FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
UNIVERSIDADE ESTADUAL PAULISTA
CÂMPUS DE JABOTICABAL

META-ANALYSIS AND CROSS-VALIDATION STUDIES OF QTLS ASSOCIATED WITH REPRODUCTIVE TRAITS IN TROPICAL BEEF CATTLE

MSc. Thaise Pinto de Melo
Zootecnista

2018

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
UNIVERSIDADE ESTADUAL PAULISTA
CÂMPUS DE JABOTICABAL

META-ANALYSIS AND CROSS-VALIDATION STUDIES OF QTLS ASSOCIATED WITH REPRODUCTIVE TRAITS IN TROPICAL BEEF CATTLE

Thaise Pinto de Melo

Advisor: Dr. Roberto Carvalheiro

Co-Advisors: Profa. Dra. Lucia Galvão Albuquerque

Profa. Dra. Marina Rufino Salinas Fortes

2018

M528m

Melo, Thaise Pinto de

Meta-analysis and cross validation studies of QTLs associated with reproductive traits in tropical beef cattle / Thaise Pinto de Melo. -- Jaboticabal, 2018

101 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientadora: Roberto Carvalheiro

Coorientadora: Lucia Galvão de Albuquerque

1. Brahman. 2. GWAS. 3. Nellore. 4. puberty. 5. QTL. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: META-ANALYSIS AND CROSS-VALIDATION STUDIES OF QTLs ASSOCIATED WITH REPRODUCTIVE TRAITS IN TROPICAL BEEF CATTLE

AUTORA: THAISE PINTO DE MELO
ORIENTADOR: ROBERTO CARVALHEIRO
COORIENTADORA: MARINA RUFINO SALINAS FORTES
COORIENTADORA: LUCIA GALVÃO DE ALBUQUERQUE

Aprovada como parte das exigências para obtenção do Título de Doutora em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:

Pesquisador Dr. ROBERTO CARVALHEIRO (Skype)
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Prof. Dr. RICARDO VIEIRA VENTURA
Departamento de Nutrição e Produção Animal/FMVZ-USP / Pirassununga/SP

Prof. Dr. YURI TANI UTSUNOMIYA
Departamento de Apoio, Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/Unesp

Prof. Dr. HENRIQUE NUNES DE OLIVEIRA
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Pós-doutorando GERARDO ALVES FERNANDES JÚNIOR
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Jaboticabal, 07 de dezembro de 2018

DADOS CURRICULARES DO AUTOR

Thaise Pinto de Melo – Nasceu em Natal, Rio Grande do Norte em 16 de Fevereiro de 1991, filha de Maria Necy de Melo e José Coelho Pinto. Iniciou sua graduação em Zootecnia na Universidade Federal do Rio Grande do Norte (UFRN) em Fevereiro de 2008 e finalizou o curso em Dezembro de 2012. Durante o período de graduação a autora obteve duas bolsas, uma de extensão (PROEX), na qual trabalhou com palestras voltadas para crianças e adolescentes, sob o tema educação ambiental, e uma bolsa de pesquisa (PROPESQ), na qual trabalhou como aluna de iniciação cientifica em um projeto na área do melhoramento genético animal com dados simulados, sob a orientação da Professora Dra Elizângela Emídio Cunha. Em Março de 2013 ingressou no curso de mestrado da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de Jaboticabal, no programa de Genética e Melhoramento Animal sob a orientação do pesquisador Dr. Roberto Carvalheiro, inicialmente a autora tinha bolsa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), e posteriormente obteve bolsa Fapesp. No mestrado trabalhou com "Genome-Wide Association Studies (GWAS)" em características reprodutivas de fêmeas Nelore. Obteve bolsa BEPE para realizar estágio sanduiche na University of Guelph, Canadá, sob supervisão do Professor Dr. Flávio Schramm Schenkel no período de Agosto a Outubro de 2014. Concluiu o mestrado em Fevereiro de 2015 e Iniciou o curso de doutorado também na Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de Jaboticabal, no programa de Genética e Melhoramento Animal sob a orientação do pesquisador Dr. Roberto Carvalheiro em Março de 2015, com bolsa Capes durante todo o curso. De Junho a Novembro de 2017 realizou estágio na The University of Queensland, Austrália, sob a supervisão da Professora Dra Marina Fortes, com bolsa Capes PDSE (Processo número: 88881.133149/2016-01). O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Sendo eu, um aprendiz
A vida já me ensinou que **besta**
É quem vive triste
Lembrando o que faltou

Magoando a cicatriz
E esquece de ser feliz
Por tudo que conquistou

(...)
E nem sempre o que você perde
É de fato um prejuízo

(...)

Mas não desanime por nada
Pois até uma topada
Empurra você pra frente

Tantas vezes parece que é o fim
Mas no fundo, é só um recomeço
Afinal, pra poder se levantar
É preciso sofrer algum tropeço

(...)

Acredite no poder da palavra desistir
Tire o D, coloque o R
Que você tem **Resistir**

Uma pequena mudança
Às vezes traz esperança
E faz a gente seguir

Continue sendo forte
Tenha fé no Criador
Fé também em você mesmo
Não tenha medo da dor

Siga em frente a caminhada
E saiba que a cruz mais pesada
O filho de Deus carregou

Bráulio Bessa

Dedico

Aos meus pais, **Maria Necy** e **José** pelo apoio, confiança e ensinamentos. E ao meu noivo **Pedro**, pelo apoio, carinho e companhia.

AGRADECIMENTOS

A Deus primeiramente, que sempre esteve comigo, me guiando, protegendo e dando forças nos momentos difíceis.

À minha família, em especial aos meus pais, pelo carinho, apoio, incentivo, e pelas conversas, que mesmo à distância, me inspiram e encorajam.

Ao meu noivo Pedro, pela companhia, amor, amizade e conselhos. E sua família, que acabou sendo nos últimos anos a minha família, pelo apoio, companhia e bons momentos compartilhados.

À Deborah, pelos quatro anos de convivência, pela amizade, parceria, paciência, histórias de vida compartilhadas. Nos dias cansativos ela me recebia com bom humor, conversas que me distraiam e me faziam esquecer um pouco os problemas.

Aos meus amigos de salinha Andrés, Ana Cristina, Dani, Diogo, Lucas, Lucio, Malane, Natália, Patrícia, Samuel e Tiago pela convivência, discussões sobre os mais variados assuntos, por terem sido sempre solícitos e dispostos a me ajudar quando eu precisei, e é claro, pelas risadas, piadas e debates no café, RU ou nos churrascos. Em especial quero agradecer à Dani, pela amizade, carinho, paciência e cuidado, pelos inúmeros festejos, alegrias que passamos juntas, por ter me ouvido, me aconselhado em momentos difíceis, e por ter me ensinado o poder do abraço.

Aos irmãos Carvalheiro Angel, Baltasar, Daiane, Elzi, Giovana, Grazyela, Henrique e Laiza pela convivência, pelas várias comemorações, bolos, presentes e a “surpresa” no dia do nosso aniversário ou despedida, pelas conversas sérias e descontraídas que tivemos. Em especial, quero agradecer à Giovana, pela amizade desde o mestrado, pelas várias conversas, e por ter me emprestado um pouquinho da família dela quando a minha estava tão distante.

À Vanessa, pela amizade e por tornar os dias pesados um pouco mais leves, pelos inúmeros festejos juntas, pela confiança e carinho.

Ao meu orientador, Dr. Roberto Carvalheiro, por todos os ensinamentos, paciência e por ter investido seu valioso tempo no meu crescimento profissional desde o mestrado. Boa parte do conhecimento que eu adquiri até aqui devo a ele.

À Profa. Dra. Lucia Galvão de Albuquerque pela co-orientação, paciência e ensinamentos, também desde o mestrado.

À Profa. Dra. Marina Fortes, por ter aceito o convite de co-orientação, me co-orientado durante o período na Austrália, por toda ajuda, paciência e incentivo que me deu, num momento difícil para mim.

Às amizades que fiz na Austrália, em especial a Ana (Lupita), que foi minha parceira durante esse período, pelas inúmeras conversas, apoio, compreensão, carinho, companhia.

Aos membros da minha banca de qualificação, pelas sugestões que fizeram para a melhoria deste trabalho, a Profa. Dra. Lucia Galvão, o Prof. Dr. Henrique Nunes, que vem participando das minhas bancas e acompanhando minha evolução desde o mestrado, o Prof. Dr. Fernando Baldi, a Dra. Daiane Cristina Skalez, e o Dr. Gerardo Alves Junior, que desde a época da salinha divide comigo experiências, conhecimento e me ajudou bastante nessa etapa final do doutorado.

Aos membros que aceitaram o convite para compor a minha banca de doutorado e avaliar este trabalho, Prof. Dr. Henrique Nunes de Oliveira, Dr. Gerardo Alves Fernandes Junior, Prof. Dr. Ricardo Ventura e Prof. Dr. Yuri Tani Utsunomiya.

À Universidade Estadual Paulista Júlio de Mesquita Filho, Campus de Jaboticabal, pelo oferecimento do curso de pós-graduação em Genética e Melhoramento Animal.

À Capes pela concessão da bolsa de doutorado no país (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001) e no exterior (Processo número: 88881.133149/2016-01).

E a todos que de alguma forma contribuíram para o meu desenvolvimento profissional e pessoal, direta ou indiretamente, mesmo aqueles que não citei aqui.

O sucesso de uma pessoa não se constrói por ela mesma apenas. A crença em um ser superior tem grande contribuição nesse processo. Também por trás do nosso sucesso estão seres que nos amam e querem bem, alguns chamam de anjos, eu chamo de família e amigos. Que Deus abençoe grandemente a vida de cada um de vocês.

Muito Obrigada!

ACKNOWLEDGMENT

Thank you to Dr Marina Fortes research group at the University of Queensland, which during the Friday's meeting shared knowledge and delicious treats.

Thank you to the CSIRO researches, Dr Laercio Porto-Neto (Juca), Dr Antonio Reverter (Toni) and Dr Ben Hayes for the patience and for have given up valuable time of your full schedule to help me in my project of research and for share valuable knowledge.

Thank you to the University of Queensland for allowing my research internship, too important for conducting the analysis of my PhD.

Thank you so much!

CONTENTS

META-ANALYSIS AND CROSS-VALIDATION STUDIES OF QTLs ASSOCIATED WITH SEXUAL PRECOCITY TRAITS IN TROPICAL BEEF CATTLE

ABSTRACT - The statistical power of an individual genome-wide association study (GWAS) to detect quantitative trait loci (QTL) is expected to be lower than combining independent GWAS results from multi-breed, i.e., from different breeds, especially for complex traits, as puberty traits. Using independent populations and methods to detect new genomic regions and validate in independent populations known regions associated with the trait of interest is a feasible strategy to identify QTLs more accurately. The aim with this study was to detect genomic regions controlling sexual precocity traits across three tropical beef cattle breeds (Nellore, Brahman and Tropical Composite - TC) by using two approaches, meta-analysis using different breeds and QTL validation across different breeds. In the meta-analysis study the traits included were age at first calving (AFC), early pregnancy (EP), and scrotal circumference (SC_N) measured at 18 months of age for Nellore, and age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI), and scrotal circumference measured at 18 months of age ($SC18_B$) for Brahman cattle. The meta-analysis was performed using a multi-trait method. A total of 108 significant single-nucleotide polymorphisms (SNPs), at an empirical threshold *P*-value of 1.39×10^{-5} (FDR < 0.05), were found in this study. Those significant SNP were distributed over 19 out of 29 autosomes, and the major of them were located on BTA14. In the meta-analysis study we identified five association regions harbouring the majority of the significant SNP (76%), namely: BTA2 (5.55%) from 95 to 96 Mb, BTA4 (5.55%) from 94.1 to 94.8 Mb, BTA14 (59.26%) from 24 to 25 Mb and 29 to 30 Mb, and BTA21 (5.55%) from 6.7 Mb to 11.4 Mb. In the across-breed validation study we used the same traits of the meta-analysis study, except AFC, and we also used the traits ability to ovulate prior to weaning the calf (PW) and SC measured at 12 and 24 months of age (SC_{12}, SC_{24}), evaluated in Brahman. We considered Nellore as validation and Brahman as discovery population and used two SNP sets: 1) SNPs inside genes previously associated with sexual precocity in Brahman cattle, and SNPs within ± 250 Kb (upstream and downstream) neighbouring regions of those genes (G1); and 2) significant SNPs from meta-analyses studies for female and male

sexual precocity traits in Brahman and their ± 250 Kb (upstream and downstream) neighbouring regions (G2). This procedure was performed in order to validate in Nellore some genomic regions that are known as affecting sexual precocity traits in Brahman, and to find new QTL affecting those traits in both breeds. For EP 144 SNP were validated in G1 and 41 SNP were validated in G2. For SC_N 14 SNP were validated in G1 and 2 SNP were validated in G2. A total of 21 candidate SNP were detected for EP, i.e., they were in common between G1 and G2 SNP sets. These SNP and their neighbour regions validated in males and females were close to QTL and genes associated with reproductive events in bovine or other species. Therefore, they are candidate regions affecting sexual precocity traits in Nellore and Brahman. Also, the results of the meta-analysis study were validated in a third breed, the TC. Those SNP that were significant in the meta-analysis under an empirical *P*-value of 1×10^{-4} and were also significant (*P*-value $\leq 1 \times 10^{-3}$) for each individual GWAS of TC traits, AGECL, PPAI and SC_{TC}, or were located ± 250 Kb (up-stream or down-stream) from those significant SNP of each GWAS were considered validated. In summary, a total of 49, 4 and 14 SNP were validated for AGECL, PPAI and SC, respectively. In general, our results indicated candidate genomic regions controlling sexual precocity in Nellore, Brahman and TC cattle, and these regions seem to harbour the same genes / QTLs controlling puberty across breeds, despite differences in the breeds' origin and in the environment and management conditions that those breeds are exposed. In summary we were able to validate and find new candidate regions that are likely controlling sexual precocity in three independent tropical beef cattle populations. This result allows us to conclude that the strategy of using multi-breed meta-analysis and across breed QTL validation is feasible to detect new QTL and to validate known genomic regions affecting the traits of interest.

Keywords: Brahman, GWAS, Nellore, puberty, QTL, SNP, validation.

ESTUDOS DE META-ANÁLISE E VALIDAÇÃO CRUZADA DE QTLS ASSOCIADOS COM CARACTERÍSTICAS REPRODUTIVAS EM BOVINOS DE CORTE DE RAÇAS TROPICAIS

RESUMO – O poder estatístico de um único estudo de associação genômica ampla (GWAS) para detectar loci de características quantitativas (QTL) é esperado ser menor do que combinando resultados de GWAS independentes de múltiplas raças, isto é, de diferentes raças, especialmente para características complexas, como as de puberdade. Utilizar populações independentes e métodos para detectar novas regiões genômicas e validar regiões previamente conhecidas como associadas com a característica de interesse em populações independentes é uma estratégia viável para identificar QTLs mais acuradamente. O objetivo deste estudo foi detectar regiões genômicas controlando características de precocidade sexual em três raças de gado de corte tropical (Nelore, Brahman e Composto Tropical - CT), utilizando duas estratégias, meta-analise utilizando diferentes raças e validação cruzada de QTL entre raças. No estudo de meta-análise as características incluídas foram idade ao primeiro parto (IPP), prenhez precoce (PP), e circunferência escrotal (CE_N) medida aos 18 meses de idade no Nelore, e a idade ao surgimento do primeiro corpo lúteo (IPCL), intervalo do primeiro anestro pós-parto (IPAPP) e circunferência escrotal medida aos 18 meses de idade ($CE18_B$) para o Brahman. A meta-análise foi realizada utilizando um método multi-característica. Um total de 108 polimorfismos de nucleotídeo único (SNP) significativos a um *P*-valor empírico de 1.39×10^{-5} (FDR < 0.05), foram encontrados neste estudo. Tais SNP significativos foram distribuídos em 19 entre 29 autossomos, e a maioria deles estava localizada no BTA14. No estudo de meta-analise foram identificadas cinco regiões abrigando a maioria dos SNP significativos (76%): BTA2 (5.55%) de 95 a 96 Mb, BTA4 (5.55%) de 94.1 a 94.8 Mb, BTA14 (59.26%) de 24 a 25 Mb e de 29 a 30 Mb, e BTA21 (5.55%) de 6.7 Mb a 11.4 Mb. No estudo de validação cruzada entre raças foram utilizadas as mesmas características do estudo de meta-analise, exceto a IPP, além das características habilidade de ovular antes do desmame do bezerro (AD) e CE medida aos 12 e 24 meses de idade (CE_{12}, CE_{24}), avaliadas no Brahman. O Nelore foi considerado a população de validação e o Brahman, a população de “Discovery”,

no qual foram utilizados dois conjuntos de SNP: 1) SNP em regiões gênicas previamente associadas com a precocidade sexual no Brahman, e SNP em regiões vizinhas a estes genes a ± 250 Kb de distância (para mais ou para menos) (G1); e 2) SNP significativos de estudos de meta-análise para características de precocidade sexual em machos e fêmeas Brahman e suas regiões vizinhas a ± 250 Kb de distância (para mais ou para menos) (G2). Este procedimento foi utilizado para validar no Nelore algumas regiões genômicas que são conhecidas por afetar características de precocidade sexual em Brahman, e para encontrar novos QTL afetando estas características em ambas as raças. Para PP 144 SNP foram validados em G1 e 41 SNP foram validados em G2. Para CE_N 14 SNP foram validados em G1 e 2 em G2.. Um total de 21 SNP candidatos foi detectado para EP, isto é, eles estiveram em ambos os conjuntos de SNP, G1 e G2. Estes SNP e suas regiões vizinhas validados em machos e fêmeas estiveram próximos a QTL ou genes associados com eventos reprodutivos em bovinos e outras espécies, e são, portanto regiões candidatas afetando características de precocidade sexual em Nelore e Brahman. Além disso, os resultados do estudo de meta-análise foram validados em uma terceira raça, o CT. Aqueles SNP que foram significativos na meta-analise sob um *P*-value empírico de 1×10^{-4} e também foram significativos (*P*-valor $\leq 1 \times 10^{-3}$) para cada GWAS individual das características do CT, lPCL, lPAPP e CE_{CT}, ou estiveram localizadas a ± 250 Kb (para mais ou para menos) dos SNP significativos de cada GWAS foram considerados validados. Em resumo, um total de 49, 4 e 14 SNP foram validados para lPCL, lPAPP e CE_{CT}, respectivamente. De modo geral, os resultados encontrados indicam regiões genômicas candidatas controlando a precocidade sexual em bovinos Nelore, Brahman e CT, e estas regiões parecem abrigar genes / QTLs que estão controlando a puberdade entre as raças estudadas aqui, apesar das diferenças na origem destas raças e nas condições ambientais e de manejo aos quais os animais estão expostos. Em resumo nós pudemos validar e encontrar novas regiões candidatas que estão provavelmente controlando a precocidade sexual em três populações de bovinos de corte tropical. Esse resultado nos permite concluir que a estratégia de utilizar meta-analise com múltiplas raças e validação de QTL entre raças é viável para detectar novos QTL e

validar regiões genômicas previamente associadas com as características de interesse.

Palavras-chave: Brahman, GWAS, Nelore, puberdade, QTL, SNP, validação.

CHAPTER 1 - GENERAL CONSIDERATIONS

1.1 Introduction

According to the Brazilian Beef Exporters Association (ABIEC) and Meat & Livestock Australia (MLA), Nellore and Brahman are the major breeds in the productive chain of beef cattle in Brazil and Australia, respectively. They are both Zebu (*Bos taurus indicus*), a subspecies of cattle well adapted to tropical regions.

Besides Brahman, another breed that has been raised in Northern Australia is the Tropical Composite (TC). This breed is approximately composed of 50% of *B. indicus* (Brahman) and 50% of a mixture of African Sanga and *Bos taurus* breeds, such as N'Dama, Senepol, Shorthorn, Hereford and Charolais (Barwick et al., 2009). Around the decades of 1950 and 1960, crossbreeding experiments showed that the use of African based composites could produce more profitable herds than a systematic rotational crossbreeding program previously conducted with Brahmans. For example, the Belmont breed, an African composite, produced an extra gross profit of $24 per head/year for animals raised on pasture and could reach a profit margin of $76 per head/year a larger profit than obtained with the traditional crossbreeding schemes (Belmont Breeders Society). This is an evidence of the competitiveness and importance of TC for the beef cattle production in Northern Australia.

Despite the importance of Zebu breeds or Zebu cross-breeds for the beef production of tropical countries, the late puberty of heifers and sires in these breeds is a bottleneck to produce profitable herds. Selecting early heifers and bulls to decrease the generation interval and increase fertility is a well-defined strategy in Zebu cattle (Nogueira, 2004; Johnston et al., 2010).

Some reproductive traits are indicators of sexual precocity in beef cattle. These traits measure how early males and females are able to mate and breed. The traits commonly used as indicator of sexual precocity are age at first calving (AFC), early pregnancy (EP), age at the first observed *corpus luteum* (AGECL), and scrotal circumference (SC) (Johnston et al., 2009; Costa et al., 2015; Irano et al. 2016). These traits are associated with puberty and their genetic improvement may reduce

the production cost (Núñez-Dominguez et al., 1991), generate a faster economic return and anticipate the age of cows and bulls in reproductive life (Perotto et al., 2006; Menegassi et al., 2011). Moreover, these traits present genetic correlation with other economically important traits (Jonhston et al., 2009; Menegassi et al., 2011, Claus et al., 2017).

Along with the discovery of single nucleotide polymorphism (SNP) markers covering the genomes of several species, Genomic Selection (GS) and Genome Wide Association Studies (GWAS) have become popular and have been largely applied to different livestock species. In cattle, GWAS have been performed for a variety of traits, including productive (Silva et al., 2017), reproductive (Fortes et al., 2012, Irano et al., 2016) and morphological (Hayes et al., 2010). These studies aimed at finding candidate genomic regions associated with the phenotype of interest and elucidating the genetic architecture of those traits.

Each individual GWAS is expected to be underpowered compared to multi-breed GWAS, namely GWAS results from different breeds (van den Berg et al., 2016), i.e. they do not present a sufficient statistic power to detect a large number of true QTL using just one breed than combining information from different breeds (Evangelou et al., 2007). Thus, combining results from independent GWAS through meta-analysis studies could improve QTL detection power and may provide more accurate GWAS results (Keightley et al., 1998).

Besides meta-analyses, cross validation studies also have been performed to confirm GWAS results in different populations. Cross validation is an important procedure before making decisions in genomic selection programs or before performing large investments to find causal mutations (Höglund et al., 2014).

1.2 Literature review

1.2.1 Brief history of QTL detection in animal breeding

Around the decades of 1920-1930 Wright (1931), Haldane (1932) and Fisher (1930) elucidated the mechanism behind the natural selection and the Mendelian Factors that produced the final phenotype of quantitative traits under selection. In this period Fisher (1919) raised the hypothesis that the Medelian factors could explain the

similarity between relatives. Those ideas constructed the current basis of scientific breeding in animals and plants (Weller, 2009).

In 1961 Neimann-Soressen and Robertson proposed the first attempt to detect QTL by using a half-sib design and the blood groups as genetic markers in a segregating dairy cattle population. The major limitation of those experiments was the necessity of using a large number of individuals to present a reasonable power of QTL detection, furthermore the number of polymorphic blood groups were quite limited. Despite the limitations in QTL detection by using this method, the authors contributed with analytical advances by proposing for the first time a χ^2 statistical test to detect QTL based on a squared sum of normal distributions and using ANOVA, and leading with the problem of multiple tests comparison (Weller, 2009). Then, in the next decade other statistical methods as the maximum likelihood (Jayakar 1970) were also used to find QTL. In 1984 Henderson proposed the mixed model equations and developed an algorithm to invert the numerator relationship matrix, a notable advance for the modern animal breeding.

After the blood groups other alternative molecular markers were discovered. As the isozymes, restriction fragment length polymorphism (RFLP), DNA microsatellites and the single nucleotide polymorphisms (SNP) (Weller, 2009). This last one was the most promissing, and the option used in nowadays, because of its high polymorphisms and its covering, being well distributed over all the genome for a number of species.

Besides the use of molecular markers, genetic maps became popular, and one of the first genetic map proposed was the map based on the probability of recombination (Morgan, 1910). In summary, this kind of map measure the probability that two loci physically close in a chromosome can be inhered together, without be separated by the recombination event during the meiosis crossover. Those markers that are sufficiently close to be inhered together are named linked markers. However, the weakness points of this kind of map is the difference of the recombination rate across the genome, and the recombination rate also range according with other factors, as sex (Weller, 2009). Furthermore the intervals of which the QTL were mapped were in general large, about 20 cM or more (Van Laere et al., 2003), hindering the identification of the underlying mutation (Goddard and Hayes, 2009).

An important concept that is the basis of the currently most used tool of QTL mapping is the linkage disequilibrium (LD). The LD measures how correlated (non-randomly) two alleles in two loci are in a population. The basic difference between linkage analysis and linkage disequilibrium is that in linkage analysis just the LD that occurs within families is considered, and in this case the LD is broken after few generations of recombination. While the LD occurs across the entire population and persists across a considerable number of generations (Hayes and Daetwyler, 2013), in other words, the LD is a linkage between markers in population level (Bush and Moore, 2012).

Along with the advent of high density SNP marker panels for several species QTL mapping studies were highly benefited by the high coverage of variants mutations across the genome. Thus, Genome-Wide Association (GWA) studies emerged as a valuable tool to investigate the genetic architecture of several species. GWAS are associations at population level between markers and QTL affecting traits of interest, which is substantiated in the concept of LD.

There are in order two types of genetic associations, the direct associations, in which genotyped SNP directly affects the phenotype, i.e., it is located in a Quantitative trait nucleotide (QTN), or the indirect association, in which the SNP is in a region in LD with the QTN. Because of these two possible associations, if a SNP is significant in a GWAS it cannot be concluded to be the causal variant, further studies should be performed to fine map the QTL position (Bush and Moore, 2012).

1.2.2 Meta-analysis

Meta-analysis studies were firstly proposed by Glass (1976) in psychotherapeutic studies, by using analytical approaches. Walling et al. (2000) performed a similar procedure to find QTL associated with growth and fatness traits in pigs. In this study these authors performed jointing analysis for discovering QTL across seven independent pig populations, by using an F-ratio statistic to determine the most significant QTL position of the model. Currently, meta-analyses are applied in several areas, especially when original data from different studies are not available to perform a unified analysis (Khatkar et al., 2004). GWAS results have been widely used in meta-analyses aiming to detect important genomic regions with more

accuracy. The majority of these studies are related with human diseases, which have combined a considerable number of GWAS results in a single meta-analysis (Evangelou et al., 2007; Pharoah et al., 2013).

The main advantage of meta-analysis is increasing the statistical power because of the increasing of the size of available information to be analysed. And the major disadvantage is the complexity of this tool, which involves complex statistical methods, which are sensitive to certain protocols. Deviations from critical conditions, which may be associated with the performing or interpretation steps of meta-analysis studies may produce biased and misleading results (Walker et al., 2008).

Several statistical methods may be applied in meta-analysis studies, as *P*-values, *Z*-scores, fixed and random effect models, Bayesian, multivariate approaches and more recently, multi-trait meta-analysis (Bolormaa et al., 2014). *P*-values and *Z*-scores methods were widely used in the 1980's. They consist in combining those statistics (*P*-values and *Z*-scores) from different studies. The most known *P*-value method is the Fisher's method, which assumes a chi-squared statistic with 2 degrees of freedom under the null hypothesis. *Z*-scores are similar to Fisher's method, however *Z*-scores takes into account the size and direction of the marker effect, differently of Fisher's method that considers just the size effect (Evangelou and Ioannidis, 2013).

