as observed in the GWAS with the full dataset remained relevant.

Discussion

Heritability estimates of cattle puberty in the literature are variable and influenced by population, breed, management and

Table 5. Effect, *P*-values and proportion of the variance explained for single-nucleotide polymorphisms (SNPs) underlying the association peak in the X chromosome observed for age at the scrotal circumference of 26 cm (AGE26)

SNP positions and distance to nearest gene are reported in base pairs and based on the UMD3 assembly of the bovine genome sequence. Most significant SNP in terms of *P*-values are shown in bold. Note that all these SNPs were associated ($P < 0.01$) in both calibration and validation datasets for AGE26, with one exception, Hapmap32639-BTA-158837 (see Supplementary Material Tables S1, S2). MAF, minor allele frequency

SNP	Chromosome X Position	MAF	AGECL Effect	AGECL *P*	AGECL %Var.	AGE26 Effect	AGE26 *P*	AGE26 %Var.	Gene	Distance
Hapmap32855-BTA-155265	69 196 959	0.18	0.65	9.32E-01	0.00	19.85	4.96E-09	1.84	LOC518911	136 499
Hapmap27896-BTA-26538	70 057 309	0.18	0.65	9.32E-01	0.00	19.85	4.96E-09	1.84	CYLC1	109 833
Hapmap32639-BTA-158837	70 234 032	0.18	2.32	7.65E-01	0.04	19.88	5.92E-09	1.82	LOC782514	122 078
Hapmap24863-BTA-154164	71 950 688	0.18	0.65	9.32E-01	0.00	20.23	3.18E-09	1.90	PLS3	144 154
Hapmap24353-BTA-19502	72 447 651	0.18	–0.65	9.32E-01	0.00	–20.23	3.18E-09	1.90	LOC784663	268 945
Hapmap31832-BTA-160852	73 379 138	0.18	0.57	9.40E-01	0.00	20.14	3.31E-09	1.88	POU3F4	153 386
Hapmap25915-BTA-161991	73 483 564	0.18	1.08	8.89E-01	0.01	20.14	3.31E-09	1.87	POU3F4	257 812
Hapmap26716-BTA-150917	75 234 781	0.18	0.65	9.32E-01	0.00	19.85	4.96E-09	1.84	APOL	0
Hapmap32363-BTA-155726	75 506 946	0.18	–0.65	9.32E-01	0.00	–19.85	4.96E-09	1.84	APOL	241 844
Hapmap31156-BTA-153166	77 369 290	0.18	0.57	9.40E-01	0.00	20.14	3.31E-09	1.88	LOC523251	599 379
Hapmap25932-BTA-30257	81 477 451	0.18	4.97	5.21E-01	0.21	21.17	3.69E-10	2.08	LOC512493	0
Hapmap60788-rs29017234	82 408 727	0.15	10.70	1.88E-01	0.85	22.25	7.91E-10	2.05	CHIC1	6068
Hapmap49114-BTA-30310	84 335 452	0.15	6.76	4.18E-01	0.34	–22.72	2.61E-10	2.13	CXCR3	19 873
Hapmap24012-BTA-158546	84 867 168	0.18	–4.37	5.52E-01	0.16	23.83	2.22E-12	2.66	SNX12	6393
Hapmap32396-BTA-164240	85 334 862	0.18	–3.54	6.39E-01	0.10	24.12	1.12E-12	2.69	KIF4A	0
ARS-BFGL-NGS-98019	85 463 982	0.17	–3.01	6.96E-01	0.07	24.03	2.24E-12	2.61	ARR3	0
BTA-30242-no-rs	**85 495 447**	**0.17**	**-5.51**	**4.67E-01**	**0.25**	**24.53**	**5.35E-13**	**2.76**	DGAT2 L3	4439
Hapmap49950-BTA-30244	85 589 749	0.13	–13.33	1.09E-01	1.13	22.56	4.15E-09	1.81	DGAT2 L6	12 013
Hapmap33084-BTA-164247	**85 833 472**	**0.17**	**-4.61**	**5.43E-01**	**0.17**	**24.51**	**6.76E-13**	**2.74**	EDA	0
Hapmap52858-rs29015876	86 556 922	0.14	–12.47	1.43E-01	1.07	24.45	4.22E-11	2.30	PJA1	47 130
Hapmap39370-BTA-108497	91 270 661	0.10	–16.49	1.10E-01	1.35	30.30	2.32E-11	2.55	LOC100141016	0
Hapmap47906-BTA-29386	93 631 688	0.14	–5.80	4.52E-01	0.23	21.70	2.71E-09	1.81	LOC784528	106 632

