

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**POLIMORFISMOS EM GENES ASSOCIADOS À
PUBERDADE E OCORRÊNCIA DE PRENHEZ PRECOCE EM
BOVINOS DE CORTE**

**Marina Mortati Dias
Biotecnologista**

2016

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

**POLIMORFISMOS EM GENES ASSOCIADOS À
PUBERDADE E OCORRÊNCIA DE PRENHEZ PRECOCE EM
BOVINOS DE CORTE**

Marina Mortati Dias

Orientador: Prof. Dr. Henrique Nunes de Oliveira

Coorientador: Prof. Dr. Fábio Ricardo Pablos de Souza

Dra. Iara Del Pilar Solar Diaz

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias (FCAV), Universidade Estadual Paulista (UNESP), Câmpus de Jaboticabal, como parte das exigências para o título de Doutor em Genética e Melhoramento Animal.

2016

D541p Dias, Marina Mortati
Polimorfismos em genes associados à puberdade e ocorrência de
prenhez precoce em bovinos de corte / Marina Mortati Dias. –
Jaboticabal, 2016
vii, 63p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de
Ciências Agrárias e Veterinárias, 2016
Orientador: Henrique Nunes de Oliveira
Co-orientadores: Fábio Ricardo Pablos de Souza, Iara Del Solar
Diaz

Banca examinadora: Lucia Galvão de Albuquerque, Vera
Fernanda Martins Hossepian de Lima, Wilson Malagó Junior,
Guilherme Costa Venturini.

Bibliografia

1. Características reprodutivas. 2. SNP. 3. DNA. 4. RNA. I. Título.
II. Jaboticabal - Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.2:636.082

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação
– Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO: POLIMORFISMOS EM GENES ASSOCIADOS À PUBERDADE E OCORRÊNCIA
DE PRENHEZ PRECOCE EM BOVINOS DE CORTE

AUTORA: MARINA MORTATI DIAS

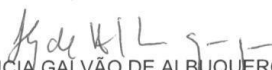
ORIENTADOR: HENRIQUE NUNES DE OLIVEIRA

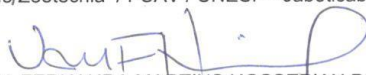
CO-ORIENTADORA: IARA DEL PILAR SOLAR DIAZ

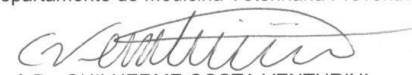
CO-ORIENTADOR: FABIO RICARDO PABLOS DE SOUZA

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


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


Profa. Dra. LUCIA GALVÃO DE ALBUQUERQUE
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal


Profa. Dra. VERA FERNANDA MARTINS HOSSEIAN DE LIMA
Departamento de Medicina Veterinária Preventiva e Reprodução Animal / FCAV / UNESP - Jaboticabal


Prof. Dr. GUILHERME COSTA VENTURINI
Departamento de Melhoramento e Nutrição Animal / Faculdade de Medicina Veterinária e Zootecnia de
Botucatu


Analista Dr. WILSON MALAGO JUNIOR
Empresa Brasileira de Pesquisa Agropecuária / EMBRAPA / São Carlos/SP

Jaboticabal, 27 de julho de 2016.

DADOS CURRICULARES DA AUTORA

MARINA MORTATI DIAS: nascida em Araraquara, São Paulo, em 7 de Junho de 1985, filha de Gabriel Dias Neto e Inês Maria Mortati Dias. Iniciou em março de 2007 o curso de graduação em Biotecnologia na Universidade Federal de São Carlos, Centro de Ciências Agrárias, campus de Araras obtendo o título de Biotecnologista em dezembro de 2010. Durante a graduação, foi bolsista PIBIC de Iniciação Científica sob orientação do Prof. Dr. Jozivaldo Prudêncio Gomes de Moraes; estagiou na EMBRAPA Pecuária Sudeste sob orientação da Dra Simone Cristina Meo Niciura e na EMBRAPA Pecuária Sul sob orientação da Dra Renata Wolf Suñé Martins da Silva. Em março de 2011, ingressou no Programa de Pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus de Jaboticabal como bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) sob orientação do Prof. Dr. Henrique Nunes de Oliveira. Obteve o título de Mestre em Genética e Melhoramento Animal em julho de 2012. Em agosto de 2012, ingressou no Programa de Pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus de Jaboticabal como bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) sob orientação do Prof. Dr. Henrique Nunes de Oliveira. Realizou doutorado sanduíche com auxílio do Programa de Doutorado Sanduíche no Exterior (PDSE) na Colorado State University (CSU), Fort Collins, CO, EUA (2014 - 2015), sob orientação do Dr Milton G. Thomas, retornando ao Brasil para conclusão do Doutorado pela FCAV, UNESP, Câmpus de Jaboticabal, São Paulo.

"Em algum lugar, algo incrível está esperando para ser descoberto"
Carl Sagan

"Se consegui ver mais longe foi por estar de pé nos ombros de gigantes"
Isaac Newton (1676)

Aos meus pais Inês Maria e Gabriel, ao meu irmão Guilherme e sua esposa, Steffany ,e minha sobrinha, Isis. Ao meu noivo e sua família, que agora são minha família também. À esta grande família que é meu porto seguro.

Dedico e Ofereço

AGRADECIMENTOS

Agradeço...

À Deus, por ter me dado tudo o que eu tenho, família, amigos, meu amor, e todas as oportunidades e até mesmo posto em meu caminho obstáculos, pois tudo o que eu vivi e aprendi é que me tornou quem sou hoje.

Ao professor Dr. Henrique Nunes de Oliveira, que mais uma vez me guiou nessa jornada. O senhor me ensinou além de melhoramento genético animal, o senhor me mostrou o que é um bom orientador, que se não der certo tentaremos de novo sem desistir, com calma e um sorriso podemos resolver tudo. Muito Obrigada!!

À minha família, por sempre, sempre, estar ao meu lado, independente das minhas escolhas. Foi a educação, o exemplo, amizade e amor que sempre encontrei dentro de casa que me faz acreditar que posso alcançar todos os meus objetivos Sem vocês eu não sou nada!

Ao meu noivo: Rondineli por ser meu melhor amigo e companheiro, por me apoiar, me entender, me ajudar e por sempre estar ao meu lado em minhas decisões. Obrigada por fazer parte da minha vida e me ajudar a ser uma pessoa melhor!

Aos meus co-orientadores Fábio e Iara, por toda ajuda, conhecimento e amizade. Mesmo de longe vocês me deram apoio e nunca se ausentaram.

Ao meu co-orientador no exterior Dr Milt Thomas pela atenção, paciência, oportunidade e por todos os ensinamentos. (*To my advisor in the sandwich period, Dr Milt Thomas, for the attention, patience, opportunity and all the*)

Aos colaboradores Dr Medrano, Dr Cánovas, Dr Riley, Dr Luna-Nevarez, Dr Enns, Dr Speidel, Dr Coleman, Dr Fernando Baldi, Dra Lúcia Galvão de Albuquerque e muitos outros que sempre me ajudaram. (*To the collaborators Dr Medrano, Dr Cánovas, Dr Riley, Dr Luna-Navarez, Dr Enns, Dr Speidel, Dr Coleman, Dr Fernando Baldi, Dr Lúcia Galvão de Albuquerque and many others that always helped me*)

Aos meus amigos do departamento da pós-graduação: Gregório, Adriana, Ana Cláudia, Fernanda, Luciana, Raphael, Natália, Ana Paula, Daniel e Dierclés. Obrigado pela convivência, risadas e ensinamentos. E o mais importante eu sei que posso contar com vocês!

Aos meus amigos que fiz durante meu período sanduíche, Mariana, Kristin, Elena, Miranda, Xi, Natalie, Ana, Miguel e Talita, que me apoiaram, me ajudaram, e

mesmo com alguns barreiras lingüísticas e culturais eu sei que são amigos para a vida.
(*To my friends that I did in the sandwich period, Mariana, Kristin, Elena, Miranda, Xi, Natalie, Ana, Miguel and, Talita, that supported me, helped me, and, even with some language problems and, cultures I know that they are my friends for life*)

Aos meus amigos de longa data: Ellen, Gisela, Vanessa, Débora (Frita), Mariana (Pirikita), André (Xuxa), Camila (Bagaça). Obrigado por estarem em minha vida e por entenderem me mesmo distante e muitas vezes ausentes eu sempre estou com vocês.

À minha banca de qualificação, Dra Lúcia Albuquerque, Dr Daniel Santos, Dr Rogério Curi e Dr. Fernando Baldi pela disponibilidade atenção e valiosas sugestões para aprimorar este trabalho

À Conexão Delta G pela doação do material biológico e disponibilidade dos dados.

À UNESP e ao Programa de Pós-Graduação, pelo apoio financeiro devido à bolsa do programa de pós-graduação junto com a **CAPES**, pela oportunidade de crescimento científico.

À Colorado State University pela atenção que tiveram comigo durante o meu período de doutorado sanduíche.

Aos funcionários da FCAV-Unesp/Jaboticabal pela ajuda principalmente da Seção de Pós-Graduação, Biblioteca e Departamento de Zootecnia.

Ao técnico do laboratório de Genética Molecular: Paulo pela ajuda.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (#2013/01108-1) pelo apoio financeiro.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos durante o período do doutorado e ao **Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)** pela bolsa de estudos durante o período do doutorado sanduíche.

à todos vocês o meu:

Muito Obrigada!!

Thank you so much!!

SUMÁRIO

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS.....	1
1. Introdução	1
2. Revisão de Literatura	2
2.1. Puberdade em Bovinos de Corte	2
2.2. Melhoramento Genético Animal da Puberdade em Bovinos de Corte.....	4
2.3. Biologia Molecular Aplicada ao Melhoramento Animal.....	5
3. Objetivo.....	8
4. Referências Bibliográficas	8
CAPÍTULO 2 - Prospecção de Polimorfismos nos Genes <i>PPP3CA</i> e <i>FABP4</i> e sua associação com a ocorrência prenhez precoce em novilhas Nelore	13
RESUMO	13
2.1. Introdução	15
2.2. Material e Métodos	16
2.2.1. Animais	16
2.2.2. Extração de DNA.....	17
2.2.3. Genes Candidatos.....	17
2.2.4. Customização de Painel, Sequenciamento MiSeq e Detecção de Variante	18
2.2.5. Frequência Alélica e Genotípica e Equilíbrio de Hardy-Weinberg	19
2.2.6. Análise de Desequilíbrio de Ligação	19
2.2.7. Controle de Qualidade dos Genótipos.....	20
2.2.8. Imputação de Genótipos e Reconstrução do Haplótipo.....	21
2.2.9. Análise Estatística	21
2.3. Resultados e Discussão	22
2.4. Conclusões	27
2.5. Referências Bibliográficas.....	27
CAPÍTULO 3 - Identificação de SNP em genes candidatos associados com Puberdade em fêmeas bovinas.....	30
1. Introdução	32
2. Material e Métodos	33
2.1. População	33
2.1.1. População de Descoberta.....	33
2.1.2. População de Comparação	35

2.1.2.1. Angus	35
2.1.2.2. Brahman	35
2.1.2.3. Nelore.....	35
2.1.2.4. Holandês	35
2.2. Genes Candidatos	36
2.4. Sequencias de RNA e Controle de Qualidade.....	38
2.5. Identificação de SNPs nos genes candidatos.....	39
2.7. Descoberta de Transcritos e Genes	40
3. Resultados e Discussão.....	40
3.1. RNA-Seq	40
3.3.2. Identificação de SNP e comparação entre Raças	40
3.3.4. Descoberta de novos Transcritos e Gene	50
3.5. Conclusões	53
3.6. Referências Bibliográficas.....	53

RESUMO: POLIMORFISMOS EM GENES ASSOCIADOS À PUBERDADE E OCORRÊNCIA DE PREENHEZ PRECOCE EM BOVINOS DE CORTE

Características reprodutivas, como a idade à puberdade, são economicamente relevantes para sistemas de produção de carne. A ocorrência de prenhez precoce aos 16 meses (OPP) é considerada um indicador da idade à puberdade e pode ser utilizada como critério de seleção. A busca por mutações potencialmente causais em genes candidatos pode ajudar a melhorar a acurácia de predição genômica se forem inseridas em painéis de marcadores genéticos. O objetivo desse estudo foi identificar polimorfismos potencialmente causais em genes candidatos associados a características reprodutivas em novilhas, por meio de dois ensaios. No primeiro, foram sequenciadas regiões dos genes *PPP3CA* e *FABP4* em 380 amostras de DNA provenientes de novilhas Nelore, por meio de equipamento MiSeq® (Illumina, Inc). Nestas sequências foram detectadas duas mutações na região do gene *PPP3CA* e 13 no gene *FABP4*. Dentre elas, 14 eram SNP e uma era deleção, sendo esta última homozigota em todas as amostras sequenciadas, o que nos leva a acreditar que possa ser uma mutação fixada na raça Nelore. Um haplótipo constituído por quatro SNP no gene *FABP4* não teve efeito significativo ao nível de $p < 0,05$ sobre a característica OPP. Todos os SNP que passaram pelo controle de qualidade também foram analisados, porém nenhum teve efeito significativo sobre a característica OPP ($p > 0,05$). No segundo ensaio, foram analisadas sequências de amostras de RNA de novilhas Brangus com o objetivo de identificar polimorfismos nas regiões de 62 genes candidatos associados a características reprodutivas. Detectou-se 1157 SNP nas regiões de 46 daqueles genes. Esses SNP foram comparados com SNP detectados a partir de sequências de amostras de RNA de outras quatro raças (Angus, Brahman, Nelore e Holandês). Cento e setenta e dois SNP foram coincidentes em todas as raças. A partir de sequências de RNA de Brangus também foram detectados 160 novos transcritos na região de 42 dos genes candidatos e cinco genes não anotados que sobrepõem a região de genes candidatos. A detecção de novos transcritos e genes é fundamental para enriquecer a anotação do genoma bovino. As mutações detectadas podem ser utilizadas para a customização de painéis de marcadores, e aplicações de seleção genômica.

Palavra chave: Características Reprodutivas, DNA, RNA, SNP, Transcritos.

ABSTRACT: POLYMORPHISMS IN GENES ASSOCIATED WITH PUBERTY AND OCCURRENCE OF EARLY PREGNANCY IN BEEF CATTLE

Fertility traits such as age at puberty are economically important for meat production system. Occurrence of early pregnancy (OPP) is considered a good indicator of age at puberty and can be used as selection criteria. The search for causal mutations in candidate genes can help to improve the accuracy of genomic prediction if used to design low-density chips. The objective of this study was to identify polymorphism in candidate genes associated with puberty in heifers. DNA sequences were obtained from 380 Nellore heifers from exons sequencing of *PPP3CA* and *FABP4* genes through MiSeq® equipment (Illumina, Inc.). From these sequences were found two SNP in the region of *PPP3CA* gene and 13 variations in the gene *FABP4*. Among these 15 variations 12 were SNP and one was a deletion, and this deletion was detected in all samples sequenced, which leads us to conclude that this may be a variation between Nellore and Hereford. One haplotype consisting of four SNP located in the *FABP4* and nine SNP, that were analyzed separately, had not a significant effect at the $p < 0.05$ on OPP characteristic. Also were analyzed RNA sequences from Brangus heifers in 62 candidate genes associated with puberty with the aim to discovery SNP in these regions. Were detected 1157 SNP in the region of 46 gene. These SNP were compared with SNP detected from RNA sequences from four breeds, Angus, Brahma, Nellore and Holstein. One hundred and seventy-two SNP were matched in all breeds. From Brangus RNA sequences were also detected 160 new transcripts in the region of 42 candidate genes and five genes not annotated that overlap the region of candidate genes. Detection of new transcripts and genes is essential to enrich the bovine genome annotation. The mutations detected can be used for the customization of low density chips and future genomics selection strategies.

Keywords: Fertility Traits, DNA, RNA, SNP, Transcripts.

LISTA DE ABREVIATURAS

ADH6: Alcohol dehydrogenase 6 (class V)

Ala: Alanina

Arg: Arginina

Asn: Asparagina

Asp: Ácido aspártico

C10H11ORF46: ADP-ribosylation factor-like 14 effector protein-like

CARTPT: CART prepropeptide

CBLN1: Cerebellin 1 precursor

CBLN2: Cerebellin 2 precursor

CBLN4: Cerebellin 4 precursor

CCK: Cholecystokinin

CDX2: Caudal type homeobox 2

CHGA: Chromogranin A (parathyroid secretory protein 1)

CHGB: Secretogranin-1

CPNE5: Copine V

Cys: Cisteína

DACH2: Dachshund homolog 2 (Drosophila)

DE: Genes diferencialmente expresos

DKK1: Dickkopf homolog 1 (Xenopus laevis)

DNA: ácido desoxiribonucleico

DPPA4: Developmental Pluripotency Associated 4

E2F3: E2F transcription factor 3

ELF5: E74-like factor 5 (ets domain transcription factor)

FABP4: fatty acid binding protein 4, adipocyte

FAM19A4: Family with sequence similarity 19 (chemokine (C-C motif)-like), member A4

FOXB2: Forkhead box B2

FOXE1: Forkhead box E1 (thyroid transcription factor 2)

FOXP4: Forkhead box N4

FSH: hormônio folículo estimulante

GATA1: GATA binding protein 1 (globin transcription factor 1)

Gln: Glutamina

Glu: Ácido glutâmico

Gly: Glicina

GnRH: hormônio liberador de gonadotrofina

GS: Seleção Genômica (*Genome-wide marker assisted selection*)

His: Histidina

IL22RA1: Interleukin 22 receptor, alpha 1

Ile: Isoleucina

INHA: Inhibin, alpha

ITIH1: Inter-alpha-trypsin inhibitor heavy chain 1

KLF1: Kruppel-like factor 1 (erythroid)

LD: Desequilíbrio de Ligação (*Linkage Disequilibrium*)

Leu: Leucina

LH: resultante da liberação do hormônio luteinizante

LHX4: LIM homeobox 4

LOC777593: Uncharacterized LOC777593

Lys: Lisina

LYSB: Intestinal lysozyme

MAF: Frequência do alelo de menor frequência (*Minor Allele Frequency*)

MAS: Seleção Assistida por Marcadores (*Marker Assisted Selection*)

MEPE: Matrix extracellular phosphoglycoprotein

Met: Metionina

MGC157266: Uncharacterized LOC507550

miRNA: micro RNA

MMD2: Monocyte to macrophage differentiation-associated 2

MOS: V-mos Moloney murine sarcoma viral oncogene homolog

MYOCD: Myocardin-like

NEBL: Nebulette

NHLH1: Nescient helix loop helix 1

NRXN3: Neurexin 3

NTS: Neurotensin

OPP: Ocorrência de Prenhez Precoce

OVGP1: Oviductal glycoprotein 1, 120kDa

PCSK1N: Proprotein convertase subtilisin/kexin type 1 inhibitor

PENK: Proenkephalin

PEPT-1: Solute carrier family 15 (oligopeptide transporter), member 1

Phe: Fenilalanina

PITX2: Paired-like homeodomain 2

POMC: Proopiomelanocortin

POU4F2: POU class 4 homeobox 2

PPP3CA: protein phosphatase 3, catalytic subunit, alpha isozyme

Pro: Prolina

PROP1: PROP paired-like homeobox 1

QTL: *Quantitative Trait Loci*

RHEBL1: Ras homolog enriched in brain like 1

RNA: ácido ribonucléico

RNAm: RNA mensageiro

SCG2: Secretogranin II

SCG5: Secretogranin V (7B2 protein)

Ser: Serina

SGC3: Secretogranin III

SIX6: SIX homeobox 6

SNP: Polimorfismo de Base Única (*Single Nucleotide Polymorphism*)

SST: Somatostatin

STMN1: Stathmin 1

SYCE1: Synaptonemal complex central element protein 1

TAC 1: Tachykinin, precursor 1

TECRL: Trans-2,3-enoyl-CoA reductase-like

TF: fatores de transcrição

Thr: Treonina

TP63: Tumor Protein p63

TS: específico de tecido

TSG101: Tumor susceptibility gene 101

TSHR: Thyroid stimulating hormone receptor

Val: Valina

VAX2: Ventral anterior homeobox 2

VGLL1: Vestigial like 1 (Drosophila)

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS

1. Introdução

A produção de gado de corte depende do desempenho reprodutivo dos rebanhos. A eficiência reprodutiva interfere na lucratividade dessa atividade (OLIVEIRA et al., 2003), e é mais relevante economicamente do que as características de crescimento (DICKERSON, 1970; BRUMATTI et al., 2011). Porém devido às baixas herdabilidades das características reprodutivas, por serem de expressão tardia e limitadas a um sexo e por dificuldades de mensuração e falhas de registros, estas características têm sido pouco utilizadas em programas de melhoramento genético (MACNEIL et al., 2006), não somente no Brasil como no mundo.

Uma das ferramentas indicadas para características com poucos QTL, e que são expressas tardiamente na vida do animal ou ainda, que são mensuradas em apenas um sexo é a seleção assistida por marcadores (Marker-assisted selection – MAS). Este tipo de seleção foi proposta com o intuito de aumentar a acurácia das avaliações genéticas e como ferramenta para selecionar as mutações desejáveis (EGGEN, 2012). Neste sentido, a mesma seria apropriada para a seleção de características reprodutivas, especialmente as características que se expressam apenas nas fêmeas (EGGEN, 2012). Entretanto, foi observado que somente uma pequena proporção da variância genética é capturada com a implementação da MAS, limitando assim, o ganho genético (SOLBERG et al., 2008).