Fixed effects meta-analyses is the most common method in GWAS meta-analyses studies, which assumes that the true effect of each allele is the same in each study. The fixed effect model most used is the weighted inverse variance (Cochran, 1954), which consists in weighting each study according to the inverse of squared standard error of each marker effect estimate (Evangelou et al., 2007). The weight (w_{ij}) for the i^{th} study (population) and the j^{th} SNP is calculated as a function of the j^{th} allele effect variance in the i^{th} population. Random effect models assume a variable effect around an average effect and incorporate the between-study variance heterogeneity component, which can be obtained by using different statistics as Cochan's Q and I^2 statistics.

There is no consensus about the best model when there is heterogeneity among studies, i.e., when there is variability between the evaluated effects among studies. Begun et al. (2012) concluded that in the presence of heterogeneity, fixed

effect models may be better to perform GWAS. This is because generally the number of independent studies in a meta-analysis is not large, producing a heterogeneity estimate with low accuracy of models with random effect, and the shape of heterogeneity generally is not well adjusted for random effect Gaussian models. However Evangelou and Ionnides (2013) considered that the fixed effect model may produce biased results.

Bolormaa et al. (2014) proposed a multi-trait meta-analysis method that showed satisfactory increased power to detect pleiotropic QTL in a beef cattle data set, composed of 9 different populations, and including 32 growth and reproductive traits. This method consists in a chi-squared statistic with *n* degrees of freedom, where *n* is the number of traits included in the model. It is calculated as a function of the SNP effects, their standard errors (combined in signed *t*-values) and a correlation matrix between each pair of traits. The null tested hypothesis is that the marker has a zero effect, i.e., does not affect any trait.

At least two studies are necessary to conduct a meta-analysis, though larger number is desirable (Hooijmans et al., 2014). The number of studies included in GWAS meta-analyses varies widely, and is higher in human than in livestock studies, due to existence of crescent human large data sets consortia. Evangelou et al. (2007) performed meta-analysis with GWAS results for Parkinson disease using three different data sets. Minozzi et al. (2012) used two Holstein cattle populations in a meta-analysis study to identify loci associated with bovine paratuberculosis. Pharoah et al. (2013) used data from 43 studies from Ovarian Cancer Association Consortium. Sahana et al., (2014) performed a meta-analysis study for clinical mastitis in three dairy cattle. Guo et al. (2015) used four populations in meta-analysis to detect significant loci for length of bones in pigs. More recently, Bouwman et al. (2018) performed a meta-analysis with 17 populations from the 1000 Bull Genome project to find genes associated with cattle stature.

In the literature, there is evidence that meta-analyses with GWAS results have improved the QTL detection ability. Pharoah et al. (2013) studied the susceptibility to ovarian cancer (EOC) in humans combining different information sources in meta-analyses studies. These authors identified three new significant loci associated with EOC and confirmed two significant loci previously reported. Bolormaa et al. (2014)

concluded that multi-trait meta-analyses approach improved power to map QTLs associated with stature, growth and reproductive traits in beef cattle. These authors raise the hypothesis that if a QTL can explain some proportion of variance for multiple traits, then multi-trait analyses could improve the power of QTL detection and the precision to map them.

1.2.3 Cross validation

There are several ways to perform a validation study and to determine the significance criteria of the SNP effects. Pryce et al. (2010) performed a cross validation study with GWAS results for milk production and fertility traits in two dairy breeds, by using two approaches, SNP-by-SNP (regression) and haplotype-by-haplotype tested in turn and in a Bayesian approach (testing all SNPs simultaneously) and determined empirical thresholds to find significant SNPs in both discovery (Holstein animals) and validation (Holstein and Jersey animals) populations. They detected SNP markers affecting milk production traits and found new promising regions. Kim et al. (2012) also performed a QTL validation study for growth and carcass traits in commercial populations of American (Michigan) and Korean (Hanwoo) cattle of two chromosomes with QTLs previously reported for those traits. They were able to confirm 18 public missense SNPs and to find 9 new SNPs. Chamberlain et al. (2012) performed a validation study of milk production traits in Holstein bulls, by using the candidate gene approach. They selected SNP in or close to candidate genes affecting those traits and were able to validate 72 of them in one or more traits. Höglund et al. (2014) performed a QTL validation study in two dairy cattle for fertility traits in females by a similar procedure used in Pryce et al. (2010). Briefly, if an association was significant in both discovery and validation population, which were independent of each other it was considered validated. They found some genomic regions in common between the two populations.

As previously demonstrated in literature, cross validation studies are efficient and a promising method to confirm QTL and to find new associations across breeds.

1.3 Objective

The objective of this study was to detect QTLs associated with sexual precocity traits in different breeds of tropical beef cattle by using two approaches: meta-analysis and validation studies.

1.4 References

ABIEC. **Rebanho Bovino Brasileiro**. Available in: <www.abiec.com.br/3_rebanho.asp>. Accessed in Oct. 2016.

Barwick SA, Wolcott ML, Johnston DJ, Burrow HM, Sullivan MT (2009) Genetics of steer daily and residual feed intake in two tropical beef genotypes, and relationships among intake, body composition, growth and other post-weaning measures. **Animal Production Science** 49:351-366.

Begun F, Ghosh D, Tseng GC, Feingold E (2012) Comprehensive literature review and statistical considerations for GWAS meta-analysis. **Nucleic Acids Research** 40:3777-3784.

Belmont Breeders Society. Available in: < http://belmontred.com.au/ >. Acessed in Nov. 2018.

van den Berg I, Boichard D, Lund MS (2016) Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. **Journal of dairy science** 99:8932-8945.

Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, Tier B, Savin K, Hayes BJ, Goddard ME (2014) A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for stature, Fatness and Reproduction in Beef Cattle. **PLOS Genetics** 10:e1004198.

Bouwman AC, Daetwyler HD, et al. (2018) Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. **Nature Genetics** 50:362-367.

Chamberlain AJ, Hayes BJ, Savin K, Bolormaa S, McPartlan HC, Bowman PJ, Van Der Jagt C, MacEachern S, Goddard ME (2012) Validation of single nucleotide polymorphisms associated with milk production traits in dairy cattle. **Journal of dairy science** 95:864-875.

Claus LAM, Koetz Júnior C, Roso VM, Borges MHF, Barcellos JOJ, Ribeiro ELDA (2017) Genetic parameters of age at first calving, weight gain, and visual scores in Nelore heifers. **Revista Brasileira de Zootecnia** 46:303-308.

Cochran WG (1954) The Combination of Estimates from Different Experiments. **Biometrics** 10:101-129.

Costa RB, Camargo GM, et al. (2015) Genome-wide association study of reproductive traits in nellore heifers using bayesian inference. **Genetics Selection Evolution** 47:67.

Evangelou E, Ioannidis JPA (2013) Meta-analysis methods for genome-wide association studies and beyond. **Nature Reviews Genetics**14:379-389.

Evangelou E, Maraganore DM, Ioannidis JPA (2007) Meta-Analysis in Genome-Wide Association Datasets: Strategies and Application in Parkinson Disease. **PLoS One** 2: e196.

Fisher RA (1919). XV — The correlation between relatives on the supposition of Mendelian inheritance. **Earth and Environmental Science Transactions of the Royal Society of Edinburgh** 52:399-433.

Fisher RA (1930) The Genetical Theory of Natural Selection. Oxford, England: Clarendon Press.

Fortes MRS, Lehnert SA, Bolormaa S, Reich C, Fordyce G, Corbet NJ, Whan V, Hawken RJ, Reverter A (2012) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science** 52:143-150.

Glass GV (1976) Primary, Secondary, and Meta-Analysis of Research. **Educational Researcher** 5:3-8.

Goddard ME, Hayes BJ (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics** 10:381-391.

Guo Y, Hou L, et al. 2015 A meta analysis of genome-wide association studies for limb bone lengths in four pig populations. **BMC genetics** 16:1-10.

Haldane, JBS (1932) The causes of Evolution. New York: Harper & Brothers.

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: e1001139.

Hayes B, and Daetwyler H (2013) **Genomic Selection in the Era of Genome Sequencing**. Guelph, Canada. 128p.

Henderson CR (1984) **Aplications of linear models in animal breeding**. University of Guelph, Guelph.

Höglund JK, Sahana G, Guldbrandtsen B, Lund MS (2014) Validation of associations for female fertility traits in Nordic Holstein, Nordic Red and Jersey dairy cattle. **BMC Genetics** 15:1-8.

Hooijmans CR, Inthout J, Ritskes-Hoitinga M, Rovers MM (2014) Meta-Analyses of Animal Studies: An Introduction of a Valuable Instrument to Further Improve Healthcare. **Institute for Laboratory Animal Research Journal** 55:418-426.

Irano N, De Camargo, et al. (2016) Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. **Plos One** 11:e0159502.

Jayakar SD (1970) On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus. **Biometrics** 26:451-464.

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** 49:399-412.

Johnston DJ, Barwick SA, Fordyce G, Holroyd RG (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. **Proceedings...** Leipzig, Germany.

Keightley PD, Morris KH, Ishikawa A, Falconer VM, Oliver F (1998) Test of candidate gene-quantitative trait locus association applied to fatness in mice. **Heredity** 81:630-637.

Khatkar MS, Thomson PC, Tammen I, Raadsma HW (2004) Quantitative trait loci mapping in dairy cattle: review and meta-analysis. **Genetics Selection Evolution** 36:163-190.

Kim KS, Kim SW, Raney NE, Ernst CW (2012) Evaluation of BTA1 and BTA5 QTL Regions for Growth and Carcass Traits in American and Korean Cattle. **Asian-Australian Journal Animal Science** 25:1521-1528.

Van Laere AS, Nguyen M, et al. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. **Nature** 425:832–836.

Meat & Livestock Australia Limited. (2011) **Developing a Dollar Index for Brahmans in the Northern Territory Live Export Trade**. Final Report. Sydney, Australia. Accessed in: April, 2018.

Menegassi SRO, Barcellos JOJ, Peripolli V, Pereira PRRX, Borges JBS, Lampert VDN (2011) Measurement of scrotal circumference in beef bulls in Rio Grande do Sul. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 63:87-93.

Minozzi G, Williams JL, Stella A, Strozzi F, Luini M, Settles ML, Taylor JF, Whitlock RH, Zanella R, Neibergs HL (2012) Meta-Analysis of Two Genome-Wide Association Studies of Bovine Paratuberculosis. **PLoS ONE** 7:e32578.

Morgan TH (1910) Chromosomes and Heredity. The American Naturalist 44:449-496.

Neimann-Soressen A, Robertson A (1961) The association between blood groups and several production characters in three Danish cattle breeds. **Acta Agriculturae Scandinavica** 11:163-196.

Nogueira GP (2004) Puberty in South American *Bos Indicus* (Zebu) cattle. **Animal Reproduction Science** 82:361-372.

Nunez-Dominguez R, Cundiff LV, Dickerson GE, Gregory KE, Koch RM (1991) Lifetime production of beef heifers calving first at two vs three years of age. **Journal of Animal Science** 69:3467-3479.

OECD. **OECD-FAO Agricultural Outlook 2011-2020**. Accessed in: Oct. 2016.

Perotto D, Miyagi AP, Souza JC, Moletta JL, Freitas JA (2006) Estudos de características reprodutivas de animais da raça Canchim, criados a pasto, no estado do Paraná, Brasil. **Archives of Veterinary Science** 11:1-6.

Pharoah PDP, Tsai YY, et al. (2013) GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. **Nature Genetics** 45:362-370.

Pryce JE, Bolormaa S, Chamberlain AJ, Bowman PJ, Savin K, Goddard ME, Hayes, BJ (2010) A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. **Journal of Dairy Science** 93:3331-3345.

Rodrigues HD, Kinder JE, Fitzpatrick LA (2002) Estradiol regulation of luteinizing hormone secretion in heifers of two breed types that reach puberty at different ages. **Biology of Reproduction** 66:603-609.

Sahana G, Guldbrandtsen B, Thomsen B, Holm LE, Panitz F, Brøndum RF, Bendixen C**,** Lund MS (2014) Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle1. **Journal of dairy science** 97:7258-7275.

Silva RMO, Stafuzza NB, et al. (2017) Genome-wide association study for carcass traits in an experimental Nelore cattle population. **PLOs One** 12:e0169860.

Walker E, Hernandez AV, Kattan MW (2008) Meta-analysis: Its strengths and limitations. **Cleveland Clinic journal of medicine** 75: 431.

Walling GA, Visscher PM, et al. (2000) Combined analyses of data from quantitative trait loci mapping studies: chromosome 4 effects on porcine growth and fatness. **Genetics** 155:1369-1378.

Weller JI. Quantitative trait loci analysis in animals. 2nd edition, CABI, 2009.

Wright S (1931) Evolution in Mendelian populations. **Genetics** 16:97-159.

CHAPTER 2 - MULTITRAIT META-ANALYSIS IDENTIFIED GENOMIC REGIONS ASSOCIATED WITH SEXUAL PRECOCITY IN TROPICAL BEEF CATTLE [a]

ABSTRACT - Multitrait meta-analyses are a strategy to produce more accurate genome-wide association studies, especially for complex phenotypes. We carried out a meta-analysis study for traits related to sexual precocity in tropical beef cattle (Nellore and Brahman) aiming to identify important genomic regions affecting these traits. The traits included in the analyses were age at first calving (AFC), early pregnancy (EP), age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI) and scrotal circumference (SC). The traits AFC, EP, and SC_N were measured in Nellore cattle, while AGECL, PPAI and SC_B were measured in Brahman cattle. Meta-analysis resulted in 108 significant single-nucleotide polymorphisms (SNPs), at an empirical threshold *P*-value of 1.39×10^{-5} (false discovery rate [FDR] < 0.05). Within 0.5 Mb of the significant SNP, candidate genes were annotated and analyzed for functional enrichment. Most of the closest genes to the SNP with higher significance in each chromosome have been associated with important roles in reproductive function. They are *TSC22D2*, *KLF7*, *ARHGAP29*, *7SK*, *MAP3K5*, *TLE3*, *WDR5*, *TAF3*, *TMEM68*, *PPP1R15B*, *NR2F2*, *GALR1*, *SUFU*, and *KCNU1*. We did not observe any significant SNP in BTA5, BTA12, BTA17, BTA18, BTA19, BTA20, BTA22, BTA23, BTA25, and BTA28. Although the majority of significant SNPs are in BTA14, it was identified significant associations in multiple chromosomes (19 out of 29 autosomes), which is consistent with the postulation that reproductive traits are complex polygenic phenotypes. Five proposed association regions harbor the majority of the significant SNP (76%) and were distributed over four chromosomes ($P < 1.39 \times 10^{-5}$, FDR < 0.05): BTA2 (5.55%) from 95 to 96 Mb, BTA4 (5.55%) from 94.1 to 94.8 Mb, BTA14 (59.26%) from 24 to 25 Mb and 29 to 30 Mb, and BTA21 (5.55%) from 6.7 Mb to 11.4 Mb. These regions harboured key genes related to reproductive function. Moreover, these genes were enriched for functional groups associated with immune response, maternal–fetal tolerance, pregnancy maintenance, embryo development, fertility, and response to stress. Further studies including other breeds and precocity traits could confirm the

[a] Article published in “Journal of Animal Science” doi: 10.1093/jas/sky289

importance of these regions and identify new candidate regions for sexual precocity in beef cattle.

Keywords: *Bos indicus*, Brahman, early puberty, GWAS, meta-analysis, Nellore

2.1 Introduction

Reproductive traits are important for the economic success of beef cattle production (Formigoni et al., 2005; Brumatti et al., 2011). Despite their economic importance, genetic gains of reproductive traits are generally slow. These traits are considered polygenic and highly influenced by environmental factors (Cardoso et al., 2015). Genomic selection has allowed higher genetic gains for fertility traits in beef and dairy cattle (Zhang et al., 2014; Garcia-Ruiz et al., 2016). Genome-wide association studies (GWAS) have detected quantitative trait loci (QTL) affecting reproductive traits aiming to better understand the genetic mechanisms controlling these traits (Hawken et al., 2012; Irano et al., 2016; Melo et al., 2017). However, these traits are controlled by many QTL of small effect and GWAS can produce high false-positive rates. Validation for the identified QTL requires further investigation. To improve QTL detection, GWAS meta-analyses corroborate evidence from independent studies of related traits by identifying polymorphisms that produce variation in common among complex traits (Bolormaa et al., 2014; Ramayo-Caldas et al., 2016). The aim of this study was to carry out meta-analysis across two independent *Bos indicus* populations to improve QTL detection for reproductive traits. Specifically, we aimed to: 1) identify genomic regions associated with sexual precocity traits measured relatively early in a cow's life and the related bull trait scrotal circumference (SC); 2) identify candidate genes in those genomic regions; and 3) investigate biological roles of the identified candidate genes. The traits measured in Nellore were age at first calving (AFC), early pregnancy (EP), and scrotal circumference (SC_N) and in Brahman were age at the first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI), and scrotal circumference (SC_B).

2.2 Material and Methods

2.2.1 Nellore Phenotypes and Genotypes

Information of animals from Alliance Nellore dataset, born in eight different farms distributed in Northwest, Southwest and Midwest of Brazil were used. Heifers were either artificially inseminated or naturally mated. Usually farmers apply two

breeding seasons, in the early breeding season females are exposed to the first mating at around 16 months of age and, the females that failed to conceive have a second chance in the later breeding season, at about 26 months of age. Females that did not conceive neither in the first or second trial were culled and were not considered in the analysis of AFC. More details about the dataset are described in Costa et al. (2015) and Irano et al. (2016).

Summary statistics of Nellore phenotypes—AFC, EP, and SC_N—are presented in Table 1. The phenotype AFC was defined as the number of days from birth to first calving; EP was defined as success (1) for heifers that had the first calf with less than 31 months of age, i.e., heifers that got pregnant in the first breeding season, or failure (0) for those heifers that had the first calf after 31 months of age. In both populations, SC was measured at about 18 months of age, at the widest point of the scrotum with a standard metal tape, in centimeters.

Table 1. Summary statistics of age at first calving (AFC), early pregnancy (EP) and scrotal circumference (SC_N) in Nellore cattle and age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI) and scrotal circumference (SC_B) in Brahman cattle.

Trait	Number of observations[1]	Mean ± SE	Minimum	Maximum
AFC (days)	1,796	1068.2 ± 118.5	778.9	1292.8
EP (%)	1,849	28.3 ± 1.1	--	--
SC_N (cm)[2]	4,248	26.57 ± 2.56	17.47	36.17
AGECL (days)	1,007	750.6 ± 142.14	394	1211
PPAI (days)	629	180.11 ± 108.71	17	484
SC_B (cm)[3]	1,203	26.51 ± 2.76	18.25	39.25

[1]Number of observations considered in analyses: animals with available genotypes and phenotypes.

[2]SC_N measured at approximately 521 days of age.

[3]SC_B measured at approximately 540 days of age.

The contemporary groups (CGs) for AFC, EP and SC_N were defined by concatenating the information of herd, year and season of birth, weaning, and yearling management groups. Groups with less than four animals for AFC and EP or less than three animals for SC_N were excluded. For AFC and SC_N, we considered

only data in the interval of CG ± 3 standard errors from the mean of each group. For EP CG without variability, i.e., presenting only females with the same categorical response were excluded. For SC the linear effect of the age of the animal at the recording was included as a covariate in the model. It was considered in the analysis just animals presenting age at recording from 10 to 24 months of age. It was excluded a total of 20 animals that presented ages out of this interval.

For Nellore traits, phenotypes were previously corrected for CG effects to avoid biased fixed effect estimates because the whole Nellore dataset contained both genotyped and non-genotyped animals. CG estimates used to pre-correct the phenotypes were obtained under a regular mixed animal model (Henderson, 1984), using the single step GBLUP method. Then, these corrected phenotypic traits for genotyped animals were used in GWAS.

Animals were genotyped with the high-density Illumina Bovine HD BeadChip Assay (Illumina, San Diego, CA) and GeneSeek Genomic Profiler Indicus HD—GGP75Ki (Neogen Corporation, Lincoln, NE), which contain 777,962 and 74,677 (73,941 in common with HD) SNP markers, respectively. Genotype imputation to Bovine HD was performed for animals genotyped with GGP75Ki using FImpute software (Sargolzaei et al., 2014), taking into account pedigree information. The imputation accuracy was expected to be higher than 0.98 (Carvalheiro et al., 2014). Quality control was performed after imputation, removing animals with call rate <0.90 and removing SNP with minor allele frequency (MAF) < 0.01, call rate < 0.95, GC score < 0.15, Hardy–Weinberg equilibrium test P-value < 10^{-5}, SNP in non-autosomal regions, and unmapped SNP. After quality control 2,923 females with genotypes for 412,993 SNP and 5,078 males with genotypes for 477,317 SNP remained for the analyses. To construct the genomic relationship matrix (G), SNP with MAF >0.1 were considered.

2.2.2 Brahman Phenotypes and Genotypes

Data from the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) were used in this study. The Beef CRC data for male and female Brahman cattle were previously described (Johnston et al., 2009; Burns et al., 2013; Corbet et al., 2013). The number of observations and descriptive statistics for each trait are presented in Table 1. The traits evaluated were AGECL, PPAI, and SC_B.

When the heifers achieved the weight of 200 kg, an ovarian ultrasound was carried out at every 4 to 6 weeks to detect the first *corpus luteum* (CL) and calculate AGECL. To get PPAI phenotypes, the number of days from calving to the first ovulation postpartum was observed. Ovarian ultrasound was used to observe CL presence after calving to indicate an ovulatory event.

For AGECL and PPAI, CGs were formed by concatenating information of cohort, year of birth, and management group. For SC_B, CG included the effects of cohort and year of birth. Age of young bull at recording was used as a covariate for SC_B.

Brahman animals were genotyped using the Illumina BovineSNP50 version 1 or 2, and posteriorly imputed for high-density panel using Beagle software (Browning and Browning, 2009). To allow imputation, representative animals of the Beef CRC population were genotyped using the HD chip from Illumina as described by Fortes et al. (2013). Quality control excluded animals with call rates < 98%, SNP with call rates < 85% and MAF < 0.02. After quality control remained 660,433 SNP for PPAI, 659,845 for AGECL and 612,992 for SC_B.

2.2.3 Genome-wide association methods and meta-analyses

A single-trait single-marker model was used to perform GWAS for all traits using the genomic mixed model proposed by Kang et al. (2010), which accounts for population substructure, fitted as follows for Nellore (1) and Brahman (2):

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{s}a + \mathbf{Zu} + \boldsymbol{\varepsilon}, \quad (1)$$
$$\mathbf{y} = \mathbf{X\beta} + \mathbf{s}a + \mathbf{Zu} + \boldsymbol{\varepsilon}, \quad (2)$$

where **y** is a vector containing the phenotypic information, corrected for the fixed effects in Nellore data (equation 1), 1 is a vector of ones, μ is an overall mean, ***X*** is an incidence matrix relating fixed effects (CG) in **β** with the phenotypes in **y**, **s** is the vector containing the genotypes coded as 0, 1, or 2 according to the number of B allele copies, a is the SNP additive genetic substitution effects, **Z** is the incidence matrix of polygenic random effects of the animals in ***u*** and ***ε*** is the vector of residuals. **u** and **ε** followed normal distribution with $\mathbf{u} \sim N(0, \mathbf{G}\sigma^2_a)$ and $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma^2_e)$, respectively, where **G** is the genomic relationship matrix for all individuals and SNP (except the SNP considered in a), calculated as described in the first method

proposed by VanRaden (2008), σ^2_a is the additive genetic variance, **I** is an identity matrix, and σ^2_e is the residual variance. The SNP effect estimates were computed using the SNP & Variation Suite v8.3.0 software (Golden Helix, Inc., Bozeman, MT) under the model previously described and the EMMAX method (Kang et al., 2010).

Multitrait meta-analyses were performed using the method described by Bolormaa et al. (2014), which were implemented in the R software (R Core Team, 2018). This method consists of a test statistic following a χ^2 distribution with *n* degrees of freedom, where *n* is the number of traits included in meta-analysis, calculated as:

$$Multi\text{-}trait\ \chi^2 = \mathbf{t'_i\ V^{-1}t_i}, \quad (3)$$

where $\mathbf{t_i}$ is a 6 × 1 vector of the i^{th} SNP effect estimates for the six traits divided by their respective standard errors, t'_i is the transpose vector of t_i, and $\mathbf{V^{-1}}$ is an inverse of the 6 × 6 correlation matrix of the correlations between the *t*-values of the six traits across the 387,971 SNP considered in the model. Just common SNP in both Nellore and Brahman panels were considered in the analysis. False discovery rate (**FDR**) (*f*) was calculated as:

$$f = \frac{m\alpha}{s}, \quad (4)$$

where *m* corresponds to the number of tests (markers) considered in the meta-analysis, α is the significant threshold, and *s* is the number of tests, where $p < \alpha$, where *p* is the correspondent observed *p*-value for each marker. To select α values producing an FDR lower than 0.05, the *p*-value rank position procedure (Benjamini and Hochberg, 1995) was used. In summary, equation 5 was changed to obtain $\alpha = fs/m$, and *s* was defined as the largest *p*-value rank position *i* that satisfies $p_i \leq f_i/m$ with $f = 0.05$ (Pereira et al., 2016).

2.2.4 Gene annotation and enrichment analysis

The genes within 0.5 Mb of significant SNP ($P < 1.39 \times 10^{-5}$) were annotated as candidate genes. These candidate genes were used to perform gene ontology (GO)

and enrichment analyses, using DAVID bioinformatics resources (Huang da et al., 2009).

2.3 Results and Discussion

2.3.1. Significant SNP

Heritabilities were previously reported for the traits and populations studied here. They were 0.08 for AFC, 0.30 for EP, 0.41 for SC_N, 0.57 for AGECL, 0.52 for PPAI, and 0.75 for SC_B (Forni and Albuquerque, 2005; Johnston et al., 2009, 2010; Corbet et al., 2013; Irano et al., 2016).

Meta-analysis resulted in 108 significant SNPs, at an empirical threshold *P*-value of 1.39×10^{-5} (Figure 1). This number of significant SNP was the maximum rank position, which *P*-value satisfied the condition $P \leq \alpha$ with $FDR = 0.05$ (Supplementary Table S1). Most of the significant SNP (59.3%) were mapped to chromosome 14, producing a strong association peak between 24.3 and 26 Mb (Figure 2), in which 14 genes were harboured (*XKR4*, *TRNAT-AUG*, *TMEM68*, *TGS1*, *LYN*, *RPS20*, *MOS*, *PLAG1*, *CHCHD7*, *SDR16C5*, *SDR16C6*, *PENK*, *LOC101907667*, *IMPAD1*). This interval harbour a region reported as a high *Bos taurus* introgression region inserted around the decade of 1990 (Koufariotis et al., 2018). Another region in BTA14, from 29.5 to 29.6 Mb presented a smaller peak, in which four SNPs were significant. All those SNPs (*rs109465877*, *rs111023138*, *rs134875123*, and *rs42534153*) were located in an intronic region of the gene *NKAIN3*, which was in a genomic window explaining 1% of the genetic variance of early pregnancy in Nellore heifers (Oliveira Júnior et al., 2017).

The peak observed in chromosome 14 is close to significant regions that were associated with Brahman and Nellore reproductive traits (Fortes et al., 2012; Fortes et al., 2013; Irano et al., 2016). Some of these regions were previously associated with related or the same traits studied here using different statistical methods. The *PLAG1* gene (mapped to 25,000,459 to 25,052,403 bp), a key gene associated with several traits of economic importance in different cattle breeds (Karim et al., 2011; Littlejohn et al., 2012; Nishimura et al., 2012; Fortes et al., 2013; Utsunomiya et al., 2013; Saatchi et al., 2014; Juma et al., 2016; Pereira et al., 2016; Fink et al., 2017;

Utsunomiya et al., 2017) is located in this region. One of the most significant SNPs (*P*-value = 6.39 × 10^{-6}), the *rs109815800* was mapped in an intronic region of *PLAG1*. This SNP was described as a QTN for bovine stature by Karim et al. (2011).

