Genetics and Molecular Research

Genome wide association study on early puberty in *Bos indicus*

A.V. Nascimento[1], M.C. Matos[2], L.O. Seno[3], A.R.S. Romero[1], J.F. Garcia[4] and A.B. Grisolia[1]

[1]Faculdade de Ciências Biológicas e Ambientais,
Universidade Federal da Grande Dourados, Dourados, MS, Brasil
[2]Centro Universitário da Grande Dourados, Dourados, MS, Brasil
[3]Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados, Dourados, MS, Brasil
[4]Departamento de Apoio, Saúde e Produção Animal,
Faculdade de Medicina Veterinária, Universidade Estadual Paulista,
Araçatuba, SP, Brasil

Corresponding author: A.B. Grisolia
E-mail: alexeiagrisolia@ufgd.edu.br

Genet. Mol. Res. 15 (1): gmr.15017548
Received August 31, 2015
Accepted November 13, 2015
Published February 5, 2016
DOI http://dx.doi.org/10.4238/gmr.15017548

ABSTRACT. The aim of this study was to evaluate a genome wide association study (GWAS) approach to identify single nucleotide polymorphisms (SNPs) associated with fertility traits (early puberty) in Nellore cattle (*Bos indicus*). Fifty-five Nellore cows were selected from a herd monitored for early puberty onset (positive pregnancy at 18 months of age). Extremes of this phenotype were selected; 30 and 25 individuals were pregnant and non-pregnant, respectively, at that age. DNA samples were genotyped using a high-density SNP chip (>777.000 SNP). GWAS using a case-control strategy highlighted a number of significant markers based on their proximity with the Bonferroni correction line. Results indicated that chromosomes 5, 6, 9, 10, and 22 were associated with the traits of interest.

The most significant SNPs on these chromosomes were rs133039577, rs110013280, rs134702839, rs109551605, and rs41639155. Candidate genes, as well as quantitative trait loci (QTL) previously reported in the Ensembl and Cattle QTLdb databases, were further investigated. Analysis of the regions close to the SNP on chromosomes 9 and 10 revealed that four QTL had been previously classified under the reproduction category. In conclusion, we have identified SNPs in close proximity to genes associated with reproductive traits. Moreover, U6 spliceosomal RNA was present on three different chromosomes, which is possibly associated with age at first calving, suggesting that it might be a strong candidate for future studies.

Key words: *Bos indicus*; GWAS; U6 spliceosomal RNA; Early puberty

INTRODUCTION

Indicine beef breeds (*Bos indicus*), and crosses with taurine (*Bos taurus*), are extensively raised in tropical regions. The late onset to puberty, which is generally observed in *B. indicus* when compared with *B. taurus*, directly affects profit in the cattle industry since it can postpone the start of the reproductive life of a cow for one year. Hence, the development of tools for the identification of more efficient cattle in this regard is of great commercial interest (Vaiciunas et al., 2008).

Among the production traits, age at first calving (AFC) stands out due to its association with reproductive performance. It is also an easily measurable phenotype and marks the entry of the heifer into the production system. A younger AFC is highly desirable as the cow is able to produce an increased number of calves during her reproductive life, which reduces the costs of heifer replacement, while allowing stronger selection intensity, and narrowing the generation intervals (Pirlo et al., 2000; Sasaki et al., 2013).

There is a scarcity of information originating from genome wide association studies (GWAS) for early puberty in *B. indicus*; therefore, this study aimed to associate single nucleotide polymorphisms (SNPs) associated with early puberty and to explore the genomic regions around them to identify prospective functional markers that influence reproductive traits in Nellore cattle.

MATERIAL AND METHODS

Animals and phenotypic data

Based on the records of AFC, 55 *B. indicus* (Nellore) heifers from a commercial farm located in Mato Grosso do Sul State, Brazil were selected and two groups were formed, namely pregnant (P), containing females who had positive pregnancy at 18 months (N = 30), and non-pregnant (NP) for those not pregnant at 36 months of age (despite being exposed to a bull in repeated breeding seasons) (N = 25). The phenotype was considered as a binary characteristic, where the P and NP assumed the values of 1 and 0, respectively.

Genotypic data

Samples were genotyped using the Illumina® BovineHD Genotyping BeadChip assay,

which contains probes to test more than 777,000 SNPs with an average 3.4-kb gap and a 2.68-kb median gap between markers. SNP genotyping and analysis was performed in a commercial laboratory (Deoxi Biotecnologia, Araçatuba, Brazil) as per the protocol recommended by the manufacturer.