environmental factors, as well as the use of different phenotypic measurements (Martin *et al*. 1992; Cammack *et al*. 2009). Heritability for age at puberty in the Beef CRC Brahman heifers (AGECL, $h^2 = 0.57$) has been reported previously (Johnston *et al*. 2009); we confirmed these results by estimating a very similar value ($h^2 = 0.56$). These heritabilities are within the range, from 0.20 to 0.67, of other estimates for age at puberty (Arije and Wiltbank 1971; Smith *et al*. 1976; Laster *et al*. 1979). Reported heritability estimates for scrotal circumference range from 0.39 to 0.75 (Lunstra *et al*. 1988; Martinez-Velazquez *et al*. 2003; Corbet *et al*. 2009). Our result for puberty in bulls (AGE26) was closer to the higher end of that range ($h^2 = 0.78$). Overall and in agreement with our results, the heritability of puberty in cattle ranges from moderate to high, allowing for improvement through genetic selection.

Pubertal traits measured in heifers (AGECL) and bulls (AGE26) had a positive and favourable genetic correlation. This result confirms previous reports that showed that bigger scrotal circumference in bulls was correlated with early puberty in their female relatives (Martin *et al*. 1992; Burns *et al*. 2011). Previous evidence and our results point to the practicality of selecting for early pubertal cattle because both heifers and bulls can be selected with beneficial correlated effects.

Considering the genetic correlation, genes and SNPs associated with puberty in heifers were likely to be relevant for puberty in bulls, and *vice versa*. On BTA14, a large number of SNPs associated with puberty were identified in both bulls and heifers. The SNPs on BTA14 are located close to *RP1*, *XKR4* and *TOX*. These genes are not known to affect puberty in any species. However, the region on BTA14 includes another 20 annotated genes and presents homology to Human chromosome HSA8q12, which is associated with height (Gudbjartsson *et al*. 2008). Among those 20 annotated genes, *PENK*, *RPS20*, *SNORD54* and *PLAG1* are plausible functional candidates. *PENK* has a role in *GnRH* regulation (Rosie *et al*. 1992; Taylor *et al*. 2007). *RPS20* and a small RNA (*SNORD54*) were recently associated with calving ease (Pausch *et al*. 2011). The alleles that lower calving-ease have a positive effect on growth traits (Pausch *et al*. 2011). Calving ease is a trait influenced by the size of the calf and so, similarly to stature, it reflects frame size. Cattle with smaller frame size achieve puberty earlier than those with larger frame size (Vargas *et al*. 1999). Taken together, this evidence could point to *RPS20* and *SNORD54* as pleiotropic genes affecting puberty and calving-ease mainly through their effect on growth. Another recent study argued that *PLAG1* was the relevant gene underlying this region on BTA14 and affecting bovine stature (Karim *et al*. 2011). Considering the physiological link between growth and puberty, it is possible to hypothesise that the association region in BTA14 may relate to a gene with many pleiotropic effects or a functional polymorphism that affects more than one gene. There is already evidence for a functional SNP in this region affecting the expression levels of multiple genes (Karim *et al*. 2011).

The two most significant SNPs for AGE26 were located on the X chromosome and not on BTA14. The genes nearest to the most significant SNPs, *EDA* and *DGAT2 L3*, do not appear to

play a role in reproduction, given the current evidence from literature. A formal discussion of the functional link between these genes and cattle puberty is beyond our present objective. However, the association peak on Chromosome X does include the following three genes that are positional and functional candidate genes: *AR* (androgen receptor), *TAF1* and *TAF9B* (both TATA box binding protein (TBP)-associated factors). These genes are candidate genes for several reasons. The androgen receptor is known to play a role in sexual development, specifically, it affects spermatogenesis, testis localisation and testis size in mice models (Verhoeven *et al*. 2010). In pigs, the *AR* is considered a candidate gene for reproduction and performance traits (Trakooljul *et al*. 2004). Manipulation of fetal androgen exposure alters the timing of puberty in sheep (Jackson *et al*. 2008). The genes *TAF1* and *TAF9B* encode transcription factors that form the TFIID complex, a regulator of cell cycle and differentiation. Testis-specific *TAF* proteins have been reported as relevant for spermatid differentiation in *Drosophila* (Hiller *et al*. 2004). Taken together, this evidence and our results support *AR*, *TAF1* and *TAF9B* as candidate genes for scrotal development and puberty in Brahman bulls.