Fortes et al. (2012) also found a region in chromosome 14 associated with AGECL and the age in which bull reach 26 cm of SC — phenotypes that represent the age at puberty for both females and males — for the same Brahman population used in this study. One of the most significant SNPs also was close to *SDR16C6* and *PLAG1* genes.

Despite this evident peak in chromosome 14, the most significant SNP (*P*-value = 3.21 × 10^{-14}) of the meta-analysis was located in chromosome 9 at 75.6 Mb. This SNP is in an intronic region of the *MAP3K5* gene, which was identified as an estrogen receptor element in human (Mirkin et al., 2005).

Figure 1. Manhattan plot of the meta-analysis for sexual precocity traits in Nellore and Brahman cattle. The y-axis represents the log inverse *P*-values for SNP associations and the x-axis represents the position in base pairs from chromosome 1 to 29. The blue line indicates genome-wide significance, a *P*-value cutoff of $P < 1.39 \times 10^{-5}$, which is equivalent of an FDR lower than 5%.

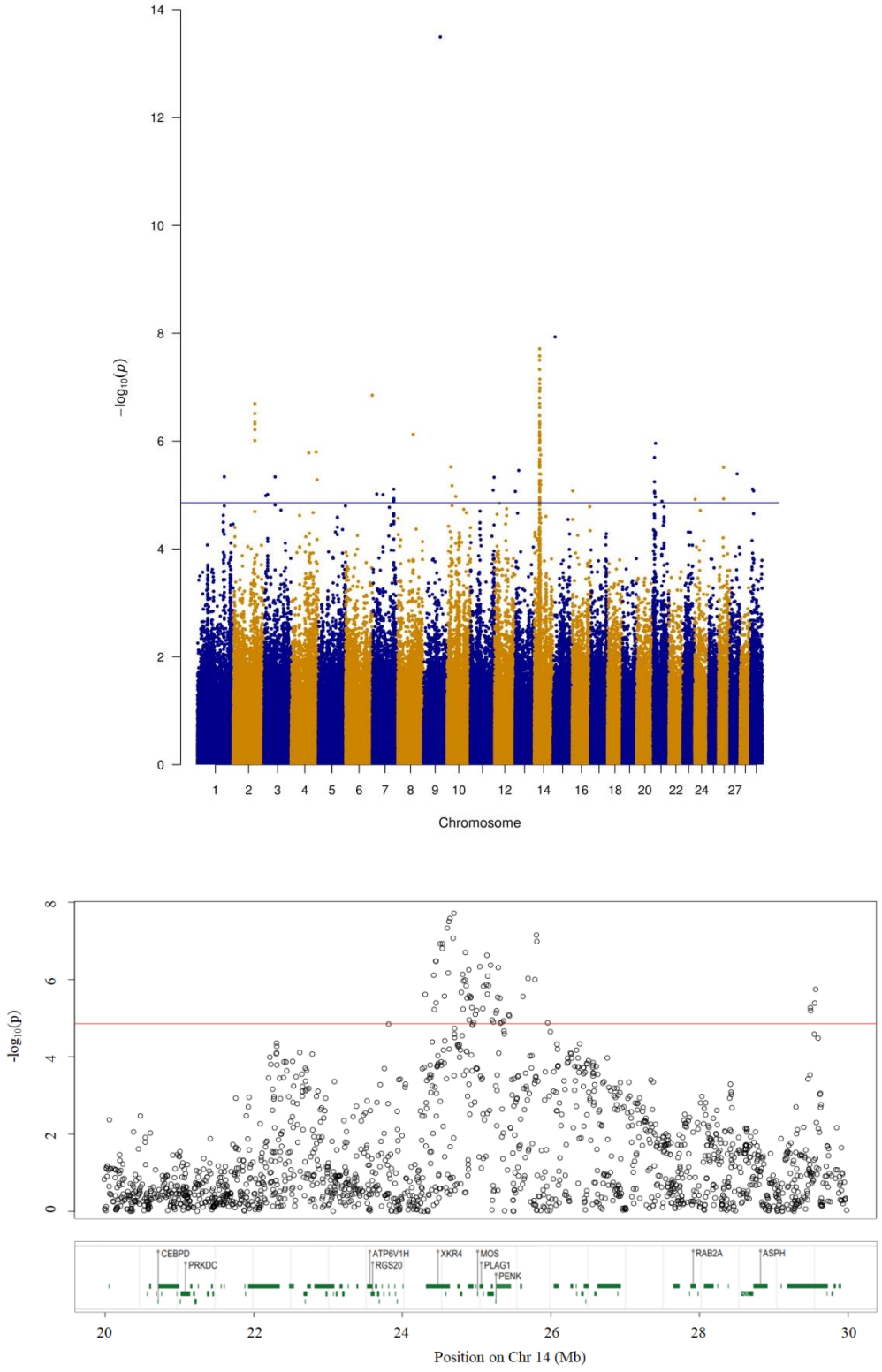

Figure 2. Zoom in BTA14, region of the strongest peak. SNPs above the red line were significant ($P < 1.39 \times 10^{-5}$). In the bottom, genes mapped in this region (total = 87).

In Table 2 are presented the most significant SNP of each chromosome, its position and the closest gene. In BTA1, *TSC22D2* emerged as a candidate gene in our meta-analysis study. This gene was high-expressed in human oocytes (Kakourou et al., 2013). In BTA2, *KLF7* was the closest gene to the top SNP (*rs134877457*). It is associated with growth traits in Chinese cattle (Ma et al., 2011); presents an important role in mouse neurogenesis (Laub et al., 2001; Caiazzo et al., 2010) and olfactory neurogenesis (Kajimura et al., 2007); and plays a role in adipogenesis regulation in humans (Kawamura et al., 2006).

Table 2. Number of significant SNP ($P < 1.39 \times 10^{-5}$) per *Bos taurus* autosome (BTA), the most significant SNP in each autosome (top SNP), its position, and its distance from the closest gene in base pairs.

BTA	Number of SNP	Top SNP	Position, bp	Gene Symbol	Distance, bp
1	1	*rs110366479*	118,606,312	*TSC22D2*	5,852
2	6	*rs134877457*	95,917,958	*KLF7*	14,648
3	3	*rs109478958*	49,436,731	*ARHGAP29*	0
4	3	*rs136050748*	110,440,534	*Uncharacterized Protein (ENSBTAG00000 048097)*	530,702
6	1	*rs43492923*	118,436,689	*7SK – misc RNA*[1]	2,828
7	6	*rs135631400*	94,710,749	*7SK – misc RNA*[1]	510,435
8	1	*rs136590180*	68,307,110	*Uncharacterized Protein (ENSBTAG00000 012266)*	16,590
9	1	*rs110257163*	75,613,158	*MAP3K5*	0
10	3	*rs110458186*	16,769,219	*TLE3*	132,436
11	2	*rs108980439*	104,936,387	*WDR5*	0
13	2	*rs136854801*	16,097,639	*TAF3*	0
14	64	*rs109748092*	24,710,609	*TMEM68*	718
15	1	*rs110493922*	9,064,376	*Uncharacterized Protein (ENSBTAG00000 047425)*	141,535
16	1	*rs109871859*	1,929,989	*PPP1R15B*	36,732
21	7	*rs110478544*	11,435,365	*NR2F2*	628,702
24	1	*rs133759831*	2,276,768	*GALR1*	129,893
26	2	*rs135708259*	23,405,679	*SUFU*	47,301
27	1	*rs135961785*	31,926,152	*KCNU1*	0

29	2	*rs109184359*	9,174,027	*Uncharacterized Protein (ENSBTAG00000 046374)*	1,952

[1]miscRNA = Miscellaneous RNA

In BTA3 the top SNP (*rs109478958*) suggested *ARHGAP29* as a candidate gene. Its protein was upregulated in bull semen with good quality, compared with bad-quality semen (Singh, et al., 2018). The top SNPs for chromosomes 6 (*rs43492923*) and 7 (*rs135631400*) were close to *7SK*, a small nuclear ribonucleoprotein-coding gene (snRNA). This snRNA plays a role in DNA transcription and was found close to a significant SNP for bull fertility in chromosome 10 (Suchocki and Szyda, 2015); has a critical role in the control of primordial germ cell proliferation of mouse embryos (Okamura et al., 2012); and was detected in seminal plasma, ejaculated sperm, and epididymal sperm in pigs with a significant biological role in spermatogenesis (Chen et al., 2017).

In BTA9, the gene *MAP3K5* was the closest to the top SNP (*rs110257163*) and was previously associated with regulation of apoptosis in porcine ovarian granulosa cells; the expression of *MAP3K5* was increased when the levels of FSH hormone increased in those cells (Sirotkin et al., 2008). Also, this gene was a strong candidate associated with growth and reproduction traits in transcriptome analysis of bovine pituitary gland (Pareek et al., 2016), and it presented some interactions with estrogen receptors in rat primary cortical neurons (Singer et al., 1999).

In BTA10, *TLE3* (closest to the top SNP *rs110458186*) was expressed in mouse mesoderm embryonic cells (Pfeffer et al., 2017). The expression of this gene was associated with FSH hormone increased levels in bovine granulosa cells (Nivet et al., 2018) and with estrogen receptors in human breast cancer cell line (Jangal et al., 2014). This gene was associated with mouse embryo survival. *TLE3* mutants presented placenta defects (Gasperowicz et al., 2013) and this gene acted as a co-regulator of adipogenesis in mouse (Villanueva et al., 2011). This gene also affects the embryonic stem cell differentiation (Laing et al., 2015).

In BTA11, the candidate gene *WDR5* presents an important role in porcine early embryo development (Ding et al., 2017) and was essential for vertebrate

development (Wysocka et al., 2005; Gori et al., 2006). In BTA13, the most significant SNP (*rs136854801*) was located in an intronic region of the *TAF3* gene. This gene harbors an SNP marker associated with the major histocompatibility complex (MHC) in an Italian human population (Pistis et al., 2013). The importance of the MHC for reproduction was previously discussed in Ziegler et al. (2010). Besides that, this gene was associated with endoderm differentiation in mouse embryonic stem cells (Liu et al., 2011), was significantly upregulated in human oocyte, and was enriched in the estrogen receptor signalling pathway (Kocabas et al., 2006).

In BTA14, *TMEM68* was identified as a candidate gene. This gene was expressed in human epididymis (Belleannee et al., 2012), was upregulated in heifer blastocyst (Carter, et al., 2010) and was associated with feed intake and growth traits in cattle (Lindholm-Perry et al., 2012). In BTA16, *PPP1R15B* was the gene closest to the top SNP (*rs109871859*). It was expressed in bovine cumulus oocyte cells (Abd El Naby et al., 2013) and in bovine granulosa cells under FSH stimulation (Nivet et al., 2018). *PPP1R15* gene knockout mouse embryos survived, but had several problems as grown retardation (Harding et al., 2009). In BTA21, the emergent candidate, *NR2F2*, was previously associated with Leydig cell steroidogenesis in mice testis cells (Mendoza-Villarroel et al., 2014). Bauersachs et al. (2006) concluded that this gene was associated with embryo implantation in cows. In BTA24 the gene closest to the top SNP (*rs133759831*) was *GALR1*. This gene was differentially expressed in ovine hypothalamus in pre-pubertal age compared with fetal and adult ages, suggesting a regulation role in the beginning of reproduction (Whitelaw et al., 2009). Also, it was located near a significant SNP associated with oligozoospermia and azoospermia in humans (Aston and Carrell, 2009), was associated with the regulation of ovarian steroids in rats (Mitchell et al., 2004), and interacted with testosterone in male rat brain (Bouret et al., 2000). In BTA26, *SUFU* was mapped close to the top SNP (*rs135708259*). This gene was a candidate gene associated with spermatogenesis in mouse, and it was found in spermatocytes (Szczepny et al., 2005). In BTA27, the *KCNU1* was the closest gene to the top SNP (*rs135961785*). This gene presents an important role in semen quality (Santi et al., 2010), which could be observed in a knockout mouse experiment (*SLO3*, null mutation). Also Zeng et al. (2011) reported the importance of *SLO3* for male fertility, by evaluating its role

in the normal morphology and sperm motility in mouse. This gene was enriched in the pathway “Vascular smooth muscle contraction,” which was activated by progesterone treatment in neonatal mouse uterus, inhibiting its development (Filant et al., 2012).

On chromosomes 4, 8, 15, and 29, the top SNPs were close to uncharacterized proteins, and their identification is presented in Table 2. No significant SNPs were observed in BTA5, BTA12, BTA17, BTA18, BTA19, BTA20, BTA22, BTA23, BTA25, and BTA28. Evidence from the current meta-analysis seems to corroborate with previous knowledge of gene function for the discussed genes. Although the majority of the significant SNPs were located in chromosome 14, significant SNPs were also mapped in multiple chromosomes, which is consistent with the idea that reproductive traits are complex polygenic phenotypes.

2.3.2 Functional Annotation and Enrichment Analyses

A total of 389 genes were located within 0.5 Mb of the 108 significant SNP ($P < 1.39 \times 10^{-5}$). From these genes, functional annotation and enrichment analyses were performed and 333 genes were identified by DAVID software (Huang da et al., 2009), using bovine genome as a background gene list. Enrichment analyses identified 48 functional gene groups as significant in DAVID resources, using high astringency criterion. The top five functional gene groups, with higher enrichment score (ES), are presented in Table 3.

Groups 1 (S100/CaBP-9k-type, calcium binding) and 2 (domain: EF-hand 1) were related to calcium binding pathway. S100 proteins are a family protein expressed just in vertebrates. They participate in several biological events, as regulation of proliferation, differentiation, apoptosis, and inflammation (Donato et al., 2013). Calbindin-D9k (CaBP-9k-type) is a vitamin-D-dependent Ca^{2+} binding protein, member of the S100 family. This protein is mainly found in intestinal tissue, but was also found in other tissues as uterus and placenta (Mathieu et al., 1989; Krisinger et al., 1995; Emam et al., 2016). All the genes clustered in group 1 were members of S100A family. This gene family was expressed in ovary, prostate, and testis of human and rats (Wicki et al., 1996) and in thyroid gland, placenta, and prostate (Cannon et al., 2011). Their roles are generally associated with immunological

function, promoting the sterility of reproductive tissues against pathogens (Germeyer et al., 2009; Teijeiro and Marini, 2015); however, they were also related to embryo implantation process (Gray et al., 2006; Smits et al., 2018), to response to temperature stress (Landriscina et al., 2001), cell proliferation, and apoptosis (Jin et al., 2011).

Table 3. Top 5 most significant functional gene groups enriched in DAVID v.6.7 bioinformatics resources (https://david.ncifcrf.gov/home.jsp). Gene Ontology (GO) terms, enrichment scores (ES) and false discovery rates (FDR) are reported for each functional group; genes in each group are also reported.

Functional groups	Top GO term, code ~ name	Ontology	ES	FDR	Genes
1	IPR001751~ S100/CaBP-9k-type, calcium binding		7.06	5.2E-6	*S100A13, S100A1, S100A14, S100A16, S100A2, S100A3, S100A4, S100A7*
2	UP_SEQ_FEATURE~ domain: EF-hand 1		2.79	1.2E-1	*S100A13, S100A1, S100A14, S100A16, S100A2, S100A4, CAPS, NCS1, S100A7*
3	GO:0016477~ Cell migration	Biological Process	2.59	2.2E0	*FCER1G, cdk5, DBH, MP14, NOS3, NR2F2, pex7, LOC515718*
4	GO:0019864~ IgG Binding	Molecular Function	1.99	1.7E0	*FCER1G, FCGR3, FCGR2B*
5	GO:0051047~ Positive regulation of secretion	Biological Process	1.65	1.6E1	*FCER1G*, *CHRNB2*, *cdk5*, *serp1*

The majority of genes clustered in group 2 were also in group 1 (Table 3). This functional group, EF-hand 1, presents a group of proteins that share a similar structure, which is commonly found in calcium binding proteins. Besides of some genes in common with group 1, other two genes were enriched in this group, *CAPS*, *NCS1*. *CAPS* gene was expressed in human epididymis and endometrium during

follicular phase (Thimon et al., 2007; Blockeel et al., 2011), while *NCS1* was expressed in bovine blastocysts cultured in vitro under stimulation of thyroid hormones, presenting an important role in mammalian neuron-memory development (Ashkar et al., 2016).

Group 3 (Cell migration) is a fundamental process in the development and maintenance of multicellular organisms. Several reproductive events depend on this process, as embryo formation (Jovanović et al., 2010), immunological responses (Sanchez-Madrid and del Pozo, 1999), spermatogenesis (Smith et al., 2012), and oogenesis (Rorth et al., 2000). Genes clustered in this group were associated with daughter pregnancy rate, cow conception rate (Ortega et al., 2016), development of nervous system (Kapur et al., 1991; Kanaani et al., 2005; Jessberger et al., 2009), embryo lethality (Nourizadeh-Lillabadi et al., 2010), postnatal mortality and growth retardation (Pallares and Gonzalez–Bulnes, 2010), immunity (Bogdan, 2015), embryo differentiation (Rosa and Brivanlou, 2011), and perinatal mortality (Brites et al., 2003).

The group 4 clustered genes associated with the immunoglobulin G (IgG). This is the most common isotype in human blood and extracellular fluid (Janeway et al., 2007), and the most abundant antibody class present in healthy individuals. This antibody was found in the mucus of woman reproductive tract in Fahrbach et al. (2013), which could suggest a role in the fortification of mucus barriers in the female reproductive tract. Besides that, mouse experiments were performed with a specific monoclonal IgM and demonstrated that females immunized with this antibody presented gestational failure (Sthoeger et al., 1993). A similar effect was observed in another mouse experiment, by Ornoy et al. (2003), where the administration of a purified IgG in pregnant mouse reduced the yolk sac and the embryo growth. This pathway enriched the genes *FCER1G*, *FCGR3A*, and *FCGR2B*, which were associated with daughter pregnancy rate, cow conception rate (Ortega et al., 2016), embryo–maternal recognition in cattle (Mamo et al., 2012), and infertility in female mouse (Wetendorf et al., 2017).

The fifth top group was “Positive regulation of secretion.” As group 3, this group is related to a large variety of biological mechanisms, mainly mediated by regulatory hormones. Genes clustered in this group were associated with positive regulation of

immune response (Macen et al., 1993; Kiba, 2016), neurogenesis (Jessberger et al., 2009), and response to stress (Faria et al., 2012).

David results presented gene clusters that were mostly associated with immune events. Earlier studies with *Bos indicus* cattle have found genes playing roles in immune responses, adaptability, and reproduction (Bahbahani et al., 2018). Fayemi (2005) observed that the presence of antibodies (IgG) in bulls' sperm was related with fertility. Our results are in agreement with previous studies with *Bos indicus* cattle, pointing the importance of the immune competence to guarantee reproductive success for this cattle subspecies.

Besides most of significant SNPs were located in chromosome 14, our meta-analysis identified genomic regions harbouring significant SNP associations in the majority of the autosome chromosomes (19 out of 29), as expected for polygenic traits. The genes mapped within these associated regions are plausible candidate genes because of their position and due to previous evidence suggesting their functional association with reproduction, as discussed above. The candidate genes identified could be grouped in functional categories that seem to affect immune system and consequently, mammalian reproduction.

In summary, five regions distributed over four chromosomes presented the majority of the most significant SNP (76%): BTA2 (5.55%) from 95 to 96 Mb, BTA4 (5.55%) from 94.1 to 94.8 Mb, BTA14 (59.26%) from 24 to 25 Mb and 29 to 30 Mb, and BTA21 (5.55%) from 6.7 Mb to 11.4 Mb. Considering that two independent populations were studied here, the animals were from different breeds, born and bred in different geographic regions, the our results suggest that the multi-trait meta-analysis is a practicable strategy to identify genomic regions affecting traits of interest across populations. The genes identified as candidates in the meta-analysis are probably affecting physiological mechanisms that control sexual precocity in both breeds, because we found key genes expressed in reproductive tissues and playing important roles in immune response (including in reproductive tract), maternal–fetal tolerance, pregnancy maintenance, embryo development, fertility, and response to stress, or were located in previously reported QTL regions for reproductive traits in bovine. Also, we found some candidate genes associated with important hormones for the regulation of puberty in mammalian, as the FSH, estrogens, and testosterone.

The importance of these hormones for sexual precocity is expected to be conserved among mammalians.

In short, the genes presented here are both positional and functional candidates for sexual precocity in Bos indicus cattle. Future works could finemap and validate the identified genomic regions and elucidate which genes are in fact harbouring the mutations that could explain the proposed QTL.

2.4. Conclusions

We were able to identify five genomic regions that are strong candidates to be affecting sexual precocity in Nellore and Brahman populations. Besides these two populations are independent and the animals are born and bred in different geographic regions, those five regions emerged as candidates for traits studied here. The genes identified as candidates here are more likely to affect physiological mechanisms that control immune response, maternal–fetal tolerance, pregnancy maintenance, embryo development, fertility, and response to stress. Future works could fine-map and validate the identified genomic regions and elucidate which genes are in fact harbouring the mutations that could explain the proposed QTLs.

References

Abd El Naby WA, Hagos TH, et al. (2013) Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. **Zygote** 21:31-51.

Ashkar FA, Revay T, Rho N, Madan P, Dufort I, Robert C, King WA (2016) Thyroid hormones alter the transcriptome of in vitro-produced bovine blastocysts. **Zygote** 24:266-276.

Aston KI, Carrell, DT (2009) Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. **Journal of andrology** 30:711-725.

Bahbahani H, Salim B, Almathen F, Al Enezi F, Mwacharo JM, Hanotte O (2018) Signatures of positive selection in African Butana and Kenana dairy zebu cattle. **PloS one** 13:e0190446.

Bauersachs S, Ulbrich SE, Gross K, Schmidt SEM, Meyer HHD, Wenigerkind H, Vermehren M, Sinowatz F, Blum H, Wolf E (2006) Embryo-induced transcriptome changes in bovine endometrium reveal species-specific and common molecular markers of uterine receptivity. **Reproduction** 132:319-331.

Belleannée C, Calvo E, Thimon V, Cyr DG, Légaré C, Garneau L, Sullivan R (2012) Role of microRNAs in controlling gene expression in different segments of the human epididymis. **PloS one** 7:e34996.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the royal statistical society. Series B (Methodological)** 57:289-300.

Blockeel C, Van Vaerenbergh I, Fatemi HM, Van Lommel L, Devroey P, Bourgain C (2010) Gene expression profile in the endometrium on the day of oocyte retrieval after ovarian stimulation with low-dose hCG in the follicular phase. **Molecular human reproduction** 17:33-41.

Bogdan C (2015) Nitric oxide synthase in innate and adaptive immunity: an update. **Trends in immunology** 36:161-178.

Bolormaa S, Pryce JE, Reverter A, Zhang YD, Barendse W, Kemper K, Tier B, Savin K, Hayes B, Goddard ME (2014) A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. **PLoS Genetics** 10:e1004198.

Bouret S, Prevot V, Croix D, Howard A, Habert-Ortoli E, Jegou S, Vaudry H, Beauvillain JC, Mitchell V (2000) Expression of *GalR1* and *GalR2* galanin receptor messenger ribonucleic acid in proopiomelanocortin neurons of the rat arcuate nucleus: effect of testosterone. **Endocrinology** 141:1780-1794.

Brites P, Motley AM, et al. (2003) Impaired neuronal migration and endochondral ossification in *Pex7* knockout mice: a model for rhizomelic chondrodysplasia punctata. **Human molecular genetics** 12:2255-2267.

Browning BL, Browning SR (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. **The American Journal of Human Genetics** 84:210-223.

Brumatti RC, Ferraz JBS, Eler JP, Formigonni IB (2011) Development of selection index in beef cattle under focus of a bio-economic model. **Archivos de Zootecnia**. 60:205-213.

Burns BM, Corbet NJ, Corbet DH, Crisp JM, Venus BK, Johnston DJ, Li Y, McGowan MR, Holroyd RG (2013) Male traits and herd reproductive capability in tropical beef cattle. 1. Experimental design and animal measures. **Animal Production Science**. 53:87-100.

Caiazzo M, Colucci-D'Amato L, Esposito MT, Parisi S, Stifani S, Ramirez F, di Porzio U (2010) Transcription factor *KLF7* regulates differentiation of neuroectodermal and mesodermal cell lineages. **Experimental cell research** 316:2365-2376.

Cannon BR, Zimmer DB, Weber DJ (2011) *S100A1* (S100 calcium binding protein A1). **Atlas of genetics and cytogenetics in oncology and haematology** 15:873-876.

Cardoso RC, Alves BRC, Sharpton SM, Williams GL, Amstalden M (2015) Nutritional programming of accelerated puberty in heifers: Involvement of pro-opiomelanocortin neurones in the arcuate nucleus. **Journal of Neuroendocrinology** 27:647-657.

Carter F, Rings F, et al (2010) Effect of elevated circulating progesterone concentration on bovine blastocyst development and global transcriptome following endoscopic transfer of in vitro produced embryos to the bovine oviduct. **Biology of Reproduction** 83:707-719.

Carvalheiro R, Boison SA, et al. (2014) Accuracy of genotype imputation in Nelore cattle. **Genetics Selection Evolution** 46:69

Chen C, Wu H, et al. (2017) Comparative profiling of small rnas of pig seminal plasma and ejaculated and epididymal sperm. **Reproduction** 153:785-796.

Corbet NJ, Burns BM, Johnston DJ, Wolcott ML, Corbet DH, Venus BK, Li Y, McGowan MR, Holroyd RG (2013) Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits. **Animal Production Science** 53:101-113.

Costa RB, Camargo GM, et al. (2015) Genome-wide association study of reproductive traits in nellore heifers using bayesian inference. **Genetics Selection Evolution** 47:67.

Ding B, Cao Z, et al. (2017) *WDR5* in porcine preimplantation embryos: expression, regulation of epigenetic modifications and requirement for early development. **Biology of Reproduction** 96:758-771.

Donato R, Cannon BR, Sorci G, Riuzzi F, Hsu K, Weber DJ, Geczy CL (2013) Functions of S100 proteins. **Current molecular medicine** 13:24-57.

Emam MA, Abouelroos ME, Gad FA (2016) Expression of calbindin-D9k and vitamin D receptor in the uterus of Egyptian buffalo during follicular and luteal phases. **Acta Histochemica** 118:471-477.

Fahrbach KM, Malykhina O, Stieh DJ, Hope TJ (2013) Differential binding of IgG and IgA to mucus of the female reproductive tract. **PloS one** 8:e76176.

Faria D, Lentze N, et al. (2012) Regulation of ENaC biogenesis by the stress response protein *SERP1.* **Pflügers Archiv-European Journal of Physiology** 463:819-827.

Fayemi O (2005) Sperm antibodies and reproductive efficiency in the zebu cattle in south-western Nigeria. **Pakistan Veterinary Journal** 25:111-114.

Filant J, Zhou H, Spencer TE (2012) Progesterone inhibits uterine gland development in the neonatal mouse uterus. **Biology of Reproduction**. 86:1-9.

Fink T, Tiplady K, Lopdell T, Johnson T, Snell RG, Spelman RJ, Stephen RD, Littlejohn MD (2017) Functional confirmation of PLAG1 as the candidate causative gene underlying major pleiotropic effects on body weight and milk characteristics. **Scientific Reports**. 7:44793.

Formigoni IB, Ferraz JBS, Silva JAIIV, Eler JP, Brumatti RC (2005) Economic value for stayability and heifer pregnancy at 14 months in beef cattle herds. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**. 57:220-226.

Forni S, Albuquerque LG (2005) Estimates of genetic correlations between days to calving and reproductive and weight traits in Nelore cattle. **Journal of Animal Science** 83:1511-1515.