Genomic data were subjected to quality control (QC), using the software R v. 3.0.2 (R Development Core Team, 2012) where all non-autosomal markers were removed. In order to evaluate possible duplicated samples, identity by state (IBS) analysis was performed and samples with IBS ≥95% were removed. The following criteria were used for excluding SNPs from the dataset: 1) minor allele frequency, markers with allelic frequencies lower than or equal to 2%; 2) Hardy-Weinberg equilibrium, SNPs with a P value $<10^{-5}$ in the Fisher exact test; and 3) call rate, polymorphism that is not found in at least 98% of the population. QC per individual was performed by call rate, and samples with less than 90% of determined genotypes were discarded.

Statistical analysis

The mmscore function (GenABEL package) was used as a score to test for association between traits and genetic polymorphisms. The residues were estimated after adjusting for a contemporary group. The analysis was conducted using the GenABEL v.1.8-0 package (Aulchenko et al., 2007) implemented in R.

Inflation factor λ was estimated to verify the reliability of the findings. The results are presented in a Manhattan plot with the cologarithm of P values, using Bonferroni correction for multiple tests. The α was defined as 0.05/N, with N being the number of tested markers (Benjamini and Hochberg, 1995).

Markers selected for candidate gene and quantitative traits loci (QTL) investigation considered the proximity between the Bonferroni correction line and the presence of other markers in the region. Association of genetic markers with phenotypes can occur indirectly, since a SNP or a group of SNPs may be in linkage disequilibrium with genomic regions that can cause phenotypic changes (Qanbari et al., 2010).

Exploration of candidate genomic regions

A 500-kb window was explored on each side of the chosen marker using the BioMart software (Kinsella et al., 2011) from the Ensembl database in order to identify genes based on the reference genome assembly UMD v3.1. The Cattle QTLdb database (Hu et al., 2013) was used to determine whether the most significant SNPs were inside QTL regions of bovine species already cataloged in the literature.

RESULTS AND DISCUSSION

Quality control of the original 786,798 SNPs provided a final list 473,792 markers for further analysis in the 55 individuals.

The Bonferroni test is often used in GWAS to correct for multiple tests and to set a significance index (Frommlet et al., 2012). However, increasing the number of markers and decreasing the population size can make the test restrictive. Using the Bonferroni test, none of the tested SNPs in this population reached the significance threshold determined after correction (α =

1.055 x 10^{-7}). Therefore, the markers closest to the established threshold were selected.

GWAS results depicted in the Manhattan plot (Figure 1) indicated the formation of diffuse peaks spread mainly on chromosomes 5, 6, 9, 10, and 22. Many regions presented an association with AFC, possibly because it is a polygenic trait (Hyeong et al., 2014).

The SNPs showing the most significant association were rs133039577, rs110013280, rs134702839, rs109551605, and rs41639155. We found a number of genes that are in close proximity to these SNPs and the outcome of this search is summarized in Table 1.

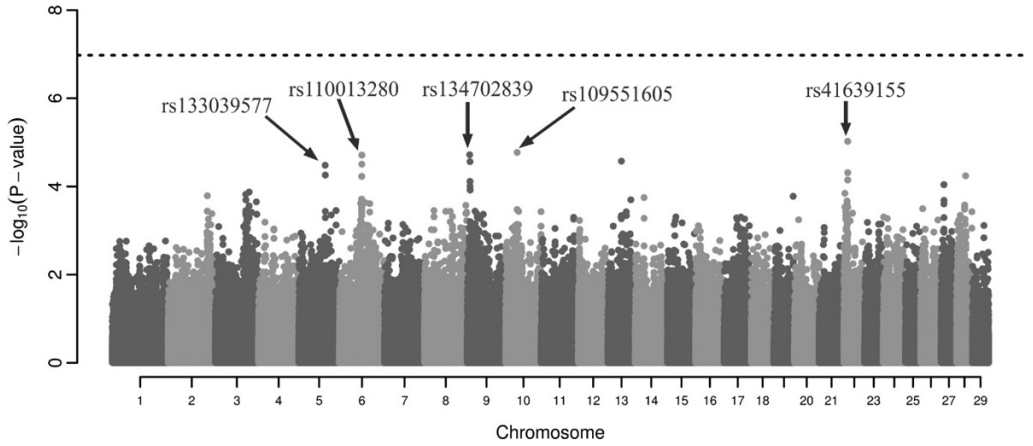

Figure 1. Manhattan plot of genome-wide-$\log_{10}$ (P values) for age at first calving in *Bos indicus* (Nellore) cattle. The horizontal line represents the Bonferroni significance threshold (α = 1.055 x 10^{-7}).