Visando superar as limitações da MAS, Meuwissen et al. (2001) propuseram a utilização da chamada seleção assistida por marcadores considerando todo o genoma (*Genome-wide marker assisted selection* ou seleção genômica, GS). Essa seleção considera um conjunto muito denso de marcadores cobrindo todo o genoma o que, potencialmente, permitiria explicar grande parte da variância genética. Em painéis densos, existe grande probabilidade de que haja marcadores em desequilíbrio de ligação com grande parte dos QTL responsáveis pela expressão da característica, de forma que os efeitos detectados em amostras da população (população referência) sejam

válidos para a população. A GS vem sendo utilizada com sucesso, especialmente em bovinos leiteiros. Porém os custos da GS ainda são elevados.

Uma alternativa para reduzir o custo da GS é a customização de painéis de baixa densidade, que podem contribuir para a avaliação genética dos reprodutores com boa relação custo benefício (SNELLING et al., 2012). Esses painéis, quando customizados com mutações causais, são viáveis pois, aumentam persistentemente as acurácias dos valores genômicos ao longo das gerações (HAYES et al., 2014), uma vez que o efeito detectado da mutação não é dependente do desequilíbrio de ligação (FORTES et al., 2014)

2. Revisão de Literatura

2.1. Puberdade em Bovinos de Corte

A puberdade pode ser definida como o primeiro estro acompanhado pelo desenvolvimento de um corpo lúteo capaz de se manter durante um ciclo estral completo (KINDER; DAY; KITTOCK, 1987). Em muitos casos, após a puberdade ocorrem ciclos estrais curtos que podem ou não serem acompanhados pela ovulação (DAY et al., 2010). Este é um processo gradual de maturação que se inicia durante o desenvolvimento fetal e continua pelo período pré e peripubertal (KINDER; DAY; KITTOCK, 1987), podendo ser atingida entre os 6 e 24 meses de idade (MORAN; QUIRKE; ROCHE, 1989). Portanto, a puberdade é consequência de uma série de eventos fisiológicos, sendo dependente de fatores ambientais, genéticos e raciais.

O processo fundamental para o início da puberdade é a integração do eixo hipotálamo-hipófise-gônada. O estímulo hormonal necessário para a ocorrência da puberdade está relacionada com o aumento na secreção do hormônio liberador de gonadotrofina (GnRH) resultante da liberação do hormônio luteinizante (LH) e do hormônio folículo estimulante (FSH) (RODRIGUES, WISE 1989 citado por DAY et al., 2010). Na pré-puberdade, a secreção do GnRH é muito sensível ao *feedback* negativo do estradiol (KINDER; DAY; KITTOCK, 1987). Durante a peripuberdade ocorre um declínio no *feedback* negativo do estradiol e, com isso, inicia-se uma maior secreção do

GnRH e consequentemente do LH (DAY et al. 1987 citado por DAY et al., 2010). Essa diminuição do feedback negativo coincide com o aumento da deposição de gordura (FONSECA-ALANIZ et al., 2006). Isso deve-se ao fato de que o tecido adiposo é um dos principais locais do metabolismo dos esteroides sexuais (KERSHAW; FLIER, 2004).

Segundo Joubert (1963 citado por NRC, 1996) as fêmeas atingem a puberdade depois de atingirem a maturidade fisiológica. Baseado nessa, e em outras pesquisas, o NRC propõe que o peso alvo para a ocorrência da puberdade é de 60% do peso adulto para *Bos taurus taurus* e de 65% para *Bos taurus indicus* além do peso alvo, as novilhas devem possuir uma nutrição adequada pois, a restrição alimentar provoca uma diminuição dos pulsos de LH (SCHILLOZ; HALLS; HILEMAN, 1992). Essa diminuição da frequência dos pulsos de LH pode atrasar o início da puberdade, sendo que o restabelecimento do balanço energético positivo poderia estimulá-lo (DODE, 2002).

Além da influência hormonal e do efeito de manejo ainda há o efeito de raças. Animais *Bos taurus taurus*, em clima temperado, quando a nutrição não é fator limitante a puberdade pode ser atingida de 10 a 13 meses (HERRING, 2014). A idade a puberdade em gado Zebu pode variar de 16 a 40 meses (DAY; NOGUEIRA, 2013), de modo geral o gado Zebu atinge a puberdade de 6 a 12 meses mais tarde que os *Bos taurus taurus* (SARTORI et al., 2010). Essa grande diferença, provavelmente, deve-se a diferenças ambientais, nutricionais e genéticas.

O clima tropical é outro fator que pode afetar a puberdade das novilhas. Apesar do Zebu estar termicamente adaptado, os intervalos estrais e ovulatórios tem maior duração durante períodos com temperaturas mais elevadas (GAMA FILHO et al., 2007). Mezzadra et al. (1993) estudaram a influência da variação sazonal e de raças (Brahman, Hereford e novilhas mestiças) sobre a puberdade. Estes autores sugerem provável influência direta do fotoperíodo sobre a regulação da atividade do ciclo estral, sendo que esta influência foi detectada em novilhas com maior proporção de genes *Bos indicus*, fato não observado nos animais com maior influência taurina.

Além de todos os fatores levantados anteriormente, como por exemplo as diferenças nutricionais e do desenvolvimento corporal entre *Bos taurus taurus* criados em clima temperado e *Bos taurus indicus* criados em clima tropical, o gado Zebu teve uma menor pressão de seleção sobre a característica de puberdade do que o *Bos taurus* (ELER et al., 2002). Portanto é necessário uma maior atenção para a puberdade nos programas de Melhoramento Genético Animal

2.2. Melhoramento Genético Animal da Puberdade em Bovinos de Corte.

Nos programas de seleção de bovinos de corte em desenvolvimento no Brasil houve grande ênfase nas características de crescimento. Isso deve-se à disponibilidade de dados e relativa simplicidade para sua análise (EVANS et al., 1999). Dickerson (1970) já afirmava que as características de produção podem ser importantes para o aumento da lucratividade por animal porém, as características reprodutivas são o fator limitante para a eficiência da produção. Brumatti et al., (2011) evidenciaram que as características reprodutivas foram de 4 a 13 vezes mais importantes economicamente do que as características de crescimento.

Segundo Van Melis et al., (2010), o lucro obtido pelos produtores de gado de corte é diretamente influenciado pelas características reprodutivas, que determinam a quantidade de produtos que pode ser disponibilizada no mercado e afetam o intervalo de gerações e a intensidade de seleção. Portanto, é interessante que se promovam ações para melhorar o desempenho reprodutivo dos rebanhos. Mesmo que, estas sejam de difícil mensuração, medidas apenas em um dos sexos, tardiamente na vida do animal e, em geral, as características reprodutivas possuem baixa herdabilidade, dificultando o processo de seleção e aumentando o intervalo de gerações.

Dentre as características reprodutivas a idade a puberdade pode ser destacada, não apenas por seu efeito direto nos custos dos sistemas de produção mas também por sua importância para programas de melhoramento genético. A observação e registro da ocorrência da puberdade em novilhas, de forma sistemática é difícil. Entretanto, a prenhez de novilha exposta

precocemente tem sido considerada o critério de seleção mais adequado, pois é mais fácil de ser observado, além de indicar a puberdade precoce é também indicador da fertilidade. Apesar das características reprodutivas, em geral possuem baixa herdabilidade, com relação a ocorrência de prenhes precoce as estimativas de herdabilidade, medida como característica dicotômica, e analisada como característica de limiar, entre 14 a 18 meses, apresentam estimativas de herdabilidade que variam de 0,42 a 0,57 em animais *Bos taurus indicus* (ELER et al., 2004; SHIOTSUKI et al., 2009; BOLIGON; ALBUQUERQUE, 2011; SANTANA et al., 2012; VALENTE et al., 2015), indicando variabilidade genética suficiente para seleção de fêmeas com puberdade precoce. Em *Bos taurus taurus* a herdabilidade para idade a puberdade foi estimada em 0,52 (GARGANTINI et al., 2005).

Segundo Shiotsuki et al. (2009) a seleção para prenhez aos 16 meses resulta em animais com menor altura. Boligon e Albuquerque (2011), concluíram que as fêmeas selecionadas para puberdade aos 16 meses poderão ter peso corporal menor. Apesar da redução no peso, a seleção para ocorrência de prenhez precoce contribui para a habilidade de permanência da vaca no rebanho, aumentando assim a vida útil do animal (BUZANSKAS et al., 2010; SANTANA et al., 2012).

Pelo fato da ocorrência de prenhez ser uma característica de limiar todas as fêmeas desafiadas são incluídas nas análises, inclusive as novilhas que não ficaram prenhes, e também pelo fato desta característica possuir herdabilidade de média a alta magnitude, o que possibilita importantes ganhos quando utilizadas como critério de seleção nos programas de melhoramento genético animal por seleção.

2.3. Biologia Molecular Aplicada ao Melhoramento Animal

Com o desenvolvimento do método da reação em cadeia de polimerase (PCR) em 1983 (MULLIS, 1990) e dos métodos de sequenciamento de DNA, surgiu a possibilidade de mapear regiões genômicas relacionadas com características de interesse em várias espécies. Essas informações podem ser utilizadas para auxiliar a seleção de animais superiores para serem pais da próxima geração, antes mesmo da expressão de seu fenótipo (REGITANO;

COUTINHO, 2001). Porém para esta identificação dos indivíduos é necessário o conhecimento de alguma referência no material genético.

Os marcadores moleculares são variações existentes no genoma, e são considerados uma ferramenta importante para a identificação de indivíduos antecipadamente. Sendo que os marcadores moleculares podem ser associadas à características de interesse econômico. As características de interesse econômico geralmente apresentam padrão de manifestação quantitativo, pois são controladas por vários genes, cujos efeitos contribuem para a expressão do fenótipo (BOURDON, 2000). Afim de detectar os marcadores moleculares que possam ser associados com a característica, são estudados os genes candidatos que podem ser selecionados de duas maneiras: 1) genes que possuem efeito biológico sobre a característica, 2) genes detectados a partir de estudos de mapeamento de *QTLs* (*Quantitative Trait Loci*) (MONTALDO; MEZA-HERRERA, 1998).

Os genes candidatos são sequenciados, seus polimorfismos detectados e então são feitas análises para verificar a associação com a característica em questão. Se uma mutação é associada à variação na característica de interesse é possível realizar seleção direta. Nos estudos com genes candidatos, geralmente, busca-se por mutações potencialmente causais, ou em desequilíbrio de ligação (LD) com as mutações potencialmente causais. Mutações causais são polimorfismos não-sinônimos, que causam uma troca de aminoácidos, alterando a sequência de aminoácidos na proteína. algumas vezes causando a alteração na estrutura terciária da proteína, sendo possível a perda da função proteica (FORTES et al., 2014) ocasionando na variação fenotípica, que é detectada no estudo de associação. Já as mutações sinônimos são polimorfismos que não afetam a sequência de aminoácidos. Mas podem afetar a taxa de transcrição ou o volume de transcritos (FLINT; WOOLLIAMS, 2008).

De Camargo et al. (2014, 2015) utilizaram os genes candidatos *JY-1* e *NCOA2*, em estudo sobre probabilidade de prenhez precoce em novilhas Nelore, pois ambos possuem ação biológica relacionada com características reprodutivas. Estes estudos associaram as mutações de único polimorfismo (SNP) potencialmente causais encontradas, em ambos os genes, com a

probabilidade de prenhez precoce em novilhas Nelore. Dias et al. (2015) avaliaram genes candidatos do metabolismo de lipídeos associados a probabilidade de prenhez precoce na mesma população e encontraram haplótipos nos genes *FABP4* e *PPP3CA* associados com a característica.

Outro método que os marcadores moleculares podem ser utilizados é por meio do uso de painéis de SNP. O estudo que busca por associação entre SNP do painel e a característica de interesse é conhecido como estudo de associação ampla do genoma (*Genome Wide Association Studies – GWAS*), já que milhares de marcadores espalhados pelo genoma são testados ao mesmo tempo. Porém, estes SNP estão uniformemente distribuídos ao longo do genoma e, normalmente, em região intergênica, sendo que estes SNP são dependentes de estarem em desequilíbrio de ligação (LD) (FORTES et al., 2014).

A fim de validar nos genes associados com a características a partir de estudos de GWAS novas metodologias foram incorporadas. Cánovas et al. (2014) utilizaram dados de estudos de GWAS e analisou em conjunto com dados de expressão gênica, a partir do sequenciamento de moléculas de RNA. E, com isso, observaram genes potencialmente influentes sobre a puberdade em bovinos da raça Brangus. Estes autores construíram ainda uma rede de coexpressão gênica com a combinação desses resultados. Este tipo de análise aumenta o conhecimento de como os genes se relacionam e como afetam a característica.

Snelling et al. (2013) começam a introdução do artigo com um pergunta: "Por que a compreensão de polimorfismos funcionais é importante para a seleção genômica?". Muitos avanços já foram feitos com as metodologias estatísticas utilizando os dados de produção e de pedigree e a partir da incorporação de genótipos provenientes de todo o genoma aumentou a acurácia de predição (SNELLING et al., 2013). Porém, a vantagem dos polimorfismos causais (ou funcionais) é que estes não são dependentes do desequilíbrio de ligação e podem ser usados para seleção por meio das gerações e em diferentes raças (FORTES et al., 2014). Por essa razão, o uso do conhecimento do polimorfismos causais poderá contribuir para a avaliação e para todo o entendimento do funcionamento do genoma.

Robertson (1970) relatou que o desenvolvimento da biologia molecular até então, havia mudado radicalmente a visão da natureza da função gênica, porém, a biologia molecular tinha pouco impacto sobre o melhoramento animal. Apenas 39 anos foram necessários da publicação deste artigo para a publicação do genoma completo do bovino (ELSIK et al., 2009). Com o rápido desenvolvimento de tecnologias foi possível acelerar e melhorar a qualidade das pesquisas e muito mais há de se desvendar com os avanços das tecnologias de sequenciamentos e os estudos no nível "omics". Isso nos indica que a biologia molecular será, cada vez mais, empregada no melhoramento genético animal como uma ferramenta para a seleção de animais.

3. Objetivo

O objetivo geral desse estudo foi identificar polimorfismos potencialmente causais a partir de sequências de DNA e RNA em genes candidatos relacionados com a ocorrência de prenhez precoce (OPP) e a puberdade em diferentes raças bovinos de corte.

Os objetivos específicos foram:

- 1) Verificar a existência de polimorfismos em regiões de exon dos genes *PPP3CA* e *FABP4* e, uma vez identificados estes SNP, avaliar sua possível influência na característica OPP;
- 2) A partir da análise do transcriptoma de oito tecidos envolvidos no processo fisiológico da puberdade (hipotálamo, glândula pituitária, fígado, útero, endométrio, ovário, adiposo, e músculo *Longissimus dorsi*) de bovinos Brangus, identificar SNP em regiões do genoma correspondentes a 62 genes associados com a puberdade em bovinos Brangus e Nelore;
- 3) Verificar se os SNP identificados na raça Brangus podem ser encontrados em outras raças (Angus, Brahman, Holandês, e Nelore);
- 4) Pesquisar diferentes splicing alternativos nos genes candidatos e também se há outros genes que sobrepõem a região dos genes candidatos.

4. Referências Bibliográficas

ANUALPEC. **Anuário da pecuária brasileira**. 19. ed. Brasil: p.34, 2015.

BOLIGON, A. A.; ALBUQUERQUE, L. G. Genetic parameters and relationships of

heifer pregnancy and age at first calving with weight gain, yearling and mature weight in Nelore cattle. **Livestock Science**, v. 141, n. 1, p. 12–16, 2011.

BOURDON, R. M. **Understanding animal breeding**. 2. ed. New Jersey: Prendice-Hall Inc, 2000.

BRUMATTI, R. C.; FERRAZ, J. B. S.; ELER, J. P.; FORMIGONNI, E. I. B. Desenvolvimento de índice de seleção em gado corte sob o enfoque de um modelo bioeconômico. **Archivos de Zootecnia**, v. 60, n. 230, p. 205–213, 2011.

BUZANSKAS, M. E.; GROSSI, D. A.; BALDI, F.; BARROZO, D.; SILVA, L. O. C.; JÚNIOR, R. A. A. T.; MUNARI, D. P.; al. ENCAR, M. M. Genetic associations between stayability and reproductive and growth traits in Canchim beef cattle. **Livestock Science**, v. 132, n. 1-3, p. 107–112, 2010.

CÁNOVAS, A.; REVERTER, A.; DEATLEY, K. L.; ASHLEY, R. L.; COLGRAVE, M. L.; FORTES, M. R. S.; ISLAS-TREJO, A.; LEHNERT, S.; PORTO-NETO, L.; RINCÓN, G.; SILVER, G. A.; SNELLING, W. M.; MEDRANO, J. F.; THOMAS, M. G. Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. **PLoS ONE**, v. 9, n. 7, p. 1–17, 2014.

DAY, L. M.; GASSER, C. L.; GRUM, D. E.; PIRES, A. V. Fatores que afetam a idade na puberdade em novilhas de corte. In: PIRES, A. V. (Ed.). **Bovinocultura de corte**. 1. ed. Piracicaba: FEALQ, 2010. p. 637–652.

DAY, M. L.; NOGUEIRA, G. P. Management of age at puberty in beef heifers to optimize efficiency of beef production. **Animal Frontiers**, v. 3, p. 6–11, 2013.

DE CAMARGO, G. M. F.; COSTA, R. B.; ALBUQUERQUE, L. G.; REGITANO, L. C. A.; BALDI, F.; TONHATI, H. Polymorphisms in TOX and NCOA2 genes and their associations with reproductive traits in cattle. **Reproduction, Fertility and Development**, v. 27, n. 3, p. 523, 2015.

DE CAMARGO, G. M. F.; COSTA, R. B.; DE ALBUQUERQUE, L. G.; REGITANO, L. C. de A.; BALDI, F.; TONHATI, H. Association between JY-1 gene polymorphisms and reproductive traits in beef cattle. **Gene**, v. 533, n. 2, p. 477–480, 2014.

DIAS, M. M.; SOUZA, F. R. P.; TAKADA, L.; FEITOSA, F. L. B.; COSTA, R. B.; DIAZ, I. D. P. S.; CARDOSO, D. F.; TONUSSI, R. L.; BALDI, F.; ALBUQUERQUE, L. G.; OLIVEIRA, H. N. Study of lipid metabolism-related genes as candidate genes of sexual precocity in nelore cattle. **Genetics and Molecular Research**, 2015.

DICKERSON, G. Efficiency of animal production - molding the biological components. 1970.

DODE, M. A. . Aspectos fisiológicos da fecundação e estabelecimento da prenhez. In: SERENO, J. R. B.; LIMA, E. C. N. Z. (Ed.). **Eficiência no manejo reprodutivo: sucesso no rebanho de cria**. Campo Grande: Embrapa Gado de Corte, 2002. p. 117–134.

EGGEN, A. The development and application of genomic selection as a new breeding paradigm. **Animal Frontiers**, v. 2, n. 1, p. 10–15, 2012.

ELER, J. P.; SILVA, J. A. I. I. V.; EVANS, J. L.; FERRAZ, J. B. S.; DIAS, F.; GOLDEN, B. L. Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle The online version of this article ,al.ong with updated information and services , is located on the World Wide Web at : Additive genetic relationships bet. **Journal of Animal Science**, p. 2519–2527, 2004.

ELER, J. P.; SILVA, J. A. I. I. V.; FERRAZ, J. B. S.; DIAST, F.; OLIVEIRA, H. N.; EVANS, J. L.; GOLDEN, B. L. Genetic evaluation of the probability of pregnancy at 14 months for Nellore heifers. **Journal of Animal Science**, v. 80, n. 4, p. 951–954, 2002.

ELSIK, C. G.; TELLAM, R. L.; WORLEY, K. C. The Genome Sequence of Taurine Cattle: A window to ruminant Biology and Evolution. **Science**, v. 324, p. 522–527, 2009.

EVANS, J. L.; GOLDEN, B. L.; BOURDON, R. M.; LONG, K. L. Additive Genetic Relationships Between Heifer Pregnancy and Scrotal Circumference in Hereford Cattle. **Journal of Animal Science**, v. 77, p. 2621–2628, 1999.

FAO. **Meat Consumption**. Disponível em:
<<http://www.fao.org/ag/againfo/themes/en/meat/background.html>>.

FLINT, A. P. F.; WOOLLIAMS, J. A. Precision animal breeding. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 363, p. 573–590, 2008.

FONSECA-ALANIZ, M. H. O Tecido Adiposo Como Centro Regulador do Metabolismo. **Arq Bras Endocrino Metab**, v. 50, n. 2, p. 216–229, 2006.

FORTES, M. R.; PORTO-NET, L. R.; DEATLE, K. L.; REVERTE, A.; THOMAS, M. G.; MOOR, S. S.; LEHNER, S. A.; SNELLIN, W. M. Genetic Markers in Transcription Factors of Differentially Expressed Genes Associated with Post-partum Anoestrus Predict Pregnancy Outcome in an Independent Population of Beef Cattle. In: Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, Vancouver, Canada. **Anais...** Vancouver, Canada: 2014.

GAMA FILHO, R. V. da; FONSECA, F. A.; UENO, V. G.; FONTES, S.; GUALBERTO, J. L. Sazonalidade na dinâmica folicular ovariana e produção embrionária em novilhas da raça Guzerá. **Braz. J. vet. Res. anim. Sci.**, v. 44, n. 6, p. 422–427, 2007.