Fortes MRS, Kemper K, et al. (2013) Evidence for pleiotropism and recent selection in the Plag1 region in australian beef cattle. **Animal Genetics** 44:636-647.

Fortes MRS, Lehnert SA, Bolormaa S, Reich C, Fordyce G, Corbet NJ, Whan V, Hawken RJ, Reverter A (2012) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science** 52:143-150.

Garcia-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-Lopez FJ, Van Tassell C P (2016) Changes in genetic selection differentials and generation intervals in us Holstein dairy cattle as a result of genomic selection. **Proceedings of the National Academy of Sciences** 113:E3995-E4004.

Gasperowicz M, Surmann-Schmitt C, Hamada Y, Otto F, Cross JC (2013) The transcriptional co-repressor *TLE3* regulates development of trophoblast giant cells lining maternal blood spaces in the mouse placenta. **Developmental biology** 382:1-14.

Germeyer A, Sharkey AM, et al. (2008) Paracrine effects of uterine leucocytes on gene expression of human uterine stromal fibroblasts. **Molecular Human Reproduction** 15:39-48.

Gori F, Friedman LG, Demay MB (2006) *Wdr5*, a WD-40 protein, regulates osteoblast differentiation during embryonic bone development.**Developmental biology** 295: 498-506.

Gray CA, Abbey CA, Beremand PD, Choi Y, Farmer JL, Adelson DL, Thomas TL, Bazer FW, Spencer TE (2006) Identification of endometrial genes regulated by early pregnancy, progesterone, and interferon tau in the ovine uterus. **Biology of Reproduction** 74:383-394.

Harding HP, Zhang Y, Scheuner D, Chen JJ, Kaufman RJ, Ron D (2009) *Ppp1r15* gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2α) dephosphorylation in mammalian development. **Proceedings of the national academy of sciences** 106:1832-1837.

Hawken RJ, Zhang YD, et al. (2012) Genome-wide association studies of female reproduction in tropically adapted beef cattle. **Journal of Animal Science** 90:1398-1410.

Henderson CR (1984) **Aplications of linear models in animal breeding**. University of Guelph, Guelph.

Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using David bioinformatics resources. **Nature Protocols** 4:44-57.

Irano N, de Camargo GMF, et al. (2016) Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. **Plos One**. 11:e0159502.

Janeway CA, Travers P, Walport M, Shlomchik MJ (2007) **Imunobiologia: O sistema imune na saúde e na doença**. 6th ed. Artmed, Sao Paulo.

Jangal M, Couture JP, Bianco S, Magnani L, Mohammed H, Gévry N (2014) The transcriptional co-repressor *TLE3* suppresses basal signaling on a subset of estrogen receptor α target genes. **Nucleic Acids Research** 42:11339-11348.

Jessberger S, Gage FH, Eisch AJ, Lagace DC (2009) Making a neuron: *Cdk5* in embryonic and adult neurogenesis. **Trends in Neurosciencies** 32:575-582.

Jin Q, Chen H, Luo A, Ding F, Liu Z (2011) *S100A14* stimulates cell proliferation and induces cell apoptosis at different concentrations via receptor for advanced glycation end products (RAGE). **PloS one** 6:e19375.

Johnston D J, 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** 49:399-412.

Johnston DJ, Barwick SA, Fordyce G, Holroyd RG (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. **Proceedings...** Leipzig, Germany.

Jovanović M, Stefanoska I, Radojčić L, Vićovac L (2010) Interleukin-8 (*CXCL8*) stimulates trophoblast cell migration and invasion by increasing levels of matrix metalloproteinase (MMP) 2 and MMP9 and integrins α5 and β1. **Reproduction**. 139:789-798.

Juma AR, Damdimopoulou PE, Grommen SV, Van de Ven WJ, De Groef B (2016) Emerging role of *PLAG1* as a regulator of growth and reproduction. **Journal of Endocrinology** 228:R45-56.

Kajimura D, Dragomir C, Ramirez F, Laub F (2007) Identification of genes regulated by transcription factor *KLF7* in differentiating olfactory sensory neurons. **Gene** 388:34-42.

Kakourou G, Jaroudi S, Tulay P, Heath C, Serhal P, Harper JC, SenGupta SB (2013) Investigation of gene expression profiles before and after embryonic genome activation and assessment of functional pathways at the human metaphase II oocyte and blastocyst stage. **Fertility and Sterility** 99:803-814.

Kanaani J, Prusiner SB, Diacovo J, Baekkeskov S, Legname G (2005) Recombinant prion protein induces rapid polarization and development of synapses in embryonic rat hippocampal neurons in vitro. **Journal of Neurochemistry** 95:1373-1386.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. **Nature Genetics** 42:348-354.

Kapur RP, Hoyle GW, Mercer EH, Brinster RL, Palmiter RD (1991) Some neuronal cell populations express human dopamine β-hydroxylase-*lacZ* transgenes transiently during embryonic development. **Neuron** 7:717-727.

Karim L, Takeda H, et al. (2011) Variants modulating the expression of a chromosome domain encompassing *PLAG1* influence bovine stature. **Nature Genetics** 43:405-413.

Kawamura Y, Tanaka Y, Kawamori R, Maeda S (2006) Overexpression of Kruppel-like factor 7 regulates adipocytokine gene expressions in human adipocytes and inhibits glucose-induced insulin secretion in pancreatic β-cell line. **Molecular Endocrinology** 20:844-856.

Kiba T (2016) Ventromedial hypothalamic lesions downregulate multiple immune signaling pathways in rat pancreatic islets. **Neuroscience letters** 610:177-181.

Kocabas AM, Crosby J, et al. (2006) The transcriptome of human oocytes.**Proceedings of the National Academy of Sciences** 103:14027-14032.

Koufariotis L, Hayes BJ, Kelly M, Burns BM, Lyons R, Stothard P, Chamberlain AJ, Moore S (2018) Sequencing the mosaic genome of Brahman cattle identifies historic and recent introgression including polled. **Scientific Reports** 8:17761.

Krisinger J, Eung EB, Simmen RC, Leung PC (1995) Porcine calbindin-D9k gene: expression in endometrium, myometrium, and placenta in the absence of a functional estrogen response element in intron A. **Biology of Reproduction** 52:115-123.

Laing AF, Lowell S, Brickman JM (2015) Gro/TLE enables embryonic stem cell differentiation by repressing pluripotent gene expression. **Developmental Biology** 397:56-66.

Landriscina M, Soldi R, Bagalá C, Micucci I, Bellum S, Tarantini F, Prudovsky I, Maciag T (2001) *S100A13* participates in the release of fibroblast growth factor 1 in response to heat shock in vitro. **Journal of Biological Chemistry**. 276:22544-22552.

Laub F, Aldabe R, Friedrich V, Ohnishi S, Yoshida T, Ramirez F (2001) Developmental expression of mouse Krüppel-like transcription factor *KLF7* suggests a potential role in neurogenesis. **Developmental Biology** 233:305-318.

Lindholm-Perry AK, Kuehn LA, Smith TPL, Ferrell CL, Jenkins TG, Freetly HC, Snelling WM (2012) A region on BTA14 that includes the positional candidate genes *LYPLA1*, *XKR4* and *TMEM68* is associated with feed intake and growth phenotypes in cattle 1. **Animal Genetics** 43:216-219.

Littlejohn M, Grala T, et al. (2012) Genetic variation in *PLAG1* associates with early life body weight and peripubertal weight and growth in *Bos taurus*. **Animal Genetics** 43:591-594.

Liu Z, Scannell DR, Eisen MB, Tjian R (2011) Control of embryonic stem cell lineage commitment by core promoter factor, *TAF3*. **Cell** 146:720-731.

Ma L, Qu YJ, Huai YT, Li ZJ, Wang J, Lan XY, Zhang CL, Wang JQ, Chen H (2011) Polymorphisms identification and associations of *KLF7* gene with cattle growth traits. **Livestock. Science** 135:1-7.

Macen JL, Upton C, Nation N, McFadden G (1993) *SERP1*, a serine proteinase inhibitor encoded by myxoma virus, is a secreted glycoprotein that interferes with inflammation. **Virology** 195:348-363.

Mamo S, Mehta JP, Forde N, McGettigan P, Lonergan P (2012) Conceptus-endometrium crosstalk during maternal recognition of pregnancy in cattle. **Biology of Reproduction** 87:1-9.

Mathieu CL, Burnett SH, Mills SE, Overpeck JG, Bruns DE, Bruns ME (1989) Gestational changes in calbindin-D9k in rat uterus, yolk sac, and placenta: implications for maternal-fetal calcium transport and uterine muscle function. **Proceedings of the national academy of sciences** 86:3433-3437.

De Melo TP, de Camargo GMF, de Albuquerque L G, Carvalheiro R (2017) Genome-wide association study provides strong evidence of genes affecting the reproductive performance of nellore beef cows. **PLoS One**. 12:e0178551.

Mendoza-Villarroel RE, Robert NM, Martin LJ, Brousseau C, Tremblay JJ (2014) The nuclear receptor *NR2F2* activates star expression and steroidogenesis in mouse MA-10 and MLTC-1 Leydig cells. **Biology of Reproduction** 91:1-12.

Mirkin S, Arslan M, Churikov D, Corica A, Diaz JI, Williams S, Bocca S, Oehninger S (2005) In search of candidate genes critically expressed in the human endometrium during the window of implantation. **Human Reproduction** 20:2104-2117.

Mitchell V, Lecompte F, Beauvillain JC (2004) Regulation of galanin receptor GalR1 mRNA expression by ovarian steroids in oestrogen receptor alpha-immunoreactive neurones: identification of distinct populations of neurones in the preoptic area. **Journal of Neuroendocrinology** 16:138-145.

Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, Sugimoto Y, Takasuga A (2012) Genome-wide association study identified three major QTL for carcass weight including the *PLAG1-CHCHD7* QTN for stature in japanese black cattle. **BMC Genetics**. 13:40.

Nivet AL, Dufort I, Gilbert I, Sirard MA (2018) Short-term effect of FSH on gene expression in bovine granulosa cells in vitro. **Reproduction, fertility, and development**. 30:1154-1160.

Nourizadeh-Lillabadi R, Torgersen JS, Vestrheim O, König M, Aleström P, Syed M (2010) Early embryonic gene expression profiling of zebrafish prion protein (Prp2) morphants. **PLoS One** 5:e13573.

Oliveira Júnior GA, Perez BC, et al. (2017) Genomic study and Medical Subject Headings enrichment analysis of early pregnancy rate and antral follicle numbers in Nellore heifers. **Journal of Animal Science** 95:4796-4812.

Ornoy A, Yacobi S, Matalon ST, Blank M, Blumenfeld Z, Miller RK, Shoenfeld Y (2003) The effects of antiphospholipid antibodies obtained from women with SLE/APS and associated pregnancy loss on rat embryos and placental explants in culture. **Lupus** 12:573-578.

Ortega MS, Denicol AC, Cole JB, Null DJ, Hansen PJ (2016) Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. **Animal Genetics** 47:288-297.

Pallares P, Gonzalez–Bulnes A (2010) The effect of embryo and maternal genotypes on prolificacy, intrauterine growth retardation and postnatal development of Nos3–knockout mice. **Reproductive Biology** 10:241-248.

Pareek CS, Smoczyński R, Kadarmideen HN, Dziuba P, Błaszczyk P, Sikora M, Szostak A (2016) Single nucleotide polymorphism discovery in bovine pituitary gland using RNA-seq technology. **PloS one** 11:e0161370.

Pereira AGT, Utsunomiya YT, et al. (2016) Pleiotropic genes affecting carcass traits in *Bos Indicus* (Nellore) cattle are modulators of growth. **Plos One** 11:e0158165.

Pfeffer PL, Smith CS, Maclean P, Berg DK (2017) Gene expression analysis of bovine embryonic disc, trophoblast and parietal hypoblast at the start of gastrulation. **Zygote** 25:265-278.

Pistis G, Okonkwo SU, et al. (2013) Genome wide association analysis of a founder population identified *TAF3* as a gene for MCHC in humans. **PloS one** 8:e69206.

Ramayo-Caldas Y, Renand G, Ballester M, Saintilan R, Rocha D (2016) Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. **Genetics Selection Evolution** 48:37.

Rørth P, Szabo K, Texido G (2000) The level of C/EBP protein is critical for cell migration during Drosophila oogenesis and is tightly controlled by regulated degradation. **Molecular Cell** 6:23-30.

Rosa A, Brivanlou AH (2011) A regulatory circuitry comprised of miR-302 and the transcription factors *OCT4* and *NR2F2* regulates human embryonic stem cell differentiation. **The EMBO Journal** 30:237-248.

Saatchi M, Schnabel RD, Taylor JF, Garrick DJ (2014) Large-effect pleiotropic or closely linked QTL segregate within and across ten us cattle breeds. **BMC Genomics** 15:16.

Sánchez-Madrid F, del Pozo MA (1999) Leukocyte polarization in cell migration and immune interactions. **The EMBO Journal** 18:501-511.

Santi CM, Martínez-López P, de la Vega-Beltrán JL, Butler A, Alisio A, Darszon A, Salkoff L (2010) The *SLO3* sperm-specific potassium channel plays a vital role in male fertility. **FEBS letters** 584:1041-1046.

Sargolzaei M, Chesnais JP Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. **BMC Genomics** 15:478.

Singer CA, Figueroa-Masot XA, Batchelor RH, Dorsa DM (1999) The mitogen-activated protein kinase pathway mediates estrogen neuroprotection after glutamate toxicity in primary cortical neurons. **Journal of Neuroscience** 19:2455-2463.

Singh R, Sengar GS, et al. (2018) Functional proteomic analysis of crossbred (Friesian × Sahiwal) bull spermatozoa. **Reproduction in Domestic Animals** 53:588-608.

Sirotkin AV, Benčo A, Tandlmajerova A, Vašíček D, Kotwica J, Darlak K, Valenzuela F (2008) Transcription factor p53 can regulate proliferation, apoptosis and secretory

activity of luteinizing porcine ovarian granulosa cell cultured with and without ghrelin and FSH. **Reproduction** 136:611-618.

Smith BE, Braun RE (2012) Germ cell migration across Sertoli cell tight junctions. **Science** 338:798-802.

Smits K, Willems S (2018) Proteins involved in embryo-maternal interaction around the signalling of maternal recognition of pregnancy in the horse. **Scientific Reports** 8:5249.

SNP & Variation Suite ™ (Version 8.3.0). Bozeman, MT: Golden Helix, Inc.

Sthoeger ZM, Mozes E, Tartakovsky B (1993) Anti-cardiolipin antibodies induce pregnancy failure by impairing embryonic implantation. **Proceedings of the National Academy of Sciences** 90:6464-6467.

Suchocki T, Szyda J (2015) Genome-wide association study for semen production traits in Holstein-Friesian bulls. **Journal of Dairy Science** 98:5774-5780.

Szczepny A, Jans DA, Hime G, Loveland KL (2005) Expression of components of the hedgehog signalling pathway during murine spermatogenesis. **Reproduction, Fertility and Development** 17:98-98.

Teijeiro JM, Marini PE (2015) *S100A7* in the Fallopian tube: a comparative study. **Zygote** 23:229-236.

Thimon V, Koukoui O, Calvo E, Sullivan R (2007) Region-specific gene expression profiling along the human epididymis. **Molecular Human Reproduction** 13: 691-704.

Okamura D, Maeda I, et al. (2012) Cell cycle gene-specific control of transcription has a critical role in proliferation of primordial germ cells**. Genes & Development** 26:2477-2482.

Utsunomiya YT, do Carmo AS, et al. (2013) Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. **BMC Genetics** 14:52.

Utsunomiya YT, Milanesi M, et al. (2017) A PLAG1 mutation contributed to stature recovery in modern cattle. **Scientific reports** 7:17140.

VanRaden PM (2008) Efficient methods to compute genomic predictions. **Journal of Dairy Science** 91:4414-4423.

Villanueva CJ, Waki H, et al. (2011) *TLE3* is a dual-function transcriptional coregulator of adipogenesis. **Cell Metabolism** 13:413-427.

Wetendorf M, Wu SP, et al. (2017) Decreased epithelial progesterone receptor A at the window of receptivity is required for preparation of the endometrium for embryo attachment. **Biology of Reproduction** 96:313-326.

Whitelaw CM, Robinson JE, Chambers GB, Hastie P, Padmanabhan V, Thompson RC, Evans NP (2009) Expression of mRNA for galanin, galanin-like peptide and galanin receptors 1–3 in the ovine hypothalamus and pituitary gland: effects of age and gender. **Reproduction** 137:141-150.

Wicki R, Schäfer BW, Erne P, Heizmann CW (1996) Characterization of the Human and Mouse cDNAsCoding for *S100A13*, a New Member of the S100 Protein Family. Biochem. **Biochemical and biophysical research communications** 227:594-599.

Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) *WDR5* associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. **Cell** 121:859-872.

Zeng XH, Yang C, Kim ST, Lingle CJ, Xia XM (2011) Deletion of the *Slo3* gene abolishes alkalization-activated K^+ current in mouse spermatozoa. **Proceedings of the National Academy of Sciences** 108:5879-5884.

Zhang YD, Johnston DJ, Bolormaa S, Hawken RJ, Tier B (2014) Genomic selection for female reproduction in Australian tropically adapted beef cattle. **Animal Production. Science** 54:16-24.

Ziegler A, Santos PSC, Kellermann T, Uchanska-Ziegler B (2010) Self/nonself perception, reproduction, and the extended MHC. **Self/Nonself** 1:176-191.

ACROSS-BREED VALIDATION STUDY CONFIRMS AND IDENTIFY NEW LOCI ASSOCIATED WITH SEXUAL PRECOCITY IN BRAHMAN AND NELLORE CATTLE

ABSTRACT: Cross validation studies are important to validate across breed mutations affecting economic important traits from Genome-wide association studies (GWAS). Genic and pleiotropic regions are strong candidates to be harbouring QTL segregating across breeds. Thus, the aim of this study was to identify candidate regions and genes potentially associated with sexual precocity in *Bos indicus* cattle, through a validation study. An across-breed approach was used to perform the validation, by using Nellore as validation population, and Brahman as the discovery population. The SNP selected in Brahman to validate in Nellore were from two origins: 1) SNP in gene regions previously reported as affecting reproductive traits in Brahman cattle (G1), and 2) significant SNP from a meta-analysis of Brahman sexual precocity traits (G2). In the validation population two traits indicators of sexual precocity in female and male of beef cattle were evaluated: early pregnancy (EP) and scrotal circumference (SC). Two statistical methods were used to perform GWAS in the validation population, regression and Bayes C, which was applied to confirm those significant associations in regression analysis. SNP that presented *P*-value < 1×10^{-3} in the regression method and Bayes Factor ≥ 3 for Bayes C method were considered significant. Significant SNP (for EP or SC) or their ± 250 kb vicinity region that were in at least one discovery SNP set (G1 or G2) were considered validated. For EP 144 SNP were validated in G1 and 41 in G2, for SC 14 SNP were validated in G1 and 2 SNP in G2, and for EP 21 candidate SNP were detected, i.e., they were in both G1 and G2 SNP sets. These validated SNP were located close to previous reported QTL for reproductive traits or close to genes playing roles in reproductive functions of cattle and other species. These SNP and their ± 250 kb surrounding region are strong candidate regions to be affecting sexual precocity in Nellore and Brahman cattle.

Keywords: *Bos indicus*, discovery population, reproductive traits, validate SNP, tropical beef cattle

3.1 Introduction

Across-breed validation studies are commonly used to validate quantitative trait loci (QTL) for several traits. In Genome-wide association studies (GWAS) those significant markers under an empirical *P*-value for the same or correlated traits in different breeds are likely segregating across-breed and that region may harbor some important genes affecting those populations. Karlsson et al. (2007) used this approach to validate single nucleotide polymorphisms (SNP) markers in different dog breeds. They notice that this strategy was highly efficient to fining mapping across breeds.

The probability of finding common genomic regions controlling correlated traits in different breeds that share common ancestry is expected to be higher than in breeds with very distinct genetic origin. This is because different breeds do not share a recent common ancestor, and as more distinct the origin of the breeds as further will be their common ancestors. This association affects the linkage disequilibrium pattern at long-ranges, and by consequence affects the QTL mapping (Goddard and Hayes, 2009). Thus, as Nellore and Brahman are both *Bos indicus* breeds, and Brahman was originally developed by three base breeds, Gir, Guzerat and Nellore (Briggs, 1980), the likelihood of both breeds share genomic regions associated with associated traits is higher than in unrelated breeds.

Several strategies have been used to conduct across breed validation studies. Pryce et al. (2010) used two dairy cattle breeds to validate QTL for milk production and fertility traits. They distributed Holstein bulls in a discovery population and younger Holstein bulls and Jersey bulls in a validation population. SNP that were detected as significant at an empirical threshold *P*-value in discovery and validation populations were considered validated. Also validating fertility traits in dairy cattle, Höglund et al. (2014) used three dairy cattle breeds to validate genome associations across breeds. They used one breed as discovery population and the other two breeds as validation populations. They argued that using two populations simultaneously to validate significant associations is a powerful strategy to decrease the risk of false positive association.

Genic regions are strong candidate regions to present QTL segregating across related breeds because it is expected that the metabolic pathways in which these genes are involved are conserved across breeds. Also, regions with pleiotropic effect across related traits could result in higher number of true positive validated associations across breeds, because genes that are controlling multiple traits in a breed might preserve similar pattern of pleiotropic effect in another related breed (Saatchi et al., 2014).

The aim of this study was to validate in a Nellore population genomic regions associated with sexual precocity that were reported for Brahman by using as discovery population two SNP sets pre-selected in Brahman, 1) from gene regions previously reported as significant for reproductive traits in Brahman, and 2) from significant associations detected in a meta-analysis study of sexual precocity traits in Brahman.

3.2 Material and Methods

3.2.1 Discovery population

Discovery population was composed by Brahman animals. Phenotypes were provided by Cooperative Research Centre for Beef Genetic Technologies (Beef CRC). Brahman phenotypes included the female traits age when the first *corpus luteum* (CL) was observed (AGECL), first postpartum anoestrus interval (PPAI), ability to ovulate prior to weaning the calf (PW), and the male traits scrotal circumference (SC) measured at 12, 18 and 24 months of age (SC_{12}, SC_{18}, SC_{24}).

The AGECL was defined as the number of days from the heifer birth to the first CL detected. To verify the presence of CL, an ovarian ultrasound was carried out at every 4 to 6 weeks after heifers achieved 200 Kg of weight. PPAI, measured in days, was calculated as the difference between the date of the first observed ovulation postpartum and the date of the calving. To detect the presence of CL, that is an indicator of the ovulation, ovarian ultrasounds were carried out following the first calving. PW, a binary trait was defined as 0 for females that had success to ovulate before weaning her calf or 1 for those females that failed. Details about animals and

phenotypes are described in Johnston et al. (2009), Johnston et al. (2010), Burns et al. (2013) and Corbet et al. (2013).

For females, contemporary groups (CG) were defined by the concatenation of cohort, year of birth and management group information. For males, CG were formed by combining effects of cohort and year of birth. The age of young bulls at recording was considered a covariate for male traits.

Animals were genotyped with the Illumina BovineSNP50 V1 and V2. Genotypes were imputed for high-density panel using Beagle software (Browning and Browning, 2009) and a reference population of representative animals of the Beef CRC population genotyped using the high-density Illumina Bovine HD Assay (Illumina, San Diego, CA, USA), as described by Fortes et al. (2013a). Quality control excluded samples with call rate < 98%, SNPs in non-autosomal regions, with call rate < 85% and minor allele frequency (MAF) < 0.02. The number of SNPs after quality control was 624,676 for AGECL, 625,231 for PPAI, 625,231 for PW, 612,953 for SC_{12}, 612,992 for SC_{18} and 612,981 for SC_{24}. Details about genotypes and imputation are described in Fortes et al. (2013a).

3.2.2 Validation population

Data from Nellore animals were used as validation population. Phenotypic information was obtained from Alliance Nellore dataset. The animals considered in this study were born in eight farms distributed over Midwest, Southeast and Northeast of Brazil. In general, two breeding seasons are applied during the year, where the females are either artificially inseminated or naturally mated. The heifers are exposed in the early breeding season at around 16 months of age. After 60 days of the early breeding season, pregnancy is confirmed and those females that failed in conceiving in the first breeding season had a second opportunity at around 2 years old.

Nellore phenotypes used here are early pregnancy (EP) and scrotal circumference (SC_N). The EP, a binary trait, assumed the value of 2 for heifers that had success in calving before 31 months of age; and 1 for heifers that failed in calving before that age. And SC was the measure, in cm, from the widest point of the scrotum. Details about animals and phenotypes used in this study are described by

Costa et al. (2015) and Irano et al. (2016). Summary statistics of all traits for both populations are presented in Table 1.

Fixed effects were concatenated in CG and included information of herd, year and season of birth, weaning and yearling management groups. For SC, it was excluded phenotypic information that was not in the interval of ± 3 SD from the mean of each CG, and CG with less than 3 animals. For EP, were excluded CG without variability, in which all females had the same categorical response, and CG with less than four animals. For SC the animal's age at the recording was included as a covariate in the model. It was included in the analysis only animals which age at recording ranged from 10 to 24 months of age. A total of 20 animals that presented ages out of this interval were removed.

As Nellore dataset presented both genotyped and non-genotyped animals, for this population phenotypes were pre-corrected for the fixed effect CG to avoid biased fixed effect estimates. A regular mixed animal model (Henderson, 1984) was used to pre-correct phenotypes. Pre-corrected EP was expressed in a continuous scale.

Markers were genotyped using the high-density Illumina Bovine HD Assay (Illumina, San Diego, CA, USA) and GeneSeek Genomic Profiler Indicus HD - GGP75Ki (Neogen Corporation, Lincoln, NE, USA), which contain 777,962 and 74,677 SNPs, respectively (FAPESP, Process Number 2009/16118-5). Imputation of the smaller panel (GGP75Ki) to the high-density panel was performed by using the software FImpute (Sargolzaei, et al., 2014). It were excluded animals with call rate < 0.90, SNPs with MAF < 0.01, call rate < 0.95, GC score < 0.15, Hardy-Weinberg equilibrium test *P*-value < 10^{-5} and in non-autosomal regions. After quality control remained 1,796 females genotyped for 412,876 SNP and 4,261 males genotyped for 512,063 SNP.

Table 1. Summary statistics of the traits early pregnancy (EP) and scrotal circumference (SC) in Nellore cattle, and age at first corpus luteum (AGECL), first postpartum anoestrus interval (PPAI), ability to ovulate prior to weaning the calf (PW), and SC measured at 12, 18 and 24 months of age (SC_{12}, SC_{18}, SC_{24}) in Brahman cattle.

Trait	Number of observations	Mean ± SE

EP (%)	1,849	28.3 ± 1.10
SC (cm)	4,248	26.57 ± 2.56
AGECL (days)	1,007	750.6 ± 142.14
PPAI (days)	629	180.11 ± 108.71
PW (%)	629	52.78 ± 2.00
SC_{12} (cm)	1,098	21.40 ± 2.41
SC_{18} (cm)	1,203	26.51 ± 2.76
SC_{24} (cm)	1,098	29.89 ± 2.86

3.2.3 Cross Validation

To select significant SNP from the discovery population two selection approaches were adopted: 1) SNP in gene regions that were previously associated with sexual precocity traits in a Brahman population (G1), or 2) significant SNP from two independent meta-analyses, one for female and other to male Brahman traits (G2). Details about these meta-analyses are further described.