Table 1. Summary of candidate genes influencing reproductive traits.

Gene	SNP	Ensembl ID	Chromosome	Distance from SNP (kb)	Description
U6	rs133039577	ENSBTAG00000042848	5	66.5	U6 spliceosomal RNA
RELL1	rs110013280	ENSBTAG00000004329	6	364.1	*Bos taurus* RELT-like 1 (RELL1), mRNA
LMBRD1	rs134702839	ENSBTAG00000016164	9	375.1	*Bos taurus* LMBR1 domain containing 1 (LMBRD1), mRNA
U6	rs109551605	ENSBTAG00000042361	10	330.7	U6 spliceosomal RNA
LRRFIP2	rs41639155	ENSBTAG00000016760	22	0	Leucine rich repeat (in FLII) interacting protein 2
U6	rs41639155	ENSBTAG00000042942	22	289.9	U6 spliceosomal RNA

The *LRRFIP2* gene on chromosome 6 encodes a protein named Flightless I (FLiI), which is involved in cell mobility, contraction, and adherence. Mutations in this gene are either embryonically lethal or lead to the occurrence of degenerated embryos (Kopecki and Cowin, 2008). More importantly, regarding reproductive aspects, FLiI leads to hormonal co-activation of estrogen, which is fundamental for placental development during pregnancy (Seward et al., 2008).

The gene *RELL1* is located on chromosome 6, near SNP rs110013280 and is expressed at high levels in the placenta and at low levels in the ovary of mammals (Cusick et al., 2006). Experiments in pigs have indicated its association with number of stillbirths, which is an important indicator of reproductive performance in pigs (Onteru et al., 2011).

The protein LMBD1, which is encoded by the gene *LMBRD1*, is involved in the transport and metabolism of cobalamin (B_{12} vitamin). A mutation in its sequence can lead to the accumulation of vitamin B12 in lysosomes, preventing their conversion into co-factors required for metabolic reactions (Gailus et al., 2010; Rutsch et al., 2011). The correlation between vitamin B12 and

reproductive performance in mammals has been demonstrated in several studies. For example, in rats, the deficiency of vitamin B12 has been linked to testicular atrophic changes and defects in spermatogenesis (Watanabe et al., 2007).

The U6 spliceosomal RNA gene is found in the cell nucleus and is a non-coding RNA that forms the main spliceosomal complex. Functional loss of the coilin protein, which is involved in U6 biogenesis, is associated with fertility problems in rats (Xu et al., 2005; Lemm et al., 2006). The U6 spliceosomal RNA gene was identified on chromosomes 5, 10, and 22, possibly indicating broad involvement with the reproductive trait under study. Four QTL classified under the reproduction category were found on chromosomes 9 and 10 in regions near to the most significant SNP: calving ease (mother), inseminations per conception, fertility index, and fertilization rate.

In conclusion, the data generated in this study provide insight into the molecular mechanisms that regulate sexual precocity in *Bos indicus* (Nellore) cattle. We have reported specific regions on chromosomes 5, 6, 9, 10, and 22 that may be involved in the early onset of puberty. Among the most significant SNPs, proximity with U6 spliceosomal RNA was found in three out of the five explored chromosomes suggesting its potential role in the early onset of puberty in cattle. This preliminary study clearly demonstrates the feasibility of using a GWAS approach for the study of candidate genes and/or QTL associated with fertility in cattle. Nevertheless, more studies using larger populations are needed to validate the present findings and to better reveal how genes interact in the manifestation of the early onset of puberty phenotype.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors would like to thank Agropecuária Jorge Ferreira for assistance in sample acquisition. This research was supported by: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

Aulchenko YS, Ripke S, Isaacs A and van Duijn CM (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23: 1294-1296. http://dx.doi.org/10.1093/bioinformatics/btm108

Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289-300.