GARGANTINI, G.; CUNDIFF, L. V; LUNSTRA, D. D.; VLECK, L. D. Van. Genetic Relationships Between Male and Female Reproductive Traits in Beef Cattle. **The Professional Animal Scientist**, v. 21, n. 2005, p. 195–199, 2005.

HAYES, B. J.; MACLEOD, I. M.; DAETWYLER, H. D.; BOWMAN, P. J.; CHAMBERLIAN, A. J.; JAGT, C. J.; CAPITAN, V. A.; PAUSCH, H.; STOTHARD, P.; LIAO, X.; SCHROOTEN, C.; MULLAART, E.; FRIES, R.; GULDBRANDTSEN, B.; LUND, M. S.; BOICHARD, D. A.; VEERKAMP, R. F.; VANTASSELL, C. P.; GREDLER, B.; DRUET, T.; BAGNATO, A.; VILKKI, J.; DEKONING, D. J.; SANTUS, E.; GODDARD, M. E. Genomic Prediction from Whole Genome Sequence in Livestock: the 1000 Bull Genomes Project B.J. In: Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, Vancouver, Canada. **Anais...** Vancouver, Canada: 2014.

HERRING, A. D. **Beef cattle production systems**. Tarxien Malta: Gutenberg Press Ltd, 2014.

KERSHAW, E. E.; FLIER, J. S. Adipose Tissue as an Endocrine Organ. **The Journal of Clinical Endocrinology & Metabolism**, v. 89, n. 6, p. 2548–2556, 2004.

KINDER, J. E.; DAY, M. L.; KITTOCK, R. J. Endocrine regulation of puberty in cows and ewes. **Journal of reproduction and fertility. Supplement**, v. 34, p. 167–186, 1987.

LENSTRA, J. A.; BRADLEY, D. G. Systematics and phylogeny of gibbons. In: FRIES, R. R.; RUVINSKY, A. (Ed.). **The Genetics of Cattle**. New York, USA: Ed CABI Publishing, 1999. Ip. 15–31.

MACNEIL, M. D.; GEARY, T. W.; PERRY, G. A.; ROBERTS, A. J.; al. EXANDER, L. J. Genetic partitioning of variation in ovulatory follicle size and probability of pregnancy in beef cattle. **Journal of Animal Science**, v. 84, n. 7, p. 1646–1650, 2006.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819–1829, 2001.

MEZZADRA, C.; HOMSE, A.; SAMPEDRO, D.; ALBERIOL, R. Pubertal traits and seasonal variation of the sexual activity in Bhraman, Hereford and crossbred heifers. **Theriogenology**, v. 40, p. 987–996, 1993.

MONTALDO, H. H.; MEZA-HERRERA, C. a. Use of molecular markers and major genes in the genetic improvement of livestock. v. 1, n. 2, p. 1–7, 1998.

MORAN, C.; QUIRKE, J. F.; ROCHE, J. F. Puberty in heifers: A Review. **Animal Reproduction Science**, v. 18, n. 1-3, p. 167–182, 1989.

MULLIS, K. B. The Unusual origin of the Polymerase Chain Reaction. **Scientific American**, v. 262, p. 56–65, 1990.

NOGUEIRA, G. P. Puberty in South American Bos indicus (Zebu) cattle. **Animal Reproduction Science**, v. 82-83, n. August 2004, p. 361–372, 2004.

NRC. **Nutrient requeriments of beef cattle**. 7. ed. Washington, DC: National Academy Press, 1996.

OLIVEIRA, C. M. G. .; OLIVEIRA FILHO, B. D. .; GAMBARINI, M. L. .; KUNZ, T. L. .; VIU, M. A. Índice de gestação em novilhas nelore que atingiram a puberdade em idade precoce. **Revista Brasileira de Reprodução Animal**, v. 27, n. 2, p. 237–239, 2003.

PISON, G. World population: Seven billion today, how many tomorrow? **Population and Societies**, n. 482, p. 3–6, 2011.

PSOUNI, E.; JANKE, A.; GARWICZ, M. Impact of carnivory on human development and evolution revealed by a new unifying model of weaning in mammals. **PLoS ONE**, v. 7, n. 4, 2012.

REGITANO, L. C. de A.; COUTINHO, L. . **Biologia molecular aplicada à produção animal** Brasília: EMBRAPA-CENARGEN, 2001.

ROBERTSON, A. Molecular Biology And Animal Breeding. **Genetics selection evolution GSE**, v. 2, n. 4, p. 393–402, 1970.

SANTANA, M. L.; ELER, J. P.; FERRAZ, J. B. S.; MATTOS, E. C. Genetic relationship between growth and reproductive traits in Nellore cattle. **Animal : an international journal of animal bioscience**, v. 6, n. 4, p. 565–70, 2012.

SARTORI, R.; BASTOS, M. R.; S., B. P.; GIMENES, L. U.; ERENO, R. L.; BARROS, C. M. Physiological differences and implications to reproductive management of *Bos taurus* and *Bos indicus* cattle in a tropical environment. **Society of Reproduction and Fertility Supplement**, v. 67, p. 357–375, 2010.

SCHILLOZ, K. K.; HALLS, J. B.; HILEMAN, S. M. Effects of nutrition and season on the onset of puberty in the beef heifer. **Journal of Animal Science**, v. 70, p. 3994–4005, 1992.

SHIOTSUKI, L.; SILVA, J. A.; TONHATI, H.; ALBUQUERQUE, L. G. Genetic associations of sexual precocity with growth traits and visual scores of conformation, finishing, and muscling in Nelore cattle. **Journal of animal science**, v. 87, n. 5, p. 1591–1597, 2009.

SNELLING, W.; CUSHMAN, R.; KEELE, J.; MALTECCA, C.; THOMAS, M.; FORTES, M. R. S.; REVERTE, A. Networks and pathways to guide genomic selection. **Journal of animal ...**, v. 91, p. 537–552, 2013.

SNELLING, W. M.; CUSHMAN, R. A.; FORTES, M. R. S.; REVERTER, A.; BENNETT, G. L.; KEELE, J. W.; KUEHN, L. A.; MCDANELD, T. G.; THALLMAN, R. M.; THOMAS, M. G. Physiology and endocrinology symposium: How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. **Journal of Animal Science**, v. 90, n. 4, p. 1152–1165, 2012.

SOLBERG, T. R.; SONESSON, A. K.; WOOLLIAMS, J. A.; MEUWISSEN, T. H. E. Genomic selection using different marker types and densities. **Journal of Animal Science**, v. 86, n. 10, p. 2447–2454, 2008.

VALENTE, T. S.; SANT'ANNA, A. C.; BALDI, F.; ALBUQUERQUE, L. G.; DA COSTA, M. J. R. P. Genetic association between temperament and sexual precocity indicator traits in Nellore cattle. **Journal of Applied Genetics**, v. 56, n. 3, p. 349–354, 2015.

VAN MELIS, M. H.; ELER, J. P.; ROSA, G. J. M.; FERRAZ, J. B. S.; FIGUEIREDO, L. G. G.; MATTOS, E. C.; OLIVEIRA, H. N. Additive genetic relationships between scrotal circumference, heifer pregnancy, and stayability in Nellore cattle. **Journal of Animal Science**, v. 88, n. 12, p. 3809–3813, 2010.

WILLHAM, R. L. From husbandry to science: A highly significant facet of our livestock heritage. p. 1742–1758, 1986.

CAPÍTULO 2 - Prospecção de Polimorfismos nos Genes *PPP3CA* e *FABP4* e sua associação com a ocorrência prenhez precoce em novilhas Nelore

RESUMO:

A característica ocorrência de prenhez precoce aos 16 meses (OPP) é importante para a pecuária de corte uma vez que é indicadora da idade à puberdade. Entretanto, a variabilidade existente entre as novilhas para esta característica, nem sempre é detectada, pois a maioria dos criadores não expõe as novilhas quando ainda jovens. Este fato, aliado à sua expressão apenas nas fêmeas, dificulta aplicação da seleção tradicional. O presente trabalho teve como objetivo identificar polimorfismos nos genes *FABP4* e *PPP3CA* associados com a ocorrência prenhez precoce em estudos anteriores. Foram sequenciadas regiões de exon dos genes candidatos em 380 novilhas Nelore, utilizando-se o sequenciador MiSeq®. Dados de todas as novilhas sequenciadas e de outras 1309 genotipadas com o *High Density Bovine SNP BeadChip* (Illumina, Inc) foram utilizados nas análises. Os genótipos polimórficos encontrados no sequenciamento foram imputados para as amostras não sequenciadas. Dois polimorfismos foram detectados no gene *PPP3CA* e 13 no gene *FABP4*. Todos os polimorfismos encontrados já haviam sido relatados em outras raças. Foi identificada uma deleção em região de intron no gene *FABP4*, rs134413439, em todas as amostras. Como o genoma referência bovino é baseado na raça Hereford, essa deleção parece ser uma diferença entre a raça Nelore e Hereford, e não ser um polimorfismo no Nelore. As análises de associação mostraram que nenhum SNP ou haplótipo teve efeito significativo ao nível de 5% de significância para a característica de prenhez precoce em novilhas da raça Nelore. Este estudo mostrou que os genes *PPP3CA* e *FABP4* são polimórficos na raça Nelore. Não foram encontrados polimorfismos na região dos exons 1, 2 e 3 do gene *PPP3CA* e também nenhum SNP foi detectado no exon 1 do gene *FABP4*. Não foi verificada a influência destes marcadores na característica ocorrência de prenhez precoce.

Palavras-chave: metabolismo de lipídeos, puberdade, características reprodutivas, *Bos taurus indicus*

Chapter 2: Prospecting Polymorphisms within *PPP3CA* and *FABP4* genes and their association with occurrence of early pregnancy in Nellore heifers

Abstract

The early pregnancy probability (OPP) in heifers is a trait of high economic value for Nellore. However, the variability among heifers is not always detected because most breeders do not expose heifers while still young. The aim of this study was identify polymorphisms in the genes *FABP4* and *PPP3CA* previously associated with (OPP). Three hundred and eighty heifers were sequenced in the exon regions of the candidate genes with MiSeq® sequencer. Two polymorphisms were detected in *PPP3CA* gene and 13 in the gene *FABP4*. All polymorphisms identified have already been reported in other breeds. A deletion was identified in *FABP4* gene rs134413439 in all samples. How the bovine reference is based in the Hereford breed, this deletion is a variation between Nellore and Hereford. The association analysis showed that the none of the SNP and haplotype had a significant effect ($p < 0.05$) for OPP in Nellore heifers. This study showed that the genes *PPP3CA* and *FABP4* are polymorphic in Nellore cattle. Were not detected polymorphisms in the region of exons 2 and 3 of the gene *PPP3CA* and also no SNP was detected in exon 1 of the gene *FABP4*. It was verified no influence of these genes in the trait occurrence of early pregnancy.

Keywords: lipid metabolism, puberty, fertility traits, *Bos taurus indicus*

2.1. Introdução

A redução da idade à primeira concepção e ao primeiro parto, promove melhoria na eficiência reprodutiva e tem grande impacto sobre a eficiência do sistema de produção como um todo (BERETTA; LOBATO; MIELITZ NETTO, 2001; FORTES et al., 2010; TEIXEIRA; ALBUQUERQUE; FRIES, 2002). Além disso, influenciam o progresso genético pois, diminuem o intervalo de gerações e permitem aumento da intensidade de seleção (VAN MELIS et al., 2010). A ocorrência de prenhez (OPP) é considerada um bom indicador da idade à puberdade e pode ser utilizada como critério de seleção. Porém por ser expressa tardiamente e apenas nas fêmeas, por dificuldades de mensuração e ainda falhas de registros, esta característica têm sido pouco utilizada em programas de melhoramento genético (MACNEIL et al., 2006), não somente no Brasil como no mundo.

Com o objetivo de identificar associação entre polimorfismos conhecidos de genes relacionados com o tecido adiposo e a OPP em novilhas da raça Nelore, Dias et al. (2015) utilizaram dados de 1689 novilhas precoces e não precoces genotipadas com o *High-Density Bovine SNP BeadChip* (Illumina, Inc). Cinquenta e sete genes candidatos e 443 SNP foram analisados. Os autores encontraram efeito significativo sobre a OPP de haplótipos localizados nos genes *PPP3CA* e *FABP4*, concluindo que há influencia destes sobre a a ocorrência de prenhez precoce.

A influência do tecido adiposo sobre a OPP, entre muitos fatores, deve-se ao fato de que este tecido é um dos principais locais do metabolismo dos esteróides sexuais (KERSHAW; FLIER, 2004). Varanasi et al. (1994) relataram que o gene *PPP3CA*, proteína fosfatase 3 - subunidade catalítica alfa isoenzima, foi diferencialmente expresso na diferenciação de pré-adipócitos. Além disso, MARTIN et al. (2007) observaram que o bloqueio deste gene ocasionou infertilidade em ratos machos.

O gene *FABP4* é um dos membros da família FABP que são proteínas que se ligam a ácidos graxos. Este gene está localizado no cromossomo 14 bovino, sendo que há relatos de associação deste gene com variações na porcentagem de colesterol (HOASHI et al., 2008). Michal et al. (2006), demonstraram que o *FABP4* desempenha um papel crítico na absorção,

transporte e metabolismo de ácidos graxos. Os ácidos graxos livres, transportados pelo *FABP4*, podem ser ativados e oxidados para produzir o acetil-CoA e NADH. Parte deste acetil-CoA é utilizado para a biossíntese do colesterol que, por sua vez, é precursor dos hormônios esteroides (BERG; TYMOCZKO; STRYER, 2002). Os esteroides são responsáveis pela diferenciação, crescimento e fisiologia dos órgãos reprodutivos (HAVLÍKOVÁ et al., 2006; PEARCE; JORDAN, 2004).

O presente estudo teve como objetivo verificar a existência de polimorfismos em regiões de exon dos genes *PPP3CA* e *FABP4* e, uma vez identificados estes SNP, avaliar sua possível influência na característica de ocorrência de prenhez precoce de novilhas (OPP).

2.2. Material e Métodos

2.2.1. Animais

Os dados e amostras utilizados no presente trabalho foram coletados de 2036 novilhas nascidas nas safras de 2007 e 2008, expostas a touros nas estações de monta antecipadas nos anos de 2009 e 2010, em propriedades da Agropecuária Jacarezinho Ltda, localizadas nos municípios de Cotegipe e Wanderley, estado da Bahia. Estas amostras foram originalmente coletadas para serem utilizadas em estudos genômicos (auxílio CNPq - Proc. 559631/2009-0 e Processo FAPESP #2009/16118-5). Foram coletados dados de prenhez (positiva ou negativa), além de informações referentes à genealogia, cadastro e manejo destes animais. Foram também coletados folículos pilosos da cauda de onde foi realizada a extração de DNA, no desenvolvimento dos projetos supracitados.

Nestas propriedades são realizadas duas estações de monta. Entre os meses de fevereiro e abril ocorre uma estação de monta antecipada, com duração de 60 dias, a fim de identificar animais que atinjam a puberdade precocemente. Todas as novilhas com idade de 16 a 18 meses são expostas a reprodução, independente do peso e da condição corporal. O sistema de acasalamento para as novilhas é de monta controlada. Aproximadamente 60 dias após o final desta estação de monta antecipada é realizada palpação retal nas novilhas para a avaliação de prenhez. Novilhas que não emprenharam na

estação de monta antecipada são utilizadas na estação de monta de vaca do ano seguinte, quando estão com idades próximas aos dois anos de idade.

Durante a fase de preparação dos dados foram eliminadas as informações com inconsistência decorrentes de erros de registro. Os GC foram montados visando abranger todos os efeitos comuns aos animais e procurando manter a conectabilidade entre os GC. Para OPP os GC foram definidos por fazenda, ano e estação de nascimentos de novilhas. Os GC com poucos animais e também aqueles em que não houve variação na característica (as fêmeas ficaram todas prenhes ou todas vazias) foram excluídos. Após as análises de consistência, restaram 1689 novilhas, sendo que a taxa de prenhez precoce aos 16 meses das 1689 novilhas foi de 30,52%. Destas, 380 novilhas foram selecionadas para o sequenciamento das regiões dos genes candidatos. A seleção foi com base na taxa de prenhez do GC considerando um mínimo de 30%. Assim, as 380 fêmeas selecionadas estavam em 19 GC, sendo que a taxa de prenhez precoce das 380 novilhas foi de 40,26%.

2.2.2. Extração de DNA

A extração de DNA foi feita, como parte do desenvolvimento do projeto (CNPq - Proc. 559631/2009-0), com base na metodologia Fenol- Clorofórmio- Álcool Isoamílico. Esse método consiste em submeter as amostras com os folículos pilosos (aproximadamente 40 folículos/amostra), inicialmente a incubações com TE-Tween e Proteinase-K, seguidas da extração efetiva do DNA com PCL (Fenol-Clorofórmio-Alcool Isoamílico / 25:24:1), precipitação com acetato de Sódio (0,3M) e etanol absoluto e ressuspensão do DNA em tampão TE (10 mM Tris HCl pH = 7,6 e 1mM EDTA pH = 8,0) na proporção de 10:1, sendo posteriormente armazenadas a 4 °C, como descrito em Dias et al. (2015).

2.2.3. Genes Candidatos

Baseado na posição dos haplótipos descritos por Dias et al.(2015), foram delimitadas regiões de exons que flanqueiam ou estão entre os haplótipos que tiveram efeitos significativos ao nível de 5%. Assim, foram

utilizados os quatro exons, do gene *FABP4*; e os exons 2 e 3 do gene *PPP3CA*. A Figura 1, adaptada do Ensembl (CUNNINGHAM et al., 2015), apresenta os haplótipos observados por Dias et al. (2015) e os exons selecionados para o sequenciamento.

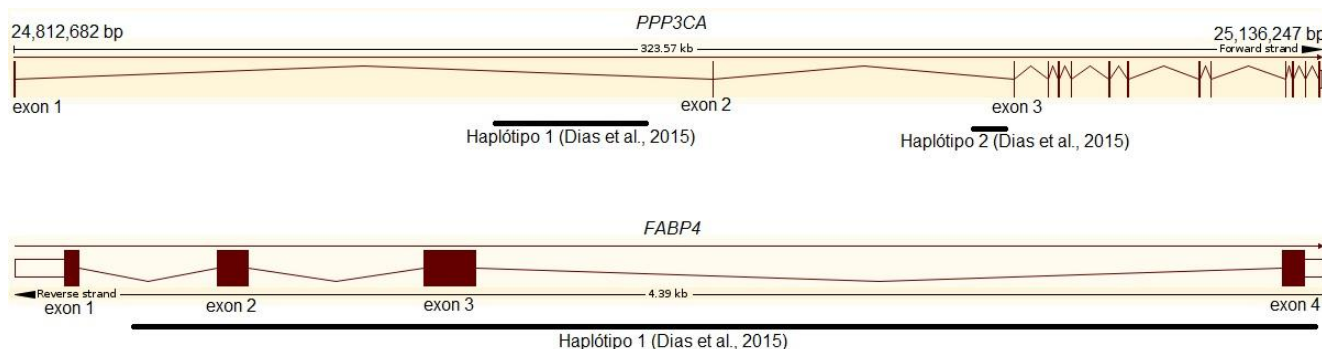


Figura 1. Imagem, adaptada do Ensembl, ilustrando as posições dos haplótipos que tiveram efeito significativo (Dias et al., 2015) e os exons selecionados para sequenciamento

2.2.4. Customização de Painel, Sequenciamento MiSeq e Detecção de Variante

Foi realizada a customização do painel TruSeq Custom Amplicon® (Illumina, Inc) no software DesignStudio (disponível em: <http://designstudio.illumina.com/>) com base na posição dos exons. Para total cobertura da região de exon, parte dos introns e região intergênica foi utilizada. O tamanho do amplicon utilizado para a customização foi de 425bp, garantindo cobertura de 100% da região delimitada de cada exon (Tabela 2).

Tabela 2. Região delimitada por cada amplicon, que contém região de exon e intron, nos genes candidatos.

Gene	Chr	Localização		Comprimento (bp)
		Início	Final	
PPP3CA	6	24,812,677	24,812,991	314
PPP3CA	6	24,985,650	24,985,860	211
PPP3CA	6	25,060,198	25,060,332	135
FABP4	14	46,833,660	46,833,886	227
FABP4	14	46,834,340	46,834,451	112
FABP4	14	46,835,034	46,835,216	183
FABP4	14	46,837,916	46,838,058	143

Um total de 250 ng de DNA foram utilizados para gerar produtos de amplificação que cobrem as regiões. Adaptadores indexados foram adicionados às extremidades dos amplicons de DNA utilizando PCR de ciclo limitado. As bibliotecas de DNA obtidas foram validadas usando Agilent Bioanalyzer 2100 (Agilent Technologies, EUA). Subsequentemente, as bibliotecas de DNA foram sequenciadas pelo sistema MiSeq® usando um Kit de MiSeq Reagente V2 (500 ciclos) (Illumina, Inc.) de acordo com as instruções do fabricante. O sequenciamento foi efetuado utilizando uma configuração de 2 × 250 paired-end, e a análise da imagem e de chamada de base (*base calling*) foram realizadas utilizando o software no instrumento MiSeq.

As sequências com formato FASTQ foram alinhadas na referência bovina UMD3.1. através do BaseSpace® (disponível em: <https://basespace.illumina.com/>) (Illumina, Inc.) usando a ferramenta TruSeq Amplicon. A detecção de variante foi realizada através da mesma ferramenta, utilizando a opção *somatic*. Foram consideradas apenas variantes com frequência do alelo de menor frequência (MAF) acima de 5% e a anotação foi realizada com base no Ensembl. Foram consideradas somente as variantes que possuíam um *Phred-score* ≥ 40.