To select SNP in genic regions (G1), it was firstly verified which genes were associated with sexual precocity traits in Brahman cattle. These genes were selected from previous studies using the same Brahman population used here (Fortes et al., 2012a; Fortes et al., 2012b; Porto-Neto et al., 2015; Fortes et al., 2016; Nguyen et al., 2017). In summary, those studies were either GWAS or transcriptional gene expression studies. SNP inside those gene regions or that were located in a region 250 Kb upstream or downstream from those gene regions were selected for the discovery dataset. A total of 2,754 genes were described in those papers in autosome chromosomes. These genomic regions of these genes were mapped by using the Biomart R package.

SNP selection based on SNP significance of meta-analysis studies (G2) was performed for female (AGECL, PPAI and PW) and male (SC_{12}, SC_{18}, SC_{24}) sexual precocity traits. These meta-analyses were performed separately for males and females to validate each trait (EP and SC) independently in Nellore.

A total of 30 SNP were in both SNP sets, i.e., in genic regions and in significant regions of the meta-analysis study for female traits. And for male traits, just one SNP was in common between these two SNP sets.

A genome mixed model proposed by Kang et al. (2010) was used to estimate SNP effects and their standard errors used in meta-analyses, as follows:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{s}a + \boldsymbol{Zu} + \boldsymbol{\varepsilon}, \quad (1)$$

where **y** is the vector of Brahman phenotypes, AGECL, PPAI, PW, SC_{12}, SC_{18}, or SC_{24}, **X** is an incidence matrix relating CG fixed effects in **β** with the phenotypes in **y**, and the age of the young bull at the measurement as a covariate for SC, **s** is a vector with genotypes coded as 0, 1 or 2 according to the number of B allele copies, a is a vector containing allelic substitution effects, **Z** is the incidence matrix of polygenic random effects of the animals in **u** and **ε** is the vector of residuals. **u** and **ε** followed normal distribution with $\mathbf{u} \sim N(0, \mathbf{G}\sigma^2_a)$ and $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma^2_e)$, respectively, and **G** is the genomic relationship matrix, that relates all individuals, calculated as in the first method described in VanRaden (2008), σ^2_a is the additive genetic variance, **I** is an identity matrix and σ^2_e is the residual variance. SNP effects were calculated by using the SNP & Variation Suite (SVS) software (Release 8.3.0, Golden Helix, Inc., 2014).

A multi-trait approach described by Bolormaa et al. (2014) was used to perform the meta-analyses. This method is a statistic test that follows a χ^2 distribution with n degrees of freedom, and n is the number of traits considered in meta-analysis. The χ^2 statistic is calculated as follows:

$$\textit{Multi-trait } \chi^2 = \mathbf{t'_i}\, \mathbf{V^{-1}} \mathbf{t_i}, \quad (2)$$

where $\mathbf{t'_i}$ is the transpose vector of $\mathbf{t_i}$, and $\mathbf{V^{-1}}$ is a *3 x 3* inverse matrix of the correlation between the *t*-values of the three traits for female and male across all SNP markers in common among these traits. The vector containing the *t*-values of the i^{th} SNP effects estimated by each independent GWAS (t_i) was computed as:

$$\mathbf{t_i} = a_i / SE(a_i), \quad (3)$$

where $SE\ (a_i)$ is the standard error of the SNP effect vector a_i.

SNP with an empirical *P*-value ≤ 10^{-3} were considered significant in the meta-analysis of the discovery population. The total number of significant SNPs detected in meta-analyses was 747 SNPs for female traits and 10 SNPs for male traits (*P*-value ≤ 0.001). This large difference may be explained because the most significant SNP for Brahman were located on chromosome X, which was removed in the current analysis.

3.2.4 Statistical Methods

To perform GWAS in Nellore validation population (with the whole HD panel) two statistical methods were used, regression (Zhang et al., 2010) and Bayes C (Habier et al., 2011), which model was implemented as:

$$y = 1\mu + \sum_{i=1}^{n} g_i b_i \delta_i + e, \quad (4)$$

Where *y* is the vector containing pre-corrected phenotypes (corrected as previously described) for Nellore traits, EP and SC, *1* is a vector of ones, μ is the overall mean, g_i is the vector with the genotypes of the animals (coded as 0, 1 or 2, according to the number of B allele copies) for the i^{th} SNP effect in b_i, and δ_i is an indicator variable, which takes value 0 or 1, and is sampled from a binomial distribution with parameters *n* and π, where *n* is the total number of SNPs and π is the proportion of SNPs with null effect in the model. It was assumed that π followed a prior beta distribution with parameters $\alpha = 10^8$ and $\beta = 10^{10}$, which in practice is equivalent to fix $\pi = 0.99$ (Legarra et al., 2014). For the variance of SNP effects (σ^2_g) and the residual variance (σ^2_e) it was assumed a scaled inverse chi-squared prior distribution.

Bayes C analyses were performed by using GS3 software (Legarra et al., 2014). It was ran a single Markov chain Monte Carlo with 500,000 iterations, a burn-in period of 50,000, and a thinning interval of 50 iterations.

Model used in regression analyses was similar to that presented in Equation 1, except that the vector containing the fixed effects and its incidence matrix ($X\beta$) were replaced by a vector of ones and the mean (1μ), as described in the Equation 4, with the same assumption for the distribution of random effects described in Equation 1. This model was implemented in the GAPIT software (Zhang et al., 2010), and consists in a regular mixed linear model with a genomic relationship matrix.

3.2.5 SNP validation criteria

To be considered validated a SNP marker should be detected as significant by both methods regression and Bayes C for each trait. We adopted a similar procedure described by Pryce et al. (2010), to determine SNP significance, SNP presenting an empirical threshold *P*-value ≤ $1x10^{-3}$ in regression analysis and Bayes factor (BF) ≥ 3 (Varona et al., 2001) in Bayes C analyses were considered significant. Bayes Factor

was used as an alternative to *P*-value to calculate SNP significance, as in Bayesian methods *P*-values are not available. It was calculated as:

$$BF = \frac{\left(\frac{p}{1-p}\right)}{\left(\frac{\pi}{1-\pi}\right)} \quad (7)$$

where p is *a posteriori* probability of an SNP present a non-zero effect and π is *a priori* probability of an SNP to be included in analysis.

SNP that were in common in both methods, i.e., were in genic regions (G1) and were significant in Brahman meta-analyses (G2) were considered validated. Also SNP in the surrounding ± 250 Kb (upstream and downstream) from these regions were clustered in G1 or G2 for each trait (Figure 1). Common SNP in G1 and G2 were considered strong candidates affecting sexual precocity in beef cattle. Figure 1 summarizes how the validation criterion was performed.

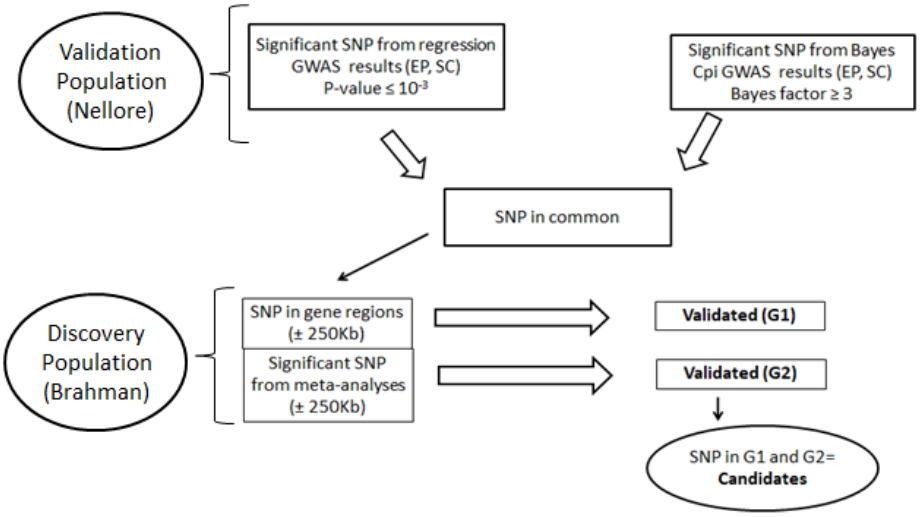

Figure 1. Scheme of across breed SNP validation for Brahman and Nellore populations. EP = Early pregnancy and SC = Scrotal circumference, both measured in Nellore, G1 = Significant SNP in Nellore in genic regions important for Brahman, G2 = Significant SNP in Nellore close to significant SNP from Brahman meta-analyses.

3.3 Results and Discussion

A total of 302 and 40 SNP were in common in both methods, linear mixed regression and Bayes C, for EP and SC, respectively. For EP a total of 144 SNP were validated in G1 and 41 SNP were validated in G2. For SC a total of 14 and 2 SNP were validated in G1 and G2 respectively, with a *P*-value ≤ 10^{-3}.

The low number of validated SNP for SC in G2 may be explained due to the most significant SNP for this trait in Brahman in all evaluated ages are located on X chromosome, which was not observed for Nellore (data not shown).

3.3.1 Validated SNP for SC

The most significant SNP (*P*-value ≤ 10^{-3}) for each chromosome validated in G1 and G2 for SC were distributed over the chromosomes 11, 12, 15, 16, 17, 19, 23 and 29 are presented in Table 2. Some of the validated SNP were close to more than one candidate gene, i.e., genes in a distance of ± 250 Kb from the validated SNP.

Table 2. Validated top SNP for SC in G1 and G2 (*P*-value ≤ 10^{-3}).

SNP	BTA[1]	Position[1] (Mb)	*P*-value	Gene[2] (symbol[3])/ SNP[4]
		G1		
BovineHD1100003302	11	9.35	$2.4x10^{-4}$	ENSBTAG00000012332 (*C2orf49*)
BovineHD1200006565	12	21.81	$2.3x10^{-4}$	ENSBTAG00000020612 (*NEK3*) ENSBTAG00000011647 (*SLC25A15*) ENSBTAG00000044105 (*FOXO1*)
BovineHD1500007531	15	28.06	$1.2x10^{-4}$	ENSBTAG00000002650 (*ZPR1*) ENSBTAG00000019770 (*APOA4*) ENSBTAG00000012398 (*APOC3*) ENSBTAG00000019764 (*APOA5*)
BovineHD1700015592	17	54.94	$4.3x10^{-5}$	ENSBTAG00000000908 (*HCAR1*)
BovineHD1900012747	19	45.37	$3.4x10^{-6}$	ENSBTAG00000013534 (*GFAP*) ENSBTAG00000006056 (*HEXIM1*) ENSBTAG00000024974

				(*HEXIM2*)
BovineHD2300015274	23	15.69	$7.3x10^{-5}$	ENSBTAG00000012384 (*TFEB*) ENSBTAG00000008977 (*TOMM6*) ENSBTAG00000010100 (*MED20*) ENSBTAG00000010106 (*CCND3*) ENSBTAG00000011339 (*TAF8*) ENSBTAG00000015301 (*MRPS10*)
BovineHD2900002704	29	9.16	$8.8x10^{-5}$	ENSBTAG00000019995 (*HIKESHI*) ENSBTAG00000007847 (*EED*)
		G2		
BovineHD1600000479	16	1.86	$1.8x10^{-4}$	BovineHD1600000503

[1]SNP were mapped to the UMD3.1 Bovine Assembly.

[2]Gene = Important genes for Brahman reproductive traits (± 250 Kb).

[3]Gene symbols were researched at Ensembl website (Cow UMD3.1).

[4]SNP = Closest significant SNP of the Brahman meta-analysis (± 250 Kb).

According to the reference Brahman studies, in which the genes were collected for selection in G1, the candidate genes found here close to the validated SNPs were either differentially expressed in pituitary gland of Brahman heifers, comparing pre-pubertal and pos-pubertal heifers (Nguyen et al., 2017), differentially expressed in pituitary gland and grouped in pathways related to reproduction (Nguyen et al., 2017) or were differentially expressed in the hypothalamus of pos versus pre pubertal Brahman heifers (Fortes et al., 2016). The axis hypothalamus-pituitary gland is responsible to control the production of estrogens hormones in ovaries. These hormones have an important role in the preparation of uterus for the reproduction.

3.3.2 Validated SNP for EP

A total of 144 SNP were validated for EP in G1. They were distributed over chromosomes 2, 3, 4, 5, 6, 7, 8, 11, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24 and 27. Most of them were located on BTA3 (14.5%). For G2, validated SNP were located on

chromosomes 6, 8, 14, 15, 17, 21 and 24, and the most of them (46.3%) were on BTA21.

Table 3 present the most significant validated SNP for each chromosome for EP in G1 and G2. Some of the candidate genes close to these validated SNP were differentially expressed in pituitary gland of Brahman heifers, comparing pre-pubertal and pos-pubertal heifers (Nguyen et al., 2017), differentially expressed in pituitary gland and grouped in pathways related to reproduction (Nguyen et al., 2017), differentially expressed in the hypothalamus of pos versus pre pubertal Brahman heifers (Fortes et al., 2017) or were differentially expressed in ovary of Brahman females (Nguyen et al., 2017).

Some of the validated SNP presented in Table 3 were validated for G1 and G2 (represented in bold). These SNP were some of the strong candidates SNP detected here for EP.

Table 3. Validated top SNP for EP in G1 and G2 (*P*-value ≤ 10^{-3}).

SNP	BTA[1]	Position[1] (Mb)	*P*-value	Gene[2] (symbol[3])/ SNP[4]
		G1		
BovineHD0200022985	2	79.97	$8.7x10^{-4}$	ENSBTAG00000007867 (*STAT1*) ENSBTAG00000001125 (*MYO1B*)
BovineHD0300002093	3	6.66	$1.1x10^{-4}$	ENSBTAG00000005976 (*HSD17B7*)
BovineHD0400028801	4	102.42	$3.9x10^{-4}$	ENSBTAG00000015802 (*CREB3L2-201*)
BovineHD0500031072	5	107.89	$3.6x10^{-4}$	ENSBTAG00000038415 (*SLC6A12*) ENSBTAG00000014525 (*SLC6A13*) ENSBTAG00000020472 (*KDM5A*)
BovineHD0600034574	6	1.74	$5.1x10^{-5}$	ENSBTAG00000044082 (*MARCH1*)/ BovineHD0600000418
BovineHD0700001070	7	3.77	$1.6x10^{-4}$	ENSBTAG00000009975 (*PBX4*) ENSBTAG00000007812 (*NDUFA13*)
BovineHD0800009146	8	30.10	$4.3x10^{-4}$	ENSBTAG00000027442 (*NFIB*)

BovineHD1100024980	11	86.98	$5.6x10^{-4}$	ENSBTAG00000004259 (*HPCAL1*)
BovineHD1300022820	13	78.85	$2.0x10^{-4}$	ENSBTAG00000014554 (*SNAI1*) ENSBTAG00000010135 (*TMEM189*)
BovineHD1400002554	14	9.18	$9.6x10^{-5}$	ENSBTAG00000001156 (*ST3GAL1*) ENSBTAG00000007823 (*TG*)
BovineHD1500007541	15	28.10	$1.8x10^{-4}$	ENSBTAG00000002650 (*ZPR1*) ENSBTAG00000019770 (*APOA4*) ENSBTAG00000012398 (*APOC3*) ENSBTAG00000007196 (*TAGLN*) ENSBTAG00000019764 (*APOA5*)
BovineHD1600013840	16	49.84	$6.5x10^{-5}$	ENSBTAG00000021576 (*LMOD1*)
BovineHD1700005155	17	17.82	$7.9x10^{-5}$	ENSBTAG00000001580 (*CLGN*)
BovineHD1800016044	18	54.75	$1.2x10^{-4}$	ENSBTAG00000013084 (*NAPA*)
BovineHD1900013552	19	48.75	$7.0x10^{-4}$	ENSBTAG00000000056 (*STRADA*)
BovineHD2100000944	21	5.30	$1.3x10^{-5}$	ENSBTAG00000007357 (*CHSY1*)/ BovineHD2100001004
BovineHD2200000079	22	0.37	$7.8x10^{-5}$	ENSBTAG00000014547 (*PGAM2*) ENSBTAG00000012073 (*VOPP1*)
BovineHD2300007469	23	27.26	$4.9x10^{-5}$	ENSBTAG00000008794 (*ATF6B*) ENSBTAG00000006864 ENSBTAG00000037533 (*C4A*) ENSBTAG00000007450 (*C2*)
BovineHD2400002342	24	8.37	$3.8x10^{-4}$	ENSBTAG00000006726 (*CCDC102B*)/ BovineHD2400002368
BovineHD2700005243	27	18.18	$8.8x10^{-4}$	ENSBTAG00000011257 (*ASAH1*)

G2

BovineHD0800012292	8	4.12	$4.0x10^{-4}$	BovineHD0800012229
BovineHD1400004742	14	16.71	$2.6x10^{-4}$	BTA-122375-no-rs
BovineHD1500008343	15	31.06	$1.1x10^{-4}$	BovineHD1500008365
BovineHD1700009553	17	34.48	$2.7x10^{-4}$	BovineHD1700009509

[1]SNP were mapped to the UMD3.1 Bovine Assembly.

[2]Gene = Important genes for Brahman reproductive traits (interval of ± 250 Kb).

[3]Gene symbols were researched at Ensembl website (Cow UMD3.1).

[4]SNP = Closest significant SNP of the Brahman meta-analysis (± 250 Kb).

SNP in bold were in common between G1 and G2.

3.3.3. Candidate SNP

It was found a total of 21 candidate SNP for female sexual precocity, i.e., SNP validated for both, G1 and G2. These SNP were located on chromosomes 6, 8, 14, 15, 21 and 24, and most of them were on BTA6 (38%). Table 4 presents the most significant candidate SNP by chromosome for EP. For SC, no candidate SNP were found.

Table 4. Candidate top SNP per chromosome for EP (*P*-value ≤ 10^{-3}).

SNP	BTA[1]	Position (Mb)[1]	*P*-value	Gene[2] (symbol)[3]/ SNP[4]
BovineHD0600034574	6	1.74	$5.1x10^{-5}$	ENSBTAG00000044082 (*MARCH1*)/ BovineHD0600000418
BovineHD0800009146	8	30.10	$4.3x10^{-4}$	ENSBTAG00000027442 (*NFIB*)/ BovineHD0800009169
BovineHD1400004742	14	16.71	$2.6x10^{-4}$	ENSBTAG00000009394 (*NSMCE2*) BTA-122375-no-rs
BovineHD1500008850	15	32.60	$2.8x10^{-4}$	ENSBTAG00000019246 (*SC5D*)/ BovineHD1500008844
BovineHD2100000944	21	5.30	$1.3x10^{-5}$	ENSBTAG00000007357 (*CHSY1*)/ BovineHD2100001004
BovineHD2400002342	24	8.37	$3.8x10^{-4}$	ENSBTAG00000006726 (*CCDC102B*)/ BovineHD2400002368

[1]SNP were mapped to the UMD3.1 Bovine Assembly.

[2]Gene = Closest gene, important for Brahman reproductive traits (± 250 Kb).

[3]Gene symbols were researched at Ensembl website (Cow UMD3.1).
[4]SNP = Closest significant SNP of the Brahman meta-analysis (± 250 Kb).

These results showed that the SNP that were validated for Nellore were not the same for Brahman for none of the Brahman traits evaluated here, but they were located in the vicinity (± 250 Kb) of the significant SNP for Brahman. Furthermore, the most significant candidate SNP of each chromosome that was significant for EP in Nellore was generally significant just for AGECL in Brahman, which suggests that EP and AGECL are more correlated than EP and PPAI or PW. The 21 SNP presented in Table 4 are plausible candidates to be affecting the sexual precocity in Nellore and Brahman, and are likely in LD with QTL segregating in both breeds, which is controlling sexual precocity phenotypes.

3.3.4 SNP in QTL regions

Some of the validated SNP here were previously mapped in the same regions in other studies that used the same populations used here for other reproductive traits or traits related with reproduction. Costa et al. (2015) found SNP associated with age at first calving and heifer rebreeding on chromosome 3 at 6 Mb in Nellore. In chromosome 5 (Table 3) the top SNP "BovineHD0500031072" (107.89 Mb) was located close to a significant signal (108 Mb) for the trait inhibin in Brahman bulls (Fortes et al., 2013b). Dias et al. (2015) found three lipid metabolism genes that were also associated with sexual precocity in Nellore cattle and were mapped close to significant regions found here: genes *UCP-1* (BTA17 at 17 Mb), *GH-1* (BTA19 at 48 Mb) and *TNF* (BTA23 at 27 Mb). In chromosome 21 the top SNP "BovineHD2100000944" (Table 3) was located in a distance of 395,444 bp away from a SNP that was significant for pregnancy outcome after fixed-time AI of Brahman heifers (Porto-Neto et al., 2015). These authors verified that this SNP was in an intronic region of the gene *LRRK1*.

Other different populations presented QTL related to sexual precocity in regions close those reported here (less than 0.5 Mb of distance). Cochran et al. (2013) found a SNP on BTA3 at 6 Mb associated with daughter pregnancy rate, heifer conception rate and cow conception rate in Holstein. This SNP was mapped into the gene

HSD17B7, which was also detected in the genes selected as important for Brahman (Table 3). Also these authors found on BTA23 at 27 Mb another SNP associated with those traits in the gene *NFKBIL1*, which is 294.03 bp away from the top SNP found in this region. Buzanskas et al. (2017) found a SNP on BTA5 at 107 MB associated with scrotal circumference at 420 days of age in Canchim cattle. On BTA14 at 16 Mb Mota et al. (2017) found a SNP window associated with age at first calving in Nellore. Parker Gaddis et al. (2016) found on chromosome 15 at 31-38 Mb an association between some markers in this region and the traits daughter pregnancy rate, cow conception rate and heifer conception rate in Holstein cattle. Oliveira Júnior et al. (2017) working with early pregnancy in Nellore found genomic 1 Mb windows on BTA18 at 54 - 56 Mb explaining more than 1% of the additive genetic variance for this trait. McClure et al. (2010) found QTL associated with scrotal circumference in Angus cattle on chromosome 19 at 45.1 Mb, which was close to the SNP BovineHD1900012747 at 45.3 Mb. Also, Sahana et al. (2010) found SNP associated with fertility index and days from first to last insemination in Holstein in this genomic region. In Brangus cows Peters et al. (2013) found a SNP associated with first service conception (FSC) on chromosome 19 at 49 Mb, which is close to a significant SNP detected here at 48.7 Mb. Fontanesi et al. (2014) also found on BTA19 at 48.7 Mb markers associated with fertility index in Holstein.

Some of the genes presented in Table 2 (male results) were related to reproductive events. On BTA11 the gene *C11H2orf49*, homologue of *C2orf49*, was down-regulated in uterus of repeat breeder cows compared with non-repeat breeder Japanese Black cows (Hayashi et al., 2017). On BTA12, the gene *NEK3* is associated with prolactin receptor in humans (Miller et al., 2005). Devi and Halperin (2014) discuss the role of prolactin hormone in some reproduction process in rodents and human. This gene was down-regulated in F3 of mice males exposed to an organochlorine insecticide with estrogenic properties, promoting a reduction in spermatogonia numbers (Gely-Pernot et al., 2018). Gene *FOXO1* showed an increasing of expression in the endometrium of women with polycystic ovary syndrome (Kohan et al., 2010). Lappas et al. (2009) suggested that *FOXO1* is related to gestation maintenance in humans. The expression of this gene could be observed

in spermatogonia and granulosa cells of several mammal species, as mouse, rat, dog, cow, rhesus, chimpanzee and human (Tarnawa et al., 2013).

On BTA15 the *ZPR1* gene is associated with embryo development; Gangwani et al. (2005) observed that mice embryos with disruption in this gene died in early embryonic development stage. *APOA4* gene was associated with follicular development in pigs oocyte (Ji et al., 2010). It was observed an increasing of this gene protein expression in the later rat embryonic development stages (Dihazi et al., 2015), and those authors suggest an association of this protein with steroid metabolic process and cholesterol synthesis. Also *APOC3* gene is related to cholesterol metabolism and Tabatabaie et al. (2011) raised the hypothesis of this gene can be affect gonadal maturity in rats. On BTA17, *HCAR1* gene is commonly expressed in adipocytes and was involved with the regulation of lipolysis in Holstein cows during the transition period from late pregnancy to lactation (Weber et al., 2016).

On BTA19 *GFAP* gene is associated with the regulation of estrous cycle in cows, mouse (Garcia-Segura et al., 1994; Stone et al., 1998). *HEXIM1* and *HEXIM2* genes are associated with cell proliferation, and participate of mouse embryo development (Yik et al., 2005). Gene *TFEB* was grouped in clusters functionally associated with placentation on the branch ancestral to Hominidae, i.e., human, chimp, gorilla and orangutan (Crosley et al., 2012). This gene was also reported as playing an important role in normal placenta vascularization in mice (Steingrímsson et al., 1998). *CCND3* gene was down-regulated in endometrium of patients with endometriosis (Ohlsson Teague et al., 2009). This gene was listed by Choudhury and Kanapp (2001) as associated with fertility, reproduction and development in human and rodents. This gene presented significant expression on bovine granulosa cells (Shimizu et al., 2013). *TAF8* gene was associated with embryo development in bovine embryos (Al Naib et al., 2012). *MRPS10* gene is associated with luteolysis in ovine corpus luteum (Xu et al., 2017). On BTA29 *HIKESHI* gene is associated with cellular stress response, as promoted by heat shock (Kose et al., 2012). The gene *EED* is associated with female puberty in rats (Lomniczi et al., 2013).

Also in Table 2 (female results) some genes were previously related to reproductive processes. On BTA2, *STAT1* gene is involved in apoptotic process in female granulosa cells (Benifla et al., 2002). Also this gene was reported to be

activated during the spermatozoa capacitation and after the fertilization, acting in the male pronucleus (Bastián et al., 2007). The mRNA of *MYO1B* gene was transcript in man reproductive tissues which sample sperms failed to promote pregnancy (García-Herrero et al., 2010). This gene expression was abundant in fish eggs after photo stimulation treatment to induce ovulation (Bonnet et al., 2007). Ortega et al. (2015) in a GWAS identify a significant SNP associated with heifer conception rate inside the gene *HSD17B7* in Holstein cows. This gene, located on BTA3 in bovine, is associated with production of stradiol, an important female hormone, in corpus luteum and with embryo development in mice (Gibori et al., 2009). This gene was related to be involved in both sex steroid hormones synthesis, in human (Osuch et al., 2012).

On BTA4 the gene *CREB3L2-201* was expressed in human placenta and in fetal tissues (Storlazzi et al., 2004). On BTA5, the *SLC6A13* gene was differentially expressed in mouse blastocysts (Giritharan et al., 2010). This gene was down-regulated in the presence of estrogen in the posterior part of adult female rats hypothalamus (Xu et al., 2008). *KDM5A* gene is involved in the regulation of spermiogenesis in mice (Lambrot et al., 2008). This gene was identified as a regulator of bovine milk fatty acid content (Pegolo et al., 2017). Boschiero et al. (2012) proposed this gene as a candidated gene associated with fat deposition in chickens.

On BTA6 the *MARCH1* gene was associated with semen production traits in Chinese Holstein bulls (Liu et al., 2017). Also this gene is associated with the Major Histocompatibility Complex I (MHC I), acting in the process of recognition of antigen presentation (Wilson et al., 2018) and was found in a candidate region affecting female sexual precocity (Table 5). On BTA7, *PBX4* gene is located in mouse testis and participates of spermatogenesis (Wagner et al., 2001). *NDUFA13* gene was associated with asthenozoospermia, a male pathology that reduce spermatozoa motility in mouse, causing infertility (Yang et al., 2017). This gene was overexpressed in blastocysts of super ovulated heifers (Gad et al., 2011).