Cusick JK, Xu LG, Bin LH, Han KJ, et al. (2006). Identification of RELT homologues that associate with RELT and are phosphorylated by OSR1. *Biochem. Biophys. Res. Commun.* 340: 535-543. http://dx.doi.org/10.1016/j.bbrc.2005.12.033

Frommlet F, Ruhaltinger F, Twaróg P and Bogdan M (2012). Modified versions of Bayesian Information Criterion for genome-wide association studies. *Comput. Stat. Data Anal.* 56: 1038-1051. http://dx.doi.org/10.1016/j.csda.2011.05.005

Gailus S, Suormala T, Malerczyk-Aktas AG, Toliat MR, et al. (2010). A novel mutation in LMBRD1 causes the cblF defect of vitamin B(12) metabolism in a Turkish patient. *J. Inherit. Metab. Dis.* 33: 17-24. http://dx.doi.org/10.1007/s10545-009-9032-7

Hu ZL, Park CA, Wu XL and Reecy JM (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res.* 41: D871-D879. http://dx.doi.org/10.1093/nar/gks1150

Hyeong KE, Iqbal A and Kim JJ (2014). A genome wide association study on age at first calving using high density single

nucleotide polymorphism chips in Hanwoo (*Bos taurus coreanae*). *Asian-australas. J. Anim. Sci.* 27: 1406-1410. http://dx.doi.org/10.5713/ajas.2014.14273

Kinsella RJ, Kähäri A, Haider S, Zamora J, et al. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)* 2011: bar030. http://dx.doi.org/10.1093/database/bar030

Kopecki Z and Cowin AJ (2008). Flightless I: an actin-remodelling protein and an important negative regulator of wound repair. *Int. J. Biochem. Cell Biol.* 40: 1415-1419. http://dx.doi.org/10.1016/j.biocel.2007.04.011

Lemm I, Girard C, Kuhn AN, Watkins NJ, et al. (2006). Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. *Mol. Biol. Cell* 17: 3221-3231. http://dx.doi.org/10.1091/mbc.E06-03-0247

Onteru SK, Fan B, Du ZQ, Garrick DJ, et al. (2012). A whole-genome association study for pig reproductive traits. *Anim. Genet.* 43: 18-26. http://dx.doi.org/10.1111/j.1365-2052.2011.02213.x

Pirlo G, Miglior F and Speroni M (2000). Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. *J. Dairy Sci.* 83: 603-608. http://dx.doi.org/10.3168/jds.S0022-0302(00)74919-8

Qanbari S, Pimentel EC, Tetens J, Thaller G, et al. (2010). The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* 41: 346-356.

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rutsch F, Gailus S, Suormala T and Fowler B (2011). LMBRD1: the gene for the cblF defect of vitamin B_{12} metabolism. *J. Inherit. Metab. Dis.* 34: 121-126. http://dx.doi.org/10.1007/s10545-010-9083-9

Sasaki S, Ibi T, Ikeda S and Sugimoto Y (2013). A genome-wide association study reveals a quantitative trait locus for age at first calving in delta/notch-like EGF repeat containing on chromosome 2 in Japanese Black cattle. *Anim. Genet.* 45: 285-287. 1-3.

Seward ME, Easley CA, 4th, McLeod JJ, Myers AL, et al. (2008). Flightless-I, a gelsolin family member and transcriptional regulator, preferentially binds directly to activated cytosolic CaMK-II. *FEBS Lett.* 582: 2489-2495. http://dx.doi.org/10.1016/j.febslet.2008.06.037

Vaiciunas A, Coutinho LL, Meirelles FV, Pires AV, et al. (2008). Leptin and hypothalamic gene expression in early-and late-maturing *Bos indicus* Nellore heifers. *Genet. Mol. Biol.* 31: 657-664. http://dx.doi.org/10.1590/S1415-47572008000400010

Watanabe T, Ebara S, Kimura S, Maeda K, et al. (2007). Maternal vitamin B12 deficiency affects spermatogenesis at the embryonic and immature stages in rats. *Congenit. Anom. (Kyoto)* 47: 9-15. http://dx.doi.org/10.1111/j.1741-4520.2006.00135.x

Xu H, Pillai RS, Azzouz TN, Shpargel KB, et al. (2005). The C-terminal domain of coilin interacts with Sm proteins and U snRNPs. *Chromosoma* 114: 155-166. http://dx.doi.org/10.1007/s00412-005-0003-y