2.2.5. Frequência Alélica e Genotípica e Equilíbrio de Hardy-Weinberg

As estatísticas descritivas foram realizadas por meio do software Plink (PURCELL et al., 2007). As frequências alélicas e genotípicas dos SNP foram calculadas por simples contagem de alelo (FALCONER; MACKAY, 1996). Possíveis desvios das frequências alélicas e genotípicas, observadas em relação ao esperado, foram testadas através do teste de chi-quadrado, com o objetivo de verificar se o locus em questão estava em equilíbrio de Hardy-Weinberg na amostra (FALCONER; MACKAY, 1996).

2.2.6. Análise de Desequilíbrio de Ligação

A estimação do desequilíbrio de ligação (LD) entre os SNP foi realizada com base na estatística r^2 (HILL; ROBERTSON, 1968) por meio do programa

Plink (PURCELL et al., 2007). Considerando dois *loci* com dois alelos para cada locus (A1/A2 e B1/B2) foi utilizada a fórmula (HILL; ROBERTSON, 1968):

$$r^2 = \frac{D^2}{[f(A1)*f(A2)*f(B1)*f(B2)]}$$

onde:

fA1, fA2, fB1 e fB2 são as frequências de alelos A1, A2, B1 e B2, respectivamente, e

$$D^2 = f(A1_B1)*f(A2_B2) - f(A1_B2)*f(A2_B1)$$

fA1_B1, fA2_B2, fA1_B2 e fA2_B1 são as frequências de haplótipos A1_B1, A2_B2, A1_B2 e A2_B1 da população, respectivamente.

(HILL, 1981)

2.2.7. Controle de Qualidade dos Genótipos

O controle de qualidade dos genótipos para as 2035 amostras genotipadas com o painel de 777.000 *SNP* do *High Density Bovine SNP BeadChip* (Illumina, Inc.) e para as 380 amostras sequenciadas a partir da tecnologia MiSeq® (Illumina, Inc.), foi realizado usando o software Plink v1.07 (PURCELL et al., 2007). Foram descartados *SNP* com taxa de genotipagem (*call rate*) <95%, MAF <0,05 e P-valor para o teste de equilíbrio de Hardy-Weinberg menor que 10^{-6} . Foram considerados os *SNP* que estavam na região de 25.000.000pb até 26.000.000pb no cromossomo 6 e na região de 46.000.000pb até 47.000.000pb no cromossomo 14, sendo que os genes candidatos estão localizados dentro destas regiões. Para o controle de qualidade das amostras foram descartados animais com *call rate* < 90%.

Dentre as 2035 fêmeas utilizadas para a imputação, 29 foram excluídas pelo controle de qualidade das amostras, isto é, apresentaram *call rate* < 90%. Dos cromossomos dos dois genes estudados foram selecionados 609 *SNP* provenientes do painel *High Density Bovine SNP BeadChip* (Illumina, Inc.). Destes foram excluídos 265 pois apresentaram MAF<0,05; 14 tiveram p-valor para o EHW< 10^{-6} e 19 tiveram *call-rate*<0,95. Assim, para a imputação dos genótipos foram utilizadas informações de 311 *SNP* sendo que 194 *SNP* estavam no cromossomo 6 e 117 *SNP* estavam no cromossomo 14. Destes, 30

estavam na região do gene *PPP3CA* e 2 na região do gene *FABP4*. Os SNP do painel localizados dentro da região dos genes também foram utilizados na análise estatística.

2.2.8. Imputação de Genótipos e Reconstrução do Haplótipo

A imputação dos genótipos polimórficos encontrados a partir do sequenciamento MiSeq® (Illumina, Inc.) para as amostras que não foram sequenciadas, mas foram genotipadas com o *High Density Bovine SNP BeadChip* (Illumina, Inc.) foi realizada por meio do software Fimpute (SARGOLZAEI; CHESNAIS; SCHENKEL, 2014). Através do software FastPhase (SCHEET; STEPHENS, 2006) foi detectada e realizada a fase de ligação e a reconstrução dos haplótipos formado por quatro SNP, sendo rs110370032 e rs110652478 localizados em exon e rs110266999 e rs109077068 localizados em 3'UTR no gene *FABP4*.

2.2.9. Análise Estatística

A característica ocorrência prenhez precoce aos 16 meses (OPP) foi definida com base na concepção e parição da novilha, desde que a mesma tenha entrado na estação de monta em torno dos 16 meses de idade. Às novilhas que pariram com menos de 31 meses foram codificadas com o valor 1 (sucesso) e, àquelas que falharam com o valor 0 (fracasso).

As análises foram realizadas utilizando modelos lineares generalizados através do procedimento PROC GLIMMIX do pacote estatístico do SAS 9.3 (SAS Institute, Inc., 2000), assumindo uma distribuição binomial para os dados com uma função de conexão logística. O teste de verossimilhança foi utilizado no estudo do efeito dos marcadores sobre a característica OPP, considerando valores de $P < 0,05$ como sendo significativos. O efeito dos genótipos identificados e imputados para as 1660 novilhas, para a característica OPP foi estimado usando o seguinte modelo estatístico:

$$Y_{ijk} = \mu + GC_i + G_j + t_k + e_{ijk}$$

Onde Y_{ijk} = fenótipos para a característica OPP para o animal; μ = média da característica na população; GC_i = efeito fixo associado ao iésimo grupo de

contemporâneos; G_j = efeito fixo associado ao j-ésimo genótipo de cada polimorfismo; t_k = efeito fixo do k-ésimo grupo de reprodutores múltiplos, e_{ijkl} = erro aleatório.

A fim de garantir a manutenção da taxa de erro dos testes estatísticos no experimento, foi aplicado o método de *Bonferroni* para comparações múltiplas.

2.3. Resultados e Discussão

No total, quinze polimorfismos, sendo 2 no gene *PPP3CA* e 13 no gene *FABP4*, foram identificados nas regiões amplificadas em 380 novilhas da raça Nelore (Tabela 3). Todos estes polimorfismos tiveram um *Phred-score* maior que 40, o que significa uma probabilidade de erro de identificação do nucleotídeo entre 0,001 e 0,01 (ILLUMINA, 2011). Apenas 4 SNP estavam em região de exon e todos se encontram no gene *FABP4*, sendo que apenas um altera o aminoácido. Este SNP, rs110757796, provoca a troca de uma Fenilalanina (Fen) por uma Leucina (Leu). Ambos os aminoácidos são apolares e assim, a estrutura terciária da proteína tem menor possibilidade de sofrer alteração, desse modo esta é uma mutação conservativa. Porém, a mudança do aminoácido pode alterar a configuração da proteína, e afetar a sua função biológica. Apesar das mutações não sinônimas causarem maiores mudanças nas proteínas, também as mutações sinônimos podem afetar a taxa de transcrição ou o volume de transcritos (FLINT; WOOLLIAMS, 2008) e assim, interferir na expressão fenotípica.

Os outros SNP estavam localizados na região não codificante (3'UTR) do gene *FABP4*; em introns dos genes *PPP3CA* e *FABP4* e em região intergênica como apresentado na Tabela 3. O SNP rs109077068 faz parte do painel do *High Density Bovine SNP BeadChip* (Illumina, Inc.) e está em bloco de haplótipo descrito por Dias et al., (2015). Os SNP localizados em região 3'UTR e em intron podem estar em local de ligação ou produção de miRNA, respectivamente. Caso estejam em região de ligação podem afetar a transcrição do gene, e caso estejam em região de produção, podem afetar a transcrição de outros genes (JIN; LEE, 2013). Os SNP localizados em região intergênica a montante do gene (*upstream*) podem estar em um sítio de ligação dos fatores de transcrição (VINSKY et al., 2013). Os SNP em região 3'UTR

também podem influenciar a estabilidade do RNA mensageiro (mRNA) (MATOULKOVA et al., 2012).

Uma deleção identificada em região de intron, rs134413439, estava presente em todos os animais, e portanto pode estar fixada na raça Nelore. O genoma de referência bovino, foi construído a partir do sequenciamento de um animal da raça Hereford (ELSIK; TELLAM; WORLEY, 2009) e assim, esta deleção pode ser uma diferença entre as duas raças, Hereford e Nelore. Não há relatos na literatura de associação desta deleção com nenhuma característica.

Tabela 3. Polimorfismos identificados nas regiões dos genes candidatos amplificados através da tecnologia MiSeq® (Illumina, Inc.)

Gene	Chr	SNP_ID	Posição	Alteração	Região	Consequência	MAF
PPP3CA	6	rs110138405	24985900	C/G	intron 2		0.09079
PPP3CA	6	rs452520974	25060456	T/C	intron 3		0.3066
FABP4	14	rs109593774	46833649	G/C	a jusante		0.1197
FABP4	14	rs110266999	46833693	T/C	3'UTR		0.1263
FABP4	14	rs109077068*	46833778	G/A	3'UTR		0.1237
FABP4	14	rs449843560	46833805	C/T	3'UTR		0.01316
FABP4	14	rs110370032	46833834	T/C	exon 4	Sinônimo	0.1316
FABP4	14	rs110383592	46834345	C/G	exon 3	Sinônimo	0.03553
FABP4	14	rs110652478	46834365	C/T	exon 3	Sinônimo	0.001316
FABP4	14	rs109014985	46834523	T/A	intron 2		0.1368
FABP4	14	rs134173517	46834560	T/G	intron 2		0.1368
FABP4	14	rs134413439	46835006	T/-	intron 2		-
FABP4	14	rs110757796	46835065	T/C	exon 2	Não-sinônimo	0.07237
FABP4	14	rs109316505	46835220	G/C	intron 1		0.006579
FABP4	14	rs109682576	46838280	T/C	a montante		0.05789

*SNP também genotipado no *High Density Bovine SNP BeadChip* (Illumina, Inc.)

Quatro SNP, rs449843560, rs110383592, rs110652478, rs109316505 foram excluídos depois do controle de qualidade pois, apresentaram $MAF < 0,05$ (Tabela 3). Pode-se observar também na Tabela 3, que apenas para o SNP rs452520974 a frequência do alelo menos frequente (MAF) é maior que 0,10, sendo o único que possui os dois alelos com frequências intermediárias. Após a imputação o único SNP excluído pelo controle de qualidade foi o rs109682576 pois teve $MAF = 0.04127764$.

As frequências alélicas e genotípicas e p-valor referente ao teste de equilíbrio de Hardy-Weinberg para os 380 animais sequenciados, (Tabela 4), mostraram que apenas o SNP rs449843560, não se encontravam em equilíbrio, de acordo com a lei de Hardy-Weinberg ($P < 0,000001$) e foi excluído das análises posteriores. O mesmo SNP já havia sido removido por apresentar MAF inferior a 0.05. Dos nove SNP restantes, apenas dois estavam presentes no gene *FABP4* estavam em região de exon.

Tabela 4: Frequência Alélica e Genotípica e Equilíbrio de Hardy-Weinberg para os SNP identificados na região dos genes candidatos

CHR	Polimorfismo	Frequência Alélica		Frequência Genotípica			P-valor (Equilíbrio de Hardy-Weinberg)
		G	C	GG	GC	CC	
6	rs110138405	0.008243	0.000068	0.00526	0.17105	0.82368	0.75490
14	rs109593774	0.014328	0.000205	0.00526	0.22895	0.76579	0.13930
14	rs110383592	0.001262	0.000002	0.00789	0.05526	0.93684	0.00828
14	rs109316505	0.000000	0.000043	0.98684	0.01316	0.00000	1
		T	C	TT	TC	CC	
6	rs452520974	0.094004	0.008837	0.07632	0.46053	0.46316	0.11780
14	rs110266999	0.015952	0.000254	0.00263	0.24737	0.75000	0.01752
14	rs449843560	0.000173	0.000000	0.01053	0.00526	0.98421	0.000000014
14	rs110370032	0.017319	0.000300	0.00526	0.25263	0.74211	0.04218
14	rs110652478	0.000002	0.000000	0.00000	0.00263	0.99737	1
14	rs110757796	0.005237	0.000027	0.00263	0.13947	0.85789	0.7082
14	rs109682576	0.000011	0.003351	0.01579	0.08421	0.90000	0.0006638
		T	A	TT	TA	AA	
14	rs109014985	0.018714	0.000350	0.00526	0.26316	0.73158	0.027750
		T	G	TT	TG	GG	
14	rs134173517	0.018714	0.000350	0.00263	0.26842	0.72895	0.004301
		G	A	GG	GA	AA	
14	rs109077068	0.015302	0.000234	0.00263	0.24211	0.75526	0.017640

* significativo ao nível de P-valor $< 10^{-6}$

Para o cálculo de LD somente foram considerados os SNP identificados a partir de sequenciamento e que passaram pelo controle de qualidade. Para os dois SNP encontrados no gene *PPP3CA*, rs110138405 e rs452520974, o r^2 foi estimado em 0.0228533. Os valores de r^2 entre os SNP, que passaram pelo controle de qualidade, para o gene *FABP4* estão apresentados na Tabela 5. Os

Tabela 5: Estimativa dos valores r^2 para os SNP encontrado no gene de *FABP4*.

SNP	rs109593774	rs110266999	rs109077068	rs110370032	rs109014985	rs134173517	rs110757796	rs109682576
rs109593774	-	0.856738	0.88079	0.790571	0.683691	0.679272	0.398607	0.000058
rs110266999		-	0.919944	0.848878	0.714244	0.732668	0.422842	0.000004
rs109077068			-	0.873585	0.736361	0.732388	0.413219	0.000004
rs110370032				-	0.676065	0.693504	0.389616	0.000085
rs109014985					-	0.974854	0.349988	0.000268
rs134173517						-	0.378712	0.000275
rs110757796							-	0.000054
rs109682576								-

LD entre os SNP são considerados de moderado a alto quando $r^2 > 0,2$ (BOHMANOVA; SARGOLZAEI; SCHENKEL, 2010), sendo mais frequentemente herdados juntos. Valores de $r^2 < 0,2$ indicam que esses SNP segregam de maneira independente. Os r^2 entre todos os SNP com o SNP rs1096682576 são menores que 0,2, desse modo estes segregam de modo independente. Qanbari et al., (2014) compararam o padrão de LD em amostras sequenciadas (n=43), genotipadas no chip de 50K e 700K (n=1293) provenientes de bovinos da raça Simental. Os autores observaram um decaimento maior no valor de r^2 em função da distância nas amostras sequenciadas do que nas amostras genotipadas. Segundo os mesmos autores, o fato de SNP estarem próximos não implica estarem ligados, podendo apresentar segregação independente. Pode se observar o decaimento do r^2 devido a distância dos SNP com o SNP rs1096682576.

A análise de associação revelou, para o haplótipo no gene *FABP4*, após a correção pelo método de *Bonferroni*, não teve efeito significativo ao nível de $p < 0,05$ para a característica de OPP em novilhas da raça Nelore.

Na análise de associação foram utilizados todos os SNP identificados, que passaram pelos controles de qualidade, considerando um de cada vez. A análise estatística, após a correção pelo método de *Bonferroni*, partir desta análise nenhum SNP teve efeito significativo ao nível de $p < 0,05$ sobre a característica OPP (Tabela 6).

Tabela 6. P-valor dos SNP detectados com efeito significativo para ocorrência de prenhez precoce (OPP) m novilhas da raça Nelore.

Gene	SNP	P-valor
PPP3CA	rs110138405	0.4459
PPP3CA	rs452520974	0.0693
FABP4	rs109593774	0.1976
FABP4	rs110266999	0.1412
FABP4	rs109077068	0.1152
FABP4	rs110370032	0.3193
FABP4	rs109014985	0.0472
FABP4	rs134173517	0.0515
FABP4	rs110757796	0.1723

Os SNP, rs110266999, rs109077068, rs110370032, rs110757796, também estão associados com o consumo alimentar residual (COHEN-ZINDER et al., 2016). Blecha et al., (2015) associou os SNP, rs109593774, rs110266999, rs109077068, rs110370032, rs110383592, rs109014985, rs134173517, com espessura de gordura e marmoreio. Apesar dos SNP do gene *PPP3CA* já serem polimorfismos presentes no banco de dados (SNPdatabase), não foram encontrados relatos na literatura sobre a associação destes com qualquer característica. Esta foi a primeira vez que houve a detecção destes polimorfismos na raça Nelore e que houve a tentativa de associar estes SNP com a ocorrência de prenhez precoce de novilhas Nelore.

2.4. Conclusões

Este estudo mostrou que os genes *PPP3CA* e *FABP4* são polimórficos na raça Nelore. Não foram encontrados polimorfismos na região dos exons 1, 2 e 3 do gene *PPP3CA* e também nenhum SNP foi detectado no exon 1 do gene *FABP4*. Não foi verificada a influência destes polimorfismos na característica ocorrência de prenhez precoce.

2.5. Referências Bibliográficas

BERETTA, V.; LOBATO, J. F. P.; MIELITZ NETTO, C. G. A. Produtividade e eficiência biológica de sistemas pecuários de cria diferindo na idade das novilhas ao primeiro parto e na taxa de natalidade do rebanho no Rio Grande de Sul. **Revista Brasileira de Zootecnia**, v. 30, n. 4, p. 1278–1286, 2001.

BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. The Biosynthesis of Membrane Lipids and Steroids. In: FREEMAN, W. H. (Ed.). **Biochemistry**. 5. ed. New York, USA: [s.n.].

BLECHA, I. M. Z. et al. Identification and evaluation of polymorphisms in *FABP3* and *FABP4* in beef cattle. **Genetics and Molecular Research**, v. 14, n. 4, p. 16353–16363, 2015.

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. S. Characteristics of linkage disequilibrium in North American Holsteins. **BMC genomics**, v. 11, n. 1, p. 421, 2010.

COHEN-ZINDER, M. et al. *FABP4* is a leading candidate gene associated with residual feed intake in growing Holstein calves. **Physiological Genomics**, v. 48, n. 5, p. 367–376, 2016.

CUNNINGHAM, F. et al. Ensembl 2015. **Nucleic Acids Research**, v. 43, n. D1, p.

D662–D669, 2015.

DIAS, M. M. et al. Study of lipid metabolism-related genes as candidate genes of sexual precocity in nellore cattle. **Genetics and Molecular Research**, 2015.

ELSIK, C. G.; TELLAM, R. L.; WORLEY, K. C. The Genome Sequence of Taurine Cattle: A window to ruminant Biology and Evolution. **Science**, v. 324, p. 522–527, 2009.

FALCONER, D.; MACKAY, T. **Introduction to quantitative genetics**. 4. ed. Edinburgh: Longman Group Limited, 1996.

FLINT, A. P. F.; WOOLLIAMS, J. A. Precision animal breeding. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 363, p. 573–590, 2008.

FORTES, M. R. S. et al. Association weight matrix for the genetic dissection of puberty in beef cattle. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 31, p. 13642–13647, 2010.

HAVLÍKOVÁ, L. et al. Determination of estradiol and its degradation products by liquid chromatography. **Journal of Chromatography A**, v. 1119, n. 1-2, p. 216–223, 2006.

HILL, W. G. Estimation of effective population size from data on linkage disequilibrium. **Genet. Res.**, v. 38, p. 209–216, 1981.

HILL, W. G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik**, v. 38, n. 6, p. 226–31, 1968.

HOASHI, S. et al. Association between fatty acid compositions and genotypes of FABP4 and LXR-alpha in Japanese black cattle. **BMC genetics**, v. 9, n. 1, p. 84, 2008.

ILLUMINA. **Quality Scores for Next-Generation Sequencing** [Http://Res.Illumina.Com/Documents/Products/Technotes/Technote_Q-Scores.Pdf](http://res.illumina.com/Documents/Products/Technotes/Technote_Q-Scores.Pdf), 2011.

JIN, Y.; LEE, C. Single Nucleotide Polymorphisms Associated with MicroRNA Regulation. **Biomolecules**, v. 3, n. 2, p. 287–302, 2013.

KERSHAW, E. E.; FLIER, J. S. Adipose Tissue as an Endocrine Organ. **The Journal of Clinical Endocrinology & Metabolism**, v. 89, n. 6, p. 2548–2556, 2004.

MACNEIL, M. D. et al. Genetic partitioning of variation in ovulatory follicle size and probability of pregnancy in beef cattle. **Journal of Animal Science**, v. 84, n. 7, p. 1646–1650, 2006.

MATOULKOVA, E. et al. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. **RNA biology**, v. 9, n. 5, p. 563–76, 2012.

MICHAL, J. J. et al. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. **Animal Genetics**, v. 37, n. 4, p. 400–402, 2006.

PEARCE, S. T.; JORDAN, V. C. The biological role of estrogen receptors alfa and beta in cancer. **Critical Reviews in Oncology/Hematology**, v. 50, n. 1, p. 3–22, 2004.

PURCELL, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. **American Journal of Human Genetics**, v. 81, n. 3, p. 559–575, 2007.

QANBARI, S. et al. Classic Selective Sweeps Revealed by Massive Sequencing in Cattle. **PLoS Genetics**, v. 10, n. 2, 2014.

SARGOLZAEI, M.; CHESNAIS, J. P.; SCHENKEL, F. S. A new approach for efficient genotype imputation using information from relatives. **BMC genomics**, v. 15, n. 1, p. 478, 2014.

SCHEET, P.; STEPHENS, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. **American journal of human genetics**, v. 78, n. 4, p. 629–44, 2006.

TEIXEIRA, R. A.; ALBUQUERQUE, L. G.; FRIES, L. E. Bio-economic comparisons among three ages at first breeding in Nelore heifers. **ARS Veterinaria**, v. 18, n. 3, p. 197–203, 2002.