Transcription of gene *NFIB*, on BTA8, is essential for lung and brain development in mice embryos (Steele-Perkins et al., 2005). This gene was located in a candidate region for female sexual precocity in Table 5. On BTA13, *SNAI1* is

essential for the normal cardiovascular system development (Lomelí et al., 2009), and *TMEM189* was upregulated in the porcine endometria in days 12 and 16 of gestation versus the other days of estrous cycle (Kiewisz et al., 2014). On BTA14 the *ST3GAL1* was identified in bovine cervical tissue (Pluta et al., 2012). *TG* gene was identified in several GWAS as associated with fat deposition in cattle (Gan et al., 2008; Anton et al., 2011). On BTA15, *TAGLN* was associated with endometriosis (dos Santos Hidalgo et al., 2011). Also this gene was detected in reproductive tract of pregnant mice and in hens (Kaftanovskaya et al., 2015; Riou et al., 2015). On BTA16 the gene *LMOD1* was regulated by androgen and progesterone receptors in the human endometrium affecting the decidualizing process (Cloke et al., 2008).

On BTA17, *CLGN* is associated with the capacity of the sperm binding mice zona pellucida during fertilization (Yamaguchi et al., 2006). *CLGN* was also expressed in porcine spermatozoa (Kempisty et al., 2008). This gene was enriched in the blastocysts of unstimulated heifers compared with those that were superovulated (Gad et al., 2011). On BTA18, mutant *NAPA* gene caused embryo lethality and infertility in mouse, due to a defect in acrosome reaction (Chae et al., 2004; Bátiz et al., 2009). *STRADA* gene, on BTA19, participates of neuronal embryo formation (Orlova et al., 2010; Crino, 2015). The gene *CHSY1* (BTA21) was overexpressed in bovine cumulus cells 6 hour after the LH surge (Assidi et al., 2010) and was located in a candidate region of this study (Table 5). *PGAM2* protein, gene on BTA22, was differentially expressed in asthenozoospermic sperm from human sperm (Jodar et al., 2017).

ATF6B gene on BTA23 was up-regulated in porcine fertilized blastocysts (Xu et al., 2014). This gene was associated with grown traits in bovine (Mao et al., 2016) and with reproductive traits in porcine (Liu et al., 2018). ENSBTAG00000006864 gene was located in a region in association with serum IGF-1 concentration in dairy cows (Gobikrushanth et al., 2018). *C4A* gene was downregulated in e low infertility risk group patients compared to intermittent infertility risk group of boys with cryptorchidism (Hadziselimovic et al., 2009). This gene is associated with recurrent spontaneous abortions in humans (Laitinen et al., 1991). *C4A* and *C2* genes are associated with the Major histocompatibility complex (MHC), which is involved in inflammatory and immune responses (Perlmutter et al., 1986; Lokki et al., 2001). On

BTA27 the *ASAH1* gene regulates the function of a steroidogenic fator in human adrenochotical cells (Lucki et al., 2012). This gene was related with embryo lethality (Li et al., 2002; Eliyahu et al., 2007).

Combining genes from both Tables 1 and 2 in an enrichment analysis using David (v. 6.7) the top 3 clusters were grouped in pathways associated with the regulation of transcription process and with lipid metabolism: “transcription regulator activity” (Enrichement Score, ES = 2.38 and *P*-value = 1.3E-3), “transcription regulation” (ES = 2.31 and *P*-value = 3.7E-3), and “Plasma lipoprotein particle remodeling” (ES = 2.08 and *P*-value = 6.1E-4). These processes are involved in a variety of biological events, including the reproductive.

In summary, we could validate genomic regions in a Brazilian Nellore cattle population that were previous known as affecting sexual precocity traits in an Australian Brahman cattle population. Also, we found new loci associated with sexual precocity in both breeds. A total of 21 SNP were found as strong candidates to be affecting sexual precocity in females of Brahman and Nellore. SNP that were significant for Brahman generally were not the same for Nellore, however they were close to significant SNP for Nellore. The reason for this could be the difference in the linkage disequilibrium pattern between these breeds, i.e., different mutations segregating across breeds can be indicating the same gene or genomic region. The different definition of female phenotypes used for both breeds may also be another explanation for the found results.

For male, fewer regions could be validated than for female and no candidate SNP was found. This could be explained because for SC in Brahman the X chromosome has an important effect, which could not be observed in Nellore population (results not shown), which led us to do not use this chromosome in our analysis.

Some genes that were found close to validated SNP seem to be affecting sexual precocity traits in females and males, suggesting the action of pleiotropic effects. Results also showed that some of the candidate SNP from both validation sets (G1 and G2), were located near genes playing roles in reproductive events in several mammal species, including bovine. Therefore, those genes are good candidates to be affecting sexual precocity in tropical beef cattle. As expected for

complex traits, sexual precocity depends on the effect of several QTL, with small effect each, to produce the final phenotype. The genomic regions described here seem to be related to the expression of the sexual precocity in Nellore and Brahman.

3.4 Conclusion

Some genomic regions that are affecting sexual precocity in Brahman seem to be also affecting it in Nellore. The strategy of using genic and significant regions from meta-analysis studies was efficient to validate genomic regions across breeds. The presence of genes that play roles in reproduction function close to the validated regions is an extra evidence to consider those regions in future studies as candidate regions. Further studies using the same phenotype across different populations could obtain higher power of QTL detection.

3.5 References

Al Naib A, Mamo S, Lonergan P (2012) Investigation of a Preferentially Upregulated Gene Cluster in Day 7 Bovine Embryos Derived from RNA Sequencing Data. **Reproduction, Fertility and Development** 25:256-256.

Anton I, Kovacs K, Holló G, Farkas V, Lehel L, Hajda Z, Zsolnai A (2011) Effect of leptin, DGAT1 and TG gene polymorphisms on the intramuscular fat of Angus cattle in Hungary. **Livestock Science** 135:300-303.

Assidi M, Dieleman SJ, Sirard MA (2010) Cumulus cell gene expression following the LH surge in bovine preovulatory follicles: potential early markers of oocyte competence. **Reproduction** 140:835-852.

Bastián Y, Zepeda-Bastida A, Uribe S, Mujica A (2007) In spermatozoa, Stat1 is activated during capacitation and the acrosomal reaction. **Reproduction** 134:425-433.

Bátiz LF, De Blas GA, Michaut MA, Ramírez AR, Rodríguez F, Ratto MH, Oliver C, Tomes CN, Rodríguez EM, Mayorga LS (2009) Sperm from hyh mice carrying a point mutation in αSNAP have a defect in acrosome reaction. **PLoS One** 4:e4963.

Benifla JL, Sifer C, Bringuier AF, Blanc-Layrac G, Camus E, Madelenat P, Feldmann G (2002) Induced apoptosis and expression of related proteins in granulosa cells from women undergoing IVF: a preliminary study. **Human Reproduction** 17:916-920.

Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, Tier B, Savin K, Hayes BJ, Goddard ME (2014) A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for stature, Fatness and Reproduction in Beef Cattle. **PLOS Genetics** v. 10.

Bonnet E, Fostier A, Bobe J (2007) Microarray-based analysis of fish egg quality after natural or controlled ovulation. **Bmc Genomics** v. 8.

Boschiero C, Jorge EC, Ninov K, Nones K, do Rosário MF, Coutinho LL, Moura ASA (2013) Association of IGF1 and KDM5A polymorphisms with performance, fatness and carcass traits in chickens. **Journal of applied genetics** 54:103-112.

Briggs HM (1980) **Modern breeds of livestock**. The Macmillian Publishing Company. 139 p.

Browning BL, Browning SR (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. **American Journal of Human Genetics** 84:210-223.

Burns BM, Corbet NJ, Corbet DH, Crisp JM, Venus BK, Johnston DJ, Li Y, Mcgowan, MR, Holroyd RG (2013) Male traits and herd reproductive capability in tropical beef cattle. 1. Experimental design and animal measures. **Animal Production Science** 53:87-100.

Buzanskas ME, do Amaral Grossi D, et al. (2017) Candidate genes for male and female reproductive traits in Canchim beef cattle. **Journal of animal science and biotechnology** v. 8.

Chae TH, Kim S, Marz KE, Hanson PI, Walsh CA (2004) The hyh mutation uncovers roles for αSnap in apical protein localization and control of neural cell fate. **Nature genetics** 36:264-270.

Choudhury SR, Knapp LA (2001) Human reproductive failure II: immunogenetic and interacting factors. **Human reproduction update** 7:135-160.

Cochran SD, Cole JB, Null DJ, Hansen PJ (2013) Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. **BMC genetics** v. 14.

Corbet NJ, Burns BM, Johnston DJ, Wolcott ML, Corbet DH, Venus BK, Li Y, Mcgowan, MR, Holroyd RG (2013) Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits. **Animal. Production Science** 53:101-113.

Costa RB, Camargo GM, et al. (2015) Genome-wide association study of reproductive traits in Nellore heifers using bayesian inference. **Genetics Selection Evolution** v. 47.

Crino PB (2015) mTOR signaling in epilepsy: insights from malformations of cortical development. **Cold Spring Harbor perspectives in medicine** 5:a022442.

Crosley EJ, Elliot MG, Christians JK, Crespi BJ (2013) Placental invasion, preeclampsia risk and adaptive molecular evolution at the origin of the great apes: evidence from genome-wide analyses. **Placenta** 34:127-132.

Devi YS, Halperin J (2014) Reproductive actions of prolactin mediated through short and long receptor isoforms. **Molecular and cellular endocrinology** 382:400-410.

Dias MM, Souza FRP, et al. (2015). Study of lipid metabolism-related genes as candidate genes of sexual precocity in Nellore cattle. **Genetics and Molecular Research** 14:234-243.

Dihazi GH, Mueller GA, Asif AR, Eltoweissy M, Wessels JT, Dihazi H (2015) Proteomic characterization of adrenal gland embryonic development reveals early initiation of steroid metabolism and reduction of the retinoic acid pathway. **Proteome science** v. 13.

Durinck S, Spellman P, Birney E, Huber W (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. **Nature Protocols** 4:1184-1191.

Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W (2005) BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. **Bioinformatics** 21:3439-3440.

Eliyahu E, Park JH, Shtraizent N, He X, Schuchman EH (2007) Acid ceramidase is a novel factor required for early embryo survival. **The FASEB journal** 21:1403-1409.

Fontanesi L, Calò DG, Galimberti G, Negrini R, Marino R, Nardone A, Ajmone-Marsan P, Russo V (2014) A candidate gene association study for nine economically important traits in Italian Holstein cattle. **Animal genetics** 45:576-580.

Fortes MRS, Lehnert SA., Bolormaa S, Reich C, Fordyce G, Corbet NJ, Whan V, Hawken RJ, Reverter A (2012a) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science** 52:143-150.

Fortes MR, Reverter A, Hawken RJ, Bolormaa S, Lehnert SA (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** v. 87.

Fortes MRS, Kemper K, et al. (2013a) Evidence for pleiotropism and recent selection in the PLAG1 region in australian beef cattle. **Animal Genetics** 44:636-647.

Fortes MRS, Reverter A, Neto LP, Kelly M, Moore SS, Lehnert SA (2013b) Genetic markers associated with male reproductive traits across 2 beef cattle breeds: Brahman and tropical composite. In: Association for Advancement of Animal Breeding and Genetics. **Proceddings…** Australia, 20:389-392.

Fortes MRS, Nguyen LT, et al. (2016) Transcriptome analyses identify five transcription factors differentially expressed in the hypothalamus of post-versus prepubertal Brahman heifers. **Journal of Animal Science** 94:3693-3702.

Gad A, Besenfelder U, et al. (2011) Effect of reproductive tract environment following controlled ovarian hyperstimulation treatment on embryo development and global transcriptome profile of blastocysts: implications for animal breeding and human assisted reproduction. **Human Reproduction** 26:1693-1707.

Gan QF, Zhang LP, Li JY, Hou GY, Gao X, Ren HY, Chen JB, Xu SZ (2008). Association analysis of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. **Journal of Applied Genetics** 49:251-255.

Gangwani L, Flavell RA, Davis RJ (2005) ZPR1 is essential for survival and is required for localization of the survival motor neurons (SMN) protein to Cajal bodies. **Molecular and cellular biology** 25:2744-2756.

García-Herrero S, Meseguer M, Martínez-Conejero JA, Remohí J, Pellicer A, Garrido N (2010) The transcriptome of spermatozoa used in homologous intrauterine insemination varies considerably between samples that achieve pregnancy and those that do not. **Fertility and sterility** 94:1360-1373.

Garcia-Segura LM, Luqín S, Párducz A, Naftolin F (1994) Gonadal hormone regulation of glial fibrillary acidic protein immunoreactivity and glial ultrastructure in the rat neuroendocrine hypothalamus. **Glia** 10:59-69.

Gely-Pernot A, Hao C, Legoff L, Multigner L, D'Cruz SC, Kervarrec C, Jégou B, Tevosian S, Smagulova F (2018) Gestational exposure to chlordecone promotes transgenerational changes in the murine reproductive system of males. **Scientific reports** v. 8.

Gibori G, Shehu A, et al. (2009) The Large Luteal Cells-Derived PRAP/HSD17B7: An Enzyme with a Split Personality. **Biology of Reproduction** v. 81.

Giritharan G, Li MW, De Sebastiano F, Esteban FJ, Horcajadas JA, Lloyd KCK, Donjacour A, Maltepe E, Rinaudo PF (2010) Effect of ICSI on gene expression and development of mouse preimplantation embryos. **Human reproduction** 25:3012-3024.

Gobikrushanth M, Purfield DC, Colazo MG, Wang Z, Butler ST, Ambrose, DJ (2018) The relationship between serum insulin-like growth factor-1 (IGF-1) concentration and reproductive performance, and genome-wide associations for serum IGF-1 in Holstein cows. **Journal of dairy science** 101:1-14.

Goddard ME, Hayes BJ (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics** 10:381-391.

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011) Extension of the Bayesian alphabet for genomic selection. **BMC bioinformatics** v. 12.

Hadziselimovic F, Hadziselimovic NO, Demougin P, Krey G, Hoecht B, Oakeley EJ (2009) EGR4 is a master gene responsible for fertility in cryptorchidism. **Sexual Development** 3:253-263.

Hayashi KG, Hosoe M, Kizaki K, Fujii S, Kanahara H, Takahashi T, Sakumoto R (2017) Differential gene expression profiling of endometrium during the mid-luteal phase of the estrous cycle between a repeat breeder (RB) and non-RB cows. **Reproductive Biology and Endocrinology** v. 15.

Henderson CR (1984) **Aplications of linear models in animal breeding**. University of Guelph, Guelph.

Höglund JK, Sahana G, Guldbrandtsen B, Lund MS (2014) Validation of associations for female fertility traits in Nordic Holstein, Nordic Red and Jersey dairy cattle. **BMC genetics** v. 15.

Irano N, Camargo GM, et al. (2016) Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. **Plos One**, v. 11.

Ji MR, Jang DM, Lee YS, Cheong HT, Yang BK, Park CK (2010) Changes of protein profiles during follicle development and in vitro oocyte maturation in the pig. **Reproduction, Fertility and Development** 23:187-187.

Jodar M, Soler-Ventura A, Oliva R, Molecular Biology of Reproduction and Development Research Group (2017) Semen proteomics and male infertility. **Journal of proteomics** 162:125-134.

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** 49:399-412.

Johnston DJ, Barwick SA, Fordyce G, Holroyd RG (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. **Proceedings...** Leipzig, Germany.

Kaftanovskaya EM, Huang Z, Lopez C, Conrad K, Agoulnik AI (2015) Conditional deletion of the relaxin receptor gene in cells of smooth muscle lineage affects lower reproductive tract in pregnant mice. **Biology of reproduction** v. 92.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. **Nature Genetics** 42:348-354.

Karlsson EK, Baranowska I, et al. (2007) Efficient mapping of mendelian traits in dogs through genome-wide association. **Nature genetics** 39:1321-1328.

Kempisty B, Antosik P, Bukowska D, Jackowska M, Lianeri M, Jaśkowski JM, Jagodziński PP (2008) Analysis of selected transcript levels in porcine spermatozoa, oocytes, zygotes and two-cell stage embryos. **Reproduction, Fertility and Development** 20:513-518.

Kiewisz J, Krawczynski K, Lisowski P, Blitek A, Zwierzchowski L, Ziecik AJ, Kaczmarek MM (2014) Global gene expression profiling of porcine endometria on Days 12 and 16 of the estrous cycle and pregnancy. **Theriogenology** 82:897-909.

Kohan K, Carvajal R, Gabler F, Vantman D, Romero C, Vega M (2010) Role of the transcriptional factors FOXO1 and PPARG on gene expression of SLC2A4 in endometrial tissue from women with polycystic ovary syndrome. **Reproduction** 140:123-131.

Kose S, Furuta M, Imamoto N (2012) Hikeshi, a nuclear import carrier for Hsp70s, protects cells from heat shock-induced nuclear damage. **Cell** 149:578-589.

Laitinen T, Lokki ML, Tulppala M, Ylikorkala O, Koskimies S (1991) Increased frequency of complement C4 'null'alleles in recurrent spontaneous abortions. **Human Reproduction** 6:1384-1387.

Lambrot R, Tian C, Jones S, Saint-Phar S, Godmann M, Kimmins S. (2008). Specialized spermiogenic distribution of histone lysine methyltransferase KMT6 and demethylase KDM5A. In: Biology of reproduction. **Proceedings…** Madison, USA: Soc Study Reproduction, 2008. p. 210-210.

Legarra A, Ricard A, Filangi O (2014) **GS3. User Manual**. France.

Li CM, Park JH, et al. (2002) Insertional mutagenesis of the mouse acid ceramidase gene leads to early embryonic lethality in homozygotes and progressive lipid storage disease in heterozygotes. **Genomics** 79:218-224.

Liu S, Yin H, Li C, Qin C, Cai W, Cao M, Zhang S (2017) Genetic effects of PDGFRB and MARCH1 identified in GWAS revealing strong associations with semen production traits in Chinese Holstein bulls. **BMC genetics** v. 18.

Liu C, Ran X, Yu C, Xu Q, Niu X, Zhao P, Wang J (2018) Whole-genome analysis of structural variations between Xiang pigs with larger litter sizes and those with smaller litter sizes. **Genomics**.

Lokki ML, Laitinen T (2001) Role of major histocompatibility complex class lll genes in recurrent spontaneous abortions. **Frontiers in Bioscience** 6:23-29.

Lomelí H, Starling, C, Gridley T (2009) Epiblast-specific Snai1 deletion results in embryonic lethality due to multiple vascular defects. **BMC research notes** v. 2.

Lomniczi A, Loche A, Castellano JM, Ronnekleiv OK, Bosch M, Kaidar G, Knoll JG, Wright H, Pfeifer GP, Ojeda SR (2013) Epigenetic control of female puberty. **Nature neuroscience** 16:281-292.

Lucki NC, Li D, Bandyopadhyay S, Wang E, Merrill AH, Sewer MB (2012) Acid ceramidase (asah1) represses steroidogenic factor-1 (sf-1)-dependent gene transcription in h295r human adrenocortical cells by binding to the receptor. **Molecular and cellular biology**.

Mao X, Sahana G, De Koning DJ, Guldbrandtsen B (2016) Genome-wide association studies of growth traits in three dairy cattle breeds using whole-genome sequence data. **Journal of animal science** 94:1426-1437.

McClure MC, Morsci NS, et al. (2010) A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. **Animal genetics** 41:597-607.

Miller SL, DeMaria JE, Freier DO, Riegel AM, Clevenger CV (2005) Novel association of Vav2 and Nek3 modulates signaling through the human prolactin receptor. **Molecular endocrinology** 19:939-949.

Mota RR, Guimarães SEF, et al. (2017) Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. **Journal of Animal Breeding and Genetics** 134:484-492.

Nguyen LT, Reverter A, Cánovas A, Venus B, Islas-Trejo A, Porto-Neto LR, Fortes MRS (2017) Global differential gene expression in the pituitary gland and the ovaries of pre-and postpubertal Brahman heifers. **Journal of Animal Science** 95:599-615.

Ohlsson Teague EMC, Print CG, Hull ML (2009) The role of microRNAs in endometriosis and associated reproductive conditions. **Human reproduction update** 16:142-165.

Oliveira Júnior G, Perez BC, et al. (2017) Genomic study and Medical Subject Headings enrichment analysis of early pregnancy rate and antral follicle numbers in Nelore heifers. **Journal of animal science** 95:4796-4812.

Orlova KA, Parker WE, Heuer GG, Tsai V, Yoon J, Baybis M, Fenning RS, Strauss K, Crino PB (2010) STRADα deficiency results in aberrant mTORC1 signaling during corticogenesis in humans and mice. **The Journal of clinical investigation** 120:1591-1602.

Ortega MS, Denicol AC, Cole JB, Null DJ, Hansen PJ (2016) Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. **Animal genetics** 47:288-297.

Osuch JR, Hsu WW, Todem D, Landgraf J, Mikucki D, Haan PS, Karmaus W (2012) Female reproductive status and circulating blood leukocyte expression of selected metabolic or signaling genes involved in sex steroid metabolism. **International journal of molecular epidemiology and genetics** 3:134-143.

Parker Gaddis K, Null DJ, Cole JB (2016) Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. **Journal of dairy science** 99:6420-6435.

Pegolo S, Dadousis C, Mach N, Ramayo-Caldas Y, Mele M, Conte G, Schiavon S, Bittante G, Cecchinato A (2017) SNP co-association and network analyses identify E2F3, KDM5A and BACH2 as key regulators of the bovine milk fatty acid profile. **Scientific reports** v. 7.

Perlmutter DH, Goldberger G, Dinarello CA, Mizel SB, Colten, HR (1986) Regulation of class Ill major histocompatibility complex gene products by interleukin-1. **Science** 232:850-852.

Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL, Silver GA, Thomas MG (2013) Heritability and Bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. **Journal of animal science** 91:605-612.

Pluta K, McGettigan PA, et al. (2012) Molecular aspects of mucin biosynthesis and mucus formation in the bovine cervix during the periestrous period. **Physiological genomics** 44:1165-1178.

Porto-Neto LR, Edwards S, Fortes MR, Lehnert SA, Reverter A, Mcgowan M (2015) Genome-wide association for the outcome of fixed-time artificial insemination of Brahman heifers in northern Australia. **Journal of Animal Science** 93:5119-5127.

Pryce JE, Bolormaa S, Chamberlain AJ, Bowman PJ, Savin K, Goddard ME, Hayes BJ (2010) A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. **Journal of dairy science** 93:3331-3345.

Riou C, Saint-Dizier M, Gérard N (2016) Sperm storage: expression of progesterone receptors, structural proteins, and heat shock proteins in the avian oviduct. **Reproduction, Fertility and Development** 28:203-203.

Saatchi M, Schnabel RD, Taylor JF, Garrick DJ (2014) Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. **BMC genomics** v. 15.

Sahana G, Guldbrandtsen B, Bendixen C, Lund MS (2010) Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. **Animal Genetics** 41:579-588.

dos Santos Hidalgo G, Meola J, Silva JCR, de Paz CCP, Ferriani RA (2011) TAGLN expression is deregulated in endometriosis and may be involved in cell invasion, migration, and differentiation. **Fertility and sterility** 96:700-703.

Sargolzaei M, Chesnais JP, Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. **BMC genomics** v. 15.

Shimizu T, Hirai Y, Miyamoto A (2013) Expression of Cyclins and Cyclin-Dependent Kinase Inhibitors in Granulosa Cells from Bovine Ovary. **Reproduction in Domestic Animals** 48:e65-e69.

Steele-Perkins G, Plachez C, Butz KG, Yang G, Bachurski CJ, Kinsman SL, Litwack D, Richards LJ, Gronostajski RM (2005) The transcription factor gene Nfib is essential for both lung maturation and brain development. **Molecular and cellular biology** 25:685-698.

Steingrímsson E, Tessarollo L, Reid SW, Jenkins NA, Copeland NG (1998) The bHLH-Zip transcription factor Tfeb is essential for placental vascularization. **Development** 125:4607-4616.

Stone DJ, Song Y, Anderson CP, Krohn KK, Finch CE, Rozovsky I (1998) Bidirectional transcription regulation of glial fibrillary acidic protein by estradiol in vivo and in vitro. **Endocrinology** 139:3202-3209.

Storlazzi CT, Mertens F, Panagopoulos I (2004) CREB3L2 (cAMP responsive element binding protein 3-like 2). **Atlas of Genetics and Cytogenetics in Oncology and Haematology** 8:316-317.

Tabatabaie V, Atzmon G, Rajpathak SN, Freeman R, Barzilai N, Crandall J (2011) Exceptional longevity is associated with decreased reproduction. **Aging (Albany NY)** 3:1202-1205.

Tarnawa ED, Baker MD, Aloisio GM, Carr BR, Castrillon DH (2013) Gonadal expression of Foxo1, but not Foxo3, is conserved in diverse Mammalian species. **Biology of reproduction** 88:1-11.

VanRaden, PM (2008) Efficient methods to compute genomic predictions. **Journal of Dairy Science** 91:4414-4423.

Varona L, García-Cortés LA, Pérez-Enciso M (2001) Bayes factors for detection of quantitative trait loci. **Genetics Selection Evolution** 33:133-152.

Wagner K, Mincheva A, Korn B, Lichter P, Pöpperl H (2001) Pbx4, a new Pbx family member on mouse chromosome 8, is expressed during spermatogenesis. **Mechanisms of development** 103:127-131.

Weber M, Locher L, Huber K, Kenéz Á, Rehage J, Tienken R, Meyer U, Dänicke S, Sauerwein H, Mielenz M (2016) Longitudinal changes in adipose tissue of dairy cows from late pregnancy to lactation. Part 1: The adipokines apelin and resistin and their relationship to receptors linked with lipolysis. **Journal of dairy science** 99:1549-1559.

Wilson KR, Liu H, Healey G, Vuong V, Ishido S, Herold MJ, Villadangos JA, Mintern JD (2018) MARCH1-mediated ubiquitination of MHC II impacts the MHC I antigen presentation pathway. **PloS one** 13:e0200540.

Xu Q, Hamada T, Kiyama R, Sakuma Y, Wada-Kiyama Y (2008) Site-specific regulation of gene expression by estrogen in the hypothalamus of adult female rats. **Neuroscience letters** 436:35-39.

Xu WH, Li ZC, Ouyang ZP, Yu B, Shi JS, Liu DW, Wu ZF (2014) RNA-Seq transcriptome analysis of porcine cloned and in vitro fertilized blastocysts. **Journal of Integrative Agriculture** 14:926-938.

Xu Y, Hutchison SM, Hernández-Ledezma JJ, Bogan RL (2017) Increased 27-hydroxycholesterol production during luteolysis may mediate the progressive decline in progesterone secretion. **MHR: Basic science of reproductive medicine** 24:2-13.

Yamaguchi R, Yamagata K, Ikawa M, Moss SB, Okabe M (2006) Aberrant distribution of ADAM3 in sperm from both angiotensin-converting enzyme (Ace)-and calmegin (Clgn)-deficient mice. **Biology of reproduction** 75:760-766.

Yang Y, Cheng L, Wang Y, Han Y, Liu J, Deng X, Chao L (2017) Expression of NDUFA13 in asthenozoospermia and possible pathogenesis. **Reproductive biomedicine online** 34:66-74.

Yik JH, Chen R, Pezda AC, Zhou Q (2005) Compensatory contributions of HEXIM1 and HEXIM2 in maintaining the balance of active and inactive P-TEFb complexes for control of transcription. **Journal of Biological Chemistry**.

Zhang Z, Ersoz E, et al. (2010) Mixed linear model approach adapted for genome-wide association studies. **Nature genetics** 42:355-360.