Results of the present study are consistent with the hypothesis that the genes underlying the associated regions on BTA14 and Chromosome X play significant roles on defining age at puberty in Brahman cattle. Candidate genes mapping to these regions are the subject of ongoing research. Information about specific genes and markers will add value to genomic selection. Therefore, identifying these associated regions contributing to genetic variation in AGECL and AGE26 should assist in the selection for early onset of puberty in Brahman cattle.

Acknowledgements

Data reported herein were generated as part of a larger study performed by the Cooperative Research Centre for Beef Genetic Technologies. Meat and Livestock Australia, Department of Employment, Economic Development and Innovation (DEEDI, Queensland government), CSIRO Livestock Industries, Australian Centre for International Agricultural Research, Northern Pastoral Group and University of Queensland supported the project. We gratefully acknowledge those who measured traits and managed cattle or DNA samples, including Richard Holroyd, Matthew Lee Wolcott, Yuandan Zhang, Mick Sullivan, Tim Grant, Brian Burns, Neil Cooper, Peggy Olsson, Debrah Corbet, Bronwyn Venus, Tracy Longhurst, Paul Williams, Warren Sim, Rob Young, Laercio R. Porto Neto, Rowan Bunch, Russell McCulloch, Blair Harrison, Eliza Collis, Mike Goddard, Brett Mason, Keith Savin and Kathryn Guthridge.

References

Arije GF, Wiltbank JN (1971) Age and weight at puberty in Hereford heifers. *Journal of Animal Science* **33**(2), 401–406.

Bagu ET, Madgwick S, Duggavathi R, Bartlewski PM, Barrett DM, Huchkowsky S, Cook SJ, Rawlings NC (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**(5), 861–873. doi:10.1016/j.theriogenology.2003.12.021

Bolormaa S, Hayes BJ, Savin K, Hawken R, Barendse W, Arthur PF, Herd RM, Goddard ME (2011) Genome-wide association studies for feedlot and growth traits in cattle. *Journal of Animal Science* **89**(6), 1684–1697. doi:10.2527/jas.2010-3079

Browning SR, Browning BL (2010) High-resolution detection of identity by descent in unrelated individuals. *American Journal of Human Genetics* **86**(4), 526–539. doi:10.1016/j.ajhg.2010.02.021

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**(3), 534–553. doi:10.1111/j.1439-0531.2011.01748.x

Cammack KM, Thomas MG, Enns RM (2009) Reproductive traits and their heritabilities in beef cattle. *The Professional Animal Scientist* **25**(5), 517–528.

Corbet NJ, Burns BM, Corbet DH, Johnston DJ, Crisp JM, McGowan M, Prayaga KC, Venus BK, Holroyd RG (2009) Genetic variation in groth, 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 NJ, Burns B, Corbet DH, Crisp JM, Johnston DJ, McGowan MR, Venus BK, Holroyd RG (2011) Bull traits measured early in life as indicators of herd fertility. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **19**, 55–58.

Fordyce G, Entwistle K, Norman S, Perry V, Gardiner B, Fordyce P (2006) Standardising bull breeding soundness evaluations and reporting in Australia. *Theriogenology* **66**(5), 1140–1148. doi:10.1016/j.theriogenology.2006.03.009

Foster DL (1994) Puberty in the sheep. In 'The Physiology of Reproduction'. 2nd edn. (Eds E Knobil, JD Neill). (Raven Press Ltd: New York)

Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, Helgadottir A, Ingason A, Steinthorsdottir V, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T, Pedersen O, Aben KK, Witjes JA, Swinkels DW, den Heijer M, Franke B, Verbeek ALM, Becker DM, Yanek LR, Becker LC, Tryggvadottir L, Rafnar T, Gulcher J, Kiemeney LA, Kong A, Thorsteinsdottir U, Stefansson K (2008) Many sequence variants affecting diversity of adult human height. *Nature Genetics* **40**(5), 609–615. doi:10.1038/ng.122

Hawken RJ, Zhang YD, Fortes MRS, Collis E, Barris WC, Corbet NJ, Williams PJ, Fordyce G, Holroyd RG, Walkley JRW, Barendse W, Johnston DJ, Prayaga KC, Tier B, Reverter A, Lehnert SA (2012) Genome-wide association studies of female reproduction in tropically adapted beef cattle. *Journal of Animal Science*, in press. doi:10.2527/jas.2011-4410

Hayes BJ, Pryce J, Chamberlain AJ, Bowman PJ, Goddard ME (2010) Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLOS Genetics* **6**(9), e1001139. doi:10.1371/journal.pgen.1001139

Hiller M, Chen X, Pringle MJ, Suchorolski M, Sancak Y, Viswanathan S, Bolival B, Lin TY, Marino S, Fuller MT (2004) Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. *Development* **131**(21), 5297–5308. doi:10.1242/dev.01314

Jackson LM, Timmer KM, Foster DL (2008) Sexual differentiation of the external genitalia and the timing of puberty in the presence of an antiandrogen in sheep. *Endocrinology* **149**(8), 4200–4208. doi:10.1210/en.2007-1382

Johnston DJ, Barwick SA, Corbet NJ, Fordyce G, Holroyd RG, Williams PJ, Burrow HM (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**(5–6), 399–412. doi:10.1071/EA08276

Johnston D, Barwick S, Fordyce G, Holroyd R (2010) Understanding the genetics of lactation anoestrus in Brahman beef cattle to enhance genetic evaluation of female reproductive traits. In '9th world congress on genetics applied to livestock production'. (German Society for Animal Science: Leipzig, Germany)

Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B, Harris BL, Keehan MD, Littlejohn MD, Spelman RJ, Georges M, Coppieters W (2011) Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. *Nature Genetics* **43**(5), 405–413. doi:10.1038/ng.814

Laster DB, Smith GM, Cundiff LV, Gregory KE (1979) Characterization of biological types of cattle (Cycle-II). 2. Post-weaning growth and puberty of heifers. *Journal of Animal Science* **48**(3), 500–508.

Lesmeister JL, Burfening PJ, Blackwell RL (1973) Date of first calving in beef cows and subsequent calf production. *Journal of Animal Science* **36**(1), 1–6.

Lopez R, Thomas MG, Hallford DM, Keisler DH, Silver GA, Obeidat BS, Garcia MD, Krehbiel CR (2006) Case study: metabolic hormone profiles and evaluation of associations of metabolic hormones with body fat and reproductive characteristics of Angus, Brangus and Brahman heifers. *The Professional Animal Scientist* **22**, 273–282.

Lunstra DD, Cundiff LV (2003) Growth and pubertal development in Brahman-, Boran-, Tuli-, Belgian Blue-, Hereford- and Angus-sired F1 bulls. *Journal of Animal Science* **81**(6), 1414–1426.

Lunstra DD, Ford JJ, Echternkamp SE (1978) Puberty in beef bulls: hormone concentrations, growth, testicular development, sperm production and sexual aggressiveness in bulls of different breeds. *Journal of Animal Science* **46**(4), 1054–1062.

Lunstra DD, Gregory KE, Cundiff LV (1988) Heritability estimates and adjustment factors for the effects of bull age and age of dam on yearling testicular size in breeds of bulls. *Theriogenology* **30**(1), 127–136. doi:10.1016/0093-691X(88)90270-1

Martin LC, Brinks JS, Bourdon RM, Cundiff LV (1992) Genetic effects on beef heifer puberty and subsequent reproduction. *Journal of Animal Science* **70**(12), 4008–4017.

Martinez-Velazquez G, Gregory KE, Bennett GL, Van Vleck LD (2003) Genetic relationships between scrotal circumference and female reproductive traits. *Journal of Animal Science* **81**(2), 395–401.

Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TPL, Sonstegard TS, Van Tassell CP (2009) Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* **4**(4), e5350. doi:10.1371/journal.pone.0005350

McGowan MR, Bertram JD, Fordyce G, Fitzpatrick LA, Miller RG, Jayawardhana GA, Doogan VJ, De Faveri J, Holroyd RG (2002) Bull selection and use in northern Australia. 1. Physical traits. *Animal Reproduction Science* **71**(1–2), 25–37. doi:10.1016/S0378-4320(02)00023-4

Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Gotz KU, Fries R (2011) Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. *Genetics* **187**(1), 289–297. doi:10.1534/genetics.110.124057

Perez-Enciso M, Misztal I (2004) Qxpak: a versatile mixed model application for genetical genomics and QTL analyses. *Bioinformatics* **20**(16), 2792–2798. doi:10.1093/bioinformatics/bth331

Perez-Enciso M, Misztal I (2011) Qxpak.5: Old mixed model solutions for new genomics problems. *BMC Bioinformatics* **12**(1), 202. doi:10.1186/1471-2105-12-202

Romano MA, Barnabe VH, Kastelic JP, de Oliveira CA, Romano RM (2007) Follicular dynamics in heifers during pre-pubertal and pubertal period kept under two levels of dietary energy intake. *Reproduction in Domestic Animals* **42**(6), 616–622. doi:10.1111/j.1439-0531.2006.00832.x

Rosie R, Thomson E, Blum M, Roberts JL, Fink G (1992) Oestrogen positive feedback reduces arcuate proopiomelanocortin messenger ribonucleic acid. *Journal of Neuroendocrinology* **4**(5), 625–630. doi:10.1111/j.1365-2826.1992.tb00212.x

Shamay A, Werner D, Moallem U, Barash H, Bruckental I (2005) Effect of nursing management and skeletal size at weaning on puberty, skeletal growth rate, and milk production during first lactation of dairy heifers. *Journal of Dairy Science* **88**(4), 1460–1469. doi:10.3168/jds.S0022-0302(05)72814-9

Siddiqui MA, Bhattacharjee J, Das ZC, Islam MM, Islam MA, Haque MA, Parrish JJ, Shamsuddin M (2008) Crossbred bull selection for bigger scrotum and shorter age at puberty with potentials for better quality semen. *Reproduction in Domestic Animals* **43**(1), 74–79.

Silva MR, Pedrosa VB, Silva JCB, Eler JP, Guimarães JD, Albuquerque LG (2011) Testicular traits as selection criteria for young Nellore bulls. *Journal of Animal Science* **89**(7), 2061–2067. doi:10.2527/jas.2010-3525

Smith GM, Fitzhugh HA, Cundiff LV, Cartwright TC, Gregory KE (1976) Genetic-analysis of maturing patterns in straightbred and crossbred Hereford, Angus and Shorthorn cattle. *Journal of Animal Science* **43**(2), 389–395.

Taylor JA, Goubillon ML, Broad KD, Robinson JE (2007) Steroid control of gonadotropin-releasing hormone secretion: associated changes in pro-opiomelanocortin and preproenkephalin messenger RNA expression in the ovine hypothalamus. *Biology of Reproduction* **76**(3), 524–531. doi:10.1095/biolreprod.106.055533

Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004) Polymorphisms of the porcine androgen receptor gene affecting its amino acid sequence and expression level. *Biochimica et Biophysica Acta (BBA) – Gene Structure and Expression* **1678**(2–3), 94–101. doi:10.1016/j.bbaexp.2004.02.004

Van Tassell CP, Smith TP, Matukumalli LK, Taylor JF, Schnabel RD, Lawley CT, Haudenschild CD, Moore SS, Warren WC, Sonstegard TS (2008) SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature Methods* **5**(3), 247–252. doi:10.1038/nmeth.1185

Vargas CA, Olson TA, Chase CC Jr, Hammond AC, Elzo MA (1999) Influence of frame size and body condition score on performance of Brahman cattle. *Journal of Animal Science* **77**(12), 3140–3149.

Verhoeven G, Willems A, Denolet E, Swinnen JV, De Gendt K (2010) Androgens and spermatogenesis: lessons from transgenic mouse models. *Philosophical Transactions of the Royal Society B Biological Sciences* **365**(1546), 1537–1556. doi:10.1098/rstb.2009.0117