VAN MELIS, M. H. et al. Additive genetic relationships between scrotal circumference, heifer pregnancy, and stayability in Nellore cattle. **Journal of Animal Science**, v. 88, n. 12, p. 3809–3813, 2010.

VINSKY, M. et al. Communication: Association analyses of a single nucleotide polymorphism in the promoter of OLR1 with growth, feed efficiency, fat deposition, and carcass merit traits. **Canadian Journal of Animal Science**, v. 93, p. 193–197, 2013.

CAPÍTULO 3 - Identificação de SNP em genes candidatos associados com Puberdade em fêmeas bovinas

RESUMO: Características reprodutivas, tais como o início da puberdade, são economicamente importantes nos sistemas de produção bovino e é interessante incluir tais características em programas de melhoramento genético animal. O presente estudo objetivou: 1) identificar SNP a partir de dados de sequenciamento de RNA em 62 genes candidatos envolvidos na puberdade em bovinos, 2) comparar a existência dos SNP identificados em outras quatro raças, Angus, Brahman, Nelore e Holandês, e 3) identificar regiões de splicing e genes que sobrepõem os genes candidatos. Sequências de RNA-Seq foram alinhadas ao genoma de referência bovino (UMD 3.1.1) e analisadas utilizando o software CLC Genomics Workbench v.8.0.1, em 5 raças separadamente: Brangus, Brahman, Nelore, Angus e Holandês. Para diminuir o erro de identificação de SNP, duas abordagens de análise com os dados provenientes de Brangus foram utilizadas. Primeiro, todas as amostras foram analisadas em conjunto (*pool*) e segundo, a descoberta de SNP foi realizada individualmente em cada amostra. Apenas SNP que foram descobertos em ambas as abordagens foram incluídos na análise de comparação. Foram considerados 1.157 SNP identificados em Brangus. Estes SNP foram então comparados com os identificados de polimorfismos nas amostras analisadas em conjunto em cada uma das outras raças. Detectamos 172 SNP que são coincidentes nas cinco raças. Esses SNP estão localizados em 13 genes. Usando recursos do ENSEMBL (VEP) verificamos a localização e o efeito das variações, 8% destes SNP estavam no exon, destes 73% são mutações sinônimas e 24% são não-sinônimos, 78% dos SNP estavam em intron e 15% estavam perto de genes candidatos. Como foram utilizadas sequências de RNA, o SNPs precisariam estar em exon ou em região UTR, por isso, fizemos uma análise para descoberta de transcritos, ou detecção de splicing alternativo, e genes que sobrepõem genes candidatos. Os resultados desta análise demonstraram 160 novos transcritos em 42 genes candidatos e cinco novos genes que sobrepõem cinco genes candidatos. Não houve alteração na localização dos SNPs mostrada nas análises realizadas pela ferramenta VEP. Assim, são necessários mais estudos a fim de verificar a localização dos SNPs detectados a partir de sequências de RNA. As sequências de RNA podem ser utilizadas para a identificação de polimorfismos. A comparação de detecção de polimorfismos em diferentes raças pode ajudar a selecionar SNP para a customização de painel de baixa densidade. Devido a quantidade de transcritos e genes descobertos pode-se concluir que mais análises são necessárias a fim de enriquecer o genoma referência bovino.

Palavras-chave: *Bos taurus taurus*, *Bos taurus indicus*, genes sobrepostos, polimorfismo, splicing alternativo.

CHAPTER 3: SNP Identification in Candidate Genes Associated with Puberty in cattle

Abstract:

Fertility traits, such as the onset of puberty, are economically important in cattle production systems and it is of interest to include such traits in genetic selection programs. The aim of the current study was to identify SNP from RNA-Seq data in 62 candidate gene involved in puberty in cattle. RNA-Seq reads were assembled to the bovine reference genome (UMD 3.1.1) and analyzed using CLC Genomics Workbench v.8.0.1 in 5 breeds, Brangus, Brahman, Nellore, Angus, and Holstein. To enhance the accuracy of SNP detection, we used two analyze approaches with the Brangus (3/8 indicine x 5/8 taurine) data. First, all the samples were pooled and second, SNP discovery was performed individually in each sample. Only SNP that were discovered with both approaches were included in the comparison analysis, were identified 1,157 SNPs in Brangus samples. These SNP were then compared with those identified in the pooled samples of the other breeds. We detected 172 SNP that were coincident in the five breeds. The shared SNP were within 13 genes. Using Ensembl Resources (VEP) we checked the effect of the variation, 8% of these SNP were in an exon, that include 73% were synonymous and 24% were missense, and 78% were in intron and 15% were near the gene. How we used RNA-Seq, the SNPs need to be exon or UTR region, so we did a transcript and gene discovery in the region of the candidate genes. The results from this analysis showed 160 novel transcripts in 42 candidates genes and five novel genes that overlapping five candidate genes. There was no change in the location of the SNPs shown in the analysis from VEP tool. Thus, further studies are needed to verify the location of the SNPs detected from RNA sequences. The RNA sequence may be used to identify polymorphisms. Due the amount of discovered genes and transcripts can be concluded that further analysis is required in order to enrich the bovine genome reference.

Keywords: bovine, Gene Discovery, SNP, Transcript Discovery.

1. Introdução

A puberdade é um importante processo fisiológico que influencia as características reprodutivas, como a idade ao primeiro parto. Este processo é definido como o período no qual o animal atinge a maturidade sexual e tornar-se capaz de se reproduzir, sendo este um desenvolvimento gradual de maturação que começa durante o desenvolvimento fetal e prossegue nas fases de pré- e peripuberais, tendo início a partir da interação do eixo hipotalâmico-pituitário-gonadal (HPG) (DORN; BIRO, 2011). Os eventos que levam à ocorrência da puberdade são semelhantes entre as duas sub-espécies de bovinos, *Bos taurus taurus* e *Bos taurus indicus*, mas acontecem tardiamente em novilhas *Bos taurus indicus* (DAY; NOGUEIRA, 2013). A maturação sexual do animal tem influência direta na eficiência reprodutiva, sendo esta considerada economicamente mais relevante e importante do que características de crescimento (DICKERSON, 1970; BRUMATTI et al., 2011).

Desde a publicação do genoma referência bovino (ELSIK; TELLAM; WORLEY, 2009) houve o desenvolvimento e aumento do conhecimento das técnicas omics, acelerando a investigação da genética bovina (DEATLEY, 2012; FORTES et al., 2012; SNELLING et al., 2013; CÁNOVAS et al., 2014c). Além disso, é importante conduzir análises com o objetivo de identificar variações no genoma a fim de conhecer os polimorfismos que influenciam a expressão da característica. A vantagem da detecção de variação no transcriptoma é a possibilidade de que esta variação influencie a transcrição e tradução do RNA mensageiro (RNAm).

O objetivo deste estudo foi a identificação de SNP a partir da análise do transcriptoma de oito tecidos que estão envolvidos no processo fisiológico da puberdade (hipotálamo, glândula pituitária, fígado, útero, endométrio, ovário, adiposo, e músculo *Longissimus dorsi*) de bovinos Brangus em regiões do genoma correspondentes a 62 genes associados com a puberdade em bovinos Brangus e Nelore e verificar se os SNP identificados na raça Brangus podem ser encontrados em outras raças (Angus, Brahman, Holandês, e Nelore). Além disso, foram pesquisados splicing alternativos nos genes candidatos e também se há genes que sobrepõem a região dos genes candidatos.

2. Material e Métodos

2.1. População

Os animais utilizados no presente trabalho fazem parte de um estudo amplo envolvendo a Universidade Estadual Júlio de Mesquita Filho (UNESP/FCAV), Colorado State University (CSU), New Mexico State University (NMSU), University of California (UCDAVIS), Texas A&M University, Instituto Tecnológico de Sonora (ITSON), University of Wyoming, Virginia Tech, University of Colorado Denver, The University of Queensland e Commonwealth Scientific and Industrial Research Organisation (CSIRO), foram utilizados diferentes grupos genéticos para identificar polimorfismos para características reprodutivas em diferentes grupos genéticos envolvendo animais *Bos taurus taurus*, *Bos taurus indicus* e cruzamento entre *Bos taurus taurus* e *Bos taurus indicus*. A população base utilizada neste estudo foi denominada de População de Descoberta, sendo esta constituída por novilhas Brangus. As raças Angus, Brahman, Nelore e Holandês foram utilizadas para comparar mutações entre as raças, sendo essas populações denominadas de População de Comparação. Na Figura 1 estão descritos, resumidamente, as análises feitas em cada população que serão melhor descritas nos itens abaixo.

2.1.1. População de Descoberta

Para a população de descoberta foram utilizados, especificamente, oito novilhas da raça Brangus (3/8 Brahman x 5/8 Angus) escolhidas entre 800 novilhas do rebanho da New Mexico State University, analisadas em estudo de GWAS para características reprodutivas. Estas oito novilhas foram escolhidas a partir de uma análise de pedigree, com o objetivo de representarem a diversidade genética do rebanho (CÁNOVAS et al., 2014c). As novilhas foram aleatoriamente designadas para dois grupos que foram então abatidos nos estados fisiológicos pre-pubertal (PRE; n = 4) e pos-pubertal (POST; n = 4). Oito tecidos foram coletados de cada novilha que incluiu: hipotálamo (HYP), glândula pituitária (PIT), fígado (LIV), útero (UTE), endométrio (END), ovário (OVA), adiposo (gordura), e músculo *Longissimus dorsi* (LDM). Durante o processamento das amostras no laboratório, duas amostras, ambas do END foram perdidas no momento da extração do RNA. A extração de RNA foi realizada em todos os tecidos e as amostras sequenciadas com o

equipamento Illumina HiSeq analyzer (Illumina, San Diego, CA) (CÁNOVAS et al., 2014c).

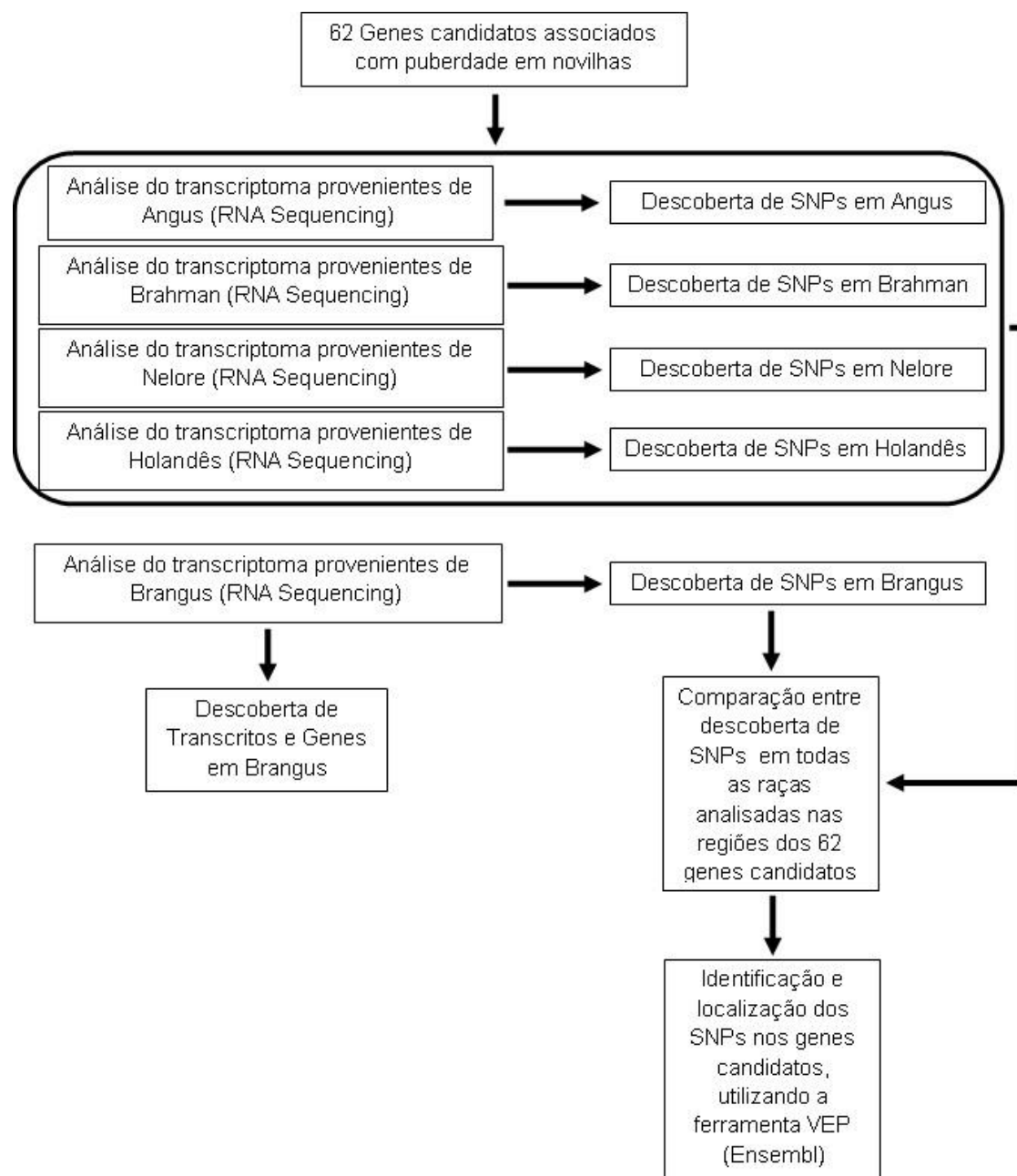


Figura 1. Fluxo para identificação de SNP e Transcritos nos 62 genes candidatos associados à puberdade em bovinos.

2.1.2. População de Comparação

2.1.2.1. Angus

Dez novilhas da raça Angus, criadas em região de altitude elevada no Colorado, EUA, diagnosticadas como saudáveis (n=5) e hipertensas (n=5) para pressão arterial pulmonar. Foram coletados seis tecidos de cada novilha (ventrículo esquerdo, ventrículo direito, artéria pulmonar, aorta, músculo *Longissimus dorsi* e, pulmão). A extração de RNA foi realizada em todos os tecidos e o sequenciamento feito em equipamento Illumina HiSeq analyzer (Illumina, San Diego, CA) (CÁNOVAS et al., 2015).

2.1.2.2. Brahman

Foram utilizadas doze novilhas Brahman, originárias da Austrália, foram abatidos nos estados fisiológicos pré-pubertal (n = 6) e pos-pubertal (n = 6) a fim de estudar. As amostras foram coletadas em seis tecidos diferentes, que incluem: o hipotálamo, hipófise, fígado, útero, ovário esquerdo, ovário direito. A extração de RNA foi realizada em todos os tecidos e o sequenciamento feito em equipamento Illumina HiSeq analyzer (Illumina, San Diego, CA) (dados não publicados).

2.1.2.3. Nelore

Quarenta e oito machos Nelore, provenientes do Brasil. A extração de RNA e estudos de genes diferencialmente expressos foram realizados para estudos de qualidade de carne (Processo FAPESP #2011/21241-0 e #2009/16118-5, dados não publicados). Estes animais foram selecionados de um grupo de 120 animais com diferença de idade ao abate de 90 dias, no máximo, sendo filhos de seis touros diferentes. Apenas o músculo *Longissimus thoracis* foi coletado. Os dados de sequenciamento de RNA foram gerados na HiScanSQ (Illumina, San Diego, CA).

2.1.2.4. Holandês

Doze bezerros recém-nascidos da raça Holandesa, criados em região de alta altitude do Colorado, EUA, diagnosticados como saudáveis (n = 6) ou hipertensiva (n = 6) para pressão arterial pulmonar. As amostras para a análise do transcriptoma foram derivadas de cultura de fibroblastos (CÁNOVAS et al., 2014a, 2014b). A

extração de RNA foi realizada a partir da cultura de tecidos e o sequenciamento feito em equipamento Illumina HiSeq analyzer (Illumina, San Diego, CA).

2.2. Genes Candidatos

Foram selecionados 62 genes candidatos associados com a puberdade e a ocorrência de prenhez precoce, com base em estudos de transcriptoma (CÁNOVAS et al., 2014c), neuropeptidoma (DEATLEY, 2012), e de associação (DIAS et al., 2015). Estes 62 genes incluem: 25 genes diferencialmente expressos em animais PRE e POST e que abrigam SNPs associados a características reprodutivas em estudo de GWAS (Tabela 1) (CÁNOVAS et al., 2014c); 20 genes que expressam fatores de transcrição (TF) altamente evidenciados (Tabela 2) para a expressão da puberdade (CÁNOVAS et al., 2014c); 18 genes de neuropeptídeos relacionados com a reprodução (DeAtley, 2012) (Tabela 3), e 2 genes, *PPP3CA* e *FABP4*, relatados a partir de um estudo associação com a probabilidade de prenhez precoce aos 16 meses (P16) em Nelore (DIAS et al., 2015), sendo a característica P16 considerada um bom indicador para a puberdade. Três dos genes foram detectados em mais de um dos estudos relatados (*ELF5*, *POU4F2* e *PENK*). As informações dos genes quanto à localização genômica, com base na referência bovina UMD 3.1, foram extraídas do Ensembl (CUNNINGHAM et al., 2015) e estão apresentadas nos Apêndices 1, 2, 3 e 4.

Tabela 1. Genes que contêm SNP associados com a fertilidade de novilhas em estudo de GWAS e também genes diferencialmente expressos (DE) ou específico de tecido (TS) em análise de RNA-Seq. (Adaptado de CÁNOVAS et al., 2014c).

Gene	Tecido	SNP	Chr.	Posição (bp)	Distance (bp)	Localização do SNP no gene
<i>DPPA4</i>	HYP	rs41571491	1	54,523,895	5,517	a jusante
<i>TP63</i>	ADIPOSE	rs109034747	1	78,088,393	0	Intron
<i>INHHA</i>	OVA	rs136318374	2	108,231,291	1,460	a jusante
<i>IL22RA1</i>	END	rs29019426	2	129,348,719	2,978	a montante
<i>RHEBL1</i>	PIT	rs1107702	5	30,912,029	7,398	a montante
<i>LYSB</i>	END	rs110181782	5	44,556,197	0	Intron
<i>ADH6</i>	END	rs41612964	6	26,799,276	0	Intron
<i>MEPE</i>	PIT	rs41650773	6	38,286,952	0	Intron
<i>TECRL</i>	LDM	rs109759346	6	81,577,343	0	Intron
<i>LOC777593</i>	END	rs109621209	7	61,683,533	0	Intron
<i>MGC157266</i>	END	rs136828274	8	62,927,382	8,411	a montante
<i>C10H11ORF46</i>	HYP	rs109705635	10	4,882,326	0	Intron
<i>NRXN3</i>	END	rs43651752	10	91,751,584	0	Intron
<i>TSHR</i>	END	ss117964119	10	93,528,030	0	Intron
<i>NEBL</i>	END	ss61500113	13	22,476,667	0	Intron
<i>MOS</i>	OVA	rs110243083	14	24,973,324	2,465	a montante
<i>PENK</i>	PIT	rs134428213	14	25,218,861	0	Exon
<i>ELF5*</i>	UTE	rs133233558	15	65,840,403	0	Intron
<i>POU4F2*</i>	END	rs41838669	17	11,828,433	1,818	a montante
<i>FAM19A4</i>	LIV	rs136177962	22	32,820,574	0	Intron
<i>ITIH1</i>	OVA	rs110723544	22	48,663,639	0	Exon
<i>CPNE5</i>	END	rs109264326	23	10,690,330	0	Intron
<i>MMD2</i>	UTE	rs135797510	25	39,622,260	0	Intron
<i>DKK1</i>	UTE	rs29024937	26	6,860,529	4,762	A jusante
<i>SYCE1</i>	LIV	rs41629412	26	50,517,416	0	Intron

Abreviações incluem: HYP (hipotálamo); PIT (glândula pituitária); UTE (útero); OVA (ovário); FAT (adiposo); LIV (fígado); LDM (músculo *Longissimus dorsi*).* Os genes também estão entre os 20 mais relevantes fatores de transcrição.

Tabela 2: Vinte genes que expressam fatores de transcrição (TF) altamente evidenciados, sendo este os mais relevantes para a expressão da puberdade e seus respectivos tecidos de expressão máxima em novilhas Brangus pré e pós-púberes. (Adaptado de CÁNOVAS et al., 2014c).

Gene	Tecido de máxima expressão	Gene	Tecido de máxima expressão
VAX2	Hipotálamo e/ou Glândula Pituitária	CDX2	Útero e/ou Endométrio
FOXE1		ELF5	
FOXB2		FOXN4	
LHX4		VGLL1	
PITX2		TSG101	
POU4F2		MYOCD	Tecido Adiposo
PROP1		KLF1	
SIX6		GATA1	
NHLH1		E2F3	Longissimus dorsi músculo
DACH2		OVGP1	Ovário

Tabela 3. Genes que expressam neuropeptídeos no hipotálamo e hipófise de novilhas Brangus Pré e Pós Puberdade relatados por DeAtley (2012).