Chapter 4 – Associations from a multi-breed meta-analysis are validated in a Tropical Composite cattle population for early puberty traits

ABSTRACT – The age that animals achieve puberty is an important factor to maximize the profitability of beef cattle herds. The objective of the present study was to validate associations for sexual precocity traits from a meta-analysis study with Nellore and Brahman populations in a Tropical Composite (TC) cattle population. For that purpose, the significant genomic regions that were identified in the meta-analysis study were validated for each sexual precocity trait in TC, including: age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI), and scrotal circumference measured at 18 months of age for TC (SC_{TC}). Markers that were significant in the meta-analysis (*P*-value ≤ 10^{-4}) and in each individual GWAS (*P*-Value ≤ 10^{-3}), or were located 250 Kb apart from those significant SNP from the meta-analysis were considered validated. By this procedure we were able to validate a total of 49 SNP for AGECL, 4 for PPAI, and 14 SC_{TC}. For AGECL, 8 candidate genes (*COL8A1*, *PENK*, ENSBTAG00000047425, *BPNT1*, *ADAMTS17*, *CCHCR1*, *SUFU* and ENSBTAG00000046374) were close to the most significant SNP in each chromosome. For PPAI 3 candidate genes were detected: *PCBP3*, *KCNK10* and *MRPS5*. And for SC_{TC} 8 candidate genes were identified: *SNORA70*, *TRAC*, *ASS1*, *BPNT1*, *LRRK1*, *PKHD1*, *PTPRM* and ENSBTAG00000045690. Several candidate regions presented here were previous reported as significant in other studies for cattle puberty traits. The majority of the candidate genes were related to biological processes involved in the reproduction mechanism, as maintenance of gestation, expressed in reproductive tissues, and presenting differential expression according to the estrous cycle stage. Validating multi-breed meta-analyses results in other independent populations is a feasible approach to find QTL segregating across breeds. Our results were able to confirm associations discovered in a meta-analysis study with Nellore and Brahman in TC, and also associations that were reported for other breeds. Some QTL controlling early puberty seem to be segregating across Nellore, Brahman and Tropical Composite beef cattle populations.

Key-words: across-breed, GWAS, precocity, SNP, validation

4.1 Introduction

The profitability of beef cattle herds depends on several factors, such as nutrition, health, herd management and genetics. This last component is strictly associated with the reproductive performance of sires and dams. Selection for prolific and precocious animals produces more calves in less time, reduces the generation interval and promote faster economic return and genetic gains (Perotto et al., 2006; Menegassi et al., 2011; Zhang et al., 2014).

Zebu (*Bos indicus*) cattle has been bred in several tropical countries for dairy or beef, mainly because of their heat tolerance, tick resistance and greater ability to tolerate poor feed (Chan et al., 2010). Nellore is the most representative breed of the beef industry in Brazil. Also, Brahman and Tropical Composite (TC) are important breeds for Australia. These countries are in the top 10 rank of beef producers in the word. According to USDA report (2018), in 2017 Brazil was the second largest beef producer in the world, with a total of 9,550 tons of carcass produced, and Australia is the seventh largest producer (2,149 tonnes produced in 2017). However in the rank of exports, Brazil is the largest beef exporter (1,856 tons) in the world, and Australia is in the third position (1,486 tonnes), behind India. Figure 1 and 2 presents the rank of the higher beef cattle producers and exporters in the word, respectively.

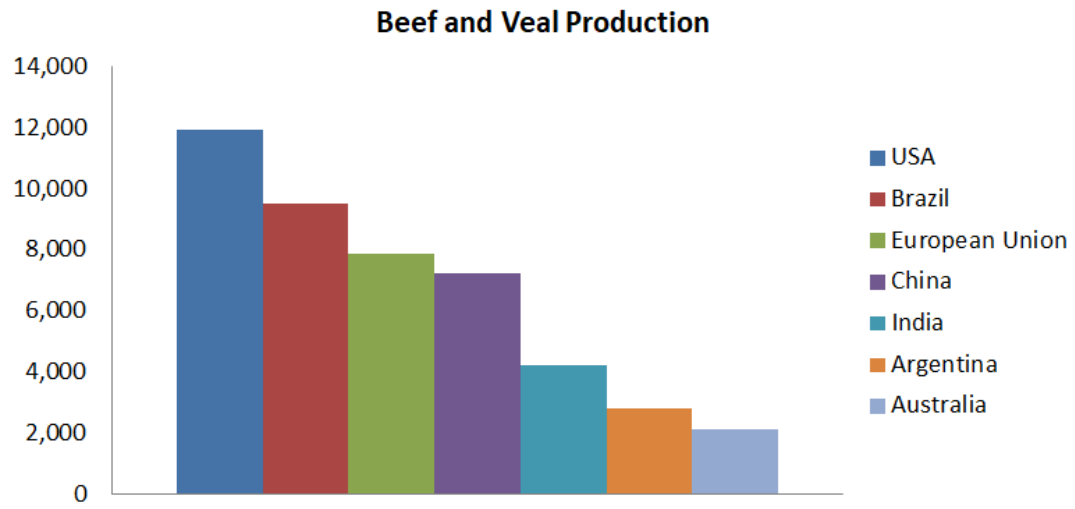

Figure 1. 2017 Rank of Beef and Veal Production in tons. (Adapted from USDA 2018 report)

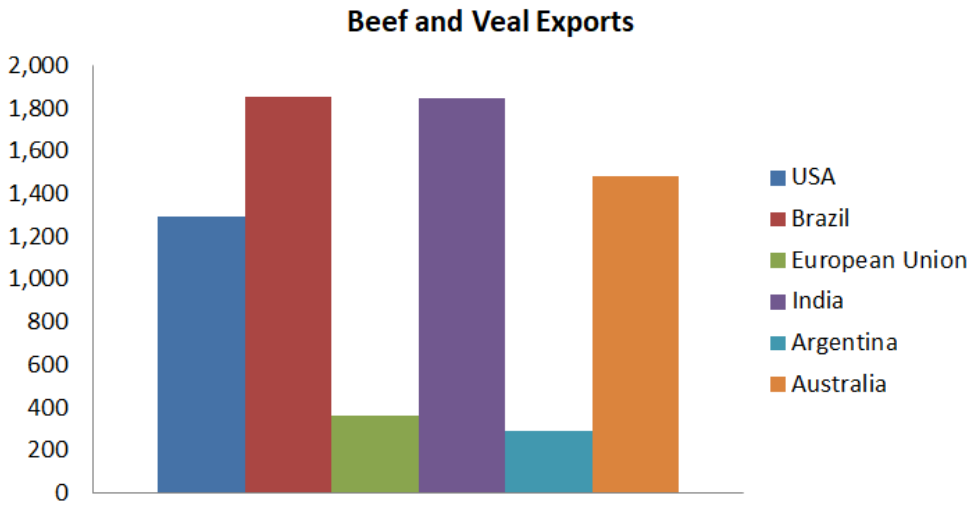

Figure 2. 2017 Rank of Beef and Veal Exports in tons. (Adapted from USDA 2018 report)

In Brazil the predominant subspecies of cattle is the Zebu, corresponding to approximately 80% of the total of bovine herds. The main breed in this country is Nellore. In the Northern territory of Australia, *Bos indicus* cattle are also predominant, and Brahman and TC are some of the most important breeds for the beef production chain of this region. TC is a breed formed by approximately 50% of *Bos indicus* (Brahman) and 50% of a mixture of African Sanga and *Bos taurus* breeds and cross breeds (N'Dama, Senepol, Shorthorn, Hereford and Charolais) (Barwick et al., 2009).

Although Zebu cattle (*Bos indicus*) have been selected in the last years they are commonly reported as later than Taurine cattle (*Bos Taurus*) to achieve puberty, which is an opportunity to study deeper sexual precocity traits in this subspecies, understand their genetic architecture and promoting their genetic improvement.

Genome-wide association studies are a feasible approach to understand the genetic architecture of several traits, and have been widely performed in livestock species to investigate the genetic background of economic important traits (Hayes et al., 2010; Fortes et al., 2011; Costa et al., 2015; Irano et al., 2016). According to Keightley et al. (1998) combining results from independent genome-wide surveys in a single analysis may improve the QTL detection power. More recently, some studies have demonstrated that multi-breed GWAS may improve the power of mapping QTL compared with within-breed GWAS (van den Berg et al., 2016).

As reported by Höglund et al. (2014), is important to perform cross validation studies before making decisions in genomic selection programs or before performing large investments to detect causal mutations. Across breed validation studies are a valuable tool to validate GWAS results across different populations. According to Saatchi et al. (2014) and van den Berg (2016), if a mutation is segregating across breeds, then is very likely that this marker is in high linkage disequilibrium with the causal variant.

The objective of the present study was to validate associations for sexual precocity traits from a meta-analysis study using Nellore and Brahman cattle in a Tropical Composite beef cattle population.

4.2 Material and Methods

4.2.1 Animals and Genotypes

Phenotypes and genotypes used in the meta-analysis study are described in Melo et al. (2018). In summary, data from Nellore and Brahman breeds were used to perform the meta-analysis. Nellore information came from Alliance Nellore dataset (Brazil) and Brahman data was provided by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) (Australia).

Nellore traits used to perform the meta-analysis study were age at first calving (AFC), early pregnancy (EP) and scrotal circumference measured at about 18 months of age (SC_N). Brahman traits considered in meta-analysis were age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI) and scrotal circumference measured at 18 months of age (SC_B).

Contemporary groups (CGs) for AFC, EP, and SC_N were formed by herd, year and season of birth, and weaning and yearling management groups. For AGECL and PPAI, CGs were formed by the information of cohort, year of birth, and management group and for SC_B CGs included the effects of cohort and year of birth. Age of young bull at recording was used as covariate for SC_N and SC_B.

Nellore animals were genotyped by using the high-density Illumina Bovine HD Assay (Illumina, San Diego, CA) and the GeneSeek Genomic Profiler Indicus HD—GGP75Ki (Neogen Corporation, Lincoln, NE), which have 777,962 and 74,677 SNP markers, respectively. The imputation was performed using FImpute software

(Sargolzaei et al., 2014), taking into account pedigree information. A total of 2,923 females with genotypes for 412,993 SNP and 5,078 males with genotypes for 477,317 SNP remained for the meta-analysis.

Brahman animals were genotyped using Illumina BovineSNP50 version 1 or 2, and further imputed for the high-density panel by using the Beagle software (Browning and Browning, 2009) taking into account representative animals of the Beef CRC population which were genotyped using high-density Illumina Bovine HD Assay (Fortes et al., 2013). About 1,000 females genotyped for 659,845 (AGECL) and 660,433 SNP (PPAI), and 1,200 males genotyped for 465,644 SNP (SC_B) were available for the meta-analysis.

TC animals used in this study were bred by the Cooperative Research Centre for Beef Genetics Technologies (or Beef CRC), in Australia. The experimental design, the origin, the phenotypes and the breed composition of these TC animals were described before by Barwick et al. (2009), Johnston et al. (2009), and Burns et al. (2013). The TC traits used in the validation study were the same as described for Brahman: AGECL, PPAI and SC_{TC}. A summary of TC phenotypes is provided in Table 1.

Table 1. Summary statistics of the traits age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI) and scrotal circumference (SC_{TC}) measured in Tropical Composite cattle.

Trait	Number of observations[1]	Mean ± SE
AGECL (days)	1,097	592.8 ± 218.4
PPAI (days)	1,097	106.8 ± 113.4
SC_{TC} (cm)	1,719	29.8 ± 2.8

[1]Samples with available genotype and phenotype.

Illumina BovineSNP50 genotypes for 1,097 and 1,719 female and male TC animals respectively, were imputed to high-density with the Beagle software (Browning and Browning, 2009). To allow imputation, representative animals of the Beef CRC population were genotyped using the HD chip from Illumina as described by Fortes et al. (2013). Quality control excluded animals with call rates lower than 98%, SNP with call rates lower than 85% and SNPs with MAF lower than 0.02. After

quality control, 683,285 SNP remained for PPAI, 683,835 for AGECL and 682,757 for SC in TC animals.

For AGECL and PPAI, CGs were formed by concatenating information of cohort, year of birth, and management group. For SC_{TC}, CG included the effects of cohort and year of birth.

4.2.2 GWAS and meta-analyses

The model used to perform the GWAS was a genomic mixed model proposed by Kang et al. (2010). This model was used for all three breeds, but for Nellore the fixed effects were pre-adjusted, because the Nellore dataset contained both genotyped and non-genotyped animals. Details about this adjustment and the GWAS for Nellore and Brahman are further described in Melo et al. (2018). For TC the model was fitted as follows:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{s}a + \mathbf{Zu} + \boldsymbol{\varepsilon}, \quad (1)$$

where **y** is a vector with the phenotypes, **X** is an incidence matrix relating fixed effects (CGs) in **β** with the phenotypes in **y**, and the young bull age at the measurement as covariate for SC_{TC}, **s** is the vector with the genotypes coded as 0, 1, or 2 according to the number of B allele copies, a is the vector with the SNP allelic substitution effects, **Z** is the incidence matrix of the polygenic random effects of the animals in **u** and **ε** is the vector of residuals. **u** and **ε** followed a normal distribution with $\mathbf{u} \sim N(0, \mathbf{G}\sigma^2_a)$ and $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma^2_e)$, respectively, where **G** is the genomic relationship matrix for all individuals and SNP (exception for those SNP considered in a), and was calculated by the first method proposed by VanRaden (2008), σ^2_a is the additive genetic variance, **I** is an identity matrix, and σ^2_e is the residual variance. The SNP effect estimates were obtained with the SNP & Variation Suite v8.3.0 software (Golden Helix, Inc., Bozeman, MT) using the model previously described and the EMMAX method (Kang et al., 2010).

The meta-analysis was performed by using the multi-trait meta-analysis method proposed by Bolormaa et al. (2014). Details about this procedure are described in Melo et al. (2018). In summary, this meta-analysis method is a statistic test that follows a χ^2 distribution with n degrees of freedom, where n is the number of traits

included in meta-analysis. The chi-squared multi-trait meta-analysis is calculated as follow:

$$\textit{Multi-trait } \chi^2 = \mathbf{t'_i V^{-1} t_i}, \quad (2)$$

where $\mathbf{t_i}$ is a vector 6 × 1 of the i^{th} SNP effect estimates for the six traits for Nellore and Brahman divided by their respective standard errors, $\mathbf{t'_i}$ is the transpose vector of $\mathbf{t_i}$, and $\mathbf{V^{-1}}$ is an inverse of the 6 × 6 correlation matrix of the correlations between the *t*-values of the six traits across the 387,971 SNP considered in the model. The *t*-values vector was computed as:

$$\mathbf{t_i} = \frac{a_i}{SE(a_i)}, \quad (3)$$

where $SE(a_i)$ is the SE of a_i (SNP effect vector). Just common SNP in both Nellore and Brahman panels were considered in the analysis.

4.2.3 Validation Procedure

To be considered validated in TC, significant SNP in the meta-analysis, i.e., under an empirical *P*-value of 1×10^{-4} should also be significant for each single GWAS in TC (*P*-value ≤ 1×10^{-3}), or the SNP should be in a region of ± 250 Kb (up-stream or down-stream) away from those significative SNP of each GWAS.

4.3 Results and Discussion

A total of 296 SNP presented a *P*-value < 1×10^{-4} in the meta-analysis. For AGECL, PPAI and SC_{TC} were found 874, 842 and 918 significant SNP (*P*-value ≤ 1×10^{-3}).

A total of 49 SNP were validated for AGECL, they were distributed over chromosomes 1, 14, 15, 16, 21, 23, 26 and 29. For PPAI 4 SNP located on BTA1, BTA10 and BTA11 were validated. And for SC_{TC} 14 SNP were validated, distributed over chromosomes 8, 10, 11, 16, 21, 23, 24 and 27. The top SNP of each chromosome for AGECL, PPAI and SC_{TC} are presented in Table 2. For AGECL the BTA14 was the chromosome that presented most of the validated SNP (49%), followed by the BTA21 (35%). For PPAI most of the validated SNP were located on BTA10 (50%), and for SC_{TC} the majority of the validated SNPs were located on

chromosomes 10 and 21. Many of the top SNP for all traits were very close or inside to a candidate gene, i.e., those genes closest to the top SNP of each chromosome. The complete list of validated SNP for each trait is presented in Supplemental Table S2.

Table 2. Number of validated SNP ($P < 1 \times 10^{-3}$) per *Bos taurus* autosome (BTA), the most significant SNP in each autosome (top SNP), its position, and its distance from the closest gene (± 250 Kb) in base pairs.

BTA[1]	Number of SNP	Top SNP	Position, bp[1]	Gene Symbol	Distance, bp
			AGECL		
1	1	BovineHD0100012377	43,458,869	*COL8A1*	83,067
14	24	BovineHD1400007314	25,241,366	*PENK*	18,375
15	1	BovineHD1500002332	9,064,376	*Uncharacterized Protein (ENSBTAG00000047425)*	141,535
16	1	BovineHD1600006716	24,282,102	*BPNT1*	0
21	17	BovineHD2100001408	6,839,238	*ADAMTS17*	0
23	2	BovineHD2300007697	27,782,568	*CCHCR1*	0
26	2	BovineHD2600006035	23,405,679	*SUFU*	47,301
29	1	BovineHD2900002706	9,174,027	*Uncharacterized Protein (ENSBTAG00000046374)*	1,952
			PPAI		
1	1	BovineHD0100042609	147,345,274	*PCBP3*	0
10	2	BovineHD1000029240	100,990,620	*KCNK10*	0
11	1	ARS-BFGL-NGS-68030	1,901,787	*MRPS5*	0
			SC_{TC}		
8	1	BovineHD0800024244	81,614,440	*SNORA70*	50,413
10	4	BovineHD1000007157	22,121,623	*TRAC*	0

11	1	BovineHD11000 29283	100,818,65 7	*ASS1*	0
16	1	BovineHD16000 06716	24,282,102	*BPNT1*	0
21	4	BovineHD21000 01045	5,655,578	*LRRK1*	0
23	1	BovineHD23000 06383	23,921,232	*PKHD1*	0
24	1	BovineHD24000 11464	41,304,824	*PTPRM*	0
27	1	ARS-BFGL-NGS-118334	41,297,038	*Uncharacterized Protein (ENSBTAG0000004 5690)*	240,958

[1] SNP were mapped to the UMD3.1 Bovine Assembly.

4.3.1 QTL around validated SNP

To reinforce the importance of the validated SNP here, it was investigated the presence of QTL associated with reproductive traits reported in literature that were close to the region of the validated SNP for TC. Some important genomic regions detected here were previous detected in other studies using the same populations for other reproductive or related traits. For example, Fortes et al. (2012) found several significant SNP associated with age at puberty in TC and Brahman on BTA21 at 6 Mb. Costa et al. (2015) detected a significant SNP associated with heifer rebreeding in the same Nellore population used in the meta-analysis on BTA10 at 100 Mb. Dias et al. (2015) identified that some lipid-metabolism candidate genes among them the *TNF* gene (BTA23 at 27 Mb), were related with sexual precocity. Porto-Neto et al. (2015) discovery important associations on BTA8 at 81 Mb, BTA21 at 5 Mb and BTA11 at 1 Mb related with pregnancy outcome after fixed-time AI in Brahman heifers.

In addition, we consulted the QTLdatabase tool set (Hu et al., 2006) to find studies with other populations that reported significant SNP in the same regions detected here. McClure et al. (2010) found a significant marker on BTA10 at 100 Mb associated with scrotal circumference in Angus. Also these authors found significant markers on BTA27 at 41 Mb and BTA29 at 9 Mb. An important SNP in the gene *NFKBIL1* located on BTA23 at 27.5 Mb was associated with daughter pregnancy rate, heifer conception rate and cow conception rate in Holstein (Cochran et al.,

2013). On BTA23 at 28.0 Mb Höglund et al. (2014) validated SNP for the length in days of the interval from calving to first insemination in three breeds of dairy cattle. Furthermore these authors found on BTA11 at 2.1 Mb a SNP associated with 56-day non-return rate of cows. Both regions were close to those found here (< 0.5 Mb). Liu et al. (2017) in a meta-analysis with Holstein cattle for female fertility traits found a QTL region on BTA14 at 25 Mb. Mota et al. (2017) found a significant 1 Mb genomic window for age at first calving in Nellore cattle on BTA14 at 24.6 Mb, which is close to our findings and is also close to the *PLAG1* gene. Oliveira Júnior et al. (2017) found important genomic regions for heifer pregnancy on BTA14 at 25.0 Mb and for antral follicles on BTA15 from 8-10 Mb in Nellore heifers. And Soares et al. (2017) found an important 1 Mb window on BTA14 at 24 and 26 Mb associated with SC12, SC18 and SC24.

4.3.2 Genes surrounding the candidate regions

The majority of the genes close to the top SNP of each chromosome identified here was associated with reproductive event in mammals and hens, expressed in reproductive tissues or were related with grown traits in cattle. A summary of candidate genes and the biological process that they are related to are shown in Figure 3.

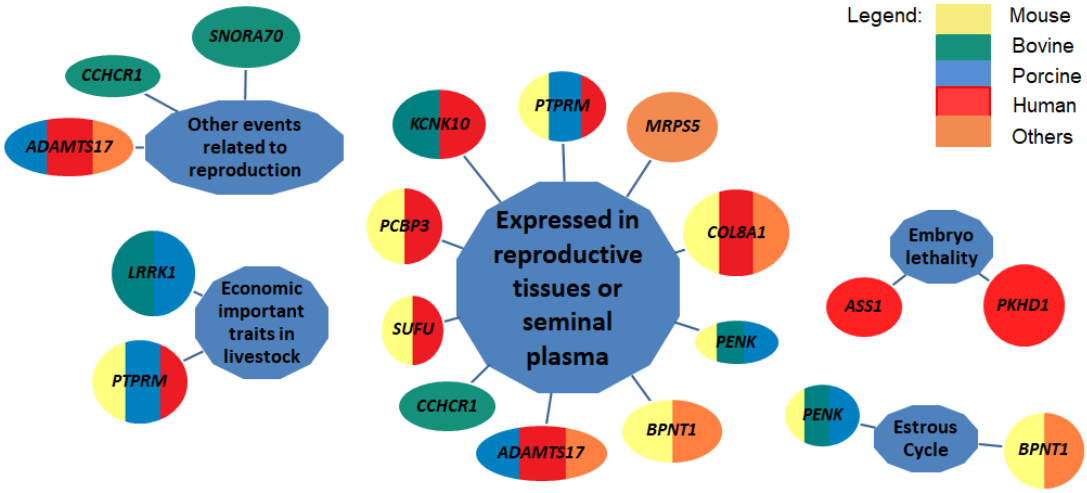

Figure 3. Summary of candidate genes and their related processes.

The candidate genes for AGECL were *COL8A1*, *PENK*, ENSBTAG00000047425, *BPNT1*, *ADAMTS17*, *CCHCR1*, *SUFU* and

ENSBTAG00000046374. The gene *COL8A1* (collagen type VIII alpha 1 chain) was expressed in epididymis and efferent ducts tissues in male mouse embryos (Snyder et al., 2010) and was differentially expressed in ovine granulosa cells between the follicular and luteal phases (Talebi et al., 2018). This gene was up-regulated in human pregnant myometrium (Rehman et al., 2003). The expression of *PENK* (proenkephalin) increased during the dioestrus in bovine endometrium (Bauersachs et al., 2005). Substantial levels of this gene were detected in mouse endometrium during the early implantation (Zhu and Pintar, 1998). Also this gene was expressed in porcine granulosa cells (Staszkiewicz et al., 2007). The gene *BPNT1* (3'(2'), 5'-bisphosphate nucleotidase 1) was associated with the modulation of sperm function in the testis of mice exposure to cigarette smoking (Xu et al., 2013). This gene was up-regulated in wild type pregnant mouse compared to females presenting loss of function in another gene, the *SMTNL1* that is also important in the regulation of pregnancy (Bodoor et al., 2011).

The levels of *BPNT1* protein was high differentially expressed in equine uterus between oestrus and luteal phase (Soleilhavoup et al., 2015). The expression of *ADAMTS17* (ADAM metallopeptidase with thrombospondin type 1 motif 17) is quickly induced by the presence of estrogens, having implications in breast cancer cell grown (Jia et al., 2014). Also the expression of this gene was down-regulated in the hypothalamus of hens exposure to heat stress, being recovered 2 hours after the acute heat stress exposure (Tu et al., 2016). This gene was up-regulated in the endometrium of pregnant pigs in the 12th day of pregnancy vs the pig endometrium in the 12th day of of the estrous cycle (Kim et al., 2012). *CCHCR1* (coiled-coil alpha-helical rod protein 1) was differentially expressed in bovine endometrium after *n*-3 polyunsaturated fatty acid supplementation (Waters et al., 2014). This gene was identified in a top genomic region associated with cell-mediated immune response in Holstein cows (Thompson-Crispi et al., 2014). *SUFU* (SUFU negative regulator of hedgehog signaling) is associated with the proliferation of Sertoli cells during the spermatogenesis (Procópio et al., 2017). The association of this gene with spermatogenesis was also reported in mouse testis (Szczepny et al., 2005). Despite the genes ENSBTAG00000047425 and ENSBTAG00000046374 have not been

related with reproductive events, they were candidate genes for the meta-analysis of reproductive traits using Brahman and Nellore (Melo et al., 2018).

Candidate genes for PPAI were: *PCBP3*, *KCNK10* and *MRPS5*. The protein of *PCBP3* (poly(rC) binding protein 3) gene was abundantly expressed around rats spermatids prior to alongation phase of spermiogenesis (Chapman et al., 2013). This gene was up-regulated in the endometrium of women with endometriosis (Aghajanova et al., 2010). The gene *KCNK10* (potassium two pore domain channel subfamily K member 10) was associated with semen volume in Holstein-Friesian bulls (Hering et al., 2014) and was also involved in the volume regulation of human spermatozoa and associated with the regulation of potassium channels (Barfield et al., 2006). This gene was expressed at the membrane of oocytes and blastocysts of Korean cattle (Hur et al., 2009). And *MRPS5* (mitochondrial ribosomal protein S5) gene was expressed in primate primordial oocyte cells (Arraztoa et al., 2005).

Candidate genes for SC_{TC} were: *SNORA70*, *TRAC*, *ASS1*, *BPNT1*, *LRRK1*, *PKHD1*, *PTPRM* and ENSBTAG00000045690. The gene *SNORA70* (small nucleolar RNA, H/ACA box 70) was detected in genomic regions under positive signature selection in Holstein, N'Dama and Hanwoo cattle (Taye et al., 2017) and in Creole cattle, affecting QTLs associated with feed intake, conformation, weight, reproduction, and milk production traits (Pitt et al., 2018). *ASS1* (argininosuccinate synthase 1) gene was associated with abnormalities in brain in the 23^{th} day of gestation resulting in abortion in humans (Pangalos et al., 2016). Other authors related the association of this gene with complications during the gestation (Enquobahrie et al., 2008; Wu et al., 2014). *LRRK1* (leucine rich repeat kinase 1) was identified in a GWAS associated with six reproductive traits in Large White pigs (Suwannasing et al., 2018). This gene was detected in a significant region of a GWAS for pregnancy outcome after fixed-time AI evaluated in the same Brahman population used here (Porto-Neto et al., 2015). Also, this gene was associated with muscle development in cattle (Widmann et al., 2013). *PKHD1* (PKHD1, fibrocystin/polyductin) was related to a severe form of polycystic kidney disease, which may affect embryo development and survival (Onuchic et al., 2002; Gigarel et al., 2008). The gene *PTPRM* (protein tyrosine phosphatase, receptor type M) was associated with pig muscling (Li et al., 2011). This gene was differentially expressed

in the blood of women with early pregnancy and pre-enclampsia diagnostics (Enquobahrie et al., 2011), and was expressed in rat spermatogonia cells in pre-pubertal and pubertal stages (Ryser et al., 2011).

Genes *TRAC* and ENSBTAG00000045690 were not previously associated with reproductive functions. The genes ENSBTAG00000047425, *SUFU* and ENSBTAG00000046374 were candidates for the Nellore- Brahman meta-analysis study (Melo et al., 2018), and *PENK* gene was located in a peak region of BTA14 of this meta-analysis. The gene *BPNT1* was candidate for AGECL and SC_{TC}.