Gene	Quantidade de Isoforma por Peptídeo	Gene	Quantidade de Isoforma por Peptídeo
CART	1	Penk	5
CBLN4	1	POMC	7
CBLN1	4	TAC 1	3
CBLN2	2	SCG5	1
CCK	1	CHGB	7
CHGA	3	SCG2	2
NTS	1	SCG3	1
PCSK1N	9	SST	1
PEPT-1	2	STMN1	1

2.4. Sequências de RNA e Controle de Qualidade

Sequências (*reads*) obtidas a partir do RNA-Seq das cinco raças foram analisadas separadamente com o software CLC Genomics Workbench 8.0.1 (CLC Bio, Aarhus, Dinamarca). Para o alinhamento, foi utilizado o genoma de referência bovino UMD 3.1.78 (ftp://ftp.ensembl.org/pub/release-84/fasta/bos_taurus/dna/). O mesmo software foi utilizado para o controle de qualidade de acordo com os procedimentos descritos por Cánovas et al, (CÁNOVAS et al., 2014d). No Brangus foram realizados dois alinhamentos. O primeiro foi utilizando todas as amostras

como um conjunto (*pool*, todos os tecidos em todos os animais) e o segundo foi realizado cada amostra separadamente. A razão pelo qual duas abordagens foram utilizadas foi para diminuir o erro de identificação de SNP. Para cada uma das outras raças, Angus, Brahman, Holandês, e Nelore, o alinhamento foi realizado para o *pool* de amostras.

2.5. Identificação de SNPs nos genes candidatos

Depois do alinhamento das sequências, a descoberta dos SNPs foi executada no software CLC Genomics Workbench 8.0.1 (CLC Bio, Aarhus, Dinamarca). Para isso, utilizou-se a ferramenta "*Fixed Ploidy Variant Detection*". Esta ferramenta permite a identificação de variantes e o descarte de variantes cuja representação na leitura é devida a erros de sequenciamento ou artefatos de mapeamento. Os filtros de qualidade e de significância foram usados como descrito por Cánovas et al. (2010).

Nas amostras provenientes da raça Brangus duas abordagens foram realizadas, em primeiro lugar, foi realizada a detecção de variantes na análise do *pool* de amostras. Esta análise permite que o alinhamento ocorra com mais acurácia. Na segunda abordagem foi realizada a detecção de variantes em cada amostra de cada animal, considerando apenas alinhamentos com cobertura de cinco vezes. Depois de realizadas as duas abordagens, os resultados foram comparados e apenas os SNPs detectados em ambas as análises foram considerados. Para cada uma das outras raças a análise foi realizada apenas com as amostras em conjunto (*pool*). Em seguida, as descobertas de variantes em cada raça foram comparadas para verificar quais SNPs estavam presentes em todas as raças.

Recursos do Ensembl (CUNNINGHAM et al., 2015) foram utilizados para localizar SNPs reportados no banco de dados (*SNP database*) nas regiões 5'UTR, de exons, e 3'UTR dos genes candidatos (Apêndice - Tabela 5, 6, 7, e 8). Os SNPs encontrados nas análises RNA-Seq foram comparados com aqueles reportados no banco de dados, com o objetivo de verificar, dentro do conjunto da descoberta, quantos eram novos e quantos já eram variantes conhecidas. Foi utilizada a ferramenta "*Variant Effect Predictor*" (VEP) do Ensembl para classificar as variações identificadas.

2.7. Descoberta de Transcritos e Genes

O mapeamento de grande intervalo (*Large Gap Mapping*) foi realizada com o software CLC Genomics Workbench 8.0.1 (CLC Bio, Aarhus, Dinamarca) para mapear sequências de RNA sem a necessidade de anotações de transcritos anteriores. Para esta análise foram utilizadas amostras de RNA-Seq provenientes da raça Brangus.

A ferramenta "*Transcript Discovery*" do mesmo software foi usada para identificar prováveis regiões de genes, locais de *splicing* e exons não anotados. Os parâmetros utilizados permitem comparar os transcritos encontrados com anotações existentes, e assim é possível enriquecer as anotações de genes conhecidos e anotar novos gene. Em ambas as análises os parâmetros *default* do software foram utilizados na análise.

3. Resultados e Discussão

3.1. RNA-Seq

Duas amostras provenientes do hipotálamo da raça Brangus não passaram pelos critérios do controle de qualidade, e foram removidas da análise. Desta forma, 60 das 62 amostras sequenciadas foram utilizados. Em média, 30 milhões de sequências (*reads*) foram obtidas a partir de cada amostra proveniente de Brangus, Angus, Brahman e Holandês. Para a raça Nelore foram obtidos, em média, 24 milhões de sequências. Estas sequências foram montadas e mapeados a partir do genoma bovino UMD 3.1.78.o qual apresenta 24,618 genes anotados.

3.3.2. Identificação de SNP e comparação entre Raças

Recursos do Ensembl (CUNNINGHAM et al., 2015) foram utilizados para consultar SNPs reportados no banco de dados (*SNP database*) dentro dos genes candidatos. No banco de dados, estão anotados 231,770 SNP em todas as regiões dos 62 genes candidatos (5'UTR, exon, intron, 3'UTR) conforme relatado nos Apêndices 5, 6, 7 e 8. Deste total, 1.061 estão em região de 5 'UTR; 218.036 estão em região de intron; 1.0493 estão em região de exon; e 2.180 estão em região de

3'UTR. Do total de SNPs reportados em exon 8.023 são mutação não-sinônimas e 2.470 são mutações sinônimas.

A primeira análise para a identificação de polimorfismos em Brangus, realizada com todas as amostras em conjunto, revelou 9.357 SNPs em 56 dos 62 genes candidatos. Nas regiões dos genes *CDX2*, *FOXB2*, *FOXE1*, *LYSB*, *MOS*, e *NHLH1* não foram encontrados nenhum SNPs.

A segunda análise, realizada com cada amostra separadamente, revelou 2.175 SNPs em 50 genes candidatos. Os genes em que não tiveram SNP identificados nesta segunda análise foram *CDX2*, *DACH2*, *DPPA4*, *FOXB2*, *FOXE1*, *GATA1*, *IL22RA1*, *LYSB*, *MEPE*, *MOS*, *NHLH1* e *RHEBL1*. A média e o erro padrão da quantidade de SNP por amostra foi de $186,11 \pm 11,28$, com mínimo de no número de SNP foi de 64 e o máximo de 424. A diferença em número de SNP entre as amostras é devido ao fato de alguns genes possuírem expressão em um único tecido e também alguns transcritos possuírem especificidade com relação ao tecido. Oitocentos e quarenta SNPs eram específicos da amostra, isto significa que estes SNP apenas existiam em uma amostra. Nove polimorfismos identificados foram observados em todas as amostras, desse modo esses polimorfismos devem ser uma variação entre as raças Brangus e Hereford, uma vez que a identificação foi realizada na raça Brangus e o genoma referência bovino é baseado na raça Hereford.

Apenas os SNP identificados em ambas as abordagens foram mantidos, permanecendo 1.157 SNPs em 46 genes candidatos, desse modo houve diminuição na taxa de erro da detecção dos SNP. Os genes que não tiveram SNP identificados foram *CDX2*, *DACH2*, *DPPA4*, *FOXB2*, *FOXE1*, *GATA1*, *IL22RA1*, *LYSB*, *MEPE*, *MOS*, *NHLH1* e *RHEBL1*. Os genes *ADH 6*, *NTS*, *PCSK1N* e *VAX2* embora tenham tido SNP identificados em ambas as análises, esses SNPs não eram coincidentes, com isso esses genes foram excluídos. Apenas os SNP resultantes da união das duas abordagens foram utilizados nas outras análises, dentre elas a comparação da existência destes SNP nas outras raças.

A ferramenta VEP do site Ensembl (CUNNINGHAM et al., 2015) foi utilizada para classificar as variações identificadas com relação a alteração de aminoácidos (Figura 2). Verificou-se que 699 (60,4%) SNPs eram desconhecidos e 458 (39,6%)

eram variantes já conhecidas. Estes SNPs estão presentes em 48 genes. Os resultados gerados pela análise VEP posicionaram SNPs, dos genes candidatos SCG5 e SCG2, também em regiões de dois genes que não constavam entre os candidatos, *LYSMD2* e *OBSL1*. Um resultado não esperado foi que, tendo como base a anotação do genoma referência, 78% dos 1.157 SNP identificados foram mapeados em regiões de introns. Como os SNP foram detectados a partir de sequências de RNA, esse resultado não seria possível, uma vez que os introns não são transcritos em RNA mensageiro. Bohmanova et al., (2010) demonstraram a partir de um algoritmo baseado no decaimento do LD nos SNP provenientes do painel Illumina BovineSNP50K BeadChip alguns SNP são mapeados erroneamente, sendo a localização dos SNP fundamental para estudos genômicos. Possivelmente, não somente problemas de anotação existem no genoma referência, é provavelmente também que existam exons que não anotados e possíveis splicing alternativos. Devido a este resultado, foi realizada uma análise para a descoberta de transcritos.

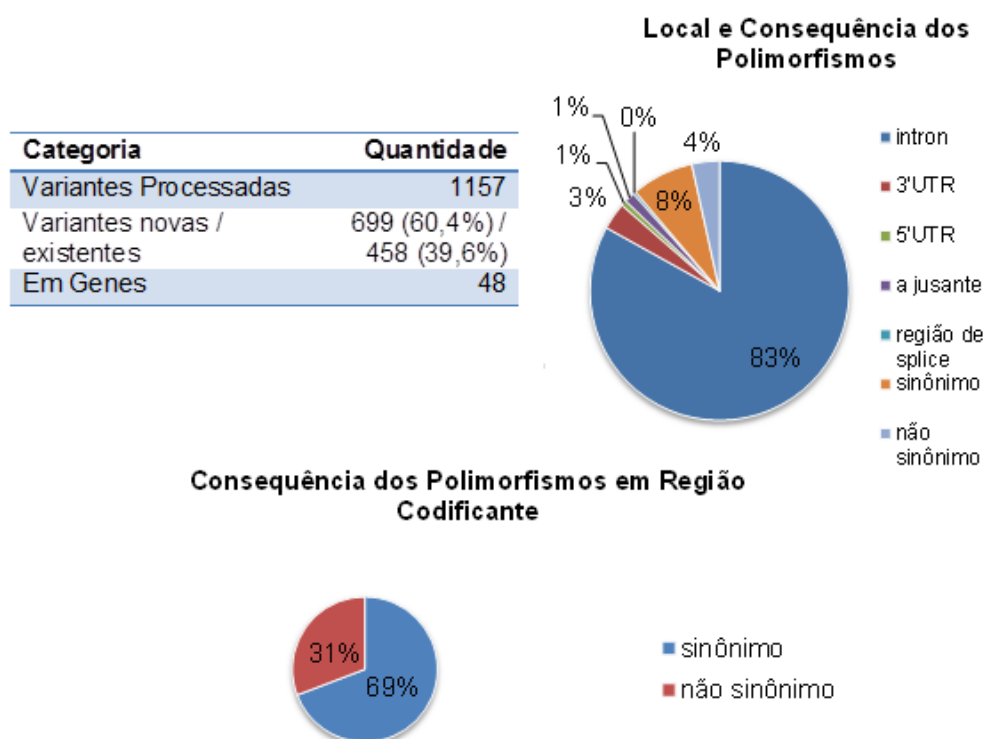


Figura 2. Sumário dos resultados estatísticos provenientes da ferramenta "Variant Effect Predict" (VEP) do Ensembl utilizando os 1.157 SNPs detectados em novilhas da raças Brangus.

A ferramenta VEP demonstrou também que cinco SNP estão na região de splicing (*splice region*), ou seja, estes polimorfismos se encontram nas extremidades, dentro de um a três pares de base do exon ou três a oito pares de bases do intron (CUNNINGHAM et al., 2015), sendo que um dos SNP está em região de (*splice donor*), o que significa que o SNP está em uma das duas últimas bases na extremidade 5' de um intron, influenciando assim a região do splicing (CUNNINGHAM et al., 2015)..

Os SNP que estão localizados na região não codificante (3'UTR) dos gene candidatos pode estar em local de ligação de miRNA, podendo afetar a transcrição do gene (JIN; LEE, 2013). Além disso, SNP em região 3'UTR podem também influenciar a estabilidade do RNA mensageiro (MATOULKOVA et al., 2012) uma vez que esta região influencia na formação da cauda Poli A e a expressão do gene na etapa pós-transcricional (KUERSTEN; GOODWIN, 2003). Já os SNP localizados em região de intron podem estar em local de produção de miRNA, podendo afetar a transcrição de outros genes (JIN; LEE, 2013). Os SNP localizados em região intergênica a montante do gene (*upstream*) podem estar em um sítio de ligação dos fatores de transcrição (VINSKY et al., 2013).

A partir da anotação do genoma referência, a ferramenta VEP do Ensembl foi capaz de alocar 130 SNP em região de exon. Destes, 78 são SNP já reportados no SNPdatabase e 52 são SNP novos. Apenas 40 dos SNP em exon alteram o aminoácido. Na Tabela 4 pode-se ver qual é a mudança de aminoácido e as propriedades bioquímicas do aminoácido alterado. Quando as propriedades bioquímicas do aminoácido são diferentes, pode haver alteração na estrutura terciária da proteína, alterando assim a configuração da proteína, e podendo afetar sua função biológica.

Tabela 4. Mutações não-sinônimas detectadas nos genes candidatos associados com a puberdade na raça Brangus.

(continua)

CHR	Posição	Gene	Mudança de Aminoácido	Mudança nas Propriedades Bioquímica do aa	Codon	Variante Existente
1	80251288	SST	Lys/Glu	polar básico/polar ácido	AAA/GAA	rs17870997
2	112550874	SCG2	Ser/Pro	polar neutro/apolar	TCG/CCG	rs43316973
3	32058726	OVGP1	His/Arg	polar básico/polar básico	CAC/CGC	rs41570791
3	32058831	OVGP1	His/Arg	polar básico/polar básico	CAC/CGC	rs41570792
4	14921388	TAC1	Val/Leu	apolar/apolar	GTG/TTG	rs378353665
4	14922541	TAC1	Ile/Leu	apolar/apolar	ATT/CTT	rs211390627
7	13786612	KLF1	Ser/Gly	polar neutro/apolar	AGC/GGC	-
7	13786627	KLF1	Ala/Pro	apolar/apolar	GCA/CCA	-
7	13786637	KLF1	Arg/Thr	polar básico/polar neutro	AGG/ACG	-
7	13786644	KLF1	Leu/Pro	apolar/apolar	CTG/CCG	-
7	13786663	KLF1	Pro/Ala	apolar/apolar	CCG/GCG	-
7	13786679	KLF1	Ser/Thr	polar neutro/polar neutro	AGC/ACC	-
7	13788755	KLF1	Ala/Pro	apolar/apolar	GCC/CCC	-
7	13788756	KLF1	Ala/Asp	apolar/polar ácido	GCC/GAC	-
7	13788762	KLF1	Thr/Lys	polar neutro/polar básico	ACG/AAG	-
7	13788769	KLF1	Asp/Glu	polar ácido/polar ácido	GAC/GAA	-
7	13788782	KLF1	Arg/Gly	polar básico/apolar	CGG/GGG	rs456791503
7	13788783	KLF1	Arg/Gln	polar básico/polar neutro	CGG/CAG	-
10	93571223	TSHR	Gln/Arg	polar neutro/polar básico	CAG/CGG	rs207552184
13	48446641	CHGB	Asn/Asp	polar neutro/polar ácido	AAT/GAT	rs110261417
13	48446975	CHGB	Thr/Met	polar neutra/apolar	ACG/ATG	rs109400236
13	48447215	CHGB	His/Arg	apolar/polar básico	CAC/CGC	rs110013334
13	48447590	CHGB	Pro/Arg	apolar/polar básico	CCG/CGG	rs111002720
13	48447685	CHGB	Arg/Gly	polar básico/apolar	CGT/GGT	rs380061510
13	48448222	CHGB	Met/Val	apolar/apolar	ATG/GTG	rs110931509
17	11830384	POU4F2	Ile/Met	apolar/apolar	ATC/ATG	-
17	11830386	POU4F2	Leu/Phe	apolar/apolar	CTC/TTC	rs475409192
17	11830478	POU4F2	Ser/Ile	polar neutro/apolar	AGC/ATC	-
17	11830492	POU4F2	His/Gln	polar básico/polar neutro	CAC/CAG	-
19	31767321	MYOCD	Ile/Ser	apolar/polar neutro	ATT/AGT	-
21	58170306	CHGA	Thr/Asn	polar neutro/polar neutro	ACC/AAC	rs109042853
21	58170321	CHGA	Pro/Leu	apolar/apolar	CCG/CTG	rs520684424
21	58172131	CHGA	Ser/Phe	polar neutro/apolar	TCT/TTT	rs109145556
21	58172484	CHGA	Pro/Ala	apolar/apolar	CCC/GCC	rs517883631

Tabela 4. Mutações não-sinônimas detectadas nos genes candidatos associados com a puberdade na raça Brangus

(continuação)

CHR	Posição	Gene	Mudança de Aminoácido	Mudanças nas Propriedades Bioquímica do aa	Codon	Variante Existente
22	15195074	CCK	Met/Val	apolar/apolar	ATG/GTG	rs42891945
24	5619434	CBLN2	Val/Met	apolar/apolar	GTG/ATG	rs42038885
26	6855119	DKK1	Thr/Ser	polar neutro/polar neutro	ACT/AGT	rs210866503
X	20142496	VGLL1	Cys/Phe	apolar/apolar	TGT/TTT	-
X	20142501	VGLL1	Glu/Lys	polar ácido/polar básico	GAA/AAA	-

A identificação de SNP em amostras da raça Angus revelaram 2.742 SNP em 38 genes. Os genes que tiveram menos SNP foram *PCSK1N* e *Penk*, ambos com um SNP. O gene com o maior número de polimorfismos identificados foi o *NRXN3*, com 441 SNPs. A média de SNP e desvio padrão por gene foram $72,16 \pm 16,33$. Os genes em que não foram identificados nenhum SNP em Angus foram: *CARTPT*, *CBLN1*, *CCK*, *CDX2*, *DACH2*, *DKK1*, *E2F3*, *FOXB2*, *FOXE1*, *GATA1*, *IL22RA1*, *ITIH1*, *LHX4*, *LYSB*, *MEPE*, *MOS*, *NHLH1*, *OVGP1*, *PITX2*, *POU4F2*, *PROP1*, *SIX6*, *SST* e *TAC1*.

Em amostras provenientes de Brahman, a identificação de SNP revelou 4.316 SNP em 55 genes. A média o desvio padrão de SNP por gene foram $78,44 \pm 14,79$. Os genes que tinham menos SNP foram *CCK* e *GATA1*, ambos com um SNP. O gene com mais polimorfismos identificados foi o *MMD2* com 467 SNP. Os genes em que não foram detectados quaisquer SNP foram: *CDX2*, *DPPA4*, *E2F3*, *FOXB2*, *FOXE1*, *MOS*, e *NHLH1*.

A análise realizada a partir de amostras de Nelore revelou 1.027 SNPs em 31 genes. A média e desvio padrão de SNP por gene foram $33,13 \pm 7,49$. O gene que tinham menos SNP foi o *Penk* com apenas uma mutação detectada. A quantidade máxima de SNPs foi de 213 no gene *NRXN3*. Os genes em que não foram identificados SNP foram: *CARTPT*, *CBLN1*, *CBLN2*, *CCK*, *CDX2*, *CHGA*, *CHGB*, *DACH2*, *DKK1*, *DPPA4*, *FOXB2*, *FOXE1*, *GATA1*, *IL22RA1*, *INHA*, *ITIH1*, *LHX4*, *LYSB*, *MEPE*, *MOS*, *NHLH1*, *PCSK1N*, *POU4F2*, *PROP1*, *SCG2*, *SIX6*, *SST*, *STMN1*, *TAC1*, *TP63* e *VGLL1*.

A última análise de descoberta de SNP foi em amostras da raça Holandesa. A análise revelou 728 SNP em 31 genes com média e desvio padrão de $23,48 \pm 5,91$. Quatro genes tiveram apenas um SNP identificado, os genes foram: *INHA*, *SCG3*, *SST* e *STMN1*. No gene *NRXN3* foi detectado a maior quantidade de SNP, 171. Os genes que não tiveram SNP identificados foram: *CARTPT*, *CBLN1*, *CBLN2*, *CBLN4*, *CCK*, *CDX2*, *CHGA*, *CHGB*, *DACH2*, *DKK1*, *DPPA4*, *FOXB2*, *GATA1*, *IL22RA1*, *ITIH1*, *LHX4*, *LYSB*, *MEPE*, *MOS*, *NHLH1*, *NTS*, *PCSK1N*, *Penk*, *PITX2*, *POU4F2*, *PROP1*, *RHEBL1*, *SCG2*, *SIX6*, *TAC1* e *VGLL1*.

Somente os estudos delineados para as raças Brangus e Brahman tinham como o objetivo estudar diferenças entre animais Pré e Pós-púbere. Estas são as únicas raças que tiveram RNA extraído de tecidos relacionados com a reprodução. Possivelmente os tecidos de onde o RNA foi extraído influenciaram a quantidade de SNP detectados, sendo que alguns genes candidatos são tecidos específicos, não sendo observados na outras raças.

Depois de realizada a identificação de SNPs em cada raça foi feita a comparação dos SNPs encontrados. Cento e setenta e dois SNPs foram comuns em todas as raças, como mostrado no diagrama de Venn (Figura 3). Estes SNPs estão localizados em 13 genes, que incluem: *CPNE5*, *FAM19A4*, *FOXN4*, *KLF1*, *LOC777593*, *MGC157266*, *NEBL*, *NRXN3*, *PEPT-1*, *PPP3CA*, *SCG5*, *TSG101* e *TSHR*.

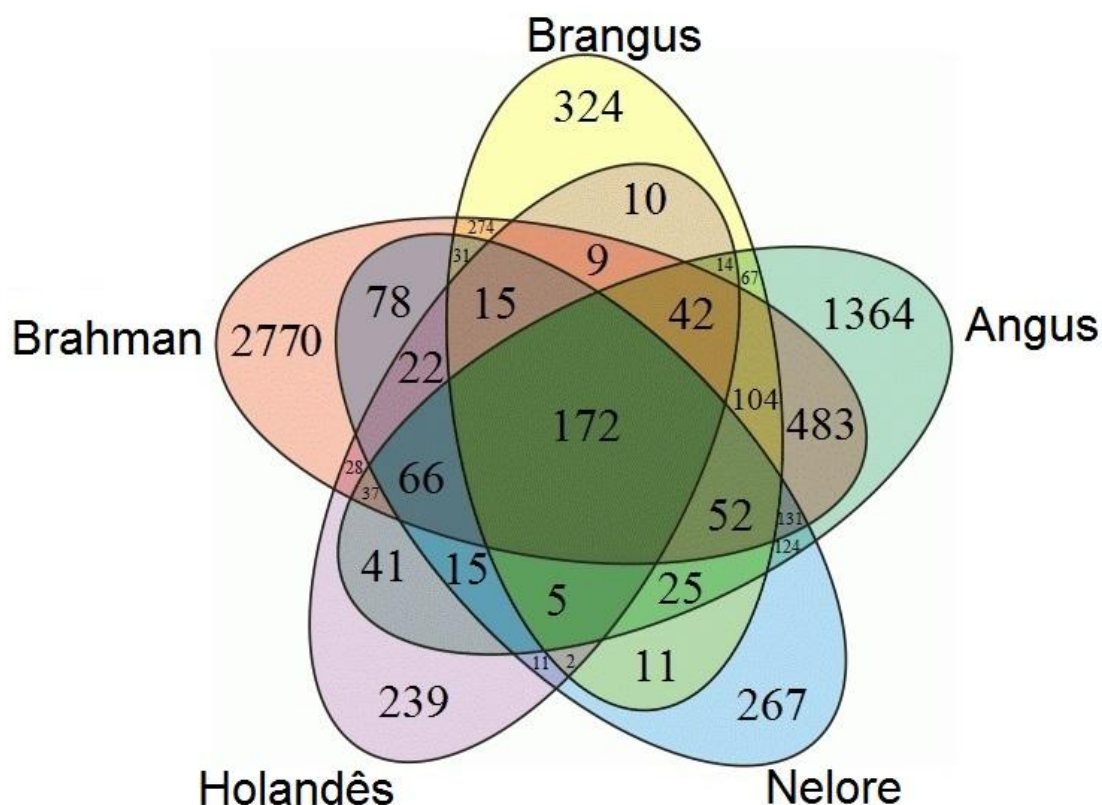


Figura 3: Diagrama de Venn comparando SNPs descobertos em Brangus, Angus, Brahman, Nelore e Holstein.

O gene *CPNE5* é da família das copinas, e podem estar envolvidos com o transporte pela membrana devido as suas propriedades de ligação com os fosfolipídios (DAMER et al., 2007). Este gene também pode mediar processos celulares, pois conferem regulação de cálcio para várias vias de sinalização (TOMSIG; SNYDER; CREUTZ, 2003). Wang et al (2015) relataram que este gene influencia a obesidade em humanos. Outros genes que também estão associados com a obesidade são o *NRXN3* (HEARD-COSTA et al., 2009) e o *PEPT-1* (ZANNI et al., 2015). A importância desta associação é devido a obesidade ter impacto no desenvolvimento da puberdade como observado em humanos (SOLORZANO; MCCARTNEY, 2010).

O *FOXN4* codifica uma proteína da superfamília *Forkhead box* (Fox) que são evolutivamente conservadas e atuam como reguladores de transcrição e, controlando um amplo espectro de processos biológicos (MYATT; LAM, 2007). Os

Fatores de transcrição desta superfamília atuam como importantes reguladores no desenvolvimento, homeostase e reprodução (THACKRAY, 2014).

Para a síntese de hormônios na glândula da tireoide, é necessário a ligação dos hormônios estimulantes da tireoide (TSH) ao seu receptor, o TSHR (ROCHA et al., 2007). Além disso o TSHR controla produção endócrina nesta glândula (BONGER et al., 2009). Mutinati et al. (2010) detectaram, por imunohistoquímica, a presença de TSHR em células maduras do corpo lúteo de bovinos, sugerindo uma ação dessa proteína no ovário.

O gene *PPP3CA* está associado a via metabólica de sinalização do MAPK (SANGIOVANNI; LEE, 2013). Além disso, há relatos de que quando este gene é bloqueado ocorre infertilidade em ratos machos (MARTIN et al., 2007; MIYATA et al., 2015). Até o momento não houve nenhuma associação deste gene com características reprodutivas em fêmeas.

Dentre os 13 genes que possuem SNPs coincidentes em todas as raças, os genes *LOC777593*, *MGC157266*, *NEBL*, *KLF1* e *TSG101* apenas são associados à puberdade no estudo de expressão gênica realizado por Cánovas et al. (2014c). Da mesma forma o gene *SCG5* somente teve associação com a puberdade no estudo realizado por DeAtley (2012).

Estes resultados podem ser úteis para a elaboração de um painel de SNP de baixa densidade para a puberdade para todas as raças utilizadas neste estudo. O painel de baixa densidade pode ser uma abordagem para reduzir os custos para a seleção baseado em genótipo (HABIER; FERNANDO; DEKKERS, 2009). Uma estratégia para selecionar SNPs para projetar um painel de baixa densidade é utilizar um conjunto de polimorfismos biologicamente relevantes (FORTES et al., 2014). Os SNP detectados neste estudo estão em exon, 3'UTR ou 5'UTR, uma vez que usamos sequências a partir de RNA-Seq. Com isso, esses SNP podem afetar a expressão da proteína sintetizada no gene no qual esses polimorfismos foram identificados. Painéis de baixa densidade que incluem mutação potencialmente causal possuem vantagem sobre aqueles que incluem apenas marcadores, uma vez que este tipo de mutação não é dependente do LD, e também podem ser utilizados para a seleção ao longo de gerações e em várias raças (FORTES et al., 2014).

A mesma ferramenta do Ensembl, VEP, foi usada com os 172 SNPs, coincidentes entre as raças (Figura 4). Os resultados do VEP mostraram que 138 (80,2%) SNPs dos 172 não foram previamente reportados e 34 (19,8%) já são conhecidos. Esta análise também mostrou que esses SNPs estão mapeados na região de 16 genes. Esta ferramenta indicou que 5 SNP que estão na região gene *LOC777593*, também estão localizados a montante do gene *SPINK7*, sendo que este gene não foi considerado como candidato neste estudo. Todos os SNP mapeados na região do gene *KLF1* também estão localizados a montante do gene *DNASE1*, e 4 SNPs também estão localizados a jusante do gene *MAST1*, sendo que estes dois últimos genes não foram considerados como candidatos nesse estudo. Novamente o resultado mostrou que 77% dos 172 SNPs estão localizados em região de intron. É importante ressaltar que estes resultados também fazem parte dos resultados descritos na Figura 3, uma vez que os 172 SNPs analisados aqui são um subconjunto dos 1157 SNPs.

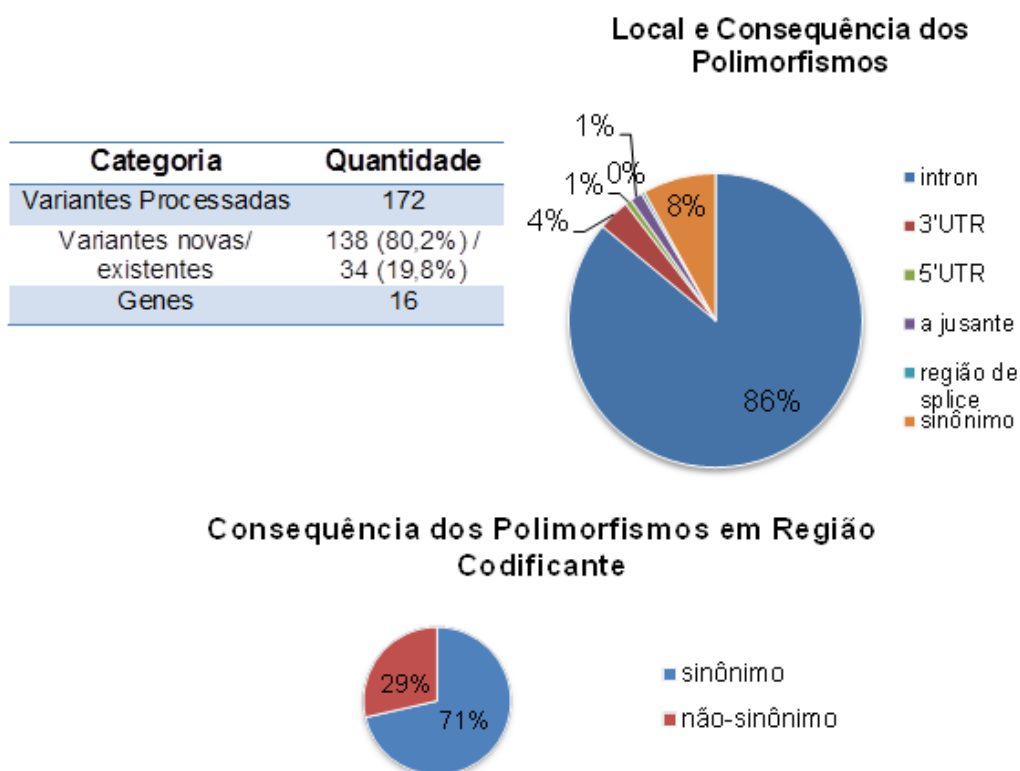


Figura 4. Sumário dos resultados estatísticos provenientes da ferramenta "Variant Effect Predict" (VEP) do Ensembl usando os 172 SNPs que coincidentes em toda as raças.

Os polimorfismos que, segundo a ferramenta VEP, se encontram em região de exon são 22, sendo que 16 são mutação sinônimos e 6 são mutações não sinônimas. Dentre as mutações sinônimas somente uma já é reportada no SNPdatabase e todas as outras, inclusive as mutação não sinônimas, eram ainda não conhecidas. Na Tabela 5, encontram-se todas as informações referentes a esses polimorfismos. Vale ressaltar que essas informações também estão na Tabela 4.

Tabela 5. Mutações não-sinônimas detectadas no gene candidato, KLF1, associados com a puberdade na raça Brangus

CHR	Posição	Mudança de Aminoácido	Mudança da Estrutura Bioquímica	Codons	Variante Existente
7	13786627	Ala/Pro	apolar / apolar	GCA/CCA	-
7	13788769	Asp/Glu	polar ácido / polar ácido	GAC/GAA	-
7	13786663	Pro/Ala	apolar / apolar	CCG/GCG	-
7	13786637	Arg/Thr	polar básico / polar neutra	AGG/ACG	-
7	13786612	Ser/Gly	polar neutra / apolar	AGC/GGC	-
7	13786679	Ser/Thr	polar neutra / polar neutra	AGC/ACC	-

3.3.4. Descoberta de novos Transcritos e Gene

Com a análise "*Transcript Discovery*" do software CLC Genomics Workbench 8.0.1 (CLC Bio, Aarhus, Dinamarca) foi possível identificar 160 novos transcritos em 42 dos 62 genes candidatos (Tabela 6). Nos genes *CBLN1*, *CDX2*, *DKK1*, *DPPA4*, *ELF5*, *FABP4*, *FOXB2*, *FOXE1*, *GATA1*, *INHA*, *KLF1*, *LYSB*, *NHLH1*, *NTS*, *PCSK1N*, *Penk*, *PEPT-1*, *POU4F2*, *SYCE1* e *VAX2* não foram revelados novos transcritos. Não houve alteração na localização dos SNPs mostrada nas análises realizadas pela ferramenta VEP. Assim, são necessários mais estudos a fim de verificar o mapeamento dos SNP identificados a partir de sequências de RNA e, desse modo, enriquecer a anotação do genoma referência bovino.

Tabela 6. Descoberta de transcritos em genes candidatos associados à puberdade em bovinos *Bos-indicus* influenciado.

(continua)

Gene	CHR	Número de Transcripto	Transcritos Conhecidos	Transcritos Desconhecidos
SST	1	2	1	1
TP63	1	2	1	1
STMN1	2	2	1	1
IL22RA1	2	4	1	3
SCG2	2	7	1	6
OVGP1	3	4	1	3
TAC1	4	8	3	5
RHEBL1	5	4	1	3
ADH6	6	2	1	1
MEPE	6	2	1	1
TECRL	6	2	1	1
PITX2	6	11	1	10
PPP3CA	6	8	2	6
LOC777593	7	3	1	2
PROP1	7	3	1	2
MGC157266	8	3	1	2
NRXN3	10	4	3	1
C10H11ORF46	10	3	1	2
SIX6	10	4	1	3
TSHR	10	3	1	2
SCG3	10	6	1	5
SCG5	10	23	1	22
POMC	11	10	1	9
CBLN4	13	4	1	3
CHGB	13	4	1	3
NEBL	13	15	8	7
MOS	14	2	1	1
LHX4	16	5	1	4
FOXN4	17	2	1	1
MYOCD	19	9	1	8
CARTPT	20	3	1	2
CHGA	21	2	1	1
ITIH1	22	2	1	1
CCK	22	3	1	2
FAM19A4	22	8	2	6
CPNE5	23	6	1	5
E2F3	23	7	1	6

Tabela 6. Descoberta de transcritos em genes candidatos associados à puberdade em bovinos *Bos-indicus* influenciado.

(continuação)

Gene	CHR	Número de Transcripto	Transcritos Conhecidos	Transcritos Desconhecidos
CBLN2	24	12	1	11
MMD2	25	3	1	2
TSG101	29	2	1	1
DACH2	X	2	1	1
VGLL1	X	4	1	3

Na mesma análise foram descobertos cinco novos genes que sobrepõem a região de cinco genes considerados candidatos nesse estudo (Tabela 7). Os polimorfismos identificados nos genes candidatos também podem influenciar a expressão dos genes que se encontram na outra fita do DNA. A função desses genes não é conhecida ou se estes genes afetam a puberdade em bovinos. Se existem promotores transcrpcionais em ambas as fitas, ambos os genes irão sintetizar RNA. Os genes sobrepostos não podem ser transcrito com a mesma polimerase já que esta é uni direcional, mas os genes podem compartilhar a ilha de CpG para fins de expressão coordenadas (NAKAYAMA et al., 2007). Os genes sobrepostos podem interagir com relação à regulação da expressão do gene ao nível da transcrição, processamento de RNAm, splicing e na estabilidade, ou a tradução (BOI; SOLDA; TENCHINI, 2004).

Tabela 7. Descoberta de genes sobrepostos com genes candidatos que estão associados com a puberdade em bovinos *Bos indicus* influenciados.

Gene	Crom	Desconhecido	Início	Fim	Fita
LHX4	16	No	62868725	62916031	+
Gene 3295	16	Yes	62891230	62958617	-
FAM19A4	22	No	32709906	32908710	+
Gene 4287	22	Yes	32886396	32901328	-
MEPE	6	No	38279966	38291351	-
Gene 4697	6	Yes	38289787	38338976	+
LOC777593	7	No	61639308	61734503	-
Gene 3471	7	Yes	61682521	61709568	+
Gene 1878	8	Yes	53302486	53346814	-
FOXB2	8	No	53346672	53348693	+

3.5. Conclusões

A partir dos resultados discutidos anteriormente podemos concluir que as sequência de RNA podem ser utilizadas para a identificação de polimorfismos. A comparação dos polimorfismos detectados é interessante para melhor compreensão da diversidade entre as raças. Além disso, a detecção de splicing alternativo e genes sobrepostos são importantes para o enriquecimento do genoma bovino.

3.6. Referências Bibliográficas

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. S. Characteristics of linkage disequilibrium in North American Holsteins. **BMC genomics**, v. 11, n. 1, p. 421, 2010.

BOI, S.; SOLDA, G.; TENCHINI, M. L. Shedding Light on the Dark Side of the Genome: Overlapping Genes in Higher Eukaryotes. **Current Genomics**, v. 5, n. 6, p. 509–524, 2004.

BONGER, K. M.; VAN DEN BERG, R. J. B. H. N.; KNIJNENBURG, A. D.; HEITMAN, L. H.; VAN KOPPEN, C. J.; TIMMERS, C. M.; OVERKLEEF, H. S.; VAN DER MAREL, G. A. Discovery of selective luteinizing hormone receptor agonists using the bivalent ligand method. **ChemMedChem**, v. 4, n. 7, p. 1189–1195, 2009.

BRUMATTI, R. C.; FERRAZ, J. B. S.; ELER, J. P.; FORMIGONNI, E. I. B. Desenvolvimento de índice de seleção em gado corte sob o enfoque de um modelo bioeconômico. **Archivos de Zootecnia**, v. 60, n. 230, p. 205–213, 2011.

CÁNOVAS, A.; COCKRUM, R.; BROWN, D.; RIDDLE, S.; NEARY, J. M.; HOLT, T.;

KRAFSUR, G.; GARRY, F.; MEDRANO, J. F.; ISLAS-TREJO, A.; ENNS, R. M.; SPEIDEL, S. E.; STENMARK, K.; THOMAS, M. Transcriptomics and Systems Biology of High Altitude Disease in Angus Cattle. In: Proc. XXIII. Plant Anim. Gen., **Anais...**2015.

CÁNOVAS, A.; ISLAS-TREJO, A.; STENMARK, K.; BROWN, D.; THOMAS, M.; COCKRUM, R.; MEDRANO, J. F. Transcriptomics of High Altitude Disease in Holstein Calves. In: Proc. XXII. Plant Anim. Gen, San Diego, CA. **Anais...** San Diego, CA: 2014a.

CÁNOVAS, A.; ISLAS-TREJO, A.; STENMARK, K.; BROWN, D.; THOMAS, M.; COCKRUM, R.; MEDRANO, J. F. Transcriptomics of High Altitude Disease. In: Proc. XXII. Plant Anim. Gen, San Diego, CA. **Anais...** San Diego, CA: 2014b.

CÁNOVAS, A.; REVERTER, A.; DEATLEY, K. L.; ASHLEY, R. L.; COLGRAVE, M. L.; FORTES, M. R. S.; ISLAS-TREJO, A.; LEHNERT, S.; PORTO-NETO, L.; RINCÓN, G.; SILVER, G. A.; SNELLING, W. M.; MEDRANO, J. F.; THOMAS, M. G. Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. **PLoS ONE**, v. 9, n. 7, p. 1–17, 2014c.

CÁNOVAS, A.; RINCÓN, G.; BEVILACQUA, C.; ISLAS-TREJO, A.; BRENAUT, P.; HOVEY, R. C.; BOUTINAUD, M.; MORGENTHALER, C.; VANKLOMPENBERG, M. K.; MARTIN, P.; MEDRANO, J. F. Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. **Scientific reports**, v. 4, p. 5297, 2014d.

CÁNOVAS, A.; RINCON, G.; ISLAS-TREJO, A.; WICKRAMASINGHE, S.; MEDRANO, J. F. SNP discovery in the bovine milk transcriptome using RNA-Seq technology. **Mammalian Genome**, v. 21, n. 11-12, p. 592–598, 2010.

CUNNINGHAM, F.; AMODE, M. R.; BARRELL, D.; BEAL, K.; BILLIS, K.; BRENT, S.; CARVALHO-SILVA, D.; CLAPHAM, P.; COATES, G.; FITZGERALD, S.; GIL, L.; GIRÓN, C. G.; GORDON, L.; HOURLIER, T.; HUNT, S. E.; JANACEK, S. H.; JOHNSON, N.; JUETTEMANN, T.; KÄHÄRI, A. K.; KEENAN, S.; MARTIN, F. J.; MAUREL, T.; MCLAREN, W.; MURPHY, D. N.; NAG, R.; OVERDUIN, B.; PARKER, A.; PATRICIO, M.; PERRY, E.; PIGNATELLI, M.; RIAT, H. S.; SHEPPARD, D.; TAYLOR, K.; THORMANN, A.; VULLO, A.; WILDER, S. P.; ZADISSA, A.; AKEN, B. L.; BIRNEY, E.; HARROW, J.; KINSELLA, R.; MUFFATO, M.; RUFFIER, M.; SEARLE, S. M. J.; SPUDICH, G.; TREVANION, S. J.; YATES, A.; ZERBINO, D. R.; FLICEK, P. Ensembl 2015. **Nucleic Acids Research**, v. 43, n. D1, p. D662–D669, 2015.

DAMER, C. K.; BAYEVA, M.; KIM, P. S.; HO, L. K.; EBERHARDT, E. S.; SOCEC, C. I.; LEE, J. S.; BRUCE, E. A.; GOLDMAN-YASSEN, A. E.; NALIBOFF, L. C. Copine A is required for cytokinesis, contractile vacuole function, and development in Dictyostelium. **Eukaryotic Cell**, v. 6, n. 3, p. 430–442, 2007.

DAY, M. L.; NOGUEIRA, G. P. Management of age at puberty in beef heifers to optimize efficiency of beef production. **Animal Frontiers**, v. 3, p. 6–11, 2013.

DEATLEY, K. L. **Neuropeptidome of hypothalamus and pituitary gland of pre-**

and post-pubertal brangus heifers. 2012. 2012.