A reason that could explain the low number of SNP validated for all TC traits is because this is a breed with a half percent of *Bos indicus* origin, and is less related with Brahman and Nellore, both of 100% *Bos indicus* origin. However, the strategy of using results from an meta-analysis study to validate genomic regions in other populations proved to be a feasible tool to validate QTL across breeds.

4.4 Conclusions

We were able to validate genomic regions previously associated with sexual precocity traits in an meta-analysis study using Brahman and Nellore in a third breed, the TC. Some of these regions contained genes reported as associated with reproduction in different species. Using meta-analysis results we were able to validate genomic regions in independent beef cattle populations.

4.5 References

Aghajanova L, Tatsumi K, Horcajadas JA, Zamah AM, Esteban FJ, Herndon CN, Conti M, Giudice LC (2011) Unique transcriptome, pathways, and networks in the human endometrial fibroblast response to progesterone in endometriosis. **Biology of reproduction** 84:801-815.

Arraztoa JA, Zhou J, et al. (2005) Identification of genes expressed in primate primordial oocytes. **Human Reproduction** 20:476-483.

Barfield JP, Yeung CH, Cooper TG (2005) Characterization of potassium channels involved in volume regulation of human spermatozoa. **Molecular human reproduction** 11:891-897.

Barwick SA, Johnston DJ, Burrow HM, Holroyd RG, Fordyce G, Wolcott ML, Sim WD, Sullivan, MT (2009) Erratum to: Genetics of heifer performance in'wet'and'dry'seasons and their relationships with steer performance in two tropical beef genotypes. **Animal Production Science** 49:727-727.

Bodoor K, Lontay B, Safi R, Weitzel D, Loiselle D, Wei Z, Lengyel S, McDonnell DP, Haystead TA (2011) Smoothelin-like 1 protein is a bifunctional regulator of the progesterone receptor during pregnancy. **Journal of Biological Chemistry** 286:31839-31851.

Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, Tier B, Savin K, Hayes BJ, Goddard ME (2014) A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for stature, Fatness and Reproduction in Beef Cattle. **PLOS Genetics** 10:e1004198.

Browning BL, Browning SR (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. **The American Journal of Human Genetics** 84:210-223.

Burns BM, Corbet NJ, Corbet DH, Crisp JM, Venus BK, Johnston DJ, Li Y, McGowan MR, Holroyd RG (2013) Male traits and herd reproductive capability in tropical beef cattle. 1. Experimental design and animal measures. **Animal Production Science**. 53:87-100.

van den Berg I, Boichard D, Lund MS (2016) Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. **Journal of dairy science** 99:8932-8945.

Chan EKF, Nagaraj SH, Reverter A (2010) The evolution of tropical adaptation: comparing taurine and zebu cattle. **Animal genetics** 41: 467-477.

Chapman KM, Powell HM, Chaudhary J, Shelton JM, Richardson JA, Richardson T E, Hamra FK (2013) Linking spermatid RNA binding protein and retrogene diversity to reproductive success. **Molecular & Cellular Proteomics** 12:3221:3236.

Cochran SD, Cole JB, Null DJ, Hansen PJ (2013) Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. **BMC genetics** 14:49.

Costa RB, Camargo GM, et al. (2015) Genome-wide association study of reproductive traits in nellore heifers using bayesian inference. **Genetics Selection Evolution** 47:67.

Dias MM, Souza FRP, et al. (2015) Study of lipid metabolism-related genes as candidate genes of sexual precocity in Nellore cattle. **Genetics and Molecular Research** 14:234-243.

Enquobahrie DA, Meller M, Rice K, Psaty BM, Siscovick DS, Williams MA (2008) Differential placental gene expression in preeclampsia. **American journal of obstetrics and gynecology** 199:566-e1.

Enquobahrie DA, Qiu C, Muhie SY, Williams MA (2011) Maternal peripheral blood gene expression in early pregnancy and preeclampsia. **International journal of molecular epidemiology and genetics** 2:78-94.

Fortes MR, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, Lehnert S, Boe-Hansen GB, Hawken RJ (2011) A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. **Journal of animal science** 89:1669-1683.

Fortes MR, Li Y, Collis E, Zhang Y, Hawken RJ (2012) The IGF1 pathway genes and their association with age of puberty in cattle. **Animal genetics** 44:91-95.

Fortes MRS, Reverter A, Kelly M, McCulloch R, Lehnert SA (2013) Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species. **Andrology**, 1:644-650.

Gigarel N, Frydman N, Burlet P, Kerbrat V, Tachdjian G, Fanchin R, Antignac C, Frydman R, Munnich A, Steffann J (2008) Preimplantation genetic diagnosis for autosomal recessive polycystic kidney disease. **Reproductive biomedicine online** 16:152-158.

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: e1001139.

Hering DM, Oleński K, Ruść A, Kaminski S (2014) Genome-wide association study for semen volume and total number of sperm in Holstein-Friesian bulls. **Animal reproduction science** 151:126-130.

Höglund JK, Sahana G, Guldbrandtsen B, Lund MS (2014) Validation of associations for female fertility traits in Nordic Holstein, Nordic Red and Jersey dairy cattle. **BMC Genetics** 15:1-8.

Hu ZL, Fritz ER, Reecy JM (2006) AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. **Nucleic Acids Research** 35:D604-D609.

Hur CG, Choe C, Kim GT, Cho SK, Park JY, Hong SG, Han J, Kang D (2009) Expression and localization of two-pore domain K+ channels in bovine germ cells. **Reproduction** 137:237-244.

Irano N, De Camargo, et al. (2016) Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. **Plos One** 11:e0159502.

Jia Z, Gao S, M'Rabet N, De Geyter C, Zhang H (2014) Sp1 is necessary for gene activation of Adamts17 by estrogen. **Journal of cellular biochemistry** 115:1829-1839.

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** 49:399-412.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. **Nature Genetics** 42:348-354.

Keightley PD, Morris KH, Ishikawa A, Falconer VM, Oliver F (1998) Test of candidate gene-quantitative trait locus association applied to fatness in mice. **Heredity** 81:630-637.

Kim M, Seo H, Choi Y, Shim J, Kim H, Lee CK, Ka H (2012) Microarray analysis of gene expression in the uterine endometrium during the implantation period in pigs. **Asian-Australasian journal of animal sciences** 25:1102.

Li X, Kim SW, Do KT, Ha YK, Lee YM, Yoon SH, Kim HB, Kim JJ, Choi BH, Kim KS (2011) Analyses of porcine public SNPs in coding-gene regions by re-sequencing and phenotypic association studies. **Molecular biology reports** 38:3805-3820.

Liu A, Wang Y, Sahana G, Zhang Q, Liu L, Lund MS, Su G (2017) Genome-wide association studies for female fertility traits in Chinese and Nordic Holsteins. **Scientific reports** 7:8487.

McClure MC, Morsci NS, et al. (2010) A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. **Animal genetics** 41:597-607.

Melo TP, Fortes MRS, Bresolin T, Mota LFM, Albuquerque LG, Carvalheiro R (2018) Multitrait meta-analysis identified genomic regions associated with sexual precocity in tropical beef cattle. **Journal of Animal Science** 96: 4087–4099.

Menegassi SRO, Barcellos JOJ, Peripolli V, Pereira PRRX, Borges JBS, Lampert VDN (2011) Measurement of scrotal circumference in beef bulls in Rio Grande do Sul. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 63:87-93.

Mota RR, Guimarães SEF, et al. (2017) Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. **Journal of Animal Breeding and Genetics** 134:484-492.

Oliveira Júnior GA, Perez BC, et al. (2017) Genomic study and Medical Subject Headings enrichment analysis of early pregnancy rate and antral follicle numbers in Nellore heifers. **Journal of Animal Science** 95:4796-4812.

Onuchic LF, Furu L, et al. (2002) PKHD1, the polycystic kidney and hepatic disease 1 gene, encodes a novel large protein containing multiple immunoglobulin-like plexin-transcription–factor domains and parallel beta-helix 1 repeats. **The American Journal of Human Genetics** 70:1305-1317.

Pangalos C, Hagnefelt B, Lilakos K, Konialis C (2016) First applications of a targeted exome sequencing approach in fetuses with ultrasound abnormalities reveals an important fraction of cases with associated gene defects. **PeerJ** 4:e1955.

Perotto D, Miyagi AP, Souza JC, Moletta JL, Freitas JA (2006) Estudos de características reprodutivas de animais da raça Canchim, criados a pasto, no estado do Paraná, Brasil. **Archives of Veterinary Science** 11:1-6.

Pitt D, Bruford MW, Barbato M, Orozco-terWengel P, Martínez R, Sevane N (2018) Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. **Evolutionary Applicattions** 1-18.

Porto-Neto LR, Edwards S, Fortes MR, Lehnert SA, Reverter A, Mcgowan M (2015) Genome-wide association for the outcome of fixed-time artificial insemination of Brahman heifers in northern Australia. **Journal of Animal Science** 93:5119-5127.

Procópio MS, de Avelar GF, Costa GM, Lacerda SM, Resende RR, de França LR (2017) MicroRNAs in Sertoli cells: implications for spermatogenesis and fertility. **Cell and tissue research** 370:335-346.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available in: <https://www.R-project.org/>.

Rehman KS, Yin S, Mayhew BA, Word RA, Rainey WE (2003) Human myometrial adaptation to pregnancy: cDNA microarray gene expression profiling of myometrium from non-pregnant and pregnant women. MHR: **Basic science of reproductive medicine** 9:681-700.

Ryser S, Glauser D, Vigier M, Zhang YQ, Tachini P, Schlegel W, Durand P, Irminger-Finger I (2011) Gene expression profiling of rat spermatogonia and Sertoli cells reveals signaling pathways from stem cells to niche and testicular cancer cells to surrounding stroma. **BMC genomics** 12:1-16.

Saatchi M, Schnabel RD, Taylor JF, Garrick DJ (2014) Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. **BMC genomics** 15:1-16.

Sargolzaei M, Chesnais JP Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. **BMC Genomics** 15:478.

SNP & Variation Suite ™ (Version 8.3.0). Bozeman, MT: Golden Helix, Inc.

Snyder EM, Small CL, Bomgardner D, Xu B, Evanoff R, Griswold MD, Hinton BT (2010) Gene expression in the efferent ducts, epididymis, and vas deferens during embryonic development of the mouse. **Developmental Dynamics** 239:2479-2491.

Soares ACC, Guimarães SEF, Kelly MJ, Fortes MRS, Silva FF, Verardo LL, Mota RR, Moore S (2017) Multiple-trait genomewide mapping and gene network analysis for scrotal circumference growth curves in Brahman cattle. **Journal of animal science**, 95:3331-3345.

Soleilhavoup C, Riou C, Tsikis G, Labas V, Harichaux G, Kohnke P, Druart X (2015) Proteomes of the female genital tract during the oestrous cycle. **Molecular & Cellular Proteomics** mcp-M115.

Staszkiewicz J, Skowronski MT, Kaminski T, Siawrys G, Krazinski BE, Kusmider M, Przala J, Okrasa S (2007) Expression of proopiomelanocortin, proenkephalin and prodynorphin genes in porcine theca and granulosa cells. **Animal Reproduction Science** 101:97-112.

Suwannasing R, Duangjinda M, Boonkum W, Taharnklaew R, Tuangsithtanon K (2018) The identification of novel regions for reproduction trait in Landrace and Large White pigs using a single-step genome-wide association study. **Asian-Australasian journal of animal sciences**.

Szczepny A, Jans DA, Hime G, Loveland KL (2005) 247. Expression of components of the hedgehog signalling pathway during murine spermatogenesis. **Reproduction, Fertility and Development** 17:98-98.

Talebi R, Ahmadi A, Afraz F (2018) Analysis of protein-protein interaction network based on transcriptome profiling of ovine granulosa cells identifies candidate genes in cyclic recruitment of ovarian follicles. **Journal of Animal Science and Technology** 60:1-7.

Taye M, Lee W, et al. (2017) Exploring evidence of positive selection signatures in cattle breeds selected for different traits. **Mammalian Genome** 28:528-541.

Thompson-Crispi KA, Sargolzaei M, Ventura R, Abo-Ismail M, Miglior F, Schenkel F, Mallard BA (2014) A genome-wide association study of immune response traits in Canadian Holstein cattle. **BMC genomics** 15:1-10.

Tu WL, Cheng CY, Wang SH, Tang PC, Chen CF, Chen HH, Lee YP, Chen SE, Huang SY (2016) Profiling of differential gene expression in the hypothalamus of broiler-type Taiwan country chickens in response to acute heat stress. **Theriogenology** 85:483-494.

United States Department of Agriculture, USDA (2018) **Livestock and Poultry: World Markets and Trade**. Foreign Agricultural Service/USDA, 29p.

Xu W, Fang P, Zhu Z, Dai J, Nie D, Chen Z, Qin Q, Wang L, Wang Z, Qiao Z (2013) Cigarette smoking exposure alters pebp1 DNA methylation and protein profile involved in MAPK signaling pathway in mice testis. **Biology of reproduction** 89:1-11.

Waters SM, Coyne GS, Kenny DA, Morris DG (2014) Effect of dietary n-3 polyunsaturated fatty acids on transcription factor regulation in the bovine endometrium. **Molecular biology reports** 41:2745-2755.

Widmann P, Reverter A, Fortes MR, Weikard R, Suhre K, Hammon H, Albrecht E, Kuehn C (2013) A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. **BMC genomics** 14:1-17.

Wu TF, Liu YP, Li XY Wang Q, Song JQ, Yang YL (2014) Prenatal diagnosis of citrullinemia type 1: a Chinese family with a novel mutation of the ASS1 gene. **Brain and Development** 36:264-267.

Zhang YD, Johnston DJ, Bolormaa S, Hawken RJ, Tier B (2014) Genomic selection for female reproduction in Australian tropically adapted beef cattle. **Animal Production. Science** 54:16-24.

Zhu Y, Pintar JE (1998) Expression of opioid receptors and ligands in pregnant mouse uterus and placenta. **Biology of reproduction** 59:925-932.

Chapter 5 – Final Considerations

In this study we were able to confirm and discover new QTL associated with sexual precocity in three tropical beef cattle breeds, Nellore, Brahman and Tropical Composite (TC). Some genic regions that were previously associated with puberty in Brahman could be validated here, and new regions were identified by the meta-analysis and across breed validation studies. Also, we validated the results of a meta-analysis study with Nellore and Brahman in a third breed (TC), and we could find genomic regions in common among Nellore, Brahman and TC associated with different traits indicators of sexual precocity.

The validation for male traits was less successful if compared with female traits. This fact could be explained because the X chromosome contained the most significant associations for Brahman and TC, and this importance was more pronounced for Brahman, compared with TC. That strong peak in X chromosome could not be validated in Nellore, additional analysis (results not shown) including the X chromosome for Nellore did not find a large number of significant SNP in the X chromosome for SC as observed for Brahman and TC, and the number of significant SNP decreased as well their significance compared with the analysis without X. This probably happened due to the decreasing of the sample size, when adding the X chromosome, we used just the animals genotyped with the HD panel, losing information of imputed animals. Another fact that could explain this smaller number of validated SNP for SC in Nellore is the number of Brahman samples, which was 3.5 times lower if compared to Nellore samples. This difference was smaller for female traits.

In general, the candidate regions were distributed over almost all autosomes (24 out of 29 chromosomes), which is consistent with the polygenic traits architecture, in which the phenotype is controlled by many QTL with small effect each.

Using the candidate gene approach, i.e., genic regions previous related with the trait of interest, to validate genomic regions across breeds may be more successful than using results from meta-analysis, i.e., significant SNP marker and their neighbour regions. The explanation for that is because the meta-analysis results passed through more significance thresholds, i.e., the SNP should be significant for

different traits in different breeds, and then, be significant in a third breed to be validated, as was the case here. This could hinder the validation process, and may produce a higher false negative rate. Another reason is the genomic size, meta-analysis results are punctual, i.e., the result is just a SNP more 500 Kb, while using candidate genes there are more SNP, which are harboured in the gene, more 500 Kb.

In meta-analysis studies a well definition of the phenotypes across populations is fundamental. If there are divergences in the phenotype definition among studies may decrease the power of validate those QTL observed in each individual breed. Another point that influences the success in validating across breeds is the genetic difference among the breeds included in the study: fewer regions could be validated across Nellore and TC than across Nellore and Brahman, confirming which has been showed in other studies.

Our results, in accordance with the literature, indicate that a multi-breed GWAS, using independent populations, is a feasible tool to detect and fine mapping QTL. Additionally, the results found here can be further confirmed by applying laboratory experiments, as gene knockout of strongest association regions or gene expression in cattle or other mammalian species to investigate better the biological role of the genes located in or close to the candidate regions found here. Moreover, approaches used here may also be applied to other economically important traits and breeds to validate known and discovery new QTL. Also, the use of the whole genomic sequence may improve the power of QTL detection and can also be applied with meta-analysis and across breed validation approaches.

APPENDICES

Supplementary Material

Table S1. Significant SNPs ($P < 1.39x10^{-5}$) of meta-analysis of sexual precocity traits in Nellore and Brahman cattle, sorted from the highest to smallest *P*-value.

Marker	Chromosome	Position[1] (bp)	*P*-value	log10-Pvalue
BovineHD0900021032	9	75613158	3.21E-14	13.49338074
BovineHD1500002332	15	9064376	1.17E-08	7.932398139
BovineHD1400007053	14	24323400	2.44E-06	5.61190001
BTB-01532239	14	24437778	7.78E-07	6.109208
BovineHD1400007095	14	24445514	6.01E-06	5.220991815
BovineHD1400007101	14	24466047	3.37E-07	6.472711132
BovineHD1400007102	14	24470245	4.06E-06	5.391744415
BovineHD1400007103	14	24471148	3.37E-07	6.472711132
BTB-01530788	14	24524205	1.20E-07	6.921674371
BovineHD1400007131	14	24553162	1.58E-07	6.801499507
BovineHD1400007132	14	24556301	1.19E-07	6.925391362
ARS-BFGL-NGS-53471	6	118436689	1.41E-07	6.852133469
BovineHD1400007141	14	24582124	2.70E-06	5.568749498
BovineHD1400007153	14	24621142	4.68E-08	7.329916566
BovineHD0200027559	2	95917958	2.02E-07	6.694524104
BovineHD1400007156	14	24633076	6.80E-07	6.167349969
BovineHD0200027569	2	95952415	3.06E-07	6.513624431
BTB-00557532	14	24643266	3.15E-08	7.501279391
BovineHD1400007161	14	24656389	2.64E-08	7.578821883
BovineHD1400007169	14	24699409	8.54E-08	7.068593648
BovineHD0200027573	2	95959321	4.33E-07	6.363354608
BovineHD1400007173	14	24710609	1.95E-08	7.711061425
BovineHD0200027578	2	95983401	4.80E-07	6.318577247
BovineHD1400007205	14	24808828	2.16E-06	5.665694953
BovineHD1400007210	14	24828922	7.43E-07	6.128855303
BovineHD0200027587	2	96009982	6.15E-07	6.211438512
BovineHD1400007211	14	24833259	5.87E-06	5.231645528
BovineHD1400007213	14	24839114	1.09E-06	5.960973322
BovineHD0800020507	8	68307110	7.50E-07	6.125192492
BovineHD1400007217	14	24860275	1.04E-06	5.984306257
BovineHD1400007218	14	24864286	2.01E-07	6.696253313
BovineHD1400007220	14	24874608	1.47E-06	5.83392255
BovineHD0200027563	2	95935954	9.74E-07	6.011243145
BovineHD1400007224	14	24886631	3.04E-06	5.517208278
BovineHD1400007232	14	24909247	5.65E-07	6.248253838
BovineHD1400007234	14	24913654	1.11E-05	4.954666083

BovineHD2100002841	21	11435365	1.10E-06	5.958586795
BovineHD1400007235	14	24915886	2.66E-06	5.574912933
BovineHD1400007239	14	24931388	2.98E-06	5.525355333
BovineHD1400007241	14	24939285	2.84E-06	5.547329515
BovineHD1400007244	14	24952035	5.35E-06	5.271465531
BovineHD0400031625	4	110440534	1.59E-06	5.799491665
BovineHD0400021653	4	78107837	1.66E-06	5.781165627
BovineHD1400007250	14	24975563	1.31E-05	4.882446206
BovineHD2100001390	21	6774864	2.01E-06	5.697323143
BovineHD1400007253	14	24998326	8.03E-06	5.095418045
BovineHD1400007259	14	25015640	6.39E-06	5.194234743
BovineHD1400007269	14	25058053	4.70E-07	6.328235741
BovineHD1400007277	14	25098364	5.69E-06	5.24510948
Hapmap41234-BTA-34285	14	25107556	1.50E-06	5.82519322
BovineHD1400007289	14	25147967	1.36E-06	5.867833324
BovineHD1400007290	14	25154132	2.38E-07	6.624293126
BovineHD1400007293	14	25160597	2.40E-06	5.619185135
BovineHD1400007295	14	25164603	8.21E-07	6.085821573
BovineHD1000005601	10	16769219	3.01E-06	5.520812267
BTB-02056709	14	25175950	1.45E-06	5.838739775
BovineHD1400007303	14	25204467	4.28E-07	6.368858553
BovineHD2600006035	26	23405679	3.08E-06	5.511800435
BovineHD1300004547	13	16097639	3.49E-06	5.456599158
BovineHD2700008950	27	31926152	4.05E-06	5.392089009
BovineHD1400007308	14	25225097	1.12E-05	4.950810073
BovineHD1400007314	14	25241366	1.27E-05	4.896602625
BovineHD0100033478	1	118606312	4.58E-06	5.338997868
BovineHD0300014999	3	49436731	4.61E-06	5.336623795
BovineHD1100030501	11	104936387	4.69E-06	5.328668124
BovineHD0400033218	4	114808853	5.23E-06	5.281545297
BovineHD1400007323	14	25276491	7.37E-06	5.132378166
BovineHD4100011295	14	25284162	6.47E-06	5.18924508
BovineHD2100001389	21	6772933	5.67E-06	5.246137274
BovineHD1400007328	14	25298972	2.87E-06	5.541916589
Hapmap46986-BTA-34282	14	25307116	4.97E-07	6.303646938
BovineHD1400007333	14	25329035	3.06E-06	5.514081593
BovineHD1400007335	14	25336906	1.35E-05	4.870437596
BTB-01779799	14	25351733	1.29E-05	4.890494062
BovineHD1400007343	14	25376827	1.18E-05	4.927789654
BovineHD1000007163	10	22126960	6.71E-06	5.173351835
BovineHD1400007357	14	25446793	8.18E-06	5.087184125
BovineHD2900002706	29	9174027	7.76E-06	5.110077297
BovineHD0700027589	7	94710749	7.78E-06	5.109088898
BovineHD1400007360	14	25457504	8.72E-06	5.059370291
BovineHD1100029283	11	100818657	8.13E-06	5.089887306

ARS-BFGL-NGS-529	14	25638580	2.76E-06	5.559856855
BovineHD1600000503	16	1929989	8.38E-06	5.076915156
BovineHD2900004050	29	13652433	8.39E-06	5.076419877
BovineHD1300000204	13	1054489	8.60E-06	5.065431909
BovineHD2100001392	21	6783252	8.61E-06	5.064973196
BovineHD1400007433	14	25708285	9.39E-07	6.027155596
Hapmap38507-BTA-52931	21	7400647	9.21E-06	5.035899517
BovineHD0700005350	7	19212164	9.59E-06	5.018381246
BovineHD0300005372	3	16508680	9.76E-06	5.010626935
Hapmap27055-BTA-99423	7	46700828	9.85E-06	5.006777754
BovineHD0300002672	3	8172732	1.03E-05	4.987527239
BovineHD1000011955	10	38623674	1.06E-05	4.973333066
BovineHD2100002600	21	10677966	1.08E-05	4.964679782
BovineHD1400007456	14	25800191	1.01E-06	5.995849698
BovineHD1400007462	14	25819872	7.09E-08	7.149230546
BovineHD0700027588	7	94708540	1.17E-05	4.933255951
BovineHD2600006042	26	23419423	1.18E-05	4.92876261
BovineHD1400007464	14	25826189	1.04E-07	6.983375724
BovineHD2400000603	24	2276768	1.20E-05	4.919710036
BovineHD0700027596	7	94730969	1.26E-05	4.9004636
BovineHD1400007500	14	25972263	1.33E-05	4.877276725
BovineHD1400008558	14	29505756	5.40E-06	5.267759612
BTB-00819187	21	40072235	1.30E-05	4.884841496
BovineHD1400008559	14	29508538	6.46E-06	5.190064591
BovineHD0700027421	7	94108481	1.32E-05	4.877996829
BovineHD1400008574	14	29564228	4.10E-06	5.387698904
BovineHD1400008577	14	29576218	1.80E-06	5.744559623

Table S2. Validated SNP for age at first *corpus luteum* (AGECL), first postpartum anoestrus interval (PPAI), and scrotal circumference (SC_{TC}) of Tropical Composite cattle associated with sexual precocity in Nellore and Brahman.

SNP	BTA[1]	Position, bp[1]
AGECL		
BovineHD0100012377	1	43458869
BovineHD1400007269	14	25058053
BovineHD1400007277	14	25098364
Hapmap41234-BTA-34285	14	25107556
BovineHD1400007289	14	25147967
BovineHD1400007290	14	25154132
BovineHD1400007293	14	25160597
BovineHD1400007295	14	25164603
BTB-02056709	14	25175950
BovineHD1400007303	14	25204467

BovineHD1400007308	14	25225097
BovineHD1400007314	14	25241366
BovineHD1400007323	14	25276491
BovineHD4100011295	14	25284162
BovineHD1400007328	14	25298972
Hapmap46986-BTA-34282	14	25307116
BovineHD1400007330	14	25312775
BovineHD1400007333	14	25329035
BovineHD1400007335	14	25336906
BTB-01779799	14	25351733
BovineHD1400007343	14	25376827
BovineHD1400007345	14	25383331
BovineHD1400007346	14	25389001
BovineHD1400007357	14	25446793
BovineHD1400007360	14	25457504
BovineHD1500002332	15	9064376
BovineHD1600006716	16	24282102
BovineHD2100000292	21	2329730
BovineHD2100001045	21	5655578
BovineHD2100001367	21	6698573
BovineHD2100001376	21	6721504
BovineHD2100001389	21	6772933
BovineHD2100001390	21	6774864
BovineHD2100001392	21	6783252
BovineHD2100001408	21	6839238
BovineHD2100001433	21	6915751
BovineHD2100001450	21	6972071
BovineHD2100001572	21	7260381
BovineHD2100001576	21	7293437
BovineHD2100001577	21	7294720
Hapmap38507-BTA-52931	21	7400647
BovineHD2100001045	21	5655578
BovineHD2100002586	21	10644504
BovineHD2100002600	21	10677966
BovineHD4100014975	21	10702540
BovineHD2300006383	23	23921232
BovineHD2300007697	23	27782568
BovineHD2600006035	26	23405679
BovineHD2600006042	26	23419423
BovineHD2900002706	29	9174027
PPAI		
BovineHD0100042609	1	147345274
BovineHD1000029236	10	100984427
BovineHD1000029240	10	100990620
ARS-BFGL-NGS-68030	11	1901787
SC_{TC}		

BovineHD0800024244	8	81614440
BovineHD1000007157	10	22121623
BovineHD1000007163	10	22126960
Hapmap55139-rs29022040	10	38622420
BovineHD1000011955	10	38623674
BovineHD1100029283	11	100818657
BovineHD1600006716	16	24282102
BovineHD2100002586	21	10644504
BovineHD2100002600	21	10677966
BovineHD4100014975	21	10702540
BovineHD2100001045	21	5655578
BovineHD2300006383	23	23921232
BovineHD2400011464	24	41304824
ARS-BFGL-NGS-118334	27	41297038

[1] SNP were mapped to the UMD3.1 Bovine Assembly.