DIAS, M. M.; SOUZA, F. R. P.; TAKADA, L.; FEITOSA, F. L. B.; COSTA, R. B.; DIAZ, I. D. P. S.; CARDOSO, D. F.; TONUSSI, R. L.; BALDI, F.; ALBUQUERQUE, L. G.; OLIVEIRA, H. N. Study of lipid metabolism-related genes as candidate genes of sexual precocity in nellore cattle. **Genetics and Molecular Research**, 2015.

DICKERSON, G. Efficiency of animal production - molding the biological components. 1970.

DORN, L. D.; BIRO, F. M. Puberty and its measurement: A decade in review. **Journal of Research on Adolescence**, v. 21, n. 1, p. 180–195, 2011.

ELSIK, C. G.; TELLAM, R. L.; WORLEY, K. C. The Genome Sequence of Taurine Cattle: A window to ruminant Biology and Evolution. **Science**, v. 324, p. 522–527, 2009.

FORTES, M. R.; PORTO-NET, L. R.; DEATLE, K. L.; REVERTE, A.; THOMA, M. G.; MOOR, S. S.; LEHNER, S. A.; SNELLIN, W. M. Genetic Markers in Transcription Factors of Differentially Expressed Genes Associated with Post-partum Anoestrus Predict Pregnancy Outcome in an Independent Population of Beef Cattle. In: Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, Vancouver, Canada. **Anais...** Vancouver, Canada: 2014.

FORTES, M. R. S.; SNELLING, W. M.; REVERTER, A.; NAGARAJ, S. H.; LEHNERT, S. A.; HAWKEN, R. J. Gene network analyses of first service conception in Brangus heifers : Use of genome and trait associations , hypothalamic-transcriptome information , and transcription factors 1. p. 2894–2906, 2012.

HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. Genomic selection using low-density marker panels. **Genetics**, v. 182, n. 1, p. 343–353, 2009.

HEARD-COSTA, N. L.; ZILLIKENS, M. C.; MONDA, K. L.; JOHANSSON, A.; HARRIS, T. B.; FU, M.; HARITUNIAN, T.; FEITOSA, M. F.; ASPELUND, T.; EIRIKSDOTTIR, G.; GARCIA, M.; LAUNER, L. J.; SMITH, A. V.; MITCHELL, B. D.; MCARDLE, P. F.; SHULDINER, A. R.; BIELINSKI, S. J.; BOERWINKLE, E.; BRANCATI, F.; DEMERATH, E. W.; PANKOW, J. S.; ARNOLD, A. M.; CHEN, Y.-D. I.; GLAZER, N. L.; MCKNIGHT, B.; PSATY, B. M.; ROTTER, J. I.; AMIN, N.; CAMPBELL, H.; GYLLENSTEN, U.; PATTARO, C.; PRAMSTALLER, P. P.; RUDAN, I.; STRUCHALIN, M.; VITART, V.; GAO, X.; KRAJA, A.; PROVINCE, M. A.; ZHANG, Q.; ATWOOD, L. D.; DUPUIS, J.; HIRSCHHORN, J. N.; JAQUISH, C. E.; O'DONNELL, C. J.; VASAN, R. S.; WHITE, C. C.; AULCHENKO, Y. S.; ESTRADA, K.; HOFMAN, A.; RIVADENEIRA, F.; UITTERLINDEN, A. G.; WITTEMAN, J. C. M.; OOSTRA, B. A.; KAPLAN, R. C.; GUDNASON, V.; O'CONNELL, J. R.; BORECKI, I. B.; VAN DUIJN, C. M.; CUPPLES, L. A.; FOX, C. S.; NORTH, K. E. NR3X1 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. **PLoS genetics**, v. 5, n. 6, p. e1000539, 2009.

JIN, Y.; LEE, C. Single Nucleotide Polymorphisms Associated with MicroRNA Regulation. **Biomolecules**, v. 3, n. 2, p. 287–302, 2013.

KUERSTEN, S.; GOODWIN, E. B. The power of the 3' UTR: translational control and development. **Nature Reviews Genetics**, v. 4, p. 626–637, 2003.

MARTIN, L. J.; CHEN, H.; LIAO, X.; ALLAYEE, H.; SHIH, D. M.; LEE, G. S.; HOVLAND, D. N.; ROBBINS, W. A.; CARNES, K.; HESS, R. A.; LUSIS, A. J.; COLLINS, M. D. FK506, a calcineurin inhibitor, prevents cadmium-induced testicular toxicity in mice. **Toxicological Sciences**, v. 100, n. 2, p. 474–485, 2007.

MATOULKOVA, E.; MICHALOVA, E.; VOJTESEK, B.; HRSTKA, R. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. **RNA biology**, v. 9, n. 5, p. 563–76, 2012.

MIYATA, H.; SATOUH, Y.; MASHIKO, D.; MUTO, M.; NOZAWA, K.; SHIBA, K.; FUJIHARA, Y.; ISOTANI, A.; INABA, K.; IKAWA, M. Sperm calcineurin inhibition prevents mouse fertility with implications for male contraceptive. **Science**, v. 350, n. 6259, p. 442–446, 2015.

MUTINATI, M.; DESANTIS, S.; RIZZO, A.; ZIZZA, S.; VENTRIGLIA, G.; PANTALEO, M.; SCIORSCI, R. L. Localization of thyrotropin receptor and thyroglobulin in the bovine corpus luteum. **Animal Reproduction Science**, v. 118, n. 1, p. 1–6, 2010.

MYATT, S. S.; LAM, E. W.-F. The emerging roles of forkhead box (Fox) proteins in cancer. **Nature Reviews Cancer**, v. 7, p. 847–859, 2007.

NAKAYAMA, T.; ASAI, S.; TAKAHASHI, Y.; MAEKAWA, O.; KASAMA, Y. Overlapping of genes in the human genome. **International Journal Of Biomedical Science: IJBS**, v. 3, n. 1, p. 14–19, 2007.

ROCHA, A.; GOMEZ, A.; GALAY-BURGOS, M.; ZANUY, S.; SWEENEY, G. E.; CARRILLO, M. Molecular characterization and seasonal changes in gonadal expression of a thyrotropin receptor in the European sea bass. **General and Comparative Endocrinology**, v. 152, n. 1, p. 89–101, 2007.

SANGIOVANNI, J. P.; LEE, P. H. AMD-Associated Genes Encoding Stress-Activated MAPK Pathway Constituents Are Identified by Interval-Based Enrichment Analysis. **PLoS ONE**, v. 8, n. 8, p. 1–8, 2013.

SNELLING, W.; CUSHMAN, R.; KEELE, J.; MALTECCA, C.; THOMAS, M.; FORTES, M. R. S.; REVERTE, A. Networks and pathways to guide genomic selection. **Journal of animal ...**, v. 91, p. 537–552, 2013.

SOLORZANO, C. M. B.; MCCARTNEY, C. R. Obesity and the pubertal transition in girls and boys. **Reproduction**, v. 140, n. 3, p. 399–410, 2010.

THACKRAY, V. G. Fox tales: Regulation of gonadotropin gene expression by forkhead transcription factors. **Molecular and Cellular Endocrinology**, v. 385, n. 1–2, p. 62–70, 2014.

TOMSIG, J. L.; SNYDER, S. L.; CREUTZ, C. E. Identification of targets for calcium signaling through the copine family of proteins. Characterization of a coiled-coil copine-binding motif. **Journal of Biological Chemistry**, v. 278, n. 12, p. 10048–

10054, 2003.

VINSKY, M.; ISLAM, K.; CHEN, L.; LI, C. Communication: Association analyses of a single nucleotide polymorphism in the promoter of OLR1 with growth, feed efficiency, fat deposition, and carcass merit traits. **Canadian Journal of Animal Science**, v. 93, p. 193–197, 2013.

WANG, K.; ZUO, L.; PAN, Y.; XIE, C.; LUO, X. Genetic variants in the CPNE5 gene are associated with alcohol dependence and obesity in Caucasian populations. **Journal of Psychiatric Research**, v. 71, p. 1–7, 2015.

ZANNI, E.; LAUDENZI, C.; SCHIFANO, E.; PALLESCHI, C.; PEROZZI, G.; UCCELLETTI, D.; DEVIRGILIIS, C. Impact of a complex food microbiota on energy metabolism in the model organism *caenorhabditis elegans*. **BioMed Research International**, v. 2015, 2015.

5. Apêndices

Tabela A 1. Informações sobre vinte e cinco genes que abrigam SNP associado com a fertilidade de novilhas e com expressão encontrada para ser diferencialmente expressos (DE) ou específico de tecido (TS) na análise de RNA-Seq. (Canovas et al., 2014c).

Gene	Gene ID (Ensembl)	Locus			Tamanho (bp)	Quantidade de Splicing Anotados
		Chr	Localização			
			Início	Fim		
DPPA4	ENSBTAG00000038103	1	54509507	54516574	7067	3
TP63	ENSBTAG00000015460	1	78012414	78283420	271006	1
INHA	ENSBTAG00000009972	2	108226858	108229830	2972	1
IL22RA1	ENSBTAG00000001100	2	129351698	129378160	26462	1
RHEBL1	ENSBTAG00000006557	5	30918838	30922282	3444	1
LYSB	ENSBTAG00000026323	5	44554333	44559920	5587	1
ADH6	ENSBTAG00000006991	6	26786780	26802348	15568	1
MEPE	ENSBTAG00000007568	6	38279966	38291339	11373	1
TECRL	ENSBTAG00000024826	6	81511554	81653990	142436	1
LOC777593	ENSBTAG00000033759	7	61639308	61734503	95195	1
MGC157266	ENSBTAG00000040034	8	62935793	62978422	42629	1
C10H11ORF46	ENSBTAG00000039618	10	4872818	4904741	31923	1
NRXN3	ENSBTAG00000025324	10	91597994	92223876	625882	3
TSHR	ENSBTAG00000017489	10	93417338	93571380	154042	1
NEBL	ENSBTAG00000013339	13	22382763	22496405	113642	1
MOS	ENSBTAG00000019145	14	24975950	24976948	998	1
PENK	ENSBTAG00000004924	14	25218586	25222991	4405	1
ELF5	ENSBTAG00000012460	15	65824442	65854386	29944	1
POU4F2	ENSBTAG00000005648	17	11830252	11831926	1674	1
FAM19A4	ENSBTAG00000003110	22	32732392	32877802	145410	1
ITIH1	ENSBTAG00000007843	22	48655027	48669271	14244	1
CPNE5	ENSBTAG00000007514	23	10632882	10734550	101668	1
MMD2	ENSBTAG00000021913	25	39609810	39672867	63057	1
DKK1	ENSBTAG00000014218	26	6852970	6855647	2677	1
SYCE1	ENSBTAG00000035072	26	50515052	50527036	11984	1

Ensembl, 2016

Tabela A 2. Informações de genes provenientes de vinte fatores de transcrição mais relevantes em novilhas Brangus Pré e Pós Púbere (Cánovas et al., 2014c).

Gene	Gene ID (Ensembl)	Locus			Tamanho (bp)	Quantidade de Splicing Anotados
		Chr	Localização			
			Início	Fim		
<i>OVGP1</i>	ENSBTAG000000005973	3	32047399	32059347	11948	1
<i>NHLH1</i>	ENSBTAG000000020027	3	9442357	9442758	401	1
<i>PITX2</i>	ENSBTAG000000009059	6	16000762	16006077	5315	1
<i>PROP1</i>	ENSBTAG000000007359	7	41205895	41209917	4022	1
<i>KLF1</i>	ENSBTAG000000006083	7	13785087	13789391	4304	1
<i>FOXE1</i>	ENSBTAG000000014760	8	63509720	63510841	1121	1
<i>FOXB2</i>	ENSBTAG000000032983	8	53347353	53348693	1340	1
<i>SIX6</i>	ENSBTAG000000013035	10	72953088	72955661	2573	1
<i>VAX2</i>	ENSBTAG000000010605	11	13492655	13520241	27586	1
<i>CDX2</i>	ENSBTAG000000001819	12	32317772	32322605	4833	1
<i>ELF5</i>	ENSBTAG000000012460	15	65824442	65854386	29944	1
<i>LHX4</i>	ENSBTAG000000003992	16	62868732	62915247	46515	1
<i>FOXN4</i>	ENSBTAG000000010847	17	66074400	66097412	23012	1
<i>POU4F2</i>	ENSBTAG000000005648	17	11830252	11831926	1674	1
<i>MYOCD</i>	ENSBTAG000000035706	19	31767278	31822920	55642	1
<i>E2F3</i>	ENSBTAG000000021045	23	37262021	37343620	81599	1
<i>TSG101</i>	ENSBTAG000000013563	29	26450457	26492046	41589	1
<i>DACH2</i>	ENSBTAG000000044037	X	72745652	72755858	10206	1
<i>VGLL1</i>	ENSBTAG000000001880	X	20126976	20143499	16523	1
<i>GATA1</i>	ENSBTAG000000003184	X	91982660	91985592	2932	1

Ensembl, 2016

Tabela A 3. Informações sobre genes provenientes de estudo sobre neuropeptídeos no hipotálamo e da hipófise de Pré e Pós Púbere em novilhas Brangus relatados por DeAtley (2012).

Gene	Gene ID (Ensembl)	Locus			Tamanho (bp)	Quantidade de Splicing Anotados
		Chr	Localização			
			Início	Fim		
<i>SST</i>	ENSBTAG00000017312	1	80250205	80251648	1443	1
<i>SCG2</i>	ENSBTAG00000021588	2	112549869	112555375	5506	1
<i>STMN1</i>	ENSBTAG00000013761	2	127773927	127779431	5504	1
<i>TAC 1</i>	ENSBTAG00000015356	4	14920664	14929428	8764	3
<i>NTS</i>	ENSBTAG00000005305	5	15609695	15621365	11670	1
<i>SCG5</i>	ENSBTAG000000032613	10	30077158	30135194	58036	1
<i>SGC3</i>	ENSBTAG00000017677	10	58748957	58787385	38428	1
<i>POMC</i>	ENSBTAG00000007897	11	74109410	74116893	7483	1
<i>PEPB-1</i>	ENSBTAG00000004440	12	79605255	79652128	46873	1
<i>CBLN4</i>	ENSBTAG00000017524	13	84024137	84031205	7068	1
<i>CHGB</i>	ENSBTAG00000011782	13	48435003	48449940	14937	1
<i>Penk-A</i>	ENSBTAG00000004924	14	25218586	25222991	4405	1
<i>Penk-B</i>	ENSBTAG00000004924	14	25218586	25222991	4405	1
<i>CBLN1</i>	ENSBTAG00000015194	18	17813467	17816071	2604	1
<i>CARTPT</i>	ENSBTAG00000047486	20	9791661	9793744	2083	1
<i>CHGA</i>	ENSBTAG00000009836	21	58163028	58176319	13291	1
<i>CCK</i>	ENSBTAG00000013027	22	15190718	15196325	5607	1
<i>CBLN2</i>	ENSBTAG00000005985	24	5614221	5621035	6814	1
<i>PCSK1N</i>	ENSBTAG00000013250	X	92018743	92023258	4515	1

Ensembl, 2016

Tabela A 4. Informações sobre genes provenientes de um estudo baseado em SNP de chips de alta densidade associada a puberdade em Nelore (Dias et al., 2015).

Gene	Gene ID (Ensembl)	Locus			Tamanho (bp)	Quantidade de Splicing Anotados
		Chr	Localização			
			Início	Fim		
FABP4	ENSBTAG00000037526	14	46833665	46838053	4388	1
PPP3CA	ENSBTAG00000016005	6	24812682	25136247	323565	2

Ensembl, 2016

Tabela A 5. Mutações relatadas no banco de dados de SNP (*SNP database*) extraídas do Ensembl (2016) existentes na região ou em torno de 25 genes que são tecido específico ou diferencialmente expressos que abrigam SNPs do GWAS relatados por Cánovas et al., (2014c).

Gene	Região						
	A montante	5'UTR	Exon		Intron	3'UTR	A jusante
			Não-sinônimo	sinônimo			
<i>DPPA4</i>	1295	0	289	140	1264	0	733
<i>TP63</i>	105	0	65	25	6998	72	201
<i>IL22RA1</i>	285	5	227	74	1320	88	235
<i>INHA</i>	763	4	161	61	160	8	164
<i>LYSB</i>	218	4	43	8	235	28	184
<i>RHEBL1</i>	151	0	26	12	221	1	83
<i>ADH6</i>	170	1	53	18	601	15	198
<i>MEPE</i>	112	0	25	13	294	0	186
<i>TECRL</i>	146	2	56	10	5316	67	322
<i>LOC777593</i>	131	11	14	2	3658	12	256
<i>FOXB2</i>	202	0	110	27	0	0	155
<i>FOXE1</i>	168	0	148	48	0	0	162
<i>MGC157266</i>	258	3	91	29	2422	36	194
<i>C10H11ORF46</i>	279	2	27	8	1405	32	147
<i>NRXN3</i>	1300	0	507	145	114354	33	852
<i>TSHR</i>	226	13	171	42	5803	0	145
<i>NEBL</i>	135	5	86	18	3772	0	185
<i>MOS</i>	89	0	143	36	0	0	256
<i>PENK</i>	173	42	85	31	195	21	227
<i>ELF5</i>	178	23	42	12	1173	138	254
<i>FOXN4</i>	196	0	372	90	1081	0	441
<i>POU4F2</i>	90	0	189	65	29	0	155
<i>MYOCD</i>	246	0	359	120	2130	0	474
<i>FAM19A4</i>	247	12	14	0	5533	0	234
<i>ITIH1</i>	131	6	188	56	1102	5	229
<i>CPNE5</i>	220	1	201	53	4633	0	473
<i>MMD2</i>	307	0	62	7	3656	9	184
<i>DKK1</i>	200	0	86	19	94	16	239
<i>SYCE1</i>	293	6	134	30	777	2	229
<i>GATA1</i>	132	0	189	52	164	14	86
Total	8446	140	4163	1251	168390	597	7883

Tabela A 6. Mutações relatadas no banco de dados de SNP (*SNP database*) extraídas do Ensembl (2016) existentes na região ou em torno dos vinte genes do fator de transcrição mais relevantes relatados por Cánovas et al., (2014c).

Genes	Região						
	A montate	5'UTR	Exon		Intron	3'UTR	A jusante
			Não-Sinônimo	Sinônimo			
<i>NHLH1</i>	294	0	36	21	0	0	272
<i>OVGP1</i>	221	34	88	33	779	23	260
<i>PITX2</i>	160	124	137	44	221	66	115
<i>KLF1</i>	177	6	143	37	143	32	722
<i>PROP1</i>	233	24	59	19	216	3	294
<i>FOXB2</i>	202	0	110	27	0	0	155
<i>FOXE1</i>	168	0	148	48	0	0	162
<i>SIX6</i>	138	43	90	19	93	115	112
<i>VAX2</i>	277	0	67	26	945	0	424
<i>CDX2</i>	121	0	49	14	208	0	257
<i>ELF5</i>	178	23	42	12	1173	138	254
<i>LHX4</i>	183	0	127	35	1989	0	176
<i>FOXN4</i>	196	0	372	90	1081	0	441
<i>POU4F2</i>	90	0	189	65	29	0	155
<i>MYOCD</i>	246	0	359	120	2130	0	474
<i>E2F3</i>	105	3	98	41	2684	1	413
<i>TSG101</i>	425	0	31	9	1369	1	101
<i>DACH2</i>	43	0	60	17	185	0	175
<i>GATA1</i>	132	0	189	52	164	14	86
<i>VGLL1</i>	169	20	58	17	327	0	158
Total	3758	277	2452	746	13736	393	5206

Tabela A 7. Mutações relatadas no banco de dados de SNP (*SNP database*) extraídas do Ensembl (2016) existentes na região ou em torno dos genes relacionadas com neuropeptídeos observados por DeAtley (2012).

Genes	Região						
	A montante	5'UTR	Exon		Intron	3'UTR	A jusante
			Não-sinônimo	Sinônimo			
<i>SST</i>	233	34	43	7	60	8	138
<i>SCG2</i>	120	6	48	26	123	29	228
<i>STMN1</i>	321		29	11			165
<i>TAC 1</i>	884	162	48	12	2376	204	918
<i>NTS</i>	141	2	17	5	356	27	141
<i>SCG3</i>	131	48	46	15	1285	47	184
<i>SCG5</i>	292		23	10	1978	11	158
<i>POMC</i>	223	21	244	78	368	29	383
<i>PEPT-1</i>	198		170	50	2931	56	326
<i>CBLN4</i>	173	6	41	9	184	59	310
<i>CHGB</i>	194	3	247	85	440	15	175
<i>PENK</i>	173	42	85	31	195	21	227
<i>CBLN1</i>	125		29	18	49	134	248
<i>CARTPT</i>	166	2	56	21	168	14	214
<i>CHGA</i>	203	44	221	45	691	123	344
<i>CCK</i>	154	3	48	18	226	3	227
<i>CBLN2</i>	235	19	109	36	327	209	204
<i>PCSK1N</i>	113	1	26	15	51	35	690
Total	4079	393	1530	492	11808	1024	5280

Tabela A 8. Mutações relatadas no banco de dados de SNP (*SNP database*) extraídas do Ensembl (2016) existentes na região ou em torno dos genes de um estudo baseado em SNP de chips de alta densidade associada a puberdade em Nelore (Dias et al., 2015).

Genes	Região						
	A montante	5'UTR	Exon		Intron	3'UTR	A jusante
			Não-Sinônimo	Sinônimo			
<i>FABP4</i>	146	1	26	5	181	25	270
<i>PPP3CA</i>	566	315	168	84	25318	300	924
Total	712	316	194	89	25499	325